JMIR mHealth and uHealth

Impact Factor (2022): 5.0
Volume 8 (2020), Issue 12 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

Contents

Viewpoints

Reviews

Original Papers

Corrigenda and Addendas

Viewpoint

Wearables in the SARS-CoV-2 Pandemic: What Are They Good for?

Brinnae Bent[1], MSc; Jessilyn P Dunn[1,2], PhD

[1]Department of Biomedical Engineering, Duke University, Durham, NC, United States

[2]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, United States

Corresponding Author:
Jessilyn P Dunn, PhD
Department of Biomedical Engineering
Duke University
2424 Erwin Road
Durham, NC, 27705
United States
Phone: 1 9196689798
Email: jessilyn.dunn@duke.edu

Abstract

Recently, companies such as Apple Inc, Fitbit Inc, and Garmin Ltd have released new wearable blood oxygenation measurement technologies. Although the release of these technologies has great potential for generating health-related information, it is important to acknowledge the repercussions of consumer-targeted biometric monitoring technologies (BioMeTs), which in practice, are often used for medical decision making. BioMeTs are bodily connected digital medicine products that process data captured by mobile sensors that use algorithms to generate measures of behavioral and physiological function. These BioMeTs span both general wellness products and medical devices, and consumer-targeted BioMeTs intended for general wellness purposes are not required to undergo a standardized and transparent evaluation process for ensuring their quality and accuracy. The combination of product functionality, marketing, and the circumstances of the global SARS-CoV-2 pandemic have inevitably led to the use of consumer-targeted BioMeTs for reporting health-related measurements to drive medical decision making. In this viewpoint, we urge consumer-targeted BioMeT manufacturers to go beyond the bare minimum requirements described in US Food and Drug Administration guidance when releasing information on wellness BioMeTs. We also explore new methods and incentive systems that may result in a clearer public understanding of the performance and intended use of consumer-targeted BioMeTs.

(JMIR Mhealth Uhealth 2020;8(12):e25137) doi:10.2196/25137

KEYWORDS

digital medicine; digital health; mHealth; wearables; sensors; validation; pandemic; COVID-19

Recently, several big technology companies have released novel wearables with health functionalities, such as wearables capable of measuring blood oxygenation (SpO_2). In September 2020, Apple Inc released the Apple Watch 6, which is the first Apple wearable with SpO_2 monitoring capabilities. This comes on the heels of the Fitbit software update in January 2020, which included SpO_2 monitoring to existing wearables, and the August 2018 release of the Garmin Vivosmart 4, which is one of the earliest wearable consumer-targeted biometric monitoring technologies (BioMeTs) to monitor SpO_2 at the wrist, with the reported intent of obtaining fitness measurements at high elevations. Currently, consumer-targeted BioMeTs that are intended for general wellness purposes [1] do not require medical device regulatory oversight. Instead, these BioMeTs fall under the oversight of the Federal Trade Commission Act, which prohibits unfair and deceptive acts or practices in commerce [2]. Accordingly, the majority of consumer-targeted BioMeT manufacturers do not publicly report on the performance of their sensor technologies. With the backdrop of the SARS-CoV-2 pandemic, wearables and other bodily connected sensors that are marketed as consumer-targeted wellness BioMeTs are, in practice, being used for health decision making [3-6]. For example, during the recent Apple Watch 6 release event, the new SpO_2 functionality was referred to as a "health sensor" multiple times by company executives and mentioned in the context of SARS-CoV-2 detection and studies on asthma, heart failure, and influenza [7-9]. The transparency of consumer-targeted BioMeT performance is desperately needed to avoid the misinterpretation of data or improper use, which will ultimately undermine public perception and trust in these products. We propose the following 3 recommendations

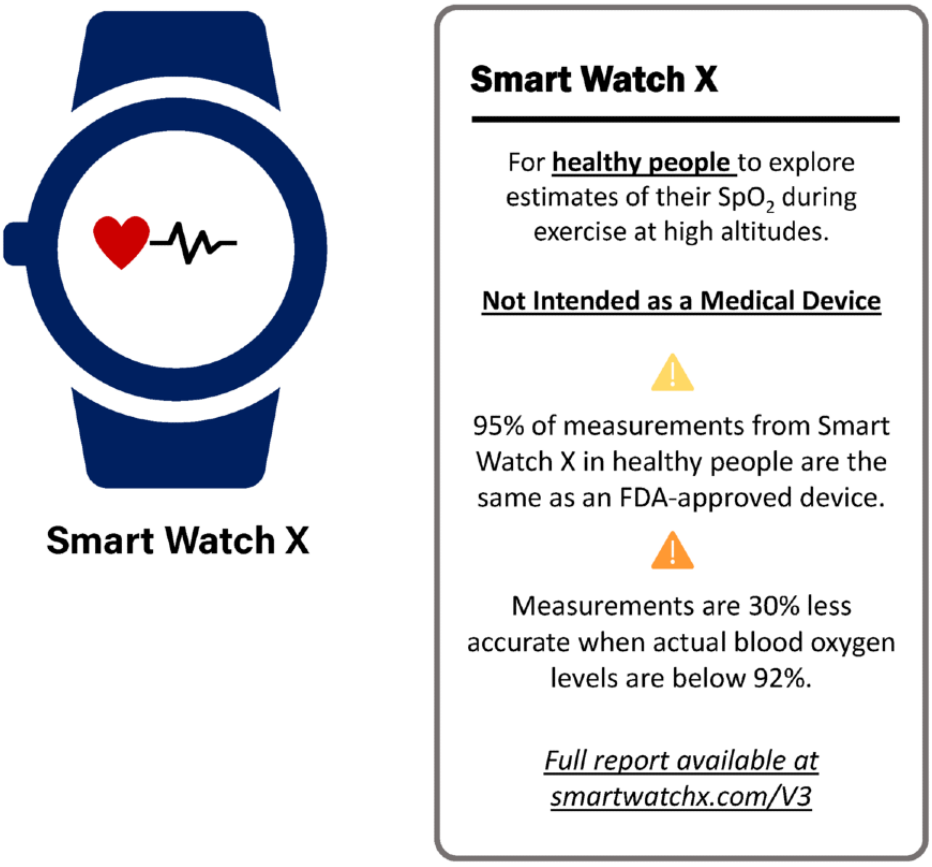
(Textbox 1): (1) consumer-targeted BioMeT companies should state the intended use of the product for consumers, patients, medical practitioners, and researchers, rather than stating what the products are not intended to be used for (ie, medical decision making); (2) consumer-targeted BioMeT companies should clarify how the data should be interpreted and move qualifying statements about how the products are not intended for health purposes from the fine print to the headlines; and (3) we advocate for clarity surrounding the performance of consumer-targeted BioMeTs to increase the trustworthiness of measurements from these products (Figure 1).

Textbox 1. Recommendations for the transparency of consumer-targeted biometric monitoring technology performance.

Recommendations

- State the use of the product for consumers, patients, medical practitioners, and researchers, rather than stating what the products are not intended to be used for.
- Clarify how data should be interpreted by moving qualifying statements about how the products are not intended for health purposes from the fine print to the headlines.
- Clarify the performance of consumer-targeted biometric monitoring technologies to increase the trustworthiness of measurements from these products.

Figure 1. Example of a BioMeT transparency label. We recommend a consumer-targeted BioMeT transparency label that would accompany all product marketing materials, packaging, and mobile applications. The transparency of the performance and intended use of consumer-targeted BioMeTs will cultivate trust and encourage the expansion of BioMeT fit-for-purpose use. BioMeT: biometric monitoring technology.



With the SARS-CoV-2 pandemic, the desire to use consumer-targeted BioMeTs to monitor potential symptoms of infection is understandable, because monitoring for signs of infection at home may reduce anxiety and increase one's confidence in their health status (ie, healthy or sick). The danger however is that many people are improperly using consumer-targeted BioMeTs to monitor for signs of illness and relying too heavily on data that does not have a sufficient evidence base, as there is no oversight by medical professionals [10,11]. Furthermore, it has become increasingly common for a single product to have differentially regulated features, which only adds to the confusion. For example, the Apple Watch received Food and Drug Administration (FDA) clearance for detecting irregular heart rhythms by using the product's electrocardiogram sensor. However, the other sensors on the Apple Watch, including the optical heart rate sensor, are unregulated [12]. It can be challenging for consumers to understand which sensors are regulated. With regard to the Apple Watch, heart rate is monitored with 2 different sensors, but only 1 of the sensors is regulated by the FDA. This challenge can be addressed through our first recommendation, which is to change the status quo from manufacturers listing what the

product is not intended for in the fine print to manufacturers clearly stating the product's intended use and specifying the general wellness category based on the FDA guidance document list [1]. For example, Garmin has set a clear target for use, which is evaluating fitness at high altitudes for mountain climbers, hikers, and runners. Other consumer-targeted BioMeT manufacturers have been less transparent about their products' target for use, resulting in confusion about the intended use for consumers, patients, clinicians, and researchers.

BioMeT users should understand the product's intended purpose, know about the product's limitations, and adhere to the instructions for wear to ensure that measurements are interpreted correctly, while also minding the effects of noise, errors, and biological variability on measurements. To support this, information should be provided in an easy-to-find and easy-to-digest format. Recently, researchers at Elektra Labs and the Digital Medicine Society have proposed the use of a "connected sensor label" (ie, similar to a nutrition facts label) to report on the objective measures of a BioMeT's validation, usability, utility, security, and data governance components [6]. To support transparent validation, the group also developed the V3 Framework, which is a systematic assessment tool for BioMeT performance [13], and very recently published evaluation criteria for using BioMeTs to monitor vital signs during the SARS-CoV-2 pandemic [14]. Perhaps the most integral component of the V3 framework is reporting results in a standardized and transparent manner [6,13,15-17]. These protocols and findings are "key tools for documenting scientific evidence needed to draw inferences on whether a technology is fit-for-purpose for the intended use and context of use" [13].

We can envision several possible scenarios to support and incentivize the open evaluation and reporting of consumer-targeted BioMeT performance (Textbox 2). Manufacturers of consumer-targeted BioMeTs are incentivized by consumer demand, and they can benefit from releasing results that have already been collected through internal product testing to build trust in their products and differentiate their products from those of their competitors. Independent third parties, including research laboratories, consumer groups, and professional societies, can also evaluate and report on the accuracy and quality of BioMeT measurements compared to reference standards. This work can be funded as an extension of the 21st Century Cures Act [18] to support the translation of these products into practice. However, while these practices are not intended to fall under FDA oversight, there is a possibility that these practices could be performed in line with the goal of the new FDA Digital Health Center of Excellence, which is to innovate regulatory approaches for providing efficient and the least burdensome oversight, while also meeting the FDA standards for safe and effective products [19]. The transparency of the performance and intended use of consumer-targeted BioMeTs will cultivate trust and encourage the expansion of BioMeT fit-for-purpose use.

Textbox 2. Nonexclusive scenarios to support and incentivize the open evaluation and reporting of consumer-targeted biometric monitoring technology performance.

Incentives for consumer-targeted biometric monitoring technology manufacturers releasing results that have already been collected through internal product testing

- Consumer demand
- Cultivating trust
- Differentiating their products from competitor products

Incentives for independent third parties (ie, research laboratories, consumer groups, and professional societies) evaluating and reporting on the accuracy and quality of biometric monitoring technology measurements compared to reference standards

- Impartial research
- Expanding biometric monitoring technology fit-for-purpose use
- Publication/exposure
- Potential funding by extension of the 21st Century Cures Act

Incentives for obtaining new regulatory definitions through the Food and Drug Administration Digital Health Center of Excellence

- Innovating regulatory approaches
- Providing efficient and the least burdensome oversight

Acknowledgments

Many thanks to Jen Goldsack, Warren Kibbe, PhD; Ryan Shaw, PhD, RN; and Lindsay Kehoe for their feedback and insights, which were integral to this work. JPD is a Whitehead Scholar and MEDx Investigator, and BB is a Duke FORGE Fellow.

Authors' Contributions

All authors contributed to the writing of the initial manuscript and to the editing process.

Conflicts of Interest

None declared.

References

1. General Wellness: Policy for Low Risk Devices, Guidance for Industry andFood and Drug Administration Staff. U.S. Food & Drug Administration. General Wellness: Policy for Low Risk Devices Guidance for Industry and Food and Drug Administration Staff Preface Public Comment; 2016. URL: https://www.fda.gov/media/90652/download [accessed 2020-12-15]
2. Federal Trade Commission Act. Federal Trade Commission. URL: https://www.ftc.gov/enforcement/statutes/federal-trade-commission-act [accessed 2020-11-19]
3. Johansson D, Malmgren K, Alt Murphy M. Wearable sensors for clinical applications in epilepsy, Parkinson's disease, and stroke: a mixed-methods systematic review. J Neurol 2018 Aug;265(8):1740-1752 [FREE Full text] [doi: 10.1007/s00415-018-8786-y] [Medline: 29427026]
4. Pevnick JM, Birkeland K, Zimmer R, Elad Y, Kedan I. Wearable technology for cardiology: An update and framework for the future. Trends Cardiovasc Med 2018 Feb;28(2):144-150 [FREE Full text] [doi: 10.1016/j.tcm.2017.08.003] [Medline: 28818431]
5. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Per Med 2018 Sep;15(5):429-448 [FREE Full text] [doi: 10.2217/pme-2018-0044] [Medline: 30259801]
6. Coravos A, Doerr M, Goldsack J, Manta C, Shervey M, Woods B, et al. Modernizing and designing evaluation frameworks for connected sensor technologies in medicine. NPJ Digit Med 2020;3:37 [FREE Full text] [doi: 10.1038/s41746-020-0237-3] [Medline: 32195372]
7. Apple Watch Series 6 delivers breakthrough wellness and fitness capabilities. Apple Inc. 2020 Sep 15. URL: https://www.apple.com/newsroom/2020/09/apple-watch-series-6-delivers-breakthrough-wellness-and-fitness-capabilities/ [accessed 2020-09-29]
8. Apple Events - September 2020. Apple Inc. 2020 Sep 15. URL: https://www.apple.com/apple-events/september-2020/?useASL=true [accessed 2020-09-29]
9. Apple September Event 2020 Full Transcript. Rev. 2020 Sep 15. URL: https://www.rev.com/blog/transcripts/apple-event-2020-transcript-september-15 [accessed 2020-09-29]
10. Seshadri DR, Davies EV, Harlow ER, Hsu JJ, Knighton SC, Walker TA, et al. Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments. Frontiers in Digital Health 2020 Jun 23;2(8):8. [doi: 10.3389/fdgth.2020.00008]
11. Natarajan A, Su HW, Heneghan C. Assessment of physiological signs associated with COVID-19 measured using wearable devices. medRxiv. Preprint posted online on August 16, 2020. [doi: 10.1101/2020.08.14.20175265]
12. Krueger AC. DEN180042 Letter. U.S. Food & Drug Administration. 2018 Sep 11. URL: https://www.accessdata.fda.gov/cdrh_docs/pdf18/DEN180042.pdf
13. Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3:55 [FREE Full text] [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
14. Manta C, Jain SS, Coravos A, Mendelsohn D, Izmailova ES. An Evaluation of Biometric Monitoring Technologies for Vital Signs in the Era of COVID-19. Clin Transl Sci 2020 Nov;13(6):1034-1044 [FREE Full text] [doi: 10.1111/cts.12874] [Medline: 32866314]
15. Godfrey A, Goldsack JC, Tenaerts P, Coravos A, Aranda C, Hussain A, et al. BioMeT and Algorithm Challenges: A Proposed Digital Standardized Evaluation Framework. IEEE J Transl Eng Health Med 2020 May 28;8:0700108 [FREE Full text] [doi: 10.1109/JTEHM.2020.2996761] [Medline: 32542118]
16. Godfrey A, Vandendriessche B, Bakker JP, Fitzer-Attas C, Gujar N, Hobbs M, et al. Fit-for-Purpose Biometric Monitoring Technologies: Leveraging the Laboratory Biomarker Experience. Clin Transl Sci. Epub ahead of print 2020 Aug 08. [doi: 10.1111/cts.12865] [Medline: 32770726]
17. Bent B, Goldstein BA, Kibbe WA, Dunn JP. Investigating sources of inaccuracy in wearable optical heart rate sensors. NPJ Digit Med 2020 Feb 10;3:18 [FREE Full text] [doi: 10.1038/s41746-020-0226-6] [Medline: 32047863]
18. H.R.34 - 114th Congress (2015-2016): 21st Centure Cures Act. Library of Congress. URL: https://www.congress.gov/bill/114th-congress/house-bill/34 [accessed 2020-12-15]
19. Digital Health Center of Excellence. U.S. Food & Drug Administration. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence [accessed 2020-09-29]

Abbreviations

BioMeT: biometric monitoring technology
FDA: Food and Drug Administration
SpO_2: blood oxygenation

Edited by L Buis; submitted 19.10.20; peer-reviewed by C Manta, J Goldsack, S Shah; comments to author 18.11.20; revised version received 20.11.20; accepted 10.12.20; published 22.12.20.

Please cite as:
Bent B, Dunn JP
Wearables in the SARS-CoV-2 Pandemic: What Are They Good for?
JMIR Mhealth Uhealth 2020;8(12):e25137
URL: http://mhealth.jmir.org/2020/12/e25137/
doi:10.2196/25137
PMID:33315580

Review

Digital and Mobile Technologies to Promote Physical Health Behavior Change and Provide Psychological Support for Patients Undergoing Elective Surgery: Meta-Ethnography and Systematic Review

Anna Robinson[1], MPharm, PG Dip; Umay Oksuz[1], MPharm; Robert Slight[2,3], MBChB, PhD, FRCSEd (C-Th); Sarah Slight[1,3], MPharm, PhD, PG Dip; Andrew Husband[1], MPharm, MSc, PhD, PG Dip

[1]School of Pharmacy, Population Health Sciences Institute, 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:
Andrew Husband, MPharm, MSc, PhD, PG Dip
Population Health Sciences Institute
Newcastle University
School of Pharmacy, King George VI Building
Newcastle upon Tyne, NE1 7RU
United Kingdom
Phone: 44 191 208 2360
Email: andy.husband@newcastle.ac.uk

Abstract

Background: Digital technology has influenced many aspects of modern living, including health care. In the context of elective surgeries, there is a strong association between preoperative physical and psychological preparedness, and improved postoperative outcomes. Health behavior changes made in the pre- and postoperative periods can be fundamental in determining the outcomes and success of elective surgeries. Understanding the potential unmet needs of patients undergoing elective surgery is central to motivating health behavior change. Integrating digital and mobile health technologies within the elective surgical pathway could be a strategy to remotely deliver this support to patients.

Objective: This meta-ethnographic systematic review explores digital interventions supporting patients undergoing elective surgery with health behavior changes, specifically physical activity, weight loss, dietary intake, and psychological support.

Methods: A literature search was conducted in October 2019 across 6 electronic databases (International Prospective Register of Systematic Reviews [PROSPERO]: CRD42020157813). Qualitative studies were included if they evaluated the use of digital technologies supporting behavior change in adult patients undergoing elective surgery during the pre- or postoperative period. Study quality was assessed using the Critical Appraisal Skills Programme tool. A meta-ethnographic approach was used to synthesize existing qualitative data, using the *7 phases of meta-ethnography* by Noblit and Hare. Using this approach, along with reciprocal translation, enabled the development of 4 themes from the data.

Results: A total of 18 studies were included covering bariatric (n=2, 11%), cancer (n=13, 72%), and orthopedic (n=3, 17%) surgeries. The 4 overarching themes appear to be key in understanding and determining the effectiveness of digital and mobile interventions to support surgical patients. To successfully motivate health behavior change, technologies should provide motivation and support, enable patient engagement, facilitate peer networking, and meet individualized patient needs. Self-regulatory features such as goal setting heightened patient motivation. The personalization of difficulty levels in virtual reality–based rehabilitation was positively received. Internet-based cognitive behavioral therapy reduced depression and distress in patients undergoing cancer surgery. Peer networking provided emotional support beyond that of patient-provider relationships, improving quality of life and care satisfaction. Patients expressed the desire for digital interventions to be individually tailored according to their physical and psychological needs, before and after surgery.

Conclusions: These findings have the potential to influence the future design of patient-centered digital and mobile health technologies and demonstrate a multipurpose role for digital technologies in the elective surgical pathway by motivating health behavior change and offering psychological support. Through the synthesis of patient suggestions, we highlight areas for digital

technology optimization and emphasize the importance of content tailored to suit individual patients and surgical procedures. There is a significant rationale for involving patients in the cocreation of digital health technologies to enhance engagement, better support behavior change, and improve surgical outcomes.

(JMIR Mhealth Uhealth 2020;8(12):e19237) doi:10.2196/19237

KEYWORDS

mobile health; mHealth; healthy lifestyle; bariatric surgery; cancer; orthopedic procedures; qualitative research; systematic review; telemedicine; mobile phone

Introduction

Background

The introduction of digital technologies has influenced many aspects of modern living, including health care [1,2]. Digital health technologies (eg, wearable activity trackers and mobile phone apps) provide opportunities for effective patient care. They can improve communication between health care providers [3], facilitate connectivity with clinicians and peers [4,5], enable remote health monitoring [6], and empower patients to play an active role in their long-term care [3,7-9].

In the context of elective surgeries, there is a strong association between preoperative physical and psychological preparedness and improved postoperative outcomes [10-12]. More specifically, improvements in physical activity levels [13], dietary intake [14], and smoking cessation [15] have been linked to improved recovery after surgery, reduced risk of complications, better tolerance of postsurgical adjunctive treatment, and prevention of long-term disease [16-19].

Although health behavior changes made in the pre- and postoperative periods can be fundamental in determining the outcomes and success of elective surgeries [19-21], there are variable amounts of support and education currently provided to patients undergoing elective surgery to motivate these health behavior changes [22-24]. A recent study evaluating patient attitudes to health behavior changes found that although preoperative patients understood the health benefits of improved behaviors, they lacked the confidence to make such changes without intervention or support [19]. Many physical and mental health interventions offered in elective care pathways use face-to-face, in-person delivery for individuals or small groups of patients. Such approaches are resource- and time-intensive for staff already working in high-pressure health care sectors [25-27]. In addition, geographic isolation, travel costs, and the time burden of attending classes can all negatively affect patient engagement with postoperative appointments [28,29]. Understanding the potential unmet needs of patients undergoing elective surgery is central to motivating health behavior changes. Integrating digital technologies within the elective surgical pathway could be one strategy to remotely deliver behavioral change advice and lifestyle support, consequently improving patient engagement and postoperative success rates [12,30].

Approach

This review uses a meta-ethnographic approach to analyze and synthesize qualitative findings. Meta-ethnography was originally developed by Noblit and Hare [31], but it has recently been used in health care–based social science research by Britten et al [32], Campbell et al [33,34], Pound et al [35], and others. It is an inductive and interpretive approach involving the translation of papers into one another. Meta-ethnographies encourage researchers to understand and transfer ideas, themes, and metaphors across different studies to gain a deeper understanding or to inform the development of broader concepts [31,36].

There are still unanswered questions relating to the optimization of digital technologies to support patients undergoing elective surgery, especially in the cohorts of bariatric, cancer, and orthopedic surgery. We seek to synthesize findings from existing qualitative research to determine whether digital technologies are effective in supporting patients undergoing elective surgery to change their health behaviors, specifically focusing on physical activity, weight, dietary intake, and mental health support (eg, cognitive behavioral therapy).

Methods

This meta-ethnographic systematic review is registered with PROSPERO (registration number CRD42020157813) and has been conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Multimedia Appendix 1).

Search Strategy and Information Sources

A comprehensive and systematic literature search was conducted in October 2019 across 6 electronic databases: MEDLINE, EMBASE, CINAHL, PsycINFO, Web of Science, and Scopus. No limit on the publication date was applied. Additional papers were identified via gray literature using Google Scholar, and we manually searched the bibliographies of all included studies. A full list of search terms is included in Multimedia Appendix 1.

Eligibility Criteria

This meta-ethnography focused on elective surgical procedures, specifically bariatric, cancer, and orthopedic surgeries. Patients undergoing these elective procedures may have improved surgical outcomes owing to pre- and postoperative health behavior changes and therefore can benefit from the support of digital health technologies. Acute, unplanned surgeries and emergency trauma procedures were excluded from this review.

Only the studies that had encompassed the use of digital health interventions to support behavior changes (such as weight changes, dietary intake, physical activity levels, and/or mental health strategies) in adult patients undergoing elective surgery (>18 years) during the pre- or postoperative period were included. There were no restrictions placed on participants' sex,

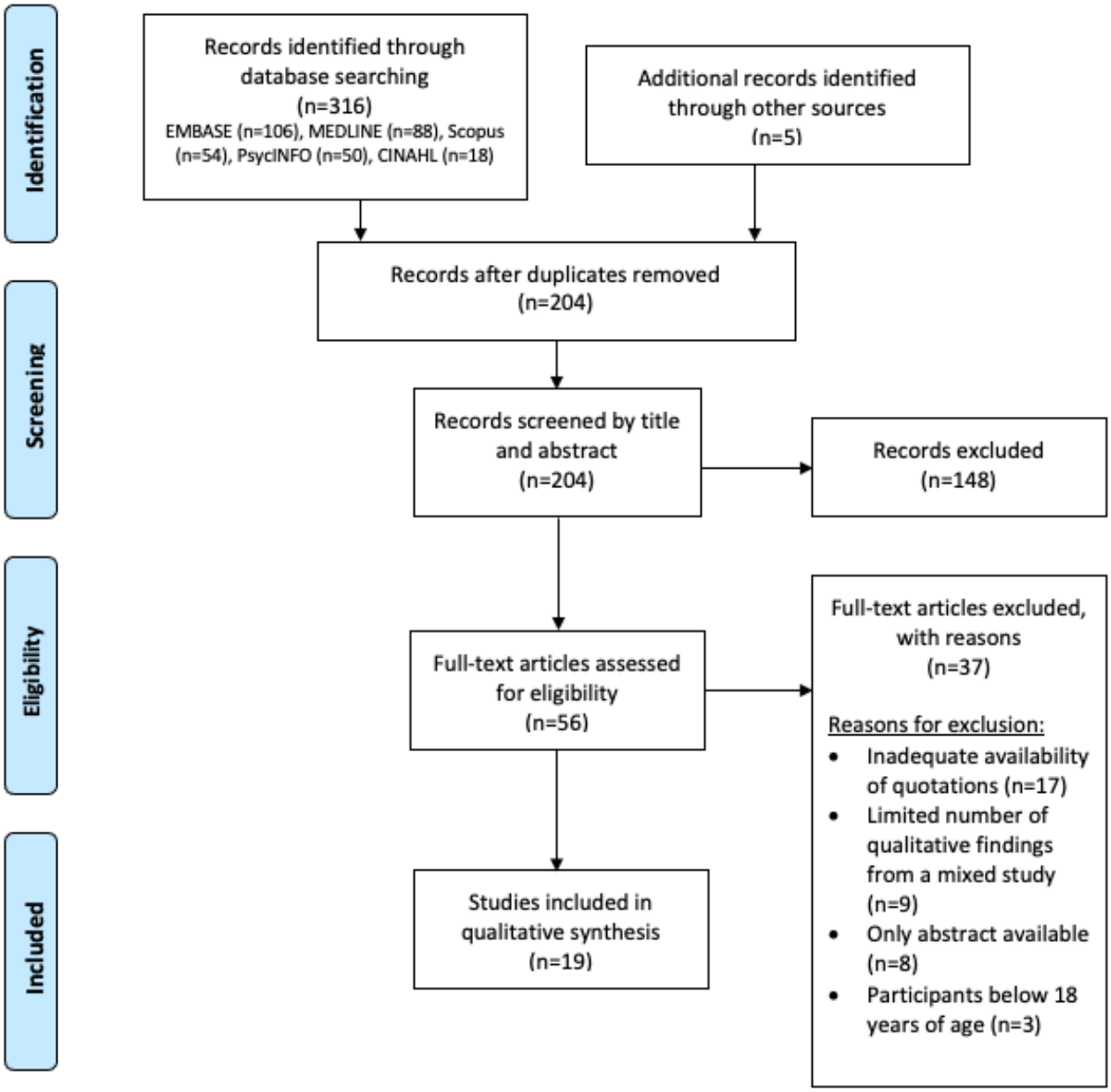
ethnicity, or nationality. The included studies must be qualitative or mixed method studies containing a significant qualitative component to analyze participant perspectives (eg, patient interviews or focus groups).

Exclusion criteria included studies employing behavior changes achieved by nondigital interventions; participants who were not scheduled to undergo an elective bariatric, cancer, and orthopedic surgery; studies where the intervention was mainly focused on perspectives of health care professionals; nonqualitative studies (eg, quantitative studies, systematic reviews, or protocols); and studies in languages other than English.

Selection of Eligible Studies

Two authors (UO and AR) reviewed the titles and abstracts from the database search. Full texts were retrieved for articles that met the inclusion criteria and for those that could not be rejected with certainty. Two authors (UO and AR, with an agreement rate of 94.7%) independently screened the full texts of eligible articles. Disagreements (on 3 of the 56 articles) were resolved through discussion with a third reviewer (AH). Figure 1 shows a PRISMA flowchart for the study selection process.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of included studies.



Reading, Data Extraction, and Quality Appraisal

Two authors (UO and AR) closely read and re-read the included studies to ensure close familiarity with the work. Data extraction was performed across the full primary study (by UO and AR) [36] and carried out using a customized data extraction form, including the study and author details, method of intervention delivery, population data, inclusion criteria, and original quotes and/or concepts developed by the authors of primary studies (within their original context). Both authors worked independently before comparing their work; disagreements were

resolved through discussion with a third reviewer (AH) where necessary. Quality appraisal was conducted independently by UO and AR using the Critical Appraisal Skills Programme (CASP) questions to understand qualitative research [37]. No papers were excluded on the grounds of quality.

Analysis and Interpretive Synthesis

Meta-ethnographic approaches were applied to this review, as determined by the 7 phases of meta-ethnography by Noblit and Hare [31]: (1) *getting started*, (2) *deciding what is relevant to the initial interest*, (3) *reading the studies*, (4) *determining how studies are related*, (5) *translating the studies into one another*, (6) *synthesizing translations*, and (7) *expressing the synthesis*.

The findings (concepts and metaphors) from the primary studies were compared to determine how they are related. Noblit and Hare [31] suggested that phase 5, where findings are translated into one another, follows something like "one case is like another, except that...". This phase of a meta-ethnographic approach is termed "reciprocal translation," and it enables the development of themes and subthemes for interpretive synthesis [31,33]. According to this, we developed 4 overarching themes (or third-order constructs) and subsequent subthemes that were consistent with the original results but also extended beyond them.

When translating the studies into one another to develop themes (and subthemes), we arranged each paper chronologically and compared the themes from paper 1 with those of paper 2, then those of paper 2 with those of paper 3, and so on. As we compared each study, we grouped similar themes and continually reviewed and refined them until they were coherent and distinctive. Two reviewers (UO and AR) were involved in the study translation at all times; however, if agreement was not reached between these, discussions with a third author (AH) helped to establish a consensus.

To adhere to recommendations for conducting meta-ethnographies, we used the term "theme" to describe the third-order construct and subthemes to describe third-order construct subthemes [36]. The development of these overarching themes enables meta-ethnographies to delve further into a topic as compared with a traditional systematic review and contribute new insights to the literature [32].

We report on the overall effectiveness of digital health technologies to support behavioral change in patients undergoing surgery through 4 established themes: (1) motivational support, (2) patient engagement with interventions, (3) the facilitation of peer networking, and (4) intervention specificity to meet patients' individual needs.

Results

Search Results

A total of 316 citations were retrieved from the database searches. A total of 5 additional records were identified through gray literature and searching references manually from relevant studies. Following the removal of duplicates (n=112), 204 papers were screened, of which 148 were excluded based on their titles and abstracts. A total of 56 full-text papers were assessed for eligibility; 38 of these were excluded due to reasons detailed in the PRISMA flowchart in Figure 1. The remaining 18 studies were included in this meta-ethnographic systematic review; of these, 68% (n=13) were qualitative and 32% (n=7) were mixed methods studies.

Study Characteristics

All 18 included papers were published between 2013 and 2019. The study was conducted in 8 different countries: United States (n=6) [38-43], United Kingdom (n=3) [44-46], Canada (n=3) [47-49], Australia (n=2) [50,51], Ireland (n=1) [52], Norway (n=1) [53], South Korea (n=1) [54], and China (n=1) [55].

The 18 studies covered 3 different surgery types: bariatric (n=2, 11%), cancer (n=13, 72%), and orthopedic (n=3, 17%) surgeries. Further study characteristics, including the method of intervention delivery and original themes extracted from the study, are detailed in Multimedia Appendix 1 [35,37-55].

A total of 3 main intervention delivery methods were identified in the 18 included studies. These included internet-based interventions (eg, emails, e-platforms, virtual reality, tele-rehabilitation) [38,40,42,44,46,48,49,53,54], mobile phone–based interventions (eg, text messages, smartphone apps) [39,45,55], and wearable interventions (eg, activity trackers) [43,47,50-52]. Only 1 study reported the use of a combination of 2 intervention methods (dual approach), including wearable- and phone-based interventions [41].

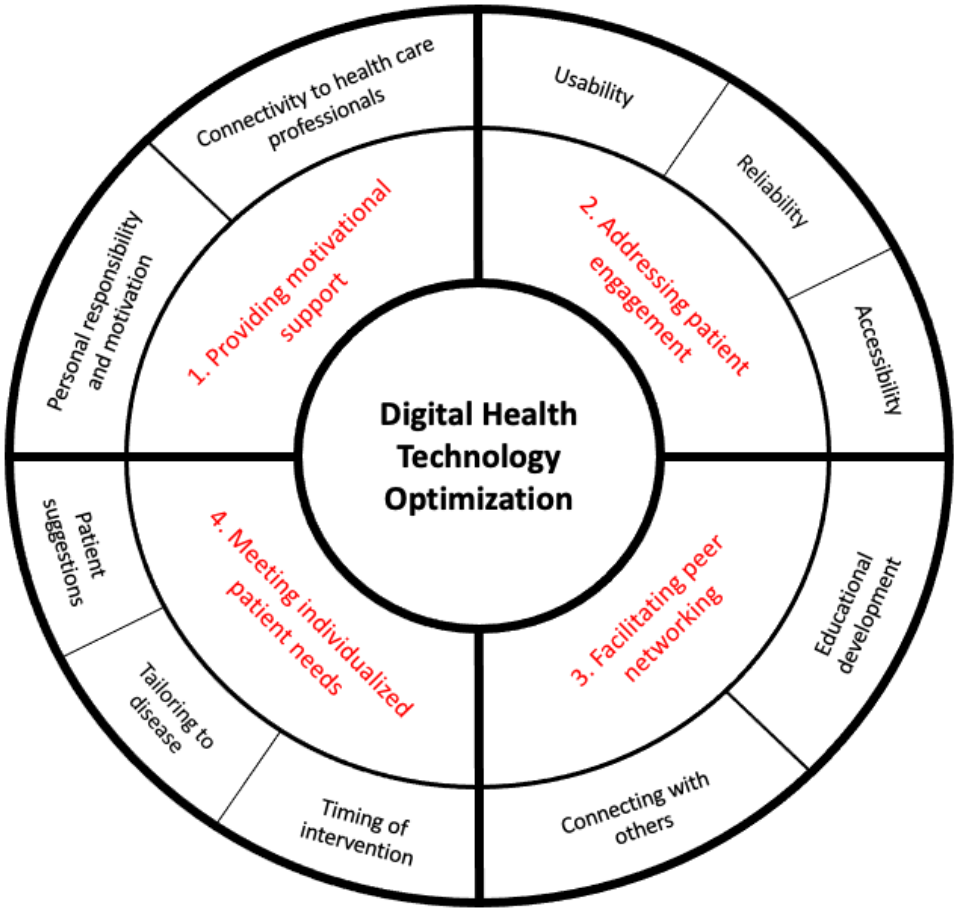
Study Quality

Multimedia Appendix 1 contains details of the quality appraisal conducted using the CASP tool for qualitative studies. Of the included studies, Shaffer et al [38], Phillips et al [39], Alberts et al [40], and Argent et al [52] were identified as having the highest quality.

Findings: Reporting Outcomes, Synthesizing Translations, and Developing Themes and Subthemes

Multimedia Appendix 1 presents the metaphors and patients' perspectives from each of the included studies. Reciprocal translation and refutation of these concepts enabled the development of 4 overarching themes and subthemes for this meta-ethnography; they are outlined in Figure 2. The 4 overarching themes and subthemes appear to be key in understanding and determining the effectiveness of digital and mobile health interventions to support behavior change in patients undergoing surgery.

Figure 2. Developed themes and subthemes for Digital Health Technology Optimization. The inner band on the diagram (red text) represents the 4 overarching themes developed by this review, and the outer band details the subsequent subthemes (black text).



The qualitative data synthesis can be found in Multimedia Appendix 1, with each table representing one of the 4 overarching themes. These tables showcase examples of direct quotations (first-order constructs) from study participants, the authors' interpretations of the original findings from the included studies (second-order constructs), and our interpretation (third-order constructs) and subthemes.

Providing Motivational Support

Personal Responsibility and Motivation

Certain features of digital and mobile health technologies increased patients' self-awareness and motivation for physical activity. Patients reported that wearable activity trackers (termed wearables) made them more aware of the importance of physical activity and helped them monitor their sedentary behavior levels, which acted as a source of motivation to engage with positive behavior change [43,47,51]. Self-regulatory features of the wearables, including goal setting and performance feedback, facilitated personal fulfilment and gave orthopedic and cancer surgical patients a sense of control and accomplishment [50-52,54]:

> *Seeing your progress, I think is very important. Seeing measurable progress, whether it's in calories burned, or minutes, or meeting a percentage of your goal [39].*

> *set goals, like mid-week if I wanna hit 150 [minutes] I should be at half that [...] and the application is on my phone and I can see what I've done [...] so it's really easy to track how well you're doing or how well you're not doing [47].*

Patients perceived feedback and text messages as methods of encouragement, motivation, and support [39]. Gell et al [41] noted that health coaching, when offered alongside daily wearable use, provided patients undergoing cancer with an increased sense of self-importance, encouraging the maintenance of physical activity:

> *If you get to say 8000 [steps] in a day, you're motivated to do those extra 2000 because you're so close. It's like "Why would I stop now?" I might as well keep going [50].*

However, findings reported the potential for this to shift to *fear of failure* with nonadherence, where prompts or reminders could turn into negative judgments [47]:

> *for now, I don't wanna [sic] be judged or evaluated or anything else… and then that will change…It's just a case of you get tired of* [judgment]

Connectivity to Health Care Professionals

Patients undergoing bariatric surgery reported an increased feeling of accountability and responsibility to adhere to treatment plans when they were being monitored through digital or mobile health technologies [53,56]. Cancer and orthopedic surgical patients reported benefits of enhanced connectivity with members of their clinical team, including the provision of

timely and personalized feedback [46,55] and the availability of instant communication for information-seeking needs [48,55]. A lower threshold for information seeking via digital technologies was reported by patients undergoing bariatric surgery, with sensitive questions being asked more readily [53]. Das and Faxvaag [53] evaluated the impact of an online forum on interactions between health care professionals and patients undergoing bariatric surgery. The authors recognized that the connectivity provided an easier access to evidence-based advice as well as offered a convenient and geographically independent platform to promote patient engagement [55]:

> *I could ask questions through the app regarding my medical condition. I could upload the lab results through your program. Then I received corresponding advice from experts. I felt followed up. When I knew more about my medical condition, I felt more likely to gain control of my life.*

> *I live far away from the hospital and I have no doctor close to me. When I had questions about my medical condition, I could not find the answer in the internet. Then I asked questions through the app. Aha, the professor or expert responded.*

Although this increased connectivity with health care professionals was reported to be beneficial in supporting postoperative recovery, patients undergoing cancer still felt that technologies should not replace traditional face-to-face appointments with clinicians. Concerns were raised by this cohort, with patients reporting that they may miss out on vital interactions, such as displays of empathy, which come from face-to-face communication.

Addressing Patient Engagement

Usability

Simplicity and ease of use were identified as prerequisites for effective engagement with digital and mobile health technologies in cancer and bariatric studies [39,40,42,43,47,48,52]. Patients in these studies reported the importance of feeling relaxed and at ease while using technology [40]. It is important to avoid complex or difficult interventions that may decrease a user's enjoyment [54]:

> *Well it was very simple. It was straightforward. It wasn't complicated…like going through chemo you have kind of a brain scramble and… just the simplest things you can't wrap your brain around sometimes [40].*

> *I would say the most important thing is the ease of use, the simplicity of it, because if it's cumbersome I will not use it [39].*

Some participants undergoing varied cancer surgeries encountered technical difficulties while operating and synchronizing devices, which affected their rates of engagement [43,50,55]:

> *Largely, I'm not wearing it because it doesn't interact with my computer very easily...why bother? I just go use my manual step counter [43].*

In studies evaluating wearable technologies, *wearability* was deemed important with references to comfort and style improving user engagement before and after cancer and orthopedic surgery [50-52]:

> *I didn't like wearing it at night. I didn't feel comfortable [50].*

> *I had the Polar first [...] I thought it was quite heavy and quite clunky but then I had the two Garmins and in the end I decided that was my favourite even though it was heavier [50].*

Reliability

The reliability of digital and mobile health technologies also affected engagement; patients recognized inaccuracies, which resulted in a lack of trust in the interventions and poor adherence to postoperative physical activity guidance [43,51,55]:

> *It seemed to register less activity than I felt I actually did because it was only measuring steps, and I was doing more than steps. I was lifting. I was bending. I was twisting. I was doing all that other sort of stuff [43].*

> *The app sometimes was unstable. It didn't work when I tried to open it. I contacted with someone in the hospital and reinstalled the app. Then I could log in. However, after a period of time, I couldn't open the app again. Finally, I gave up using your program. I haven't log in for the recent month 55.*

Accessibility

The accessibility that digital and mobile health technologies offer was perceived as beneficial by all surgical cohorts, particularly if participants were geographically, economically, or functionally isolated [40,42,48]. Digital interventions reduced the time and cost of travel to clinics, an advantage over facility-based interventions [40,48]:

> *I really like it (telerehabilitation). I found it fantastic...you know, just the fact of not having to travel when we are in pain (...) I adored it [48].*

> *Well, definitely the availability of it to anybody, no matter where you live. I know we work with a lot of rural people and after they're done here, they don't want to travel for more therapy or whatever, so something that they can do at home [40].*

Facilitating Peer Networking

Educational Development

By building a peer network, digital and mobile health technologies provide patients undergoing enhanced access to knowledge and support and, as a result, can motivate health behavior changes to improve surgical outcomes. Informational support delivered by peers was perceived as useful and relatable by patients undergoing bariatric and cancer surgeries. Patient satisfaction and reassurance were reported from the sharing of personal anecdotes and advice after bariatric surgery [49]. Strategies addressing preoperative concerns and the challenges of adhering to surgery guidelines were also shared [42,49,53]:

> *[Product name]... this is odourless and tasteless and does not clump. You can add it to hot or cold... or just sprinkle over your food. One tablespoon equals a scoop of Whey and has 30 grams of protein. It is approved by [Medical Association] and has 96% absorption... [49]*
>
> *I think it is more enjoyable to write a "diary" that everyone can read and comment on. I like to get feedback on how I do things, what I eat, and thoughts that I have about the surgery and about life after the operation, so here comes a little of everything...Hope you will read and comment [53].*
>
> *...You may want to pick up a pill crusher and a pill splitter in the drug store. The large pills such as calcium citrate, I had to crush and mix with drink in order to take them... [49]*

Connecting With Others

In addition to informational support, digital and mobile health technologies and online forums provided emotional support to patients. Studies referred to the benefits of patients undergoing cancer communicating with others who have had the same surgical procedures or experience with the same disease-related condition [40,42,50]. Peer interactions have helped patients overcome feelings of loneliness and improved individual mental well-being [40,45,55]. In preoperative peer forums, encouraging messages have motivated patients undergoing cancer surgery and bariatric surgery to lose weight and adhere to physical activity and dietary guidelines before surgery [40,42,45,49,50,55]:

> *You know that you're not alone, but when your feelings are validated just by reading someone's story, I mean that is everything [40].*
>
> *It is so important to get in touch with people who went through the same thing as you have. [...] I think that if an app for cancer survivors had a forum on it as a part of the application to motivate each other, that would be amazing [45].*
>
> *I feel better to talk to someone who is in similar situations. Cancer is not a good thing. If I always think about breast cancer alone at home, it is so easy for me to feel bad. I didn't feel alone when I talked with peers through your program [55].*

Meeting Patients' Individual Needs

Timing of the Intervention

According to patients undergoing surgical cancer, initiating and tailoring the content of a digital or mobile intervention appears to be essential to determine its effectiveness to motivate behavior change. Two papers discussed the optimal time to start an intervention within a surgical journey; some patients undergoing cancer suggested initiation should be during the preoperative period to enable preparedness and understanding of processes [42], whereas others favored postoperative provision [38]:

> *I wish I would have had something like this when I was first diagnosed... I can see this tool being useful in answering questions that have not come to mind [38].*
>
> *I had more trouble with sleep issues early on at diagnosis and in between surgeries, so it would have been helpful for me to have enrolled in the program earlier [42].*

This cohort reported a preference to start with interventions once adjuvant chemotherapy was completed, citing treatment burden and side effects as factors for disengagement at this time. Immediate postoperative issues, such as fatigue, were also noted to impact early engagement rates [45]. However, some patients appreciated low-effort strategies during the surgical journey to manage symptoms and improve relaxation [38]:

> *The very end of your treatment when you finished your chemo and...the doctor says "Ok, see you in six months." That would be the time to offer it. "Cause you feel so unwarned [sic]."*
>
> *Interestingly, there was a general agreement among cancer patients that the best time to begin an intervention is "when you recognise that you have a problem ... and that you want help."*

Tailoring According to the Disease

Participants with surgical cancer also expressed a desire for intervention tailoring according to their changing physical and psychological health needs [38,39,42,45,47,50,55], focusing on information on their disease and surgical type [39,42,55]. Puszkiewicz et al [45] noted preferences for individualization of digital interventions according to patient lifestyles rather than a disease on the whole:

> *The issues I might have as a colorectal cancer survivor are very different from the ones than someone who had breast cancer or prostate cancer.*
>
> *Anyone with any condition could use this program, which is beneficial, but it could be more beneficial [...] more tailored to the type of cancer or disease you had, to your lifestyle and fitness goals. I think it could be more fine-tuned to your circumstances, lifestyle, then that would be really helpful.*

In the virtual reality–based rehabilitation study, participants expressed positive views on the personalized task difficulty, where the varied level of difficulties helped them to choose the exercise program according to their needs, and subsequently increased their satisfaction with the intervention [45].

Patient Recommendations

Participants across all 3 surgical cohorts suggested design and technical improvements for the future development of digital and mobile interventions. Although these varied depending on the delivery method, a user-centered design was identified as a key solution to enhance and maintain engagement and to motivate behavior changes [39]:

> *I think that it needs to be aimed towards survivors. That would be the first component. There's a lot on the Internet that gives you a lot of exercises but it's not aimed towards survivors.*

Patient-reported design improvements for wearables included higher accuracy of the devices [39,50] different aesthetics (such as the tone of the prompt and color scheme) [47,50,51] and personal goal setting [50]:

> *So I'll give you a case. I filled my laundry, and it's logged I walked 2,000 steps. I did not walk 2,000 steps 43.*
>
> *I'd get a little vibration to say let's go do 250 steps, it was much more polite than MOVE 51.*
>
> *I like that the colour scheme was NOT pink! 48*

In online forums for patients undergoing bariatric surgery or cancer surgery, fear of self-disclosure was a recognized barrier that affected user engagement. Full anonymity would make it easier to share sensitive issues and ask difficult questions [49,55]:

> *On other forums, even though you don't have your name, with a nickname, you can find out who the person is anyway. You have to be very careful if you want to be anonymous 53.*

Participants also suggested adding *search* tools to locate information and save time [55], as well as the inclusion of diet recommendations and/or self-monitored food intake [55]:

> *The program can be improved by adding search engine in the Learning forum. If I search for "nausea" then all the knowledge related to nausea will come out. Search engine will help save my time [55].*
>
> *We are in a dilemma on what we should eat. The apps can provide detailed information on food choice, the time of food intake, the cooking methods, etc...Such practical information would be very helpful.*

Older users appeared more likely to experience usability issues with interventions [55]. To overcome this, patients reported preferences for *open access* so that family members or caregivers can offer support [55]:

> *I was overwhelmed by the information each time I opened it.*
>
> *Some people, like me, 40 or 50 years old. Well, this group believe the apps is a little bit troublesome. They feel challenged to use the new technology... If this program can be available for their family members, such as their son or daughter, it would be helpful.*
>
> *Many women with breast cancer come from the countryside. They are illiterate, or they cannot read and speak Mandarin... if you can open the program to other family members who can read and convey the knowledge to the women, they would also benefit.*

Discussion

Principal Findings

To our knowledge, this is the first meta-ethnographic systematic review examining the effectiveness of digital and mobile technologies to support health behavior change in patients undergoing elective surgery. Using reciprocal translation, our findings indicate 4 themes that appear to be key in determining intervention effectiveness to support health behavior change in patients undergoing surgery: (1) providing motivational support, (2) addressing patient engagement, (3) facilitating peer networking, and (4) meeting individualized patient needs. Future studies could use these findings to inform future design frameworks for specific surgical cohorts while embracing digital transformations in health care.

Limitations

Although meta-ethnographies offer an opportunity to synthesize findings to develop new or deeper understandings on a subject, the process is largely interpretive [31]; other conclusions from the same included studies may be possible but still equally as valid. It is also important to note that the focus of this meta-ethnography was solely elective cancer, bariatric, and orthopedic surgeries, and as such, the meaning of our findings may not be generalizable for acute surgeries or other specialties.

Comparisons With Previous Work

Digital and mobile technologies act as a catalyst to engage with healthy behaviors, such as loss of weight, improved dietary intake, and increased physical activity levels. Messages of positive reinforcement were viewed as useful, particularly when tailored to an individual's surgical type and readiness to make behavioral change. Existing literature suggests that individualized goal setting helps combat sedentary behavior [57-59]; personalized feedback and messages of encouragement provide a sense of accomplishment [25,39]; and visual tracking of step count has been reported as motivational [50,58]. Recent contributions to the health behavior change literature have cited the importance of empowered patient-centered strategies and use of self-regulation [60] and self-determination theoretical frameworks [61,62] to improve patient motivation. Digital technologies underpinned by behavior change theory can promote a proactive and holistic strategy to influence behavior change in a modern health care system such as the UK National Health Service [21].

In the context of patients undergoing surgical cancer, internet-delivered cognitive behavioral therapy (iCBT) was associated with numerous benefits [40,44]. Following digital intervention usage, there have been improvements related to fatigue, sleep [63], depression [64], and psychological distress [65]. In addition, our findings suggest that iCBT can also educate participants around various coping strategies to manage fears of treatment and disease recurrence [40].

Technologies enabling connectivity to health care professionals have been positively acknowledged. Two-way telemedicine consultations, emails, and text message discussions facilitated improved information delivery, real-time goal setting, psychosocial outcomes, and decision making [66-68]. Participants felt motivated, reassured, and encouraged to adhere to postoperative advice through remote monitoring. Having access to health care professionals *behind a screen* also helped patients overcome their personal barriers and raise unmet needs beyond routine clinical questioning [53,67]. From the perspective of clinicians, digital and mobile health technologies provided them with a means to monitor patient progress, which

enabled individualized advice to be given to reinforce beneficial behavior change [43,69].

Despite the benefits of digitally enabled communication, it is worth considering social norms with patient-professional relationships [70]. For some, the continuity of face-to-face appointments is essential to provide empathetic interaction and social support [70,71]. Empathy, rapport, and compassion through nonverbal behavior and body language is difficult to establish when communicating digitally. Despite this, Kairy et al [48] reported close relationships and trust between the therapist and patients when communicating *via* telerehabilitation. Perhaps, complementing traditional face-to-face appointments with digital health interventions could be a way to maintain patient-professional relationships.

Usability has been reported as a key determinant to induce and maintain health behavior change, where interventions should be easy to use as well as aesthetically and visually appealing. Patient preferences should be taken into account when it comes to the design and tailoring of interventions [50,72,73]. It is worth considering ways to overcome digital health literacy barriers to further promote usability and engagement. Additional technical support might be beneficial when targeting older adult populations to increase their engagement and thus better support health behavior change [50,74].

One reported advantage of digital interventions is the accessibility they offer [29,51,75]. Postoperative breast cancer survivors living in rural settings experienced greater depressive symptoms compared with those with shorter commutes owing to the long travel distances required to access health services [76,77]. Where telerehabilitation was implemented for postoperative orthopedic follow-ups, participants reported improved continuity of care with the same physician and improved ability to control the timing of appointments and intensity of the rehabilitation service [78].

In addition to bridging access to health services, digital and mobile health technologies are being increasingly used as networking and peer-support tools. Patients undergoing similar procedures or diagnosed with similar conditions are able to communicate and share personal experiences and coping strategies with others [53]. Peer support and behavior change have been previously reported in elective care [50,79-82], where increased social support and decreased patient isolation are associated with postsurgical success [81,83]. Although digital technologies offer opportunities to interact with peers on an educational level, concerns have been raised about the accuracy and credibility of shared information [79,84-86]. Health care professionals should caution patients when interpreting discussions on forums or online groups, given the potential detriments that may arise from following inaccurate information [53,86,87].

The optimal time point in the surgical pathway to initiate digital and mobile technologies remains uncertain, with findings suggesting that this may vary depending on the type of surgical group. Despite this, what remains clear is the potential benefit of capitalizing on a *teachable moment* to empower and educate patients about the underlying benefits of health behavior changes [88-90]. Evidence suggests that preoperative interventions based on education of lifestyle changes are significantly more effective in managing postoperative complications and patient expectations [91].

Our research has synthesized numerous design considerations that should be examined when producing future interventions to support patients undergoing surgery. It was found that internet-based interventions may benefit from adding a *search* tool to locate target information [38], the comfort of wearable technologies should be addressed [43,50], and negative connotations with using the color pink for patients undergoing cancer builds on the *cancer culture divide* [92]. The possible benefits of incorporating open-access features within interventions were also discussed. Previous research has shown that opening care access, to include relatives or caregivers, provided patients undergoing an increased sense of pre- and postoperative support [93-95]. This approach has strengthened bonds with family members, improved patient experience, resulted in effective engagement with digital interventions, and therefore supported superior outcomes in lifestyle changes [96-98]. This review synthesizes existing research to gain a deeper understanding of the ways in which digital tools can support elective surgical cohorts and identify key design features that support elective surgical patients to change their health behaviors, and thus have a greater impact on postoperative health. Considering the rapidly progressive nature of digital health interventions and digital assistive technology research, cocreation of a person-centered digital support network may help surgical patient cohorts to benefit from pre- and postoperative behavior changes on both a short- and long-term basis [19,99].

Conclusions

This meta-ethnographic synthesis developed 4 key themes that are important in determining the success of technologies to support behavior change. Our novel findings have the potential to influence the future design of patient-centered digital and mobile health technologies. This study demonstrates the important role of digital tools in the elective surgical pathway; not only can they help to motivate physical behavior change, such as improved activity levels and dietary intake, but they can also successfully provide psychological support. By synthesizing patient-informed suggestions, we have identified key areas for improvement, both to meet the general desires of patients undergoing surgery and to meet more specialized surgery-specific needs throughout the perioperative pathway. In particular, digital technologies should optimize the inclusion of tailored content specific to individual patients, with the inclusion of self-regulatory features such as goal setting to provide structured and individualized support. We believe that there is a significant rationale for involving patients in the cocreation of digital health technologies to enhance engagement, better support behavior change, and improve overall surgical outcomes for patients.

Acknowledgments

This research contributes to Miss Anna Robinson's PhD work, which is supported by the Dr WE Harker PhD Studentship from Newcastle University.

Authors' Contributions

AR led the writing of this manuscript as part of her PhD doctoral candidate. UO contributed to this work as part of her Undergraduate Master of Pharmacy degree, overseen by AH and AR. Coauthors (SS and RS) commented on various drafts of this work. All authors approved the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data, including meta-ethnographic data synthesis.
[DOCX File , 108 KB - mhealth_v8i12e19237_app1.docx]

References

1. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. Mhealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014 May 16;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
2. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J 2016 May 7;37(18):1428-1438 [FREE Full text] [doi: 10.1093/eurheartj/ehv770] [Medline: 26873093]
3. Rathbone AP, Norris R, Parker P, Lindsley A, Robinson A, Baqir W, et al. Exploring the use of WhatsApp in out-of-hours pharmacy services: a multi-site qualitative study. Res Social Adm Pharm 2020 Apr;16(4):503-510. [doi: 10.1016/j.sapharm.2019.06.019] [Medline: 31278013]
4. Vennik FD, Adams SA, Faber MJ, Putters K. Expert and experiential knowledge in the same place: patients' experiences with online communities connecting patients and health professionals. Patient Educ Couns 2014 May;95(2):265-270. [doi: 10.1016/j.pec.2014.02.003] [Medline: 24598314]
5. -. Correction to lancet infectious diseases 2020; published online april 29. Https://doi.Org/10.1016/ s1473-3099(20)30064-5. Lancet Infect Dis 2020 Jul;20(7):e148 [FREE Full text] [doi: 10.1016/S1473-3099(20)30370-4] [Medline: 32595044]
6. Tenhagen M, van Ramshorst GH, Demirkiran A, Hunfeld MA, Cense HA. Perioperative online weight monitoring in bariatric surgery with a digital internet-connected scale. Obes Surg 2016 May;26(5):1120-1126. [doi: 10.1007/s11695-016-2136-x] [Medline: 26994636]
7. Waller A, Forshaw K, Carey M, Robinson S, Kerridge R, Proietto A, et al. Optimizing patient preparation and surgical experience using ehealth technology. JMIR Med Inform 2015 Sep 1;3(3):e29 [FREE Full text] [doi: 10.2196/medinform.4286] [Medline: 26330206]
8. Ferry MD, Richards C. Biopedagogy digitalized: 'educational' relations among participants on an online weight loss surgery forum. Crit Public Health 2014 Jul 25;25(1):63-77. [doi: 10.1080/09581596.2014.940849]
9. van der Meij E, Bouwsma EV, van den 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]
10. Arora RC, Brown CH, Sanjanwala RM, McKelvie R. 'New' prehabilitation: a 3-way approach to improve postoperative survival and health-related quality of life in cardiac surgery patients. Can J Cardiol 2018 Jul;34(7):839-849. [doi: 10.1016/j.cjca.2018.03.020] [Medline: 29960613]
11. Chughtai M, Shah NV, Sultan AA, Solow M, Tiberi JV, Mehran N, et al. The role of prehabilitation with a telerehabilitation system prior to total knee arthroplasty. Ann Transl Med 2019 Feb;7(4):68. [doi: 10.21037/atm.2018.11.27] [Medline: 30963063]
12. Wynter-Blyth V, Moorthy K. Prehabilitation: preparing patients for surgery. Br Med J 2017 Aug 8;358:j3702. [doi: 10.1136/bmj.j3702] [Medline: 28790033]
13. Doiron-Cadrin P, Kairy D, Vendittoli P, Lowry V, Poitras S, Desmeules F. Feasibility and preliminary effects of a tele-prehabilitation program and an in-person prehablitation program compared to usual care for total hip or knee arthroplasty candidates: a pilot randomized controlled trial. Disabil Rehabil 2020 Apr;42(7):989-998. [doi: 10.1080/09638288.2018.1515992] [Medline: 30638076]
14. Beleigoli AM, Queiroz de Andrade A, Haueisen Diniz MD, Alvares RS, Ribeiro AL. Online platform for healthy weight loss in adults with overweight and obesity - the 'POEmaS' project: a randomized controlled trial. BMC Public Health 2018 Aug 1;18(1):945 [FREE Full text] [doi: 10.1186/s12889-018-5882-y] [Medline: 30068393]

15. Levett DZ, Edwards M, Grocott M, Mythen M. Preparing the patient for surgery to improve outcomes. Best Pract Res Clin Anaesthesiol 2016 Jun;30(2):145-157. [doi: 10.1016/j.bpa.2016.04.002] [Medline: 27396803]
16. Pecorelli N, Fiore JF, Kaneva P, Somasundram A, Charlebois P, Liberman AS, et al. An app for patient education and self-audit within an enhanced recovery program for bowel surgery: a pilot study assessing validity and usability. Surg Endosc 2018 May;32(5):2263-2273. [doi: 10.1007/s00464-017-5920-3] [Medline: 29098431]
17. van der Meij E, Huirne JA, Ten Cate AD, Stockmann HB, Scholten PC, Davids PH, et al. A perioperative ehealth program to enhance postoperative recovery after abdominal surgery: process evaluation of a randomized controlled trial. J Med Internet Res 2018 Jan 2;20(1):e1 [FREE Full text] [doi: 10.2196/jmir.8338] [Medline: 29295808]
18. van Kasteren Y, Freyne J, Hussain MS. Total knee replacement and the effect of technology on cocreation for improved outcomes and delivery: qualitative multi-stakeholder study. J Med Internet Res 2018 Mar 20;20(3):e95 [FREE Full text] [doi: 10.2196/jmir.7541] [Medline: 29559424]
19. McDonald S, Yates D, Durrand JW, Kothmann E, Sniehotta FF, Habgood A, et al. Exploring patient attitudes to behaviour change before surgery to reduce peri-operative risk: preferences for short- vs. long-term behaviour change. Anaesthesia 2019 Dec;74(12):1580-1588. [doi: 10.1111/anae.14826] [Medline: 31637700]
20. Grocott MP, Edwards M, Mythen MG, Aronson S. Peri-operative care pathways: re-engineering care to achieve the 'triple aim'. Anaesthesia 2019 Jan;74(Suppl 1):90-99 [FREE Full text] [doi: 10.1111/anae.14513] [Medline: 30604413]
21. Robinson A, Husband AK, Slight RD, Slight SP. Effectiveness of digital technologies to support surgical patients in changing their health behaviours: a systematic review. BJS Open 2020:- (forthcoming). [doi: 10.1093/bjsopen/zraa009]
22. 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]
23. Williams K, Beeken RJ, Fisher A, Wardle J. Health professionals' provision of lifestyle advice in the oncology context in the United Kingdom. Eur J Cancer Care (Engl) 2015 Jul;24(4):522-530. [doi: 10.1111/ecc.12305] [Medline: 25732397]
24. Anderson AS, Steele R, Coyle J. Lifestyle issues for colorectal cancer survivors--perceived needs, beliefs and opportunities. Support Care Cancer 2013 Jan;21(1):35-42. [doi: 10.1007/s00520-012-1487-7] [Medline: 22773297]
25. Phillips SM, Alfano CM, Perna FM, Glasgow RE. Accelerating translation of physical activity and cancer survivorship research into practice: recommendations for a more integrated and collaborative approach. Cancer Epidemiol Biomarkers Prev 2014 May;23(5):687-699 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-1355] [Medline: 24599577]
26. Bluethmann SM, Vernon SW, Gabriel KP, Murphy CC, Bartholomew LK. Taking the next step: a systematic review and meta-analysis of physical activity and behavior change interventions in recent post-treatment breast cancer survivors. Breast Cancer Res Treat 2015 Jan;149(2):331-342 [FREE Full text] [doi: 10.1007/s10549-014-3255-5] [Medline: 25555831]
27. Vehko T, Hyppönen H, Puttonen S, Kujala S, Ketola E, Tuukkanen J, et al. Experienced time pressure and stress: electronic health records usability and information technology competence play a role. BMC Med Inform Decis Mak 2019 Aug 14;19(1):160 [FREE Full text] [doi: 10.1186/s12911-019-0891-z] [Medline: 31412859]
28. 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]
29. Sudan R, Salter M, Lynch T, Jacobs DO. Bariatric surgery using a network and teleconferencing to serve remote patients in the Veterans Administration Health Care System: feasibility and results. Am J Surg 2011 Jul;202(1):71-76. [doi: 10.1016/j.amjsurg.2010.06.016] [Medline: 21601825]
30. 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]
31. Noblit GW, Hare RD. Meta-Ethnography: Synthesizing Qualitative Studies. California, USA: SAGE Publishing; 1988.
32. Britten N, Campbell R, Pope C, Donovan J, Morgan M, Pill R. Using meta ethnography to synthesise qualitative research: a worked example. J Health Serv Res Policy 2002 Oct;7(4):209-215. [doi: 10.1258/135581902320432732] [Medline: 12425780]
33. Campbell R, Pound P, Morgan M, Daker-White G, Britten N, Pill R, et al. Evaluating meta-ethnography: systematic analysis and synthesis of qualitative research. Health Technol Assess 2011 Dec;15(43):1-164 [FREE Full text] [doi: 10.3310/hta15430] [Medline: 22176717]
34. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, et al. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. Soc Sci Med 2003 Feb;56(4):671-684. [doi: 10.1016/s0277-9536(02)00064-3] [Medline: 12560003]
35. Pound P, Britten N, Morgan M, Yardley L, Pope C, Daker-White G, et al. Resisting medicines: a synthesis of qualitative studies of medicine taking. Soc Sci Med 2005 Jul;61(1):133-155. [doi: 10.1016/j.socscimed.2004.11.063] [Medline: 15847968]
36. France EF, Cunningham M, Ring N, Uny I, Duncan EA, Jepson RG, et al. Improving reporting of meta-ethnography: the eMERGe reporting guidance. BMC Med Res Methodol 2019 Jan 31;19(1):25 [FREE Full text] [doi: 10.1186/s12874-018-0600-0] [Medline: 30709371]
37. CASP Qualitative Checklist Online. Critical Appraisal Skills Programme. 2019. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2019-01-02]

38. Shaffer K, Ingersoll K, Chow P, Thorndike F, Bailey E, Shepard J, et al. Timing and tailoring of internet-based cognitive-behavioral treatment for insomnia for cancer survivors: A qualitative study. Psychooncology 2019 Sep;28(9):1934-1937 [FREE Full text] [doi: 10.1002/pon.5180] [Medline: 31319444]
39. 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 [FREE Full text] [doi: 10.1007/s11764-019-00751-3] [Medline: 30912011]
40. Alberts NM, Hadjistavropoulos HD, Titov N, Dear BF. Patient and provider perceptions of internet-delivered cognitive behavior therapy for recent cancer survivors. Support Care Cancer 2018 Feb;26(2):597-603. [doi: 10.1007/s00520-017-3872-8] [Medline: 28913732]
41. Gell NM, Tursi A, Grover KW, Dittus K. Female cancer survivor perspectives on remote intervention components to support physical activity maintenance. Support Care Cancer 2020 May;28(5):2185-2194. [doi: 10.1007/s00520-019-05038-y] [Medline: 31422476]
42. Lally RM, Eisenhauer C, Buckland S, Kupzyk K. Feasibility of synchronous online focus groups of rural breast cancer survivors on web-based distress self-management. Oncol Nurs Forum 2018 Nov 1;45(6):E111-E124 [FREE Full text] [doi: 10.1188/18.ONF.E111-E124] [Medline: 30339151]
43. Rosenberg D, Kadokura EA, Bouldin ED, Miyawaki CE, Higano CS, Hartzler AL. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. AMIA Annu Symp Proc 2016;2016:1050-1059 [FREE Full text] [Medline: 28269902]
44. Webb J, Peel J, Fife-Schaw C, Ogden J. A mixed methods process evaluation of a print-based intervention supported by internet tools to improve physical activity in UK cancer survivors. Public Health 2019 Oct;175:19-27 [FREE Full text] [doi: 10.1016/j.puhe.2019.06.013] [Medline: 31374452]
45. 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]
46. Clarke AL, Roscoe J, Appleton R, Dale J, Nanton V. 'My gut feeling is we could do more...' a qualitative study exploring staff and patient perspectives before and after the implementation of an online prostate cancer-specific holistic needs assessment. BMC Health Serv Res 2019 Feb 12;19(1):115 [FREE Full text] [doi: 10.1186/s12913-019-3941-4] [Medline: 30755188]
47. 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]
48. Kairy D, Tousignant M, Leclerc N, Côté AM, Levasseur M, Researchers TT. The patient's perspective of in-home telerehabilitation physiotherapy services following total knee arthroplasty. Int J Environ Res Public Health 2013 Aug 30;10(9):3998-4011 [FREE Full text] [doi: 10.3390/ijerph10093998] [Medline: 23999548]
49. Atwood ME, Friedman A, Meisner BA, Cassin SE. The exchange of social support on online bariatric surgery discussion forums: a mixed-methods content analysis. Health Commun 2018 May;33(5):628-635. [doi: 10.1080/10410236.2017.1289437] [Medline: 28281790]
50. 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]
51. 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;13(12):e0210039 [FREE Full text] [doi: 10.1371/journal.pone.0210039] [Medline: 30596781]
52. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Wearable sensor-based exercise biofeedback for orthopaedic rehabilitation: a mixed methods user evaluation of a prototype system. Sensors (Basel) 2019 Jan 21;19(2):- [FREE Full text] [doi: 10.3390/s19020432] [Medline: 30669657]
53. Das A, Faxvaag A. What influences patient participation in an online forum for weight loss surgery? A qualitative case study. Interact J Med Res 2014 Feb 6;3(1):e4 [FREE Full text] [doi: 10.2196/ijmr.2847] [Medline: 24509408]
54. Lee M, Suh D, Son J, Kim J, Eun S, Yoon B. Patient perspectives on virtual reality-based rehabilitation after knee surgery: importance of level of difficulty. J Rehabil Res Dev 2016;53(2):239-252 [FREE Full text] [doi: 10.1682/JRRD.2014.07.0164] [Medline: 27149529]
55. Zhu J, Ebert L, Guo D, Yang S, Han Q, Chan SW. Mobile breast cancer e-support program for chinese women with breast cancer undergoing chemotherapy (part 1): qualitative study of women's perceptions. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e85 [FREE Full text] [doi: 10.2196/mhealth.9311] [Medline: 29643056]
56. Bartlett YK, Coulson NS. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. Patient Educ Couns 2011 Apr;83(1):113-119. [doi: 10.1016/j.pec.2010.05.029] [Medline: 20599338]

57. 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 Apr 13;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
58. 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]
59. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322. [doi: 10.1093/her/cyr005] [Medline: 21321008]
60. Suls J, Mogavero JN, Falzon L, Pescatello LS, Hennessy EA, Davidson KW. Health behaviour change in cardiovascular disease prevention and management: meta-review of behaviour change techniques to affect self-regulation. Health Psychol Rev 2020 Mar;14(1):43-65. [doi: 10.1080/17437199.2019.1691622] [Medline: 31707938]
61. Ng JYY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. Perspect Psychol Sci 2012 Jul;7(4):325-340. [doi: 10.1177/1745691612447309] [Medline: 26168470]
62. Fortier MS, Duda JL, Guerin E, Teixeira PJ. Promoting physical activity: development and testing of self-determination theory-based interventions. Int J Behav Nutr Phys Act 2012 Mar 2;9:20 [FREE Full text] [doi: 10.1186/1479-5868-9-20] [Medline: 22385751]
63. Ritterband LM, Bailey ET, Thorndike FP, Lord HR, Farrell-Carnahan L, Baum LD. Initial evaluation of an Internet intervention to improve the sleep of cancer survivors with insomnia. Psychooncology 2012 Jul;21(7):695-705 [FREE Full text] [doi: 10.1002/pon.1969] [Medline: 21538678]
64. Willems RA, Bolman CA, Mesters I, Kanera IM, Beaulen AA, Lechner L. Short-term effectiveness of a web-based tailored intervention for cancer survivors on quality of life, anxiety, depression, and fatigue: randomized controlled trial. Psychooncology 2017 Feb;26(2):222-230. [doi: 10.1002/pon.4113] [Medline: 26988800]
65. Wootten AC, Abbott JM, Meyer D, Chisholm K, Austin DW, Klein B, et al. Preliminary results of a randomised controlled trial of an online psychological intervention to reduce distress in men treated for localised prostate cancer. Eur Urol 2015 Sep;68(3):471-479. [doi: 10.1016/j.eururo.2014.10.024] [Medline: 25454611]
66. 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]
67. Khan S, Dasrath F, Farghaly S, Otobo E, Riaz MS, Rogers J, Health PROMISE Consortium Group. Unmet communication and information needs for patients with IBD: implications for mobile health technology. Br J Med Med Res 2016;12(3):- [FREE Full text] [doi: 10.9734/BJMMR/2016/21884] [Medline: 28217532]
68. de Jong CC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014 Jan 16;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
69. Cook DJ, Thompson JE, Prinsen SK, Dearani JA, Deschamps C. Functional recovery in the elderly after major surgery: assessment of mobility recovery using wireless technology. Ann Thorac Surg 2013 Sep;96(3):1057-1061. [doi: 10.1016/j.athoracsur.2013.05.092] [Medline: 23992697]
70. Lelorain S, Brédart S, Dolbeault S, Sultan S. A systematic review of the associations between empathy measures and patient outcomes in cancer care. Psychooncology 2012 Dec;21(12):1255-1264. [doi: 10.1002/pon.2115] [Medline: 22238060]
71. 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]
72. 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 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
73. 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):e14 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]
74. Sanders MR, Winters P, Fortuna RJ, Mendoza M, Berliant M, Clark L, et al. Internet access and patient portal readiness among patients in a group of inner-city safety-net practices. J Ambul Care Manage 2013;36(3):251-259. [doi: 10.1097/JAC.0b013e31829702f9] [Medline: 23748274]
75. Baillot A, Boissy P, Tousignant M, Langlois M. Feasibility and effect of in-home physical exercise training delivered via telehealth before bariatric surgery. J Telemed Telecare 2017 Jun;23(5):529-535. [doi: 10.1177/1357633X16653511] [Medline: 27279467]
76. Schlegel RJ, Manning MA, Molix LA, Talley AE, Bettencourt BA. Predictors of depressive symptoms among breast cancer patients during the first year post diagnosis. Psychol Health 2012;27(3):277-293 [FREE Full text] [doi: 10.1080/08870446.2011.559232] [Medline: 22404698]

77. Weaver KE, Geiger AM, Lu L, Case LD. Rural-urban disparities in health status among US cancer survivors. Cancer 2013 Mar 1;119(5):1050-1057 [FREE Full text] [doi: 10.1002/cncr.27840] [Medline: 23096263]
78. Russell TG, Buttrum P, Wootton R, Jull GA. Internet-based outpatient telerehabilitation for patients following total knee arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2011 Jan 19;93(2):113-120. [doi: 10.2106/JBJS.I.01375] [Medline: 21248209]
79. Tanis M. Health-related on-line forums: what's the big attraction? J Health Commun 2008;13(7):698-714. [doi: 10.1080/10810730802415316] [Medline: 18958781]
80. Elfhag K, Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. Obes Rev 2005 Feb;6(1):67-85. [doi: 10.1111/j.1467-789X.2005.00170.x] [Medline: 15655039]
81. Livhits M, Mercado C, Yermilov I, Parikh JA, Dutson E, Mehran A, et al. Is social support associated with greater weight loss after bariatric surgery?: a systematic review. Obes Rev 2011 Feb;12(2):142-148. [doi: 10.1111/j.1467-789X.2010.00720.x] [Medline: 20158617]
82. Han JY, Hou J, Kim E, Gustafson DH. Lurking as an active participation process: a longitudinal investigation of engagement with an online cancer support group. Health Commun 2014;29(9):911-923 [FREE Full text] [doi: 10.1080/10410236.2013.816911] [Medline: 24345206]
83. Huber J, Muck T, Maatz P, Keck B, Enders P, Maatouk I, et al. Face-to-face vs online peer support groups for prostate cancer: a cross-sectional comparison study. J Cancer Surviv 2018 Feb;12(1):1-9. [doi: 10.1007/s11764-017-0633-0] [Medline: 28861845]
84. 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]
85. Li R, Suh A. Factors influencing information credibility on social media platforms: evidence from Facebook pages. Procedia Comput Sci 2015 Jan;72:314-328 [FREE Full text] [doi: 10.1016/j.procs.2015.12.146]
86. Koball AM, Jester DJ, Pruitt MA, Cripe RV, Henschied JJ, Domoff S. Content and accuracy of nutrition-related posts in bariatric surgery Facebook support groups. Surg Obes Relat Dis 2018 Dec;14(12):1897-1902. [doi: 10.1016/j.soard.2018.08.017] [Medline: 30274739]
87. Lindsay S, Smith S, Bellaby P, Baker R. The health impact of an online heart disease support group: a comparison of moderated versus unmoderated support. Health Educ Res 2009 Aug;24(4):646-654. [doi: 10.1093/her/cyp001] [Medline: 19251770]
88. 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]
89. Demark-Wahnefried W, Aziz NM, Rowland JH, Pinto BM. Riding the crest of the teachable moment: promoting long-term health after the diagnosis of cancer. J Clin Oncol 2005 Aug 20;23(24):5814-5830 [FREE Full text] [doi: 10.1200/JCO.2005.01.230] [Medline: 16043830]
90. 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 (Lond) 2020;9:2 [FREE Full text] [doi: 10.1186/s13741-019-0133-z] [Medline: 32042404]
91. Owen JE, Bantum EO, Gorlick A, Stanton AL. Engagement with a social networking intervention for cancer-related distress. Ann Behav Med 2015 Apr;49(2):154-164 [FREE Full text] [doi: 10.1007/s12160-014-9643-6] [Medline: 25209353]
92. Wurz A, St-Aubin A, Brunet J. Breast cancer survivors' barriers and motives for participating in a group-based physical activity program offered in the community. Support Care Cancer 2015 Aug;23(8):2407-2416 [FREE Full text] [doi: 10.1007/s00520-014-2596-2] [Medline: 25605568]
93. LaVela SL, Schectman G, Gering J, Locatelli SM, Gawron A, Weaver FM. Understanding health care communication preferences of veteran primary care users. Patient Educ Couns 2012 Sep;88(3):420-426. [doi: 10.1016/j.pec.2012.06.004] [Medline: 22766452]
94. DuBenske LL, Gustafson DH, Namkoong K, Hawkins RP, Atwood AK, Brown RL, et al. CHESS improves cancer caregivers' burden and mood: results of an eHealth RCT. Health Psychol 2014 Oct;33(10):1261-1272 [FREE Full text] [doi: 10.1037/a0034216] [Medline: 24245838]
95. Kaltenbaugh DJ, Klem ML, Hu L, Turi E, Haines AJ, Hagerty Lingler J. Using Web-based interventions to support caregivers of patients with cancer: a systematic review. Oncol Nurs Forum 2015 Mar;42(2):156-164. [doi: 10.1188/15.ONF.156-164] [Medline: 25806882]
96. Heynsbergh N, Botti M, Heckel L, Livingston PM. Caring for the person with cancer: information and support needs and the role of technology. Psychooncology 2018 Jun;27(6):1650-1655 [FREE Full text] [doi: 10.1002/pon.4722] [Medline: 29624783]
97. Northouse L, Williams A, Given B, McCorkle R. Psychosocial care for family caregivers of patients with cancer. J Clin Oncol 2012 Apr 10;30(11):1227-1234. [doi: 10.1200/JCO.2011.39.5798] [Medline: 22412124]

98. Scott K, Beatty L. Feasibility study of a self-guided cognitive behaviour therapy internet intervention for cancer carers. Aust J Prim Health 2013;19(4):270-274. [doi: 10.1071/PY13025] [Medline: 23796231]
99. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]

Abbreviations

CASP: Critical Appraisal Skills Programme
iCBT: internet-delivered cognitive behavioral therapy
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by L Buis; submitted 09.04.20; peer-reviewed by MA Bahrami, D Szinay, M Lingg, M Nieroda, V Addison; comments to author 01.09.20; revised version received 22.09.20; accepted 29.10.20; published 01.12.20.

Please cite as:
Robinson A, Oksuz U, Slight R, Slight S, Husband A
Digital and Mobile Technologies to Promote Physical Health Behavior Change and Provide Psychological Support for Patients Undergoing Elective Surgery: Meta-Ethnography and Systematic Review
JMIR Mhealth Uhealth 2020;8(12):e19237
URL: https://mhealth.jmir.org/2020/12/e19237
doi:10.2196/19237
PMID:33258787

Review

Virtual Reality Systems for Upper Limb Motor Function Recovery in Patients With Spinal Cord Injury: Systematic Review and Meta-Analysis

Amaranta De Miguel-Rubio[1], PT, MSc; M Dolores Rubio[2], PhD; Alvaro Alba-Rueda[1], PT, BSc; Alejandro Salazar[3,4,5], PhD; Jose A Moral-Munoz[4,5,6], PT, PhD; David Lucena-Anton[6], PT, PhD

[1]Department of Nursing, Pharmacology and Physiotherapy, University of Cordoba, Cordoba, Spain

[2]Department of Cell Biology, Physiology and Immunology, University of Cordoba, Cordoba, Spain

[3]Department of Statistics and Operational Research, University of Cadiz, Cadiz, Spain

[4]Institute of Research and Innovation in Biomedical Sciences of the Province of Cadiz (INiBICA), University of Cadiz, Cadiz, Spain

[5]The Observatory of Pain, University of Cadiz, Cadiz, Spain

[6]Department of Nursing and Physiotherapy, University of Cadiz, Cadiz, Spain

Corresponding Author:
Jose A Moral-Munoz, PT, PhD
Department of Nursing and Physiotherapy
University of Cadiz
Avda. Ana de Viya, 52
Cadiz, 11009
Spain
Phone: 34 956875962
Email: joseantonio.moral@uca.es

Abstract

Background: Patients with spinal cord injury (SCI) usually present with different motor impairments, including a deterioration of upper limb motor function (ULMF), that limit their performance of activities of daily living and reduce their quality of life. Virtual reality (VR) is being used in neurological rehabilitation for the assessment and treatment of the physical impairments of this condition.

Objective: A systematic review and meta-analysis was conducted to evaluate the effectiveness of VR on ULMF in patients with SCI compared with conventional physical therapy.

Methods: The search was performed from October to December 2019 in Embase, Web of Science, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Scopus, Medline, Physiotherapy Evidence Database (PEDro), PubMed, and Cochrane Central Register of Controlled Trials. The inclusion criteria of selected studies were as follows: (1) comprised adults with SCI, (2) included an intervention with VR, (3) compared VR intervention with conventional physical therapy, (4) reported outcomes related to ULMF, and (5) was a controlled clinical trial. The Cochrane Collaboration's tool was used to evaluate the risk of bias. The RevMan 5.3 statistical software was used to obtain the meta-analysis according to the standardized mean difference (SMD) and 95% CIs.

Results: Six articles were included in this systematic review. Four of them contributed information to the meta-analysis. A total of 105 subjects were analyzed. All of the studies used semi-immersive or nonimmersive VR systems. The statistical analysis showed nonsignificant results for the Nine-Hole Peg Test (SMD –0.93, 95% CI –1.95 to 0.09), muscle balance test (SMD –0.27, 95% CI –0.82 to 0.27), Motricity Index (SMD 0.16, 95% CI −0.37 to 0.68), Jebsen-Taylor Hand Function Test (JTHFT) subtests (writing, SMD –0.10, 95% CI –4.01 to 3.82; simulated page turning, SMD –0.99, 95% CI –2.01 to 0.02; simulated feeding, SMD –0.64, 95% CI –1.61 to 0.32; stacking checkers, SMD 0.99, 95% CI –0.02 to 2.00; picking up large light objects, SMD –0.42, 95% CI –1.37 to 0.54; and picking up large heavy objects, SMD 0.52, 95% CI –0.44 to 1.49), range of motion of shoulder abduction/adduction (SMD –0.23, 95% CI –1.48 to 1.03), shoulder flexion/extension (SMD 0.56, 95% CI –1.24 to 2.36), elbow flexion (SMD –0.36, 95% CI –1.14 to 0.42), elbow extension (SMD –0.21, 95% CI –0.99 to 0.57), wrist extension (SMD 1.44, 95% CI –2.19 to 5.06), and elbow supination (SMD –0.18, 95% CI –1.80 to 1.44). Favorable results were found for the JTHFT subtest picking up small common objects (SMD –1.33, 95% CI –2.42 to –0.24).

Conclusions: The current evidence for VR interventions to improve ULMF in patients with SCI is limited. Future studies employing immersive systems to identify the key aspects that increase the clinical impact of VR interventions are needed, as well as research to prove the benefits of the use of VR in the rehabilitation of patients with SCI in the clinical setting.

(JMIR Mhealth Uhealth 2020;8(12):e22537) doi:10.2196/22537

KEYWORDS

virtual reality; spinal cord injuries; neurological rehabilitation; motor function; physical therapy

Introduction

The global estimate of the spinal cord injury (SCI) prevalence is 223 to 755 per million people and the worldwide incidence is 10.4 to 83 per million people per year [1]. SCI produces a high impact on the health care system and society [2]. Patients with SCI usually present with different motor impairments, including a deterioration of upper limb motor function (ULMF), causing an important limitation in the performance of activities of daily living and a loss of quality of life [2,3].

Neurological rehabilitation benefits from virtual reality (VR), which is being used for the assessment and treatment of physical and cognitive impairments, for pain management, and even to acquire surgical skills [4]. This technology is becoming more portable, immersive, and vivid, making it more suitable for a wider range of clinical applications [5]. VR systems allow the creation of virtual environments that can be used to practice, under controlled conditions, different activities that could be hazardous in a real-world setting [6]. Different characteristics such as difficulty, intensity, exposure duration, and feedback can be adjusted to provide personalized experiences [7]. Furthermore, VR and interactive video gaming are presented as a motivational therapy that could increase patient adherence to treatment [8,9]. VR-based interventions are usually provided by commercial devices such as Nintendo Wii [10], PlayStation [11], and Xbox Kinect [12], among others. According to the level of immersion, VR systems can be divided into immersive, semi-immersive, or nonimmersive systems. Immersive systems provide a full integration into the virtual environment that is delivered through head-mounted displays and VR caves. Semi-immersive and nonimmersive systems consist of displaying the environment through a screen and these systems are usually used in video game consoles. Furthermore, VR systems can be combined with different devices such as gloves, electrical stimulation devices, and exoskeletons [13].

Several studies on the use of VR interventions have been carried out in different neurological disorders, such as stroke [14-17], cerebral palsy [18,19], Parkinson disease [20,21], and multiple sclerosis [22-24]. Nevertheless, the odds of a successful recovery are different for each disease, so the results obtained by VR interventions could be different as well. Specifically, patients with SCI usually suffer from significant participation restrictions [25], so the physical treatment should be focused on keeping the residual functionality after SCI [26].

SCI occurs with greater frequency at the cervical and thoracic levels than at lumbosacral levels. Patients with cervical and thoracic SCI can suffer loss of arm and hand function and consequently reduce significantly their autonomy and independence. However, small improvements in arm and hand function could improve the performance of the activities of daily living, independence, and quality of life, and thus recovering ULMF in patients with cervical and thoracic levels of SCI is a primary challenge [3]. In this sense, the scientific evidence through systematic reviews and meta-analyses on the potential use of VR systems to recover ULMF in patients with SCI is limited. de Araújo et al [27] stated that VR therapy could be used to improve motor function. A structured review performed by Yeo et al [28] concluded that VR therapy provides benefits on balance and posture. Conversely, a recent meta-analysis published by our group [29] suggested that VR interventions mat not be effective to improve the functional performance after SCI. Nevertheless, the previous reviews were not restricted specifically to the assessment of ULMF. We hypothesize that VR therapy could stimulate patients' attention and motivation, making the intervention more effective than conventional physical therapy (CPT). Therefore, the main objective of this systematic review and meta-analysis was to evaluate the effectiveness of VR interventions in the recovery of ULMF in patients with SCI.

Methods

Search Strategy

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [30] were followed to perform this systematic review. The search protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) database (CRD 42018093855). The literature search was performed between October and December 2019 in the following electronic databases: Embase, Web of Science, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Scopus, Medline, Physiotherapy Evidence Database (PEDro), PubMed, and Cochrane Central Register of Controlled Trials (CENTRAL). The following descriptor terms combined with Boolean operators were employed: ("spinal cord injuries" OR "spinal cord injury" OR "quadriplegia" OR "paraplegia" OR "tetraplegia") AND ("virtual reality exposure therapy" OR "virtual reality" OR "augmented reality" OR "virtual systems" OR "video games" OR "videogame" OR "exergaming" OR "exergames" OR "commercial games" OR "play-based therapy"). Medical Subject Headings (MeSH) descriptors were used in PubMed database: "virtual reality exposure therapy," "virtual reality," "spinal cord injuries," and "video games." The search was filtered to include full-text clinical trials papers. No date or language filters were applied.

Selection Criteria

The Population, Intervention, Comparison, Outcomes and Study (PICOS) design model was used to establish the article inclusion criteria: (1) population: adults with SCI; (2) intervention: VR interventions; (3) comparison: adults with SCI performing CPT; (4) outcome: outcomes specifically related to ULMF, such as muscle strength, range of motion (ROM), dexterity, grasp and pinch force, and hand function; and (5) study design: controlled clinical trials. Articles of studies which included participants with SCI and other pathologies but did not provide the outcome data for each specific population were excluded.

Study Selection Process and Data Extraction

The literature search was carried out by combining keywords in the scientific databases mentioned above and duplicated articles were excluded. Next, titles and abstracts were reviewed, and we excluded those articles that did not meet the established inclusion criteria. The remaining articles were analyzed strictly and were finally included in the systematic review. Two reviewers (ADM-R and MDR) took part independently in the study selection process, review, and systematic data extraction. A third reviewer (DLA) took part in achieving consensus in cases of dispute. The following data were extracted from the studies: (1) author and date of publication; (2) number and age of participants, levels of injury, and mean time post onset; (3) and characteristics of the interventions (intervention types in each group, outcome measures, measuring instrument) and results.

Tools for Assessing the Risk of Bias of the Studies

The Cochrane Collaboration's tool [31] and Review Manager (RevMan) software (version 5.3; The Nordic Cochrane Centre, The Cochrane Collaboration, 2014), which includes a description and evaluation of each item by means of a bias table, were used to assess the risk of bias. After assessing the risk of bias of each study, the studies were categorized as "low risk," "high risk," or "unclear risk." Two reviewers carried out the assessment independently. In cases of doubt, the final decision was determined through discussion including a third reviewer.

Statistical Analysis

The meta-analysis compared CPT with VR interventions. The studies were divided into subgroups based on the measuring instrument that was used in the study. If more than one instrument was used in the same study, we included the study in more than one subgroup. The differences in the effect size (postintervention minus preintervention) between the groups were analyzed in terms of the standardized mean difference. The confidence level was set at 95% (significance at $P<0.05$). Results are shown along with 95% CIs.

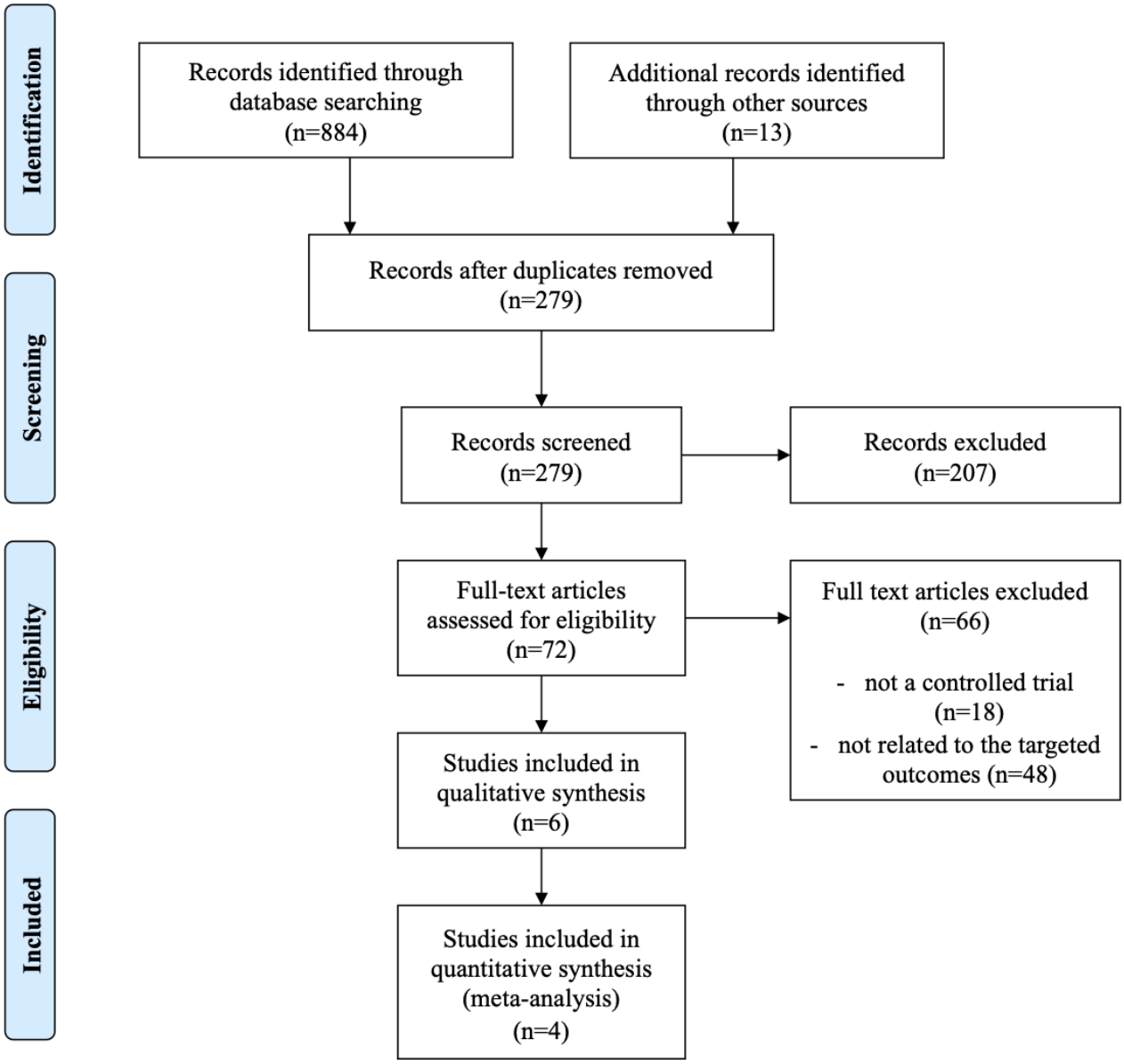
The chi-square test and the I^2 statistic (percentage of variation across studies that is due to heterogeneity) were used to test the homogeneity, using a fixed-effect model in the case of homogeneity and a random-effects model otherwise.

The analyses were performed using RevMan 5.3 software, and the results are presented in tables and forest plots.

Results

A total of 279 potentially relevant articles were retrieved after the selection process, as shown in Figure 1. A total of 6 studies were included in the systematic review. Four of them were included in the meta-analysis for statistical comparison.

Figure 1. Information flow diagram of the selection process of the systematic review and meta-analysis.



Assessment of the Risk of Bias of the Studies Included in the Review

Regarding the risk of bias of the studies included in this review, the studies conducted by Kowalczewski et al [32], Dimbwadyo-Terrer et al [3], and Prasad et al [33] presented the lowest risk of bias, as shown in Figure 2. Furthermore, concerning the risk of bias among the studies analyzed, the lowest biases were found in the selective reporting (0%) and the incomplete outcome data (0%). The highest value (85.5%) was found in the allocation concealment, as shown in Figure 3.

Figure 2. Risk of bias of the studies included in the systematic review. The green circle (+) indicates low risk of bias, the yellow circle (?) unclear risk of bias, and the red circle (-) high risk of bias.

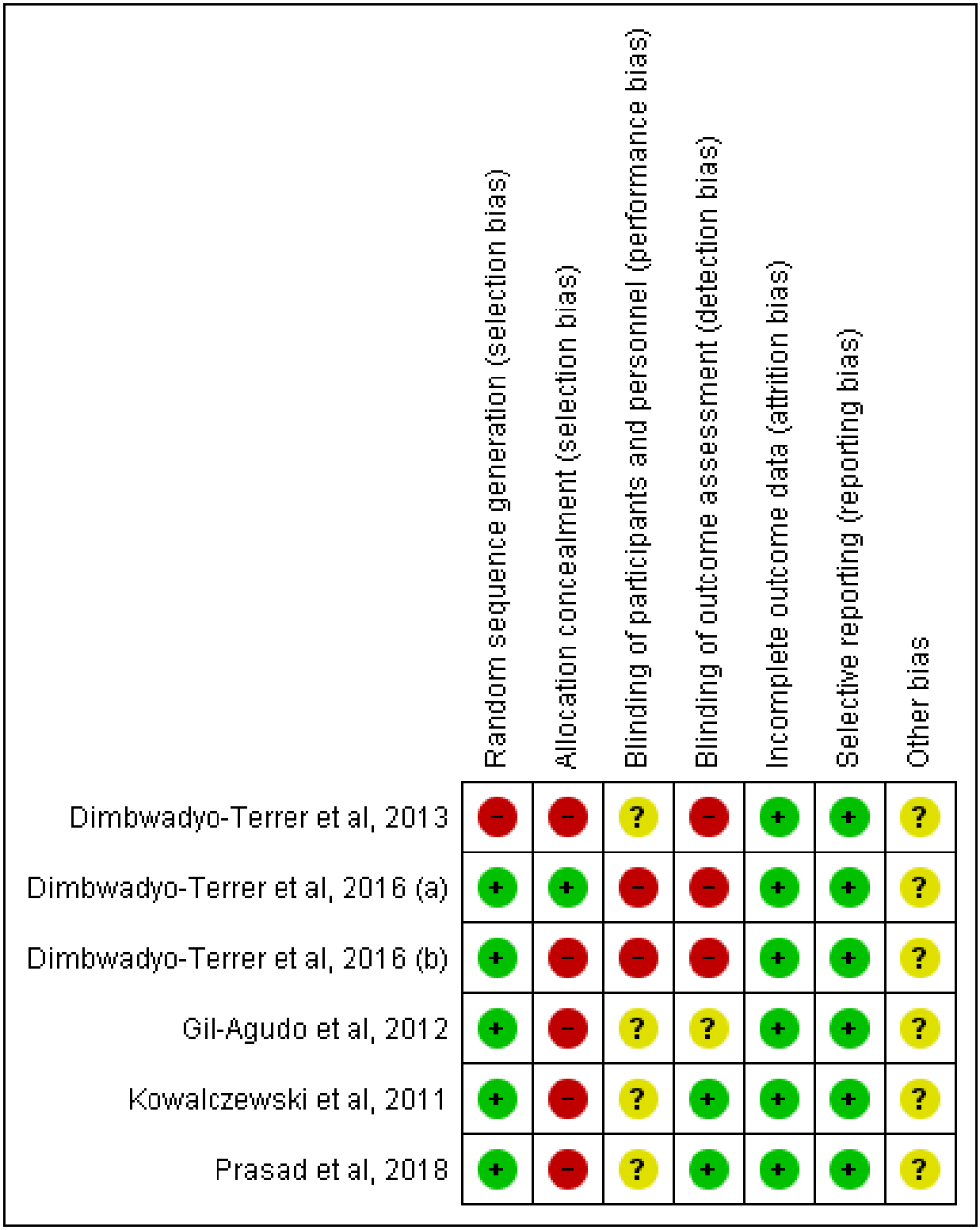
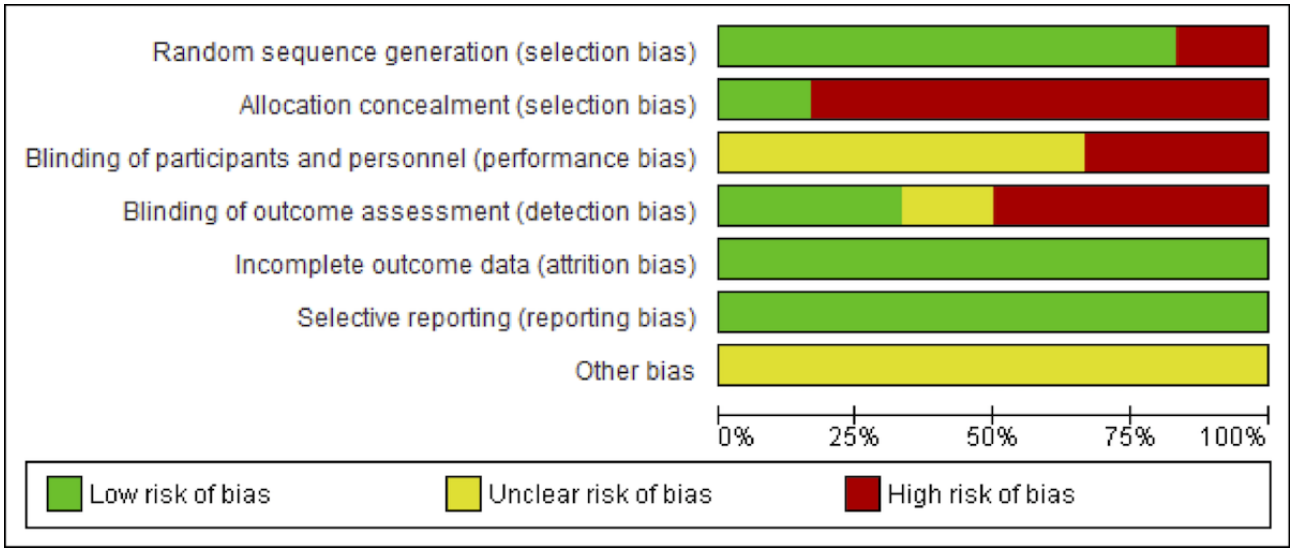


Figure 3. Overall risk of bias, with each category presented as percentages.



Data Extraction

A total of 103 subjects (control group, n=46; intervention group, n=57) took part in the different studies. A study by Dimbwadyo-Terrer et al [3] had the highest number of participants (n=31) and a different study by Dimbwadyo-Terrer et al [34] had the lowest sample size (n=9). The mean age of the participants ranged from 23.7 years [33] to 54.3 years [34]. Concerning the neurological level of injury, 4 studies [3,8,32,35] included participants with American Spinal Injury Association Impairment Scale (AIS) grades A-B injuries, while 2 studies [33,34] included participants with AIS grades A-D injuries. The main characteristics of the participants are shown in Table 1.

Table 1. Main characteristics of the participants in each study.

Study reference	Participants, n			Age (years), mean (SD)	AIS[c] grade	Level of injury	Mean time after disease onset (months)
	Total	CG[a]	IG[b]				
[32]	13	7	6	Overall: 35.9 (11.9)	A-B	C5-C7	Overall: 9.0
[8]	10	5	5	CG: 49.0 (6.1)	A-B	C5-C8	CG: 5.8
				IG: 36.2 (10.4)			IG: 4.2
[35]	18	6	12	CG: 42.0 (13.6)	A-B	C5-C8	CG: 3.6
				IG: 33.6 (14.1)			IG: 6.6
[3]	31	15	16	CG: 40.2 (13.6)	A-B	C5-C8	CG: 5.6
				IG: 34.5 (13.7)			IG: 4.3
[34]	9	3	6	CG: 44.2 (22.9)	A-D	T1-T6	CG: 5.0
				IG: 54.3 (9.9)			IG: 5.8
[33]	22	10	12	CG: 33.9 (7.1)	A-D	C5-C8	CG: 10.2
				IG: 23.7 (5.2)			IG: 15.2

[a]CG: control group.

[b]IG: intervention group.

[c]AIS: American Spinal Injury Association Impairment Scale.

Concerning the intervention protocols, the VR therapy was applied to the intervention groups via different technological devices while the comparison group performed CPT. The longest total duration of intervention and the highest intensity were achieved by Kowalczewski et al [32] (5 times/week for 6 weeks). In addition, the longest session duration (60 minutes) was achieved by Kowalczewski et al [32] and Prasad et al [33]. Conversely, the shortest intervention time and lowest intensity were achieved by Dimbwadyo-Terrer et al [34], who only performed 4 sessions (2 times/week for 2 weeks).

VR therapy was provided through different devices, such as the Rehabilitation Joystick for Computerized Exercise (ReJoyce) VR system (Saebo Inc) [32], Toyra system (National Paraplegics Hospital of Toledo and Rafael del Pino Foundation) [3,8,35], Nintendo Wii [33], and a mesh data glove [34]. All of the devices used to provide VR therapy were categorized into semi-immersive and nonimmersive VR types. The ReJoyce VR

system consists of a workstation where patients can play games shown on a screen through a segmented, jointed, spring-loaded arm. The Toyra system was used in 3 studies [3,8,35]. It reproduces the patient movements in real time through an avatar displayed on the screen, and patients can interact with different objects in the virtual environment [8]. Prasad et al [33] used the Nintendo Wii video game console. Finally, the study by Dimbwadyo-Terrer et al [34] used a data glove to interact with the virtual environment, allowing patients to manipulate virtual objects in real time.

Regarding the different deficits treated, all of the studies analyzed the effects of the VR intervention on ULMF. Moreover, the authors focused their interventions on improving upper limb ROM [8,32,35], upper limb strength [8], upper limb dexterity [33], grasp and pinch force [32], and functional performance [3,8,33-35]. Most studies reported no significant effects on the different outcomes analyzed. Only the study of Kowalczewski et al [32] showed benefits on all of the outcomes. Gil-Agudo et al [8] showed significant results on stacked checked subtest of the Jebsen-Taylor Hand Function Test (JTHFT) [36], and Dimbwadyo-Terrer et al [3] got significant benefits for muscle strength measured by the muscle balance test [37]. Table 2 shows the main characteristics of the different interventions performed by the different studies.

Table 2. Main characteristics of the interventions.

Study	Group interventions	Intensity	Session duration	Intervention duration	Outcomes	Measuring instrument	Results
[32]	CG[a]: CPT[b]; IG[c]: ReJoyce VR[d] system	5 x/wk[e]	60 min[f]	6 wks[g]	Upper limb motor function, ROM[h], functional tasks, grasp, and pinch forces	ARAT[i], ReJoyce automated hand function test	All outcomes showed statistically significant differences and clinically important improvements for IG.
[8]	CG: CPT; IG: Toyra VR system	3 x/wk	30 min	5 wks	Upper limb ROM, motor function and strength, and functional performance	NHPT[j], JTHFT[k], MI[l], BI[m], FIM[n], SCIM[o]	No significant differences were found between groups after intervention, except for JTHFT stacking checkers subtest (P=0.008).
[35]	CG: CPT; IG: Toyra VR system	4 x/wk	ND[p]	3 wks	Upper limb ROM and motor function, and functional performance	MI, MB[q], FIM, SCIM	No significant differences were found between groups after intervention.
[3]	CG: CPT; IG: Toyra VR system	3 x/wk	30 min	5 wks	Upper limb motor function and functional performance	MB, MI, FIM, SCIM, BI	No significant differences were found between groups after intervention. At follow-up, only MB was statistically improved (P=0.04).
[34]	CG: CPT; IG: VR system + CiberTouch data glove	2 x/wk	30 min	2 wks	Upper limb motor function and functional performance	NHPT, JTHFT, MB, SCIM	No significant differences were found between groups after intervention.
[33]	CG: CPT; IG: Nintendo Wii	3 x/wk	60 min	4 wks	Upper limb dexterity and motor function, and functional performance	CUE[r], BBT[s], SCIM, WHOQOL-BREF[t]	No significant differences were found between groups after intervention.

[a]CG: control group.

[b]CPT: conventional physical therapy.

[c]IG: intervention group.

[d]VR: virtual reality.

[e]x/wk: times/week.

[f]min: minutes.

[g]wks: weeks.

[h]ROM: range of motion.

[i]ARAT: Action Research Arm Test.

[j]NHPT: Nine-Hole Peg Test.

[k]JTHFT: Jebsen-Taylor Hand Function Test.

[l]MI: Motricity Index.

[m]BI: Barthel Index.

[n]FIM: Functional Independence Measure.

[o]SCIM: Spinal Cord Independence Measure.

[p]ND: not described.

[q]MB: muscle balance.

[r]CUE: Capabilities of Upper Extremity.

[s]BBT: Box and Block Test.

[t]WHOQOL-BREF: World Health Organization Quality of Life Scale, Abbreviated Version.

Instruments of Measurement Used in the Meta-Analysis

Different scales and tests were used in the studies to assess ULMF. The Nine-Hole Peg Test (NHPT) involves placing and removing pegs into 9 holes, and scores are based on the time taken to complete the activity. This scale is commonly used to measure fine manual dexterity [38]. Muscle balance (MB) tests are used to rate muscle strength, assigning a grade from 0 to 5 according to the strength of the muscle to face the gravity or the force applied by the examiner [37]. The Motricity Index (MI) measures the range and strength of active movements and each movement is rated on a point scale from 0 to 100 [39]. The JTHFT assesses the time (in seconds) spent to perform different tasks related to hand functioning, which are commonly used in activities of daily living, and it comprises 7 subtests (writing,

simulated page turning, picking up small common objects, simulated feeding, stacking checkers, picking up large light objects, and picking up large heavy objects) [36]. Finally, ROM tests consist of measuring joint mobility using a goniometer [40]. A total of 4 studies were included in the meta-analysis.

Two studies [8,34] used the NHPT to analyze ULMF. According to the I^2 statistic, 0% of variation across studies was due to heterogeneity. This homogeneity was confirmed by the chi-square test (P=0.41). A fixed-effect model was fitted. The study by Gil-Agudo et al [8] obtained the best results. We observed that VR therapy turned out to be more effective than CPT. However, the overall result of this meta-analysis was not conclusive, as shown in Figure 4.

Three studies [3,34,35] analyzed the effects of VR interventions using the results obtained from the MB test. In this group, I^2=56%, although the chi-square test (P=0.10) showed homogeneity, and a fixed-effect model was fitted. Favorable results for VR interventions were obtained in the study by Dimbwadyo-Terrer et al [34]. However, none of these results were statistically significant. The overall result of this meta-analysis was not conclusive, as shown in Figure 5.

Three studies [3,8,35] used the MI to assess ULMF. As with the studies that used the NHPT to analyze ULMF, 0% of the variation across studies was due to heterogeneity (I^2=0%), and the chi-square test confirmed that finding (P=0.89). A fixed-effect model was fitted. All the studies showed favorable results for VR interventions. However, none of these results were statistically significant. The overall result of this meta-analysis was not conclusive, as shown in Figure 6.

Figure 4. Forest plot for upper limb motor function measured by the Nine-Hole Peg Test. The green blocks indicate the weight assigned to the study, the horizontal line depicts the CI, and the black rhombus shows the overall result. IV: inverse variance; Std: standard.

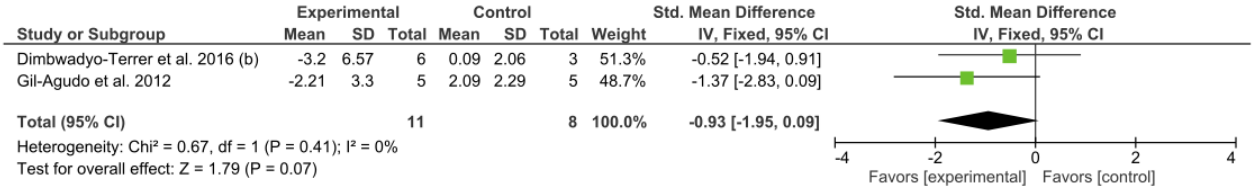


Figure 5. Forest plot for upper limb motor function measured by the muscle balance test. The green blocks indicate the weight assigned to the study, the horizontal line depicts the CI, and the black rhombus shows the overall result. IV: inverse variance; Std: standard.

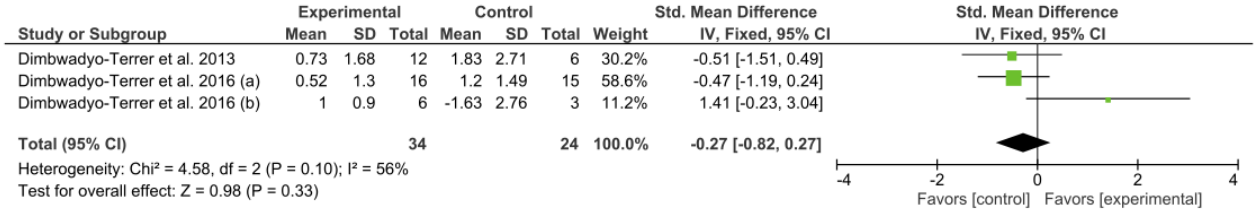
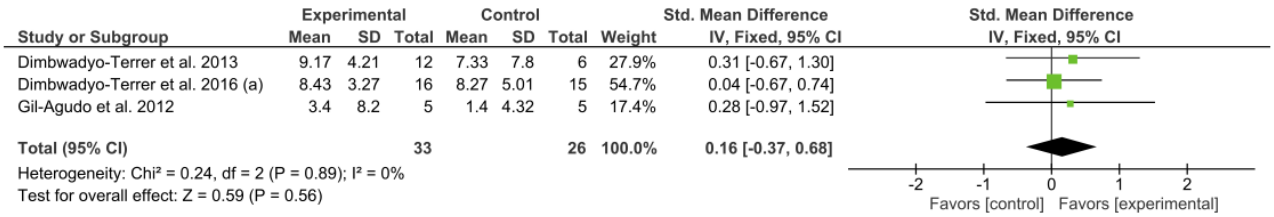


Figure 6. Forest plot for upper limb motor function measured by the Motricity Index. The green blocks indicate the weight assigned to the study, the horizontal line depicts the CI, and the black rhombus shows the overall result. IV: inverse variance; Std: standard.



The JHFT was used to measure the ULMF in two studies [8,34]. The results suggested statistically significant results for VR interventions in the "picking up small common objects" subgroup. The overall result for the remaining subgroups was not conclusive (Figure 7).

Finally, the ROM was measured in 2 of the studies [8,35]. None of the subgroups in this meta-analysis led to conclusive results, as shown in Figure 8.

Figure 7. Forest plot for upper limb motor function measured by the Jebsen-Taylor Hand Function Test. The green blocks indicate the weight assigned to the study, the horizontal line depicts the CI, and the black rhombus shows the overall result. IV: inverse variance; Std: standard.

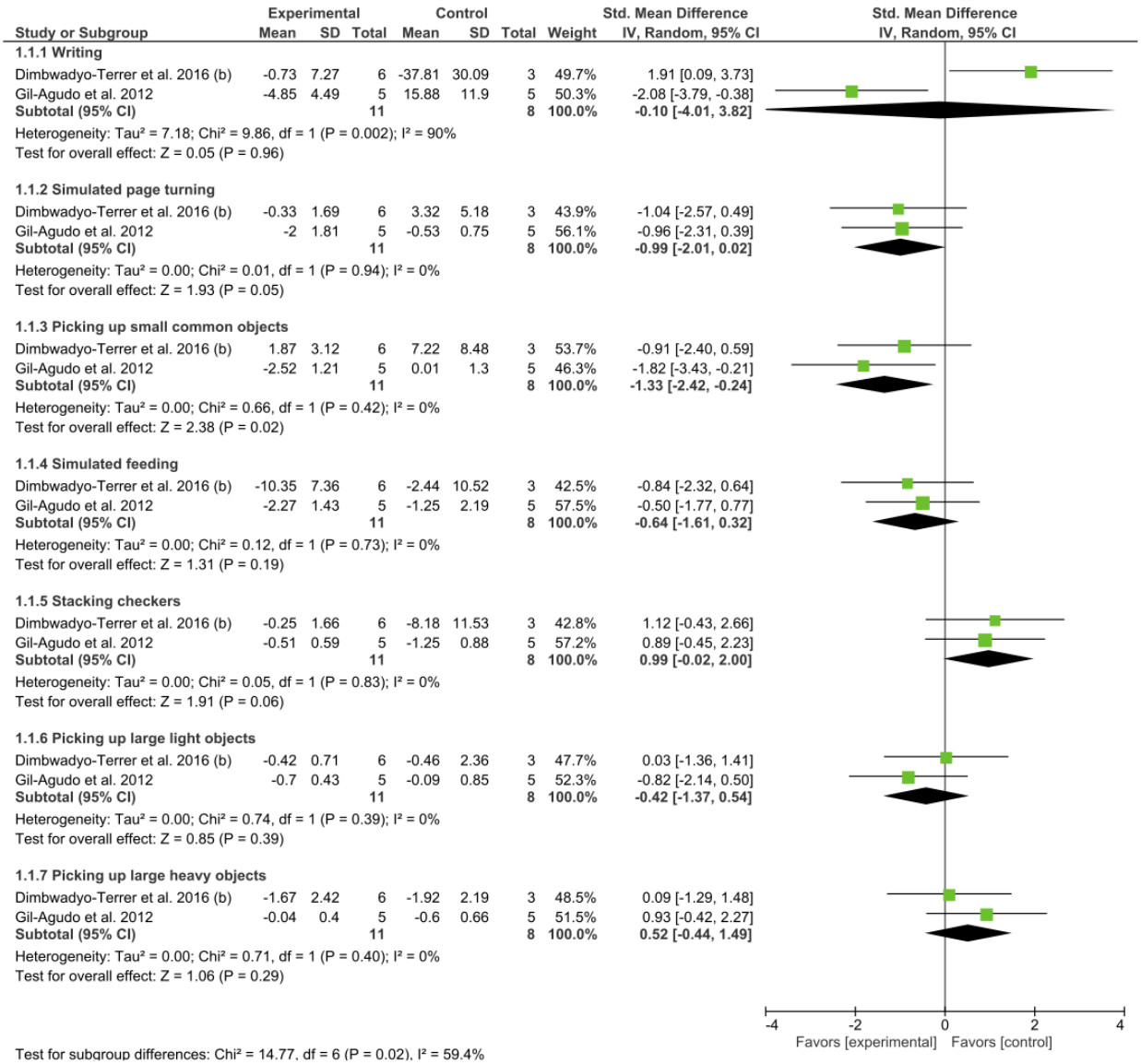
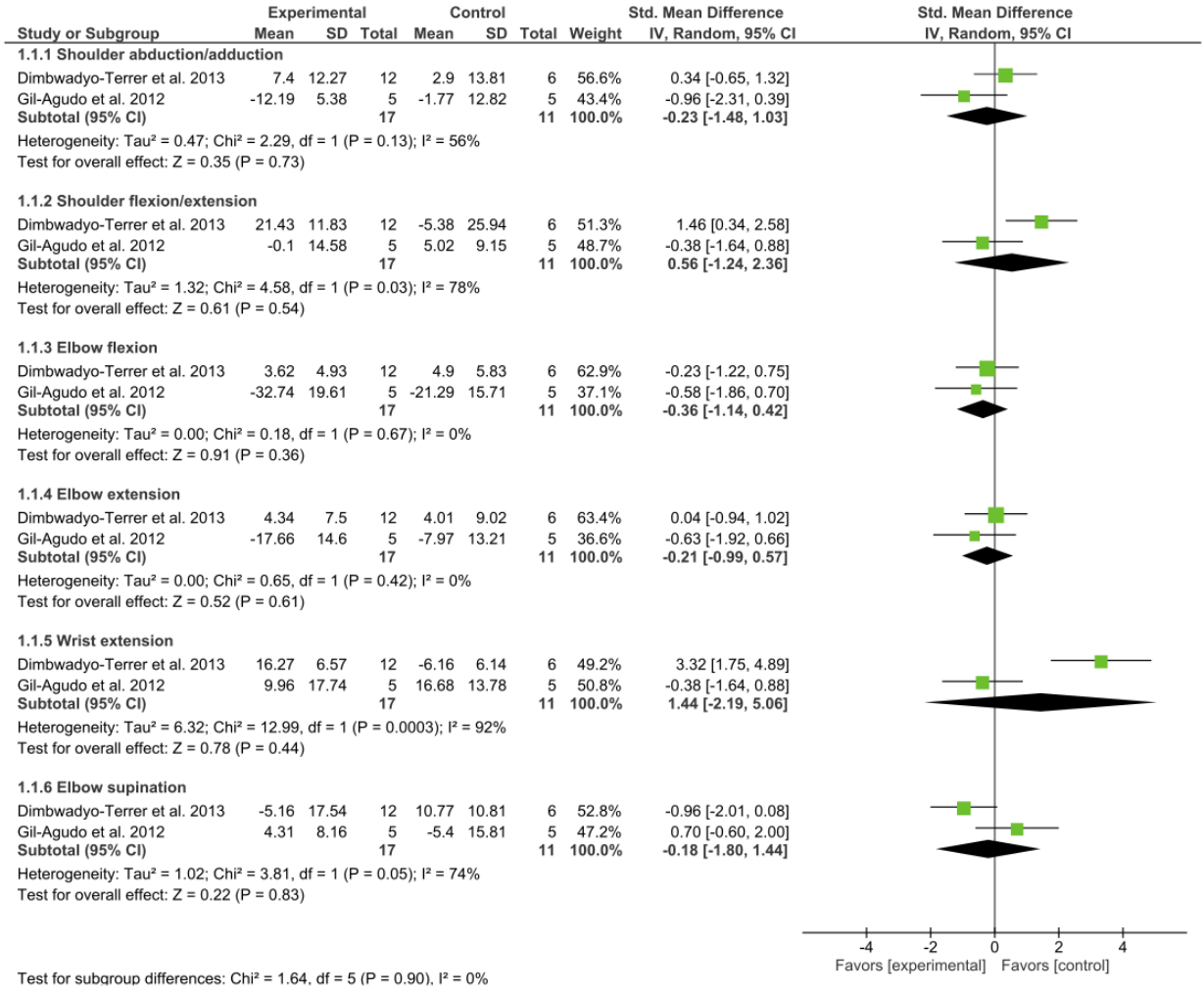


Figure 8. Forest plot for upper limb range of motion. The green blocks indicate the weight assigned to the study, the horizontal line depicts the CI, and the black rhombus shows the overall result. IV: inverse variance; Std: standard.



Discussion

Principal Findings

The present research aimed to use meta-analysis to evaluate the effectiveness of VR versus CPT on ULMF in patients with SCI. Six controlled trials were included in the systematic review and 4 of them were included in the meta-analysis. A total of 105 participants were involved in the different studies. In view of our results, we can conclude that there is not enough evidence that VR interventions are more effective than CPT in helping patients to recover ULMF after SCI.

These results match with those obtained in our previous meta-analysis [29] on functional performance recovery in patients with SCI. Furthermore, from the 6 studies included in this review, 5 [3,8,33-35] analyzed the effects of VR therapy on functional performance and none of them showed significant results. According to the International Classification of Functioning, Disability and Health (ICF), we can hypothesize that ULMF impairments influence the loss of functional performance, since impairments at the body structure and functional level can influence activity limitations and participation restrictions [41]. Conversely, our results do not match with those of Yeo et al [28] and de Araújo et al [27], who reported positive effects of VR interventions on motor function, but the reviews were not restricted specifically to assess ULMF. We suggest that the inconclusive results on ULMF revealed in the present review could have been affected by the type of VR devices used in the interventions. All of the studies performed the VR interventions through semi-immersive or nonimmersive systems, where a computer or video game console displayed the virtual environment through a screen [41]. None of the studies used immersive VR devices, which could provoke more task-focused attention than semi-immersive and nonimmersive devices [42]. Additionally, other heterogeneous factors could have influenced the results obtained, such as different tasks being performed in the VR sessions, different protocols being used for VR interventions and CPT, different session and program durations, and the participants' characteristics. Therefore, it would be desirable to unify protocols in order to clarify which factors of VR interventions may be more appropriate to achieve the intended effects.

Concerning the characteristics of participants, the injury severity was measured by AIS grades. Most studies [3,8,32,35] included participants with AIS grades A-B, while 2 studies [33,34]

included AIS grades A-D. Regarding the levels of injury, most studies [3,8,32,33,35] included patients with cervical levels of injury. Although several positive effects were found in patients with AIS A-B grades and cervical levels, we cannot conclude that the recovery of ULMF is related to the level of injury. Regarding the different effects obtained in the studies, of the 6 studies included in the present review, only the study by Kowalczewski et al [32] showed significant improvements in ULMF, ROM, functional tasks, grasp, and pinch forces. These improvements might have been seen in the study because the intervention had the longest total duration and a higher intensity (60 minutes, 5 times/week for 6 weeks).

Although VR systems have the potential to provide precise measurement of motor outcomes, provide direct feedback and safe environments [13,43], and increase patient motivation and treatment adherence [8] in clinical settings, we did not find differences between VR interventions and CPT in improving ULMF in patients with SCI. According to Morone et al [44], further research is needed in order to develop accurate user guidelines before VR systems are ready for market, to develop immersive VR systems based on personalized neurological characteristics optimizing motor learning processes [45], to implement adequate training to health care professionals [46], and to integrate this technology into neurological rehabilitation [47].

Limitations and Recommendations for Future Research

Some limitations of the study should be mentioned. The results provided by the present review should be viewed with caution because of the limited number of controlled trials analyzed. Another limitation was the small sample size used in the studies and the different injury levels of the patients. Thus, we encourage authors to use large sample sizes and to include an appropriate number of subjects in stratified groups in order to know which factors of the participants' characteristics are influencing the results. Nevertheless, these patients are usually treated in neurological institutions or centers and it is difficult to get large sample sizes. Thus, most studies include convenience samples, resulting in potential selection biases [48]. Furthermore, the heterogeneous protocols used in terms of VR devices employed, program and session durations, and CPT protocols used could affect the results obtained in this review.

In this sense, we encourage researchers to perform randomized controlled trials with higher methodological quality using larger sample sizes, and to unify VR intervention protocols in order to identify the key aspects of VR interventions that have the greatest impact on ULMF recovery after SCI. In addition, because task-focused attention is stimulated more with immersive VR devices than with semi-immersive and nonimmersive devices [42] and no positive results on ULMF were obtained using semi-immersive and nonimmersive VR devices in the studies analyzed in this review, we encourage researchers to use immersive VR devices in their future clinical trials. Finally, we urge researchers to analyze the effectiveness of the application of different CPT techniques in patients with SCI in order to provide further evidence for this topic.

Conclusions

Although the use of VR devices has expanded in neurological rehabilitation, the current evidence for using VR interventions to improve ULMF in patients with SCI is limited. Specifically, our results showed that VR may not be more effective than CPT in ULMF recovery. This may be explained by the fact that all the studies used semi-immersive and nonimmersive devices, and these devices require less task-focused attention than immersive VR devices. No solid conclusions can be drawn concerning the relationship between injury levels and severity and the effects of VR interventions.

In view of our results, it is necessary to conduct clinical trials with a high methodological quality, using larger sample sizes, and to unify VR intervention protocols in order to identify the key aspects that increase the clinical impact of VR interventions in neurological rehabilitation. Further research is needed to provide evidence for the application of VR devices to facilitate ULMF recovery in patients with SCI.

Authors' Contributions

Conceptualization and methodology: ADM-R, MDR, DL-A, and AA-R. Statistical analysis: AS. Writing—original draft preparation: ADM-R, DL-A, and JAM-M. Writing—review and editing: DL-A, JAM-M, and AS. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Wyndaele M, Wyndaele J. Incidence, prevalence and epidemiology of spinal cord injury: what learns a worldwide literature survey? Spinal Cord 2006 Jan 3;44(9):523-529. [doi: 10.1038/sj.sc.3101893]
2. Yoon SY, Leigh J, Lee J, Kim WH. Comparing Activity and Participation between Acquired Brain Injury and Spinal-Cord Injury in Community-Dwelling People with Severe Disability Using WHODAS 2.0. IJERPH 2020 Apr 27;17(9):3031. [doi: 10.3390/ijerph17093031]

3. Dimbwadyo-Terrer I, Gil-Agudo A, Segura-Fragoso A, de los Reyes-Guzmán A, Trincado-Alonso F, Piazza S, et al. Effectiveness of the Virtual Reality System Toyra on Upper Limb Function in People with Tetraplegia: A Pilot Randomized Clinical Trial. Biomed Res Int 2016;2016:6397828 [FREE Full text] [doi: 10.1155/2016/6397828] [Medline: 26885511]
4. Pourmand A, Davis S, Lee D, Barber S, Sikka N. Emerging Utility of Virtual Reality as a Multidisciplinary Tool in Clinical Medicine. Games Health J 2017 Oct;6(5):263-270. [doi: 10.1089/g4h.2017.0046] [Medline: 28759254]
5. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for Methodology of Virtual Reality Clinical Trials in Health Care by an International Working Group: Iterative Study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
6. de Lama C, González-Gaya C, Sánchez-Lite A. An Experimental Test Proposal to Study Human Behaviour in Fires Using Virtual Environments. Sensors (Basel) 2020 Jun 26;20(12):3607 [FREE Full text] [doi: 10.3390/s20123607] [Medline: 32604864]
7. Oña ED, Jardón A, Cuesta-Gómez A, Sánchez-Herrera-Baeza P, Cano-de-la-Cuerda R, Balaguer C. Validity of a Fully-Immersive VR-Based Version of the Box and Blocks Test for Upper Limb Function Assessment in Parkinson's Disease. Sensors 2020 May 13;20(10):2773. [doi: 10.3390/s20102773]
8. Gil-Agudo A, Dimbwadyo-Terrer I, Peñasco-Martín B, de los Reyes-Guzmán A, Bernal-Sahún A, Berbel-García A. Experiencia clínica de la aplicación del sistema de realidad TOyRA en la neuro-rehabilitación de pacientes con lesión medular. Rehabilitación 2012 Jan;46(1):41-48. [doi: 10.1016/j.rh.2011.10.005]
9. D'Addio G, Iuppariello L, Gallo F, Bifulco P, Cesarelli M, Lanzillo B. Comparison between clinical and instrumental assessing using Wii Fit system on balance control. In: IEEE. New York, USA: IEEE; 2014 Jun 11 Presented at: 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA); 2014; Lisboa, Portugal p. 1-5. [doi: 10.1109/MeMeA.2014.6860124]
10. Franco JR, Jacobs K, Inzerillo C, Kluzik J. The effect of the Nintendo Wii Fit and exercise in improving balance and quality of life in community dwelling elders. THC 2012 May 01;20(2):95-115. [doi: 10.3233/thc-2011-0661]
11. Khurana M, Walia S, Noohu MM. Study on the Effectiveness of Virtual Reality Game-Based Training on Balance and Functional Performance in Individuals with Paraplegia. Top Spinal Cord Inj Rehabil 2017 Jun;23(3):263-270 [FREE Full text] [doi: 10.1310/sci16-00003] [Medline: 29339902]
12. Park D, Lee D, Lee K, Lee G. Effects of Virtual Reality Training using Xbox Kinect on Motor Function in Stroke Survivors: A Preliminary Study. Journal of Stroke and Cerebrovascular Diseases 2017 Oct;26(10):2313-2319. [doi: 10.1016/j.jstrokecerebrovasdis.2017.05.019]
13. Domínguez-Téllez P, Moral-Muñoz JA, Salazar A, Casado-Fernández E, Lucena-Antón D. Game-Based Virtual Reality Interventions to Improve Upper Limb Motor Function and Quality of Life After Stroke: Systematic Review and Meta-analysis. Games Health J 2020 Feb 01;9(1):1-10. [doi: 10.1089/g4h.2019.0043] [Medline: 32027185]
14. Dominguez-Tellez P, Moral-Munoz JA, Casado-Fernandez E, Salazar A, Lucena-Anton D. [Effects of virtual reality on balance and gait in stroke: a systematic review and meta-analysis]. Rev Neurol 2019 Sep 16;69(6):223-234 [FREE Full text] [doi: 10.33588/rn.6906.2019063] [Medline: 31497866]
15. Ikbali Afsar S, Mirzayev I, Umit Yemisci O, Cosar Saracgil SN. Virtual Reality in Upper Extremity Rehabilitation of Stroke Patients: A Randomized Controlled Trial. J Stroke Cerebrovasc Dis 2018 Dec;27(12):3473-3478. [doi: 10.1016/j.jstrokecerebrovasdis.2018.08.007] [Medline: 30193810]
16. Bonuzzi GMG, de Freitas TB, Palma GCDS, Soares MAA, Lange B, Pompeu JE, et al. Effects of the brain-damaged side after stroke on the learning of a balance task in a non-immersive virtual reality environment. Physiother Theory Pract 2020 Feb 24:1-8. [doi: 10.1080/09593985.2020.1731893] [Medline: 32090670]
17. Garcia-Munoz C, Casuso-Holgado M. [Effectiveness of Wii Fit Balance board in comparison with other interventions for post-stroke balance rehabilitation. Systematic review and meta-analysis]. Rev Neurol 2019 Oct 01;69(7):271-279 [FREE Full text] [doi: 10.33588/rn.6907.2019091] [Medline: 31559625]
18. Booth ATC, Buizer AI, Meyns P, Oude Lansink ILB, Steenbrink F, van der Krogt MM. The efficacy of functional gait training in children and young adults with cerebral palsy: a systematic review and meta-analysis. Dev Med Child Neurol 2018 Mar 07;60(9):866-883. [doi: 10.1111/dmcn.13708]
19. Johansen T, Strøm V, Simic J, Rike P. Effectiveness of training with motion-controlled commercial video games for hand and arm function in people with cerebral palsy: A systematic review and meta-analysis. J Rehabil Med 2020 Jan;52(1):-10. [doi: 10.2340/16501977-2633]
20. Feng H, Li C, Liu J, Wang L, Ma J, Li G, et al. Virtual Reality Rehabilitation Versus Conventional Physical Therapy for Improving Balance and Gait in Parkinson's Disease Patients: A Randomized Controlled Trial. Med Sci Monit 2019 Jun 05;25:4186-4192. [doi: 10.12659/msm.916455]
21. Lei C, Sunzi K, Dai F, Liu X, Wang Y, Zhang B, et al. Effects of virtual reality rehabilitation training on gait and balance in patients with Parkinson's disease: A systematic review. PLoS ONE 2019 Nov 7;14(11):e0224819. [doi: 10.1371/journal.pone.0224819]
22. Moreno-Verdu M, Ferreira-Sanchez MR, Cano-de-la-Cuerda R, Jimenez-Antona C. [Efficacy of virtual reality on balance and gait in multiple sclerosis. Systematic review of randomized controlled trials]. Rev Neurol 2019 May 01;68(9):357-368 [FREE Full text] [doi: 10.33588/rn.6809.2018350] [Medline: 31017288]

23. Norouzi E, Gerber M, Pühse U, Vaezmosavi M, Brand S. Combined virtual reality and physical training improved the bimanual coordination of women with multiple sclerosis. Neuropsychol Rehabil 2020 Jan 23:1-18. [doi: 10.1080/09602011.2020.1715231] [Medline: 31971071]
24. Maggio MG, Russo M, Cuzzola MF, Destro M, La Rosa G, Molonia F, et al. Virtual reality in multiple sclerosis rehabilitation: A review on cognitive and motor outcomes. Journal of Clinical Neuroscience 2019 Jul;65:106-111. [doi: 10.1016/j.jocn.2019.03.017]
25. Sale P, Russo EF, Russo M, Masiero S, Piccione F, Calabrò RS, et al. Effects on mobility training and de-adaptations in subjects with Spinal Cord Injury due to a Wearable Robot: a preliminary report. BMC Neurol 2016 Jan 28;16(1). [doi: 10.1186/s12883-016-0536-0]
26. de Araújo AVL, Ribeiro FPG, Massetti T, Potter-Baker KA, Cortes M, Plow EB, et al. Effectiveness of anodal transcranial direct current stimulation to improve muscle strength and motor functionality after incomplete spinal cord injury: a systematic review and meta-analysis. Spinal Cord 2020 Feb 17;58(6):635-646. [doi: 10.1038/s41393-020-0438-2]
27. de Araújo AVL, Neiva JFDO, Monteiro CBDM, Magalhães FH. Efficacy of Virtual Reality Rehabilitation after Spinal Cord Injury: A Systematic Review. Biomed Res Int 2019;2019:7106951 [FREE Full text] [doi: 10.1155/2019/7106951] [Medline: 31828120]
28. Yeo E, Chau B, Chi B, Ruckle D, Ta P. Virtual Reality Neurorehabilitation for Mobility in Spinal Cord Injury: A Structured Review. Innov Clin Neurosci 2019 Jan 01;16(1-2):13-20 [FREE Full text] [Medline: 31037223]
29. De Miguel-Rubio A, Rubio MD, Salazar A, Camacho R, Lucena-Anton D. Effectiveness of Virtual Reality on Functional Performance after Spinal Cord Injury: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. JCM 2020 Jul 01;9(7):2065. [doi: 10.3390/jcm9072065]
30. Hutton B, Catalá-López F, Moher D. The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA. Medicina Clínica (English Edition) 2016 Sep;147(6):262-266. [doi: 10.1016/j.medcle.2016.10.003]
31. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
32. Kowalczewski J, Chong SL, Galea M, Prochazka A. In-Home Tele-Rehabilitation Improves Tetraplegic Hand Function. Neurorehabil Neural Repair 2011 Mar 03;25(5):412-422. [doi: 10.1177/1545968310394869]
33. Prasad S, Aikat R, Labani S, Khanna N. Efficacy of Virtual Reality in Upper Limb Rehabilitation in Patients with Spinal Cord Injury: A Pilot Randomized Controlled Trial. Asian Spine J 2018 Oct;12(5):927-934. [doi: 10.31616/asj.2018.12.5.927]
34. Dimbwadyo-Terrer I, Trincado-Alonso F, de Los Reyes-Guzmán A, Aznar MA, Alcubilla C, Pérez-Nombela S, et al. Upper limb rehabilitation after spinal cord injury: a treatment based on a data glove and an immersive virtual reality environment. Disabil Rehabil Assist Technol 2016 Aug 16;11(6):462-467. [doi: 10.3109/17483107.2015.1027293] [Medline: 26181226]
35. Dimbwadyo-Terrer I, Trincado-Alonso F, De los Reyes-Guzmán A, Bernal-Sahún A, López-Monteagudo P, Polonio-López B. Clinical, functional and kinematic correlations using the Virtual Reality System toyra® as upper limb rehabilitation tool in people with spinal cord injury. In: NEUROTECHNIX.: NEUROTECHNIX; 2013 Presented at: NEUROTECHNIX 2013 - International Congress on Neurotechnology, Electronics and Informatics; 18-20 September; Algarve, Portugal p. 81-88. [doi: 10.1007/978-3-642-34546-3_139]
36. Sears ED, Chung KC. Validity and responsiveness of the Jebsen-Taylor Hand Function Test. J Hand Surg Am 2010 Jan;35(1):30-37 [FREE Full text] [doi: 10.1016/j.jhsa.2009.09.008] [Medline: 19954898]
37. Perry J, Weiss W, Burnfield J, Gronley J. The supine hip extensor manual muscle test: a reliability and validity study. Arch Phys Med Rehabil 2004 Aug;85(8):1345-1350. [doi: 10.1016/j.apmr.2003.09.019] [Medline: 15295763]
38. Wang Y, Bohannon RW, Kapellusch J, Garg A, Gershon RC. Dexterity as measured with the 9-Hole Peg Test (9-HPT) across the age span. Journal of Hand Therapy 2015 Jan;28(1):53-60. [doi: 10.1016/j.jht.2014.09.002]
39. Demeurisse G, Demol O, Robaye E. Motor Evaluation in Vascular Hemiplegia. Eur Neurol 2008 Jan 29;19(6):382-389. [doi: 10.1159/000115178]
40. Gajdosik R, Bohannon R. Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. Phys Ther 1987 Dec;67(12):1867-1872. [doi: 10.1093/ptj/67.12.1867] [Medline: 3685114]
41. Sullivan K, Cen S. Model of disablement and recovery: knowledge translation in rehabilitation research and practice. Phys Ther 2011 Dec;91(12):1892-1904. [doi: 10.2522/ptj.20110003] [Medline: 22003166]
42. Gokeler A, Bisschop M, Myer GD, Benjaminse A, Dijkstra PU, van Keeken HG, et al. Immersive virtual reality improves movement patterns in patients after ACL reconstruction: implications for enhanced criteria-based return-to-sport rehabilitation. Knee Surg Sports Traumatol Arthrosc 2016 Jul 14;24(7):2280-2286. [doi: 10.1007/s00167-014-3374-x] [Medline: 25311052]
43. Gutiérrez Á, Sepúlveda-Muñoz D, Gil-Agudo Á, de los Reyes Guzmán A. Serious Game Platform with Haptic Feedback and EMG Monitoring for Upper Limb Rehabilitation and Smoothness Quantification on Spinal Cord Injury Patients. Applied Sciences 2020 Feb 02;10(3):963. [doi: 10.3390/app10030963]
44. Morone G, Paolucci S, Mattia D, Pichiorri F, Tramontano M, Iosa M. The 3Ts of the new millennium neurorehabilitation gym: therapy, technology, translationality. Expert Rev Med Devices 2016 Sep 13;13(9):785-787. [doi: 10.1080/17434440.2016.1218275] [Medline: 27466820]

45. Sánchez-Herrera-Baeza P, Cano-de-la-Cuerda R, Oña-Simbaña ED, Palacios-Ceña D, Pérez-Corrales J, Cuenca-Zaldivar JN, et al. The Impact of a Novel Immersive Virtual Reality Technology Associated with Serious Games in Parkinson's Disease Patients on Upper Limb Rehabilitation: A Mixed Methods Intervention Study. Sensors (Basel) 2020 Apr 11;20(8):2168 [FREE Full text] [doi: 10.3390/s20082168] [Medline: 32290517]
46. Massetti T, da Silva TD, Crocetta TB, Guarnieri R, de Freitas BL, Bianchi Lopes P, et al. The Clinical Utility of Virtual Reality in Neurorehabilitation: A Systematic Review. J Cent Nerv Syst Dis 2018 Nov 27;10:1179573518813541 [FREE Full text] [doi: 10.1177/1179573518813541] [Medline: 30515028]
47. Morone G, Tramontano M, Iosa M, Shofany J, Iemma A, Musicco M, et al. The efficacy of balance training with video game-based therapy in subacute stroke patients: a randomized controlled trial. Biomed Res Int 2014;2014:580861 [FREE Full text] [doi: 10.1155/2014/580861] [Medline: 24877116]
48. Miot H. Tamanho da amostra em estudos clínicos e experimentais. J. vasc. bras 2011 Dec;10(4):275-278. [doi: 10.1590/S1677-54492011000400001]

Abbreviations

AIS: American Spinal Injury Association Impairment Scale
CENTRAL: Cochrane Central Register of Controlled Trials
CINAHL: Cumulative Index to Nursing and Allied Health Literature
CPT: conventional physical therapy
ICF: International Classification of Functioning, Disability and Health
JTHFT: Jebsen-Taylor Hand Function Test
MeSH: Medical Subject Headings
MI: Motricity Index
NHPT: Nine-Hole Peg Test
PEDro: Physiotherapy Evidence Database
PICOS: Population, Intervention, Comparison, Outcomes and Study
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
ROM: range of motion
SCI: spinal cord injury
ULMF: upper limb motor function
VR: virtual reality

Edited by L Buis; submitted 15.07.20; peer-reviewed by R Sobrino, D Scalzitti; comments to author 02.10.20; revised version received 16.10.20; accepted 06.11.20; published 03.12.20.

Please cite as:
De Miguel-Rubio A, Rubio MD, Alba-Rueda A, Salazar A, Moral-Munoz JA, Lucena-Anton D
Virtual Reality Systems for Upper Limb Motor Function Recovery in Patients With Spinal Cord Injury: Systematic Review and Meta-Analysis
JMIR Mhealth Uhealth 2020;8(12):e22537
URL: http://mhealth.jmir.org/2020/12/e22537/
doi:10.2196/22537
PMID:33270040

Review

The Use of Virtual Reality Technologies in the Treatment of Duchenne Muscular Dystrophy: Systematic Review

Maria Rosa Baeza-Barragán[1*], BSc, MSc; Maria Teresa Labajos Manzanares[1], MD, PhD; Carmen Ruiz Vergara[1], BSc, MSc; María Jesús Casuso-Holgado[2], BSc, MSc, PhD; Rocío Martín-Valero[1*], BSc, MSc, PhD

[1]Department of Physiotherapy, Faculty of Health Sciences, University of Malaga, Malaga, Spain

[2]Department of Physiotherapy, University of Seville, Seville, Spain

[*]these authors contributed equally

Corresponding Author:
Rocío Martín-Valero, BSc, MSc, PhD
Department of Physiotherapy, Faculty of Health Sciences
University of Malaga
Arquitecto Francisco Peñalosa 3, Ampliación de Campus de Teatinos
Malaga, 29071
Spain
Phone: 34 951 95 28 58
Email: rovalemas@gmail.com

Abstract

Background: Duchenne muscular dystrophy is a serious and progressive disease affecting one in 3500-6000 live male births. The use of new virtual reality technologies has revolutionized the world of youth rehabilitation.

Objective: We performed a systematic review to study the effectiveness of the use of virtual reality systems applied in the rehabilitation of the upper limbs of individuals with Duchenne muscular dystrophy.

Methods: Between June 2018 and September 2019, we carried out a series of searches in 5 scientific databases: (1) PubMed, (2) Web of Science, (3) Scopus, (4) The Cochrane Library, and (5) MEDLINE via EBSCO. Two evaluators independently conducted the searches following the PRISMA recommendations for systematic reviews for articles. Two independent evaluators collated the results. Article quality was determined using the PEDro scale.

Results: A total of 7 clinical trials were included in the final review. These studies used new technologies as tools for physiotherapeutic rehabilitation of the upper limbs of patients with Duchenne muscular dystrophy. Collectively, the studies showed improvement in functionality, quality of life, and motivation with the use of virtual reality technologies in the rehabilitation of upper limbs of individuals with Duchenne muscular dystrophy.

Conclusions: The treatment of neuromuscular diseases has changed in recent years, from palliative symptom management to preventive methods for capacity building. The use of virtual reality is beginning to be necessary in the treatment of progressive diseases involving movement difficulties, as it provides freedom and facilitates the improvement of results in capacity training. Given that new technologies are increasingly accessible, rehabilitation and physiotherapy programs can use these technologies more frequently, and virtual reality environments can be used to improve task performance, which is essential for people with disabilities. Ultimately, virtual reality can be a great tool for physiotherapy and can be used for Duchenne muscular dystrophy rehabilitation programs to improve patient performance during training.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42018102548; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=102548

(JMIR Mhealth Uhealth 2020;8(12):e21576) doi:10.2196/21576

KEYWORDS

Duchenne muscular dystrophy; virtual reality; upper limb; physical therapy; muscular dystrophy; mutation; muscle; degeneration

Introduction

Duchenne muscular dystrophy is caused by a mutation in the gene that produces dystrophin, which is responsible for maintaining muscle properties. Duchenne muscular dystrophy is a rare disease with an incidence of 1 in 3500-6000 live male births [1]. The lack of dystrophin leads to a progressive degeneration of muscle fibers, which then become connective tissue and fat [2]. Currently, this disease has no cure. The main symptoms are muscle weakness, which progressively leads to a loss of function and independence, and, in advanced stages of the disease, a compromised respiratory system [3]. Functional tests are performed during medical assessments of children with Duchenne muscular dystrophy [4]. The Motor Function Measurement test is used to measure patients' conditions before and after virtual task training [5]. Due to its analytical simplicity, the Vignos scale is also used to evaluate functionality and overall muscle performance in neuromuscular diseases. The Egen Scale Klassification was specially developed to measure the degree of functional impairment in daily living activities experienced by those with Duchenne muscular dystrophy [5].

Over time, there have been numerous guides for interprofessional action for affected individuals and their families [2]. Various technologies are used to provide patients and professionals with reliable forms of evaluation and effective treatments [4]. For instance, the use of virtual reality during treatment can provide a fun environment for patients [5,6]. Virtual reality may use multiple devices such as glasses, game consoles, immersion systems, applications for tablets and smartphones, gloves, exoskeletons, tele-rehabilitation systems, and more [7,8]. Virtual environments can involve representations of the users (ie, avatars), communication skills, the construction of or interaction with 3D objects, and the illusion of space [8]. From the combination of the concepts of virtual reality and rehabilitation, the concept of virtual rehabilitation has emerged, a term initially coined by Thalmann and Burdea [6]. Simulations are defined as learning contexts that attempt to imitate real-life situations [7]. Games should be designed to improve learning and promote autonomy [7].

The World Health Organization defines mHealth as the practice of medicine and public health supported by mobile devices such as mobile phones, patient monitoring devices, digital personal assistants, and other wireless devices [9]. Technology, like virtual reality, is used in health care settings to encourage therapists to adapt physical exercises in order to encourage patient participation [10]. New technologies enable the creation of virtual environments that capture patients' attention by showing them interactive systems based on physical therapy exercises [11]. In this context, patients are continually challenged by constantly changing tasks, which elicit more active participation in the requested exercises and can potentially improve the intended results, thus accelerating the recovery process [6]. For example, training with commercial video games is used to help children with motor problems [12,13]. In recent years, virtual development has focused on robotic therapy (eg, the use of technology involving an exoskeleton) to improve distal movements in the hands [14,15].

Although no systematic review has been performed in children with Duchenne muscular dystrophy, a systematic review of the effectiveness of virtual reality in the manual function of children with cerebral palsy demonstrated little evidence of effectiveness [16]. Compared to those with cerebral palsy, patients with Duchenne muscular dystrophy retain musculature strength of the upper extremities for longer than they do in the lower extremities, especially in the distal muscles [17]. In those with cerebral palsy, distal muscles required for fine motor movements are the most affected in cases of hemiparesis and tetraparesis [16].

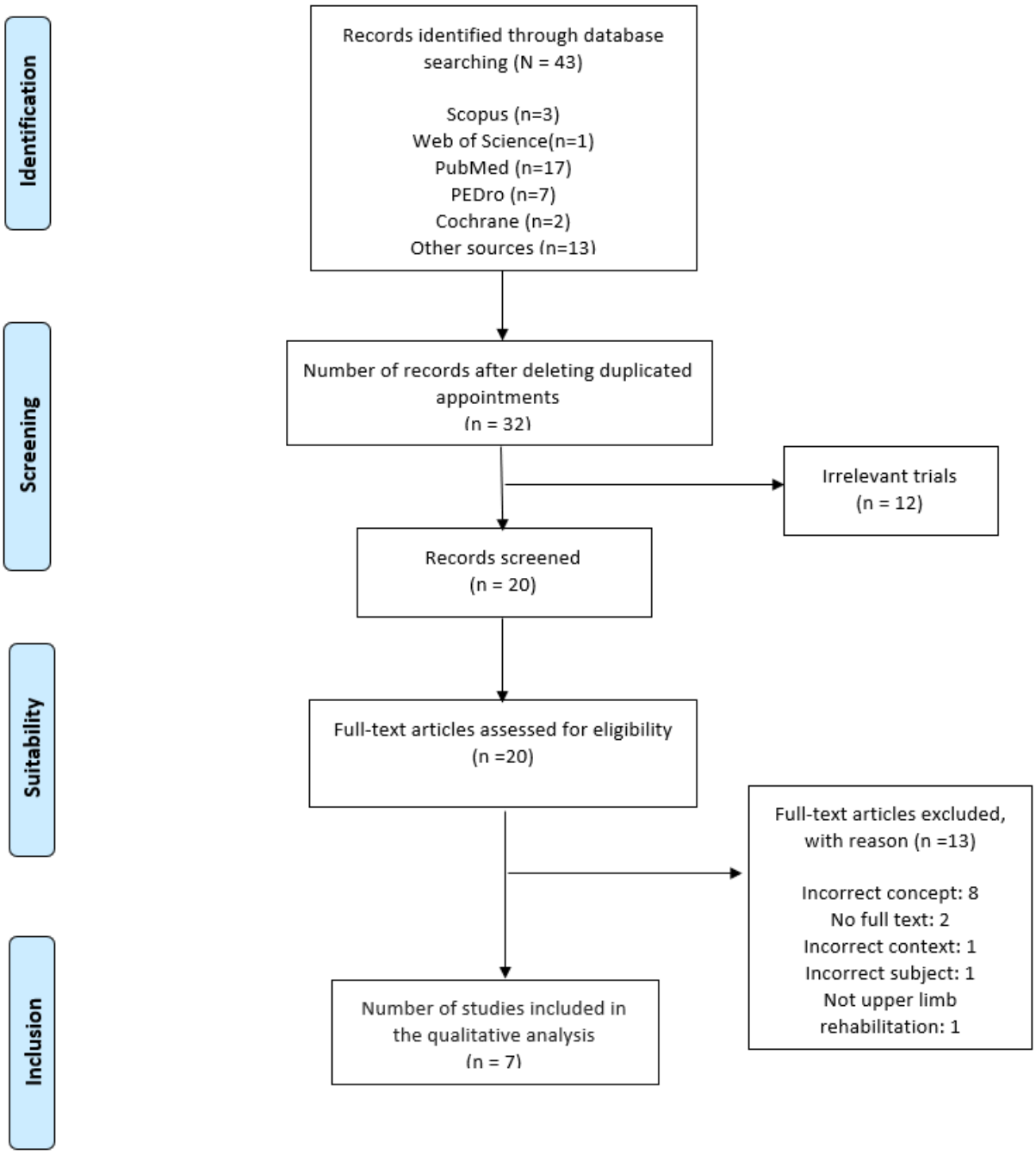
Virtual reality games in individuals with Duchenne muscular dystrophy who have considerable and progressive loss of movement could help to create safe rehabilitative environments in which to improve responsiveness or regulate treatment strategies [18]. The objective of this systematic review is to verify the effectiveness of virtual reality physiotherapy treatments in the rehabilitation of the upper limbs in individuals with Duchenne muscular dystrophy.

Methods

Search Strategy

This systematic review is registered in PROSPERO with the code CRD42018102548. We used the most important evaluation items for systematic reviews and meta-analysis guides. We carried out a series of searches following PRISMA [19] recommendations for systematic reviews in 5 scientific databases: (1) PubMed, (2) Web of Science, (3) Scopus, (4) The Cochrane Library, and (5) MEDLINE via EBSCO. The searches were carried out between June 2018 and September 2019. Complete articles in English or Spanish were required. The following keywords from the Medical Subject Headings tree were used for the search: physical therapy, physiotherapy, upper limb, VR, new technologies, Duchenne, Duchenne muscular dystrophy, physical therapy modalities, virtual reality, and virtual reality exposure therapy. See Figure 1 for a complete list of the search strategy and terms used.

Figure 1. Literature search and study selection.



Study Objective

The research question was: "Is the use of virtual reality in rehabilitation of the upper limbs effective in children with Duchenne muscular dystrophy?"

The following was derived using the PICO model [19]:

- P (patient): individuals with Duchenne muscular dystrophy.
- I (intervention): physical therapy with virtual reality.
- C (comparison): traditional physiotherapy.
- O (outcomes or results): effectiveness of use. Patient's status.

Eligibility Criteria

When searching the databases, a series of filters were chosen to limit the searches and select the articles. These were:

- Articles published between 2000 and 2019 inclusive (to ensure the relevance of the material).
- Articles whose subject of study were human.
- Randomized clinical trials.
- Articles written in English or Spanish.
- Studies based on physiotherapy applied with new technologies, including interventions with virtual reality, virtual games, and applications on tablets or smartphones.

Data Extraction and Analysis

After searching different keywords in the aforementioned list of databases and sorting articles by title and summary, relevant articles were identified for complete reading, duplicate articles were eliminated, and the inclusion and exclusion criteria were applied to the sample of definitive data (Figure 1).

Methodological Quality Assessment

We evaluated the methodological quality and internal validity of the studies using the PEDro scale. The PEDro scale (0-10) is based on the Delphi list developed by Verhagen et al [20]. Two independent evaluators (RMV and RBB) used the PEDro checklist to score each study. A study with a score of 4-5 was considered poor or acceptable, where a score below 4 was considered to indicate low methodological quality. Studies with a score below 6 were considered as having low or level 1 evidence, where a study with a score of 6-8 was considered good and a study with a score of 9-10 was considered excellent.

Results

Study Selection and PEDro Assessment

A total of 43 articles were identified for review. Duplicate articles were removed, leaving a total of 32 articles. Twelve of these were excluded as they were considered irrelevant. A total of 20 articles were screened and assessed for eligibility after reading the full text. Thirteen articles were excluded during the eligibility assessment phase: 8 because they showed an incorrect concept, 2 because they did not show the full text, 1 because it did not focus on Duchenne muscular dystrophy subjects, 1 due to incorrect context, and 1 because it did not display upper limb rehabilitation in the variables. The final complete quality assessment included 7 articles.

After deleting duplicates, 32 articles were identified through the keyword search, 20 of which were selected for screening and assessed for eligibility. After a full-text reading, only 7 were included in the qualitative analysis (Figure 1).

The 7 selected studies were published between 2009 and 2019. Of the 7 studies, 6 were carried out in Brazil and 1 in the Netherlands. Using the PEDro scale, 3 studies received a score of 7 and 1 received a score of 8. These studies were considered "good." Two studies received a score of 6 and 1 study received a score of 5. These studies were considered "acceptable." Criteria and scoring are depicted in Table 1.

Table 1. Methodological quality review of the included studies using the PEDro evaluation scale.

	Correa et al [18]	Capellini et al [5]	Heuntick et al [21]	Massetti et al [22]	Quadrado et al [23]	Malheiros et al [24]	De Freitas et al [17]
Eligibility criteria	✓	✓		✓	✓	✓	✓
Randomized	✓	✓	✓	✓	✓		✓
Allocation concealed	✓			✓	✓		
Baseline comparability		✓	✓	✓	✓	✓	✓
Subject blinding							
Therapist blinding							
Evaluator blinding			✓				
Adequate follow-up	✓	✓		✓	✓	✓	✓
Intention-to-treat analysis	✓	✓		✓	✓	✓	✓
Between-group comparisons	✓	✓	✓	✓	✓	✓	✓
Specific measurements and variability	✓	✓	✓	✓	✓	✓	✓
Total	7	7	5	8	7	6	6

Assessment of Study Design

Collectively, the studies enrolled individuals with Duchenne muscular dystrophy aged 5 to 34 years. Individuals with Duchenne muscular dystrophy were ambulatory or wheelchair-dependent. Control groups included those with typical development. The number of individuals in the intervention groups varied. The largest included 60 participants [17], and the smallest included 19 [21]. Only 2 studies compared 2 groups of individuals with Duchenne muscular dystrophy during the intervention and control periods [21,22].

Total study length ranged from 1 day to 20 weeks. Heuntick et al [21] reported training with the virtual reality task for 20 weeks at home. Correa et al [18] reported 12 cumulative days of weekly intervention. Five studies measured the time it took participants to complete tasks they had been assigned in each of the following 3 phases: (1) acquisition, (2) retention, and (3) transfer [5,17,22-24]. These 5 studies agreed with the number of attempts in each phase, with attempts being higher in the acquisition phase (20-30 trials) than retention phase (5 trials) and transfer phase (5 trials).

The most commonly used variable was Motor Function Measurement, which measures limb functionality in 3 dimensions. The scores of Dimension 1 (D1: standing and transfer), Dimension 2 (D2: axial and proximal limb), and Dimension 3 (D3: distal limb) can predict functionality

improvements [5,17,24]. However, some of the authors whose articles were selected for review do not believe that the Motor Function Measurement score is predictive of improvement, but instead believe that the time to complete the tasks assigned in each phase is predictive of improvement [22,23]. The relation between performance and Motor Function Measurement-D1 indicates that the use of virtual technology like smartphones is partly reliant on the muscles responsible for standing and transfer, as these muscles allow the head to look at the screen [5]. The dependent variable reported in studies to compare between phases is the movement time or time to perform [5,22-24]. Only one study included a motivation scale (Likert Scale) [18], and only one included a quality of life variable (Kidscreen-52) [21].

Virtual reality games can be simple, allowing individuals to adapt quickly and perform the tasks without problems [24]. Games used in the studies included in the review included musical games, virtual ball mazes, catching cubes, or labyrinths. Study participants used a variety of devices or virtual interfaces such as computers, webcams, touch screens, Kinect sensors, Leap Motion interfaces, simulated sounds, video consoles (like PlayStation II), and smartphones [5,17,18,21-24]. De Freitas et al [17] compared different interfaces with or without physical contact.

Study Outcomes

Complete summaries of the included studies and their respective intervention details are presented in Table 2 and Table 3.

Study outcomes showed that the grade of motivation improved when using virtual devices [18]. Furthermore, a better performance was seen with a smartphone if a previous learning phase was used [5]. Improvements in the quality of life and elbow extension were seen if training with virtual reality games was performed at home [21].

Although Quadrado et al [23] found that conducting a timed task in a virtual environment facilitated real-life completion of that same task for individuals with Duchenne muscular dystrophy, Massetti et al [22] found no transference of learning between environments when comparing real-life and virtual tasks. Finally, individuals with Duchenne muscular dystrophy showed better performance when using interfaces without contact (like Leap Motion and Kinect) compared to touch screen interfaces [17]. No adverse effects of the use of virtual reality were described in the studies included in our review.

Table 2. Main characteristics of participants and results of the included studies.

Author, country	Population	Results	Statistical effect
Correa et al (2009) [18], Brazil	16 individuals with DMD[a], 17-24 years old.	• In satisfaction surveys, both patients and therapists found intervention with GenVirtual beneficial. • Positive effects for patients with movement restriction.	The level of motivation using GenVirtual was greater (54%) compared to those not using the system (40%).
Capelini et al (2017) [5], Brazil	IG[b]=50 individuals with DMD; CG[c]=50 typically developing individuals. 10-34 years old.	• IG showed better performance within a short-term motor learning protocol with smartphones. • The score of MFM[d] in standing position and transfers and the first attempt in the acquisition phase predicted the degree of learning.	• Acquisition: IG=7.7-6.3 seconds; CG=4.9-4.1 seconds. • Retention: IG=6.3 seconds; CG= 4.1 seconds. • Transfer phase 2: IG retention=6.4 seconds; transfer phase 2=7.6 seconds.
Heuntinck et al (2018) [21], Brazil	IG=9 typically developing, ambulatory individuals; CG=10 wheelchair dependent individuals with DMD.	• Elbow extension increased in the IG and decreased in the CG. The CG was not assisted by a coach, while the IG was. • Elbow dimension=–0.6. • Kidscreen showed a better quality of life.	• IG SD=22.1 • Dynamometer elbow extension: IG SD=1.8, *P*=.018; CG SD=1.6, *P*=.038 • Performance of the upper limb (transfer phase 0-transfer phase 2)
Massetti et al (2018) [22], Netherlands	22 DMD individuals divided in two groups of 11. Group A started with virtual task; group B started with a real task.	• All participants decreased the movement time from the first to the last block of acquisition, more who started with the virtual task. In both virtual and real tasks, motor learning could be inferred by the short-term retention and transfer task (with increasing distance of the target). There was no transference of learning between environments. • Only the performance on acquisition phase predicted the degree of learning. MFM[e] punctuation did not.	• Real task in acquisition phase: movement time=746ms. • Virtual task in acquisition phase: movement time=1011ms.
Quadrado et al (2017) [23], Brazil	IG=32 individuals with DMD; CG=32 typically developing individuals. 12-32 years old (mean=18 years).	• Acquisition phase: for two groups the tendency was late. Absolute timing error was larger in IG group than in the CG. Only absolute timing error in the acquisition phase predicted the amount of learning; age and MFM did not. • Transfer phase: for both the CG and IG, completion of real-life tasks did not improve completion of virtual tasks. However, training in the virtual environment did improve real-life task completion.	Absolute timing error • IG 1st attempt media (movement time=255 milliseconds) • Last attempt (movement time=156 milliseconds) • CG (movement time=245 milliseconds) larger than IG. Final acquisition • (movement time=156 milliseconds) • Variable timing error transfer phase (movement time=369 milliseconds) • Final acquisition (movement time=132 milliseconds)
Malheiros et al (2015) [24], Brazil	IG=42 individuals with DMD; CG=42 typically developing individuals. 5 to 18 years old.	• Acquisition phase: a significant decrease was found in movement time between the first and last acquisition block, but only for the IG. • Movement time: first acquisition=8.4 seconds; last acquisition=5.7 seconds. • In the transfer phase movement time increased from retention to transfer: IG=5.7-6.6 seconds; CG 3.3-4.0 seconds.	• Movement time during the transfer phase was shorter than during the first acquisition in the IG. • Significant effects were found for number of attempts but not for interface type.

Author, country	Population	Results	Statistical effect
De Freitas et al (2019) [17], Brazil	IG=60 male individuals with DMD; CG=60 typically developing male individuals. 9-34 years old. Divided with cross sectional design in three groups of 20 each one.	• Acquisition phase: significant effects were found for attempt but not for the device used. • Both groups increased the number of balls touched from first to last attempt. • IG performance in all interfaces was worse in Touch Screen than Kinect and Leap Motion. • CG had better performance for touch screen. • Age did not influence the learning effects for the gaming task. • Mean score in MFM in standing position and transfers was approximately 15%. • Punctuation in Vignos score and MFM in axial and proximal limb motor function predicted the improvement of performance in IG individuals.	Acquisition phase • First attempt (M^e=70), last attempt (M=78), IG (M=57), CG (M=91) • Touch Screen: CG: M=105; IG: M=50 • Leap Motion: CG: M=86; IG: M=62 • Kinect: CG: M=81; IG: M=54 Retention Phase (R) • IG in Leap Motion. Last attempt (M=67)–retention (M=81). • CG: retention (M=112)–transfer phase 1 (M=76). • IG: retention (M=81)–transfer phase 1 (M=59).

[a]DMD: Duchenne muscular dystrophy.

[b]IG: intervention group.

[c]CG: control group.

[d]MFM: Motor Function Measurement.

[e]M: median number of balls collected.

Table 3. Summary of measures, devices used, and types of virtual reality game interventions of the included studies.

Author, Country	Measurement instruments	Device/task	Virtual reality game intervention	Duration of study
Correa et al (2009) [18], Brazil	• Likert Scale: motivation scale. • User measures: ease of use, exercise effect, and satisfaction. • Therapeutic measures: practicability of equipment and degree of patient's motivation.	GenVirtual: a virtual reality–based game that simulates the sounds of musical instruments.	Elbow and wrist extension to play 3-dimensional colored cubes that simulate sounds of musical instruments.	30 minutes of intervention given on a weekly basis with a total of 12 sessions.
Capelini et al (2017) [5], Brazil	• Vignos Scale • Egen Klassification Scale • MFM[a] in dimension 1 (standing position and transfers), dimension 2 (axial and proximal limb motor function), and dimension 3 (distal limb motor function). • Dependent variable: movement time in seconds.	• Smartphone: Nokia 500 • Game: Marble maze classic (Labyrinth)	Individuals have to direct a virtual ball through a path maze and reach the final target in the shortest time possible. Different paths were used. Three phases were used: (1) acquisition, (2) retention, (3) transfer, which was divided into transfer phase 1, transfer phase 2, and transfer phase 3.	The authors did not describe the duration of the study.
Heuntinck et al (2018) [21], Brazil	• Principal measurement was performance of the upper limb. • MFM in dimension 3 (distal limb motor function). • Kidscreen-52 (quality of life). • Muscle strength: hand held • Dynamometer • Maximum voluntary isometric contractions • Vignos Scale	• Play station II with Eyetoy using a dynamic arm support. • Gainboy (Gravity compensation of 100% in horizontal plane)	• Performance of virtual reality games with dynamic arm support (Gainboy). • Individuals were trained at home and supervised by a coach.	5 weekly sessions of 15 minutes for a total of 20 weeks.
Massetti et al (2018) [22], Netherlands	• MFM • Time to perform in milliseconds.	Kinect Sensor MoVer software which allowed the creation of different tasks, during which participants perform functional movements.	An avatar interacts with the object. The object was represented as a red cube, and the participants received visual feedback with the image of their body movement in the virtual environment. Three phases were used: (1) acquisition, (2) retention, (3) transfer, which was divided into transfer phase 1, transfer phase 2, and transfer phase 3.	2 weeks.
Quadrado et al (2017) [23], Brazil	• MFM • Constant timing error in milliseconds. • Absolute timing error in milliseconds. • Variable timing error in milliseconds.	A webcam recorded a marker on the table next to the computer keyboard. Software superimposed virtual objects over images of the real world with a webcam. 10 3-dimensional cubes were displayed on a monitor.	The participant had to press the space bar on a keyboard to reach the cube or make a hand gesture with no physical contact at the exact moment the cube turned green. Three phases were used: (1) acquisition, (2) retention, (3) transfer, which was divided into transfer phase 1, transfer phase 2, and transfer phase 3.	The authors did not describe the duration of the study.
Malheiros et al (2015) [24], Brazil	• MFM in dimension 1 (standing position and transfers), dimension 2 (axial and proximal limb motor function), and dimension 3 (distal limb motor function). • Time to performance • Age of individual	Computer maze task. Following a path on the computer screen in the shortest possible time. The authors designed one maze for the acquisition phase and another maze for the transfer phase.	Participants executed a computer maze task; all participants performed the acquisition (20 attempts) and retention (5 attempts) phases, repeating the same maze. A different maze was used to verify transfer performance (5 attempts as well).	The authors did not describe the duration of the study.

Author, Country	Measurement instruments	Device/task	Virtual reality game intervention	Duration of study
De Freitas et al (2019) [17], Brazil	• MFM in dimension 1 (standing position and transfers), dimension 2 (axial and proximal limb motor function), and dimension 3 (distal limb motor function). • Vignos Scale • Age of population	Three interfaces were used: (1) Kinect for Windows, (2) Leap Motion (a virtual interface that required nonphysical contact), (3) a touch screen monitor that required physical contact.	All participants performed tasks using the 3 interfaces. There was an acquisition phase (during which participants practiced the task), a retention phase (1 attempt after 5 minutes), and 2 transfer phases (during which participants were given an attempt to use each device not used in the acquisition phase). Participants had to touch a ball on the screen in and outside the range of movement zone, challenging the limits of individuals.	There were 3 phases of the study for a total of 30 minutes.

[a]MFM: Motor Function Measurement.

Discussion

Virtual reality technologies have been used in the study of upper limb rehabilitation in recent years and for various conditions such as strokes, cerebral palsy, and neuromuscular diseases. Shin et al [15] found that virtual reality improved treatment outcomes for the distal upper extremity, including motor impairment, hand function, and quality of life in stroke survivors. Jannink et al [13] found that the Eye Toy tool had potential to improve arm function in children with cerebral palsy. De Freitas et al [17] found that those with Duchenne muscular dystrophy benefitted from the use of virtual technologies.

Methods and Population Classification

The evaluation of study participants' functionality and overall muscle performance was mostly done using psychosocial variables and scores obtained using the Vignos Scale [17,21]. The Motor Function Measure Scale was also used by the majority of authors chosen for the review [17,21,22,24].

Although those with Duchenne muscular dystrophy can benefit from the use of virtual reality technologies, the technologies in question should be selected carefully and with mind to participants' unique abilities and needs [17,22]. The studies reviewed agree on the importance of choosing a suitable task for individuals with Duchenne muscular dystrophy due to the generalized weakness that these individuals experience, especially in the upper limbs at the level of the forearm and in the antigravity muscles of the shoulder and elbow in advanced stages [17,23]. For this reason, the pattern of movement required by a task will determine the difficulty of that task [5,17,23]. Many studies found that patients with greater muscle weakness and loss of functionality had lower speeds of movement and worse performance during assigned tasks [5,17].

According to Masseti et al [22], De Freitas et al [17] found that individuals with Duchenne muscular dystrophy exhibited better responses in virtual environments than in the real world or when using touch interfaces. On the other hand, there was no relationship between performance of virtual reality tasks and general functionality (Motor Function Measurement-total), nor between performance of virtual reality tasks and functionality of the hands (Motor Function Measurement-D3) [5,17,23]. However, these results are based on a previous study in which training did not significantly improve performance of the upper limb (performance of upper limb=3.4). This training consisted of playing PlayStation II games for 20 weeks, which challenged participants to move their upper limbs with the help of a dynamic arm support (Gainboy) that compensated for gravity [21].

The studies of Heuntick et al [21] and Massetti et al [22] compared individuals with Duchenne muscular dystrophy in both the experimental and control groups. The other 5 studies included in our review compared a Duchenne muscular dystrophy group with a typical development control group [5,17,18,23,24]. The Duchenne muscular dystrophy groups always recorded a slower result and had lower scores (time to perform) in the acquisition phases [5,18,23,24].

Metrics and Results

Regarding the metrics used and results obtained, De Freitas et al [17] reported that age did not influence participants' functionality improvements or their ability to complete virtual tasks, even with the muscle deterioration caused by the progress of the disease.

The authors of the 7 papers we reviewed debated the influence of the Motor Function Measurement variable. De Freitas et al [17] did not find a relation for Motor Function Measurement-D1 in the performance of the task, while Capelini et al [5] did find a connection when using a smartphone. This may be why individuals in the Capelini et al study had a lower mean Motor Function Measurement-D1 score than individuals in the De Freitas et al study (De Freitas et al=12%, Capelini et al=15%) [5,17]. De Freitas et al [17] also reported that participants with more severe motor impairment and fewer motor skills in the proximal musculature (which corresponds to a higher score in Vignos and a lower score in Motor Function Measurement-D2) exhibited a greater capacity to complete homework, and learning was greater. This is, in part, because the score with which the participants started was also lower and, therefore, the range of improvement after training was broader.

Quadrado et al [23] performed a transfer phase with a change in velocity, and only Quadrado et al measured the time of the

task. Massetti et al [22] used the variable "time to perform" in the same phases as Malheiros et al [24]. However, Massetti et al used a crossover model with Duchenne muscular dystrophy patients whereas Malheiros et al had an unaffected control group.

In relation to the transfer of tasks, the results of the study by Quadrado et al [23] show that the realization of the task in a virtual environment facilitated the transfer to the real environment, but that the difficulty of the task must be adjusted to be more difficult in the virtual environment in order to facilitate the transfer to the real environment. This may be attributed to the fact that the space-time organization of the virtual task is different from the real one, reducing performance in the virtual environment [22,23]. Massetti et al [22] states that there is no transfer between environments due to the difference in the complexity of tasks. For example, while Quadrado et al [23] required only the movement of the hand to press a key, Massetti et al [22] required the entire upper limb to hit real and virtual cubes. Heuntick et al [21] supported the idea that if the training is transferred from a virtual environment, the effectiveness may be greater than standard exercises with low resistance, which are normally used to work muscles in isolation rather than general movements aimed at functionality.

Limitations

The variability of the samples in the studies and the differences between the registered variables made it impossible to carry out a meta-analysis to complete the review. The correlation between the type of task assigned to the participants and the participants' cognitive demand was not measured, nor was the correlation between the type of task and the participants' quality of movement [5,22,24]. The preparation of the environment or the accuracy of the devices was not specified and may have caused measurement errors [24,25]. The amount of audio-visual stimulus in some cases caused anxiety and did not produce the desired effects in low mobility individuals compared to conventional treatments [18]. The wide range of devices used makes it difficult to compare the results of the studies. The only similar devices were smartphones [5] and computer keyboards [24]; tasks at the motor level were completely different. If we introduce virtual reality as a tool that helps with user motivation and improves the users' quality of life, these variables should be used in all studies [18,21]. Finally, A great limitation in the use of virtual reality systems is their cost, making it impossible for all users to access.

Future Directions

The application of virtual reality must go beyond measuring the time for task completion and speed of movement. We must provide studies with specific respiratory and cardiac variables that provide data concerning the metabolic effort during and after exercise and the protection of cardiac activity during exercise. Therapy times and intensities should be defined in individuals with high fatigue and difficulty breathing [26]. Virtual reality can be a great tool for physiotherapy in individuals with Duchenne muscular dystrophy.

Improvement should be measured at different levels in the upper limbs, as well as the proximal level and stabilizing muscles and at the distal and fine motor level [15]. In the Shin et al [15] study, the authors reported benefits at the distal and proximal level when they only expected to find benefits at the distal level, as reported by Capelini et al [5].

Conclusions

The treatment of neuromuscular diseases has changed in recent years, from palliative symptom management to preventive methods for capacity building. The use of new technologies such as virtual reality may be necessary in the treatment of progressive diseases involving movement difficulties, as virtual reality technologies can provide freedom and improve capacity training results. Given that new technologies are increasingly accessible, rehabilitation and physiotherapy programs more frequently employ virtual reality environments to improve task performance and promote the transfer of this practice to daily life in the real world, which is essential for people with disabilities. Ultimately, virtual reality can be used for Duchenne muscular dystrophy rehabilitation programs to improve individuals' training performance.

Acknowledgments

The University of Malaga through a predoctoral grant supports RBB, 0619438300, a PhD student at the University. This research was partially supported by ICPFA. Grant number AI-07/2019.

Conflicts of Interest

None declared.

References

1. Fagoaga J, Girabent-Farrés M, Bagur-Calafat C, Febrer A, Steffensen BF. [Translation and validation of the Egen Klassifikation scale for the Spanish population: functional assessment for non-ambulatory individuals with Duchenne's muscular dystrophy and spinal muscular atrophy]. Rev Neurol 2013 Jun 01;56(11):555-561 [FREE Full text] [Medline: 23703057]
2. Guglieri M, Bushby K. Recent developments in the management of Duchenne muscular dystrophy. Paediatrics and Child Health 2015 Nov;25(11):505-514. [doi: 10.1016/j.paed.2015.07.002]

3. Lisboa C, Díaz O, Fadic R. [Noninvasive mechanical ventilation in patients with neuromuscular diseases and in patients with chest restriction]. Arch Bronconeumol 2003 Jul;39(7):314-320. [doi: 10.1016/s0300-2896(03)75392-7] [Medline: 12846961]
4. Vilozni D, Bar-Yishay E, Gur I, Shapira Y, Meyer S, Godfrey S. Computerized respiratory muscle training in children with Duchenne muscular dystrophy. Neuromuscul Disord 1994 May;4(3):249-255. [doi: 10.1016/0960-8966(94)90026-4] [Medline: 7919973]
5. Capelini CM, da Silva TD, Tonks J, Watson S, Alvarez MPB, de Menezes LDC, et al. Improvements in motor tasks through the use of smartphone technology for individuals with Duchenne muscular dystrophy. Neuropsychiatr Dis Treat 2017;13:2209-2217 [FREE Full text] [doi: 10.2147/NDT.S125466] [Medline: 28860778]
6. Guzman D, Londoño J. Upper limb rehabilitation with virtual environments: a review. RMIB 2016 Sep 1;37(3). [doi: 10.17488/rmib.37.3.8]
7. Broeren J, Claesson L, Goude D, Rydmark M, Sunnerhagen KS. Virtual rehabilitation in an activity centre for community-dwelling persons with stroke. The possibilities of 3-dimensional computer games. Cerebrovasc Dis 2008;26(3):289-296. [doi: 10.1159/000149576] [Medline: 18667809]
8. Merchant Z, Goetz ET, Cifuentes L, Keeney-Kennicutt W, Davis TJ. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. Computers & Education 2014 Jan;70:29-40. [doi: 10.1016/j.compedu.2013.07.033]
9. Nino González JI, Fernández Morales B. Comunicación, Salud y Tecnología: mHealth. Rev Comun y Salud 2015 Jan 01;5:144-153. [doi: 10.35669/revistadecomunicacionysalud.2015.5(1).149-158]
10. Pastora-Bernal JM, Martín-Valero R, Barón-López FJ, Moyano NG, Estebanez-Pérez M. Telerehabilitation after arthroscopic subacromial decompression is effective and not inferior to standard practice: Preliminary results. J Telemed Telecare 2018 Jul;24(6):428-433. [doi: 10.1177/1357633X17706583] [Medline: 28449618]
11. Morales-Navarro NA, Cossio-Martinez AG, Guzman-Sanchez JO, Hernandez-Meneses MI, Herrera-Matias K. Management of virtual environments using the “Kinect” sensor and its application in physiotherapy. Rev Tecnol Digit 2013;3(1):1-8.
12. Winkels DGM, Kottink AIR, Temmink RAJ, Nijlant JMM, Buurke JH. Wii™-habilitation of upper extremity function in children with Cerebral Palsy. An explorative study. Developmental Neurorehabilitation 2012 Oct 03;16(1):44-51. [doi: 10.3109/17518423.2012.713401]
13. Jannink MJA, van DWGJ, Navis DW, Visser G, Gussinklo J, Ijzerman M. A low-cost video game applied for training of upper extremity function in children with cerebral palsy: a pilot study. Cyberpsychol Behav 2008 Feb;11(1):27-32. [doi: 10.1089/cpb.2007.0014] [Medline: 18275309]
14. Gonzalez-Hernandez O, Peña-Vazquez M, Vazquez-Cervantes J. Bautista-Bautista ZI (dir), Velazquez-Sanchez AT (dir). Robotic exoskeletal for upper limb rehabilitation. 2011. URL: http://tesis.ipn.mx/handle/123456789/9942 [accessed 2020-02-04]
15. Shin J, Kim M, Lee J, Jeon Y, Kim S, Lee S, et al. Effects of virtual reality-based rehabilitation on distal upper extremity function and health-related quality of life: a single-blinded, randomized controlled trial. J NeuroEngineering Rehabil 2016 Feb 24;13(1). [doi: 10.1186/s12984-016-0125-x]
16. Rathinam C, Mohan V, Peirson J, Skinner J, Nethaji KS, Kuhn I. Effectiveness of virtual reality in the treatment of hand function in children with cerebral palsy: A systematic review. Journal of Hand Therapy 2019 Oct;32(4):426-434.e1. [doi: 10.1016/j.jht.2018.01.006]
17. de Freitas BL, da Silva TD, Crocetta TB, Massetti T, de Araújo LV, Coe S, et al. Analysis of Different Device Interactions in a Virtual Reality Task in Individuals With Duchenne Muscular Dystrophy-A Randomized Controlled Trial. Front Neurol 2019;10:24 [FREE Full text] [doi: 10.3389/fneur.2019.00024] [Medline: 30761066]
18. Correa-Dionisio AG, Klein AN, Lopes RD. In: Kheng TY, editor. Augmented Reality Musical System for Rehabilitation of Patients with Duchenne Muscular Dystrophy. Rehabilitation Engineering Shangai: In Tech; 2009:13-36.
19. Hutton B, Catalá-López F, Moher D. La extensión de la declaración PRISMA para revisiones sistemáticas que incorporan metaanálisis en red: PRISMA-NMA. Medicina Clínica 2016 Sep;147(6):262-266. [doi: 10.1016/j.medcli.2016.02.025]
20. Verhagen AP, de Vet HC, de Bie RA, Boers M, van den Brandt PA. The art of quality assessment of RCTs included in systematic reviews. J Clin Epidemiol 2001 Jul;54(7):651-654. [doi: 10.1016/s0895-4356(00)00360-7] [Medline: 11438404]
21. Heutinck L, Jansen M, van den Elzen Y, van der Pijl D, de Groot IJM. Virtual Reality Computer Gaming with Dynamic Arm Support in Boys with Duchenne Muscular Dystrophy. J Neuromuscul Dis 2018;5(3):359-372. [doi: 10.3233/JND-180307] [Medline: 29991140]
22. Massetti T, Fávero FM, Menezes LDCD, Alvarez MPB, Crocetta TB, Guarnieri R, et al. Achievement of Virtual and Real Objects Using a Short-Term Motor Learning Protocol in People with Duchenne Muscular Dystrophy: A Crossover Randomized Controlled Trial. Games Health J 2018 Apr;7(2):107-115. [doi: 10.1089/g4h.2016.0088] [Medline: 29608336]
23. Quadrado VH, Silva TDD, Favero FM, Tonks J, Massetti T, Monteiro CBDM. Motor learning from virtual reality to natural environments in individuals with Duchenne muscular dystrophy. Disabil Rehabil Assist Technol 2019 Jan;14(1):12-20. [doi: 10.1080/17483107.2017.1389998] [Medline: 29124971]

24. Malheiros SRP, da Silva TD, Favero FM, de Abreu LC, Fregni F, Ribeiro DC, et al. Computer task performance by subjects with Duchenne muscular dystrophy. Neuropsychiatr Dis Treat 2015;12:41-48 [FREE Full text] [doi: 10.2147/NDT.S87735] [Medline: 26766911]
25. McPherson AC, McAdam L, Keenan S, Schwellnus H, Biddiss E, DeFinney A, et al. A feasibility study using solution-focused coaching for health promotion in children and young people with Duchenne muscular dystrophy. Dev Neurorehabil 2018 Feb;21(2):121-130. [doi: 10.1080/17518423.2017.1289271] [Medline: 28272972]
26. Alvarez MPB, Silva TDD, Favero FM, Valenti VE, Raimundo RD, Vanderlei LCM, et al. Autonomic Modulation in Duchenne Muscular Dystrophy during a Computer Task: A Prospective Control Trial. PLoS One 2017;12(1):e0169633 [FREE Full text] [doi: 10.1371/journal.pone.0169633] [Medline: 28118369]

Edited by G Eysenbach, L Buis; submitted 18.06.20; peer-reviewed by J Seco Calvo, T Guetterman; comments to author 25.09.20; revised version received 19.10.20; accepted 24.10.20; published 08.12.20.

Please cite as:
Baeza-Barragán MR, Labajos Manzanares MT, Ruiz Vergara C, Casuso-Holgado MJ, Martín-Valero R
The Use of Virtual Reality Technologies in the Treatment of Duchenne Muscular Dystrophy: Systematic Review
JMIR Mhealth Uhealth 2020;8(12):e21576
URL: http://mhealth.jmir.org/2020/12/e21576/
doi:10.2196/21576
PMID:33289679

Review

Mobile Health Apps for Medical Emergencies: Systematic Review

Alejandro Plaza Roncero[1], BSc; Gonçalo Marques[2], PhD; Beatriz Sainz-De-Abajo[1], PhD; Francisco Martín-Rodríguez[3], PhD; Carlos del Pozo Vegas[4], PhD; Begonya Garcia-Zapirain[5], PhD; Isabel de la Torre-Díez[1], PhD

[1]Department of Signal Theory and Communications, and Telematics Engineering, University of Valladolid, Valladolid, Spain
[2]Polytechnic of Coimbra, Escola Superior de Tecnologia e Gestão de Oliveira do Hospital, Oliveira do Hospital, Portugal
[3]Advanced Clinical Simulation Center, School of Medicine, University of Valladolid, Valladolid, Spain
[4]Emergency Department, Hospital Clínico Universitario de Valladolid, Valladolid, Spain
[5]eVIDA Research Group, University of Deusto, Bilbao, Spain

Corresponding Author:
Begonya Garcia-Zapirain, PhD
eVIDA Research Group
University of Deusto
Avda/ Universidades 24
Bilbao, 48007
Spain
Phone: 34 619967223
Email: mbgarciazapi@deusto.es

Abstract

Background: Mobile health apps are used to improve the quality of health care. These apps are changing the current scenario in health care, and their numbers are increasing.

Objective: We wanted to perform an analysis of the current status of mobile health technologies and apps for medical emergencies. We aimed to synthesize the existing body of knowledge to provide relevant insights for this topic. Moreover, we wanted to identify common threads and gaps to support new challenging, interesting, and relevant research directions.

Methods: We reviewed the main relevant papers and apps available in the literature. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology was used in this review. The search criteria were adopted using systematic methods to select papers and apps. On one hand, a bibliographic review was carried out in different search databases to collect papers related to each application in the health emergency field using defined criteria. On the other hand, a review of mobile apps in two virtual storage platforms (Google Play Store and Apple App Store) was carried out. The Google Play Store and Apple App Store are related to the Android and iOS operating systems, respectively.

Results: In the literature review, 28 papers in the field of medical emergency were included. These studies were collected and selected according to established criteria. Moreover, we proposed a taxonomy using six groups of applications. In total, 324 mobile apps were found, with 192 identified in the Google Play Store and 132 identified in the Apple App Store.

Conclusions: We found that all apps in the Google Play Store were free, and 73 apps in the Apple App Store were paid, with the price ranging from US $0.89 to US $5.99. Moreover, 39% (11/28) of the included studies were related to warning systems for emergency services and 21% (6/28) were associated with disaster management apps.

(JMIR Mhealth Uhealth 2020;8(12):e18513) doi:10.2196/18513

KEYWORDS

mobile health; mHealth; eHealth; Android; iOS; medical emergencies; mobile apps

Introduction

Internet and mobile computing technologies have changed people's lifestyle. With regard to mobile devices in health, mobile devices, such as personal digital assistant devices (PDAs), smartphones, and tablets, have been widely adopted by medical professionals. These devices are quickly becoming some of the main instruments for accessing clinical information, especially for young health professionals and students [1]. Several medical resources are available on the digital distribution platforms of mobile apps (Google Play Store and Apple App Store) for Android and iOS operating systems [2].

According to the World Health Organization, the development of apps for the health domain is directly or indirectly intended to maintain or improve healthy behaviors, quality of life, and people's well-being [3].

Mobile health (mHealth) refers to the practice of medicine and public health. Robert Istepanian mentioned "the emerging use of mobile communications and network technologies for health" [4]. The field of mHealth has become a subbranch of eHealth, which has to do with the use of information and communication technologies, such as computers, mobile phones, GPS, and patient monitors for health and information services. mHealth includes the use of mobile devices in the collection, delivery, and access of health information by professionals, researchers, and patients. It is an emerging and rapidly developing field, which plays a vital role in the transformation of health care to increase its quality and efficiency.

On one hand, mobile apps are specifically aimed at helping people in their own health and wellness management. On the other hand, numerous mobile apps aim to assist health care providers as tools to improve and facilitate the provision of patient care [5]. According to a recent 2019 report on global mHealth, the market can be segmented based on the following: (1) equipment/connected medical devices, (2) mHealth services, and (3) mHealth apps [6].

The main objective of this paper was to present a systematic review that addresses the study of mobile apps for health emergencies. Furthermore, this paper presents the mobile apps available for the Android and iOS operating systems [7-11]. The main contribution is synthesis of the existing body of knowledge to provide relevant insights and to identify common threads and gaps to support new challenging, interesting, and relevant research directions.

In summary, mobile apps in the health sector are continually growing, and soon, they will be able to change the concept of medicine [7]. These apps will allow patients to access their health information, have small consultations for specific issues without consulting a professional, and locate emergency services. They will also help monitor chronic patients, increase safety in taking medication, and help network with people in the same situation. Moreover, professionals have access to specific information and tools to create new relationships with patients. Prehospital medical care starts from the occurrence of the event, involves transfer, and ends at admission to the welfare institution. Moreover, it always has to be offered by a health care professional. Consequently, we consider these apps as medical emergency apps.

A search for the term "medical emergency apps" in mHealth does not provide results. All reports and studies are focused on mHealth and do not distinguish between the different branches into which this technology can be divided. Therefore, we found a lack of interest in this important domain that can improve prehospital care for patients with the use of new technologies.

The use of mobile apps can facilitate the exchange of information between health professionals in the case of a possible health emergency. Consequently, this analysis has two focuses. First, we review the current status in the literature. Second, we review the mobile apps available in the main virtual stores. Currently, owing to the proliferation of mobile apps in this domain, it is necessary to evaluate their importance to promote health care.

Following the selection of relevant studies, a statistical analysis was carried out. The results have been discussed to analyze the main contributions of each publication. Finally, the most important findings have been reported.

Methods

Overview

In this section, the methodology used in this study is defined. The search process for information extraction from the available apps in the field of health emergencies is also reported. Our study focused first on the available literature and second on the available apps in the field of health emergency.

Literature Review

The procedure for the selection of articles was the same as that followed in other previous work [8]. The articles are analyzed by reading the title and abstract to identify the most relevant papers. The number of scientific publications is very high. Therefore, we followed a protocol that allowed us to synthesize the most relevant information. Two relevant protocols for systematic reviews are Quality of Reporting of Meta-Analysis (QUOROM) [9] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [10].

On one hand, QUOROM focuses on the presentation of a meta-analysis of randomized clinical trials and includes checklists for authors, reviewers, and editors of biomedical journals, as well as a diagram of the flow that describes the whole process. On the other hand, the PRISMA protocol is an extension of QUOROM with more pedagogical purposes accompanying the checklist with extensive documentation that justifies a series of check items. Furthermore, PRISMA is applicable to all types of systematic reviews and is not limited to a meta-analysis of clinical trials. This type of protocol was introduced with the idea that clinical trial publications follow the type of standards set by each protocol, and thus, works of this type will be standardized.

Therefore, this study adopted the PRISMA methodology. This process is divided into the following four phases:

- Identification: The title is considered in the choice.
- Selection: The summary for the choice of the paper is taken into account.
- Eligibility: The content is taken into account for the choice.
- Inclusion: We finally obtain the papers with the highest potential content.

The PRISMA protocol starts with the identification phase. In this stage, we used specified keywords to identify the relevant papers in several databases. Using these series of papers, we performed a set of steps to finally obtain the papers with which we carried out our review.

The search engines on which we obtained the different papers for the analysis were as follows: IEEE Xplore, Science Direct,

PubMed, Web of Science, and Google Scholar. These databases were used since they cover the majority of papers that are within the scope of this review and include the most relevant sources. Moreover, the above-mentioned databases have been used in several systematic review papers on mHealth [8,11,12]. The papers were selected and screened by two different reviewers. Moreover, all the selected papers were included with the common agreement of all the authors.

The logic and keywords used to conduct this review were as follows: “emergency” AND “app,” “emergency” AND “mHealth,” “emergency” AND “eHealth,” and “eEmergency.”

We focused on review papers and research papers, excluding other results that these search engines offer, such as book chapters, patents, conference summaries, and news. The search was carried out starting from the year 2009, considering the papers published for 10 years until the end of 2019. The search and selection of papers were conducted during March 2020.

Review papers were included since they collect information from the most relevant sources. These papers provide clear and concise insights, which were used to carry out our analysis. Concerning the dates of publication, this systematic review considered the previous 10 years. This review only included papers in the English language because it is the universal language par excellence (Lingua Franca).

After the identification process, the papers were ordered according to relevance. Moreover, the identification process followed the PRISMA protocol that can be represented in a flow chart. Figure 1 presents the section process of a particular search engine through which the search was performed and for choosing a particular search string or logic.

After the identification process, the obtained papers from a different database were exanimated to find duplicates. Thereafter, we conducted the selection phase. The section state started with the exclusion of papers by reading the title and abstract. The eligibility phase included reading the full text of the remaining papers obtained in the selection phase. Finally, in the inclusion phase, the final number of papers included in the review were defined. This entire process is presented in Figure 2. In total, we found 28 relevant papers that met the search criteria and respected the exclusion criteria.

Figure 1. Flow diagram for the identification phase.

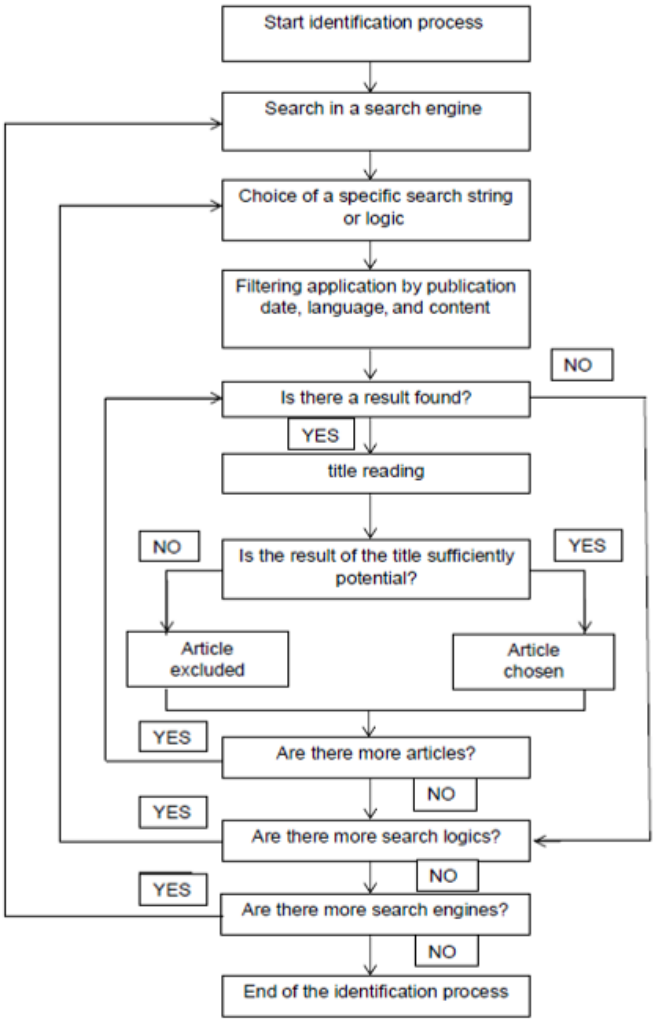
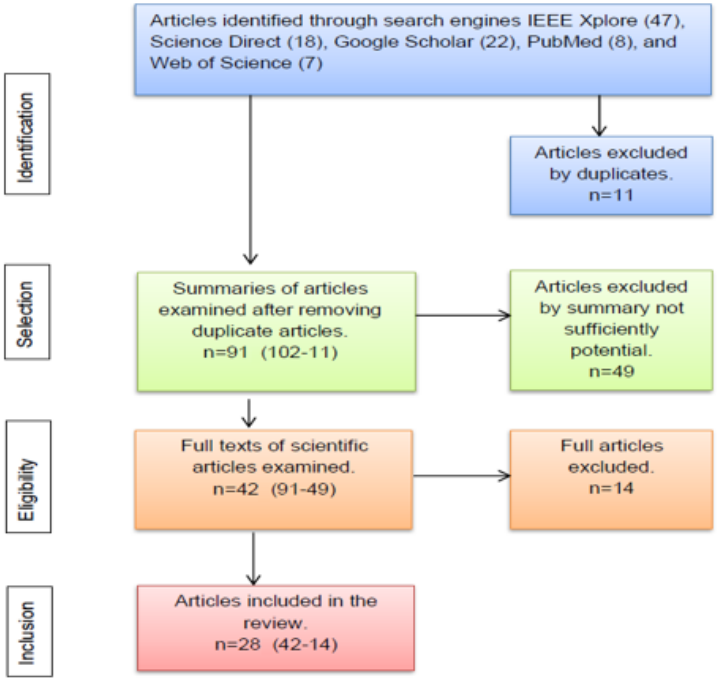


Figure 2. PRISMA flow chart.



Review of the Apps Available in Mobile Market Stores

After reviewing the available literature, a review of the mobile apps and websites available was conducted.

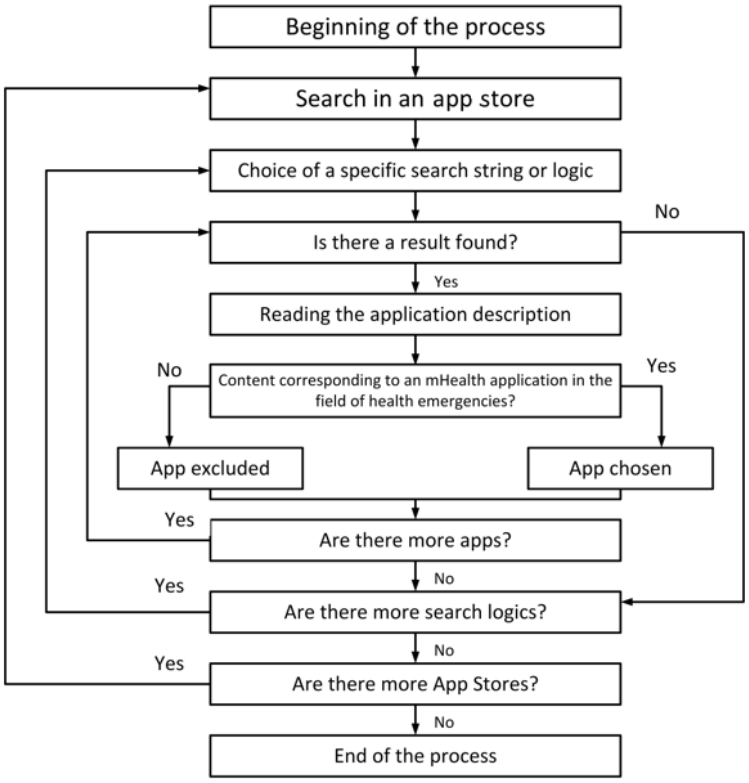
On one hand, the Google Play Store and Apple App Store were used to search for mobile apps, since they are the two app stores that have more apps. On the other hand, Google Chrome was used to perform a webpage search on eHealth. The search process on available mobile apps was conducted during March 2020.

The methodology used for this second part was very similar to the flow chart in Figure 3. The following keywords were defined to try to obtain all possible results for this analysis: “emergency” OR “emergencia,” “eEmergency,” “safety” OR “seguridad,” “alert” OR “alerta,” “disaster” OR “desastre,” “SOS,” “112,” and “blood donation” OR “donación de sangre.”

In addition, the language and country used by the different mobile apps were not taken into account. The criterion followed to choose mobile apps in terms of content was to analyze the information offered by each search engine about the app. If the app corresponded to the field of health emergency (whether designed to help health care staff or the patient), it was considered in our study.

Figure 3 presents the chart that was followed in the process of selecting the different mobile apps in our study.

Figure 3. Flowchart process for selection of apps.



Results

Overview

This section presents first an analysis of the results obtained in the literature review and second an analysis of the available mobile apps.

The number of results obtained per year after the analysis and selection conducted by the authors is presented in Table 1. The findings have been categorized according to the year of publication and the number of obtained results.

The distribution of the number of papers selected after the systematic review according to the year is presented in Table 2. The studies analyzed in this systematic review were distributed from 2013 to 2019, and most of the results (n=8) involved 2017.

Table 3 presents the title, the date of publication, and the summary of the main contributions of each paper included in the systematic review [13-40].

Table 1. Distribution of publications per year before applying the selection criteria.

Year	Number of studies
2009	21,782
2010	24,995
2011	29,327
2012	33,234
2013	37,770
2014	44,582
2015	51,111
2016	56,082
2017	63,540
2018	67,289
2019	54,735

Table 2. Distribution of publications per year after applying the methodology.

Year	Number of studies
2009	0
2010	0
2011	0
2012	0
2013	1
2014	3
2015	3
2016	4
2017	8
2018	7
2019	2

Table 3. Main contributions of each paper included.

Title	Year	Main contributions
With the Proliferation of Mobile Medical Apps, Which Ones Work Best in the Emergency Department? [13]	2015	Study on the best apps for the emergency department.
Prehospital emergency notification system [14]	2016	Proposal of a mobile app that allows emergency services to provide the hospital with information about the severity of a victim.
Vehicle-Assist Resilient Information and Network System for Disaster Management [15]	2017	Development of a network for disaster management that will be used if the internet is not available and which is made up of possible drones that collect information from a mobile app and send it to geo-distributed servers.
MHealth based ubiquitous fall detection for elderly [16]	2017	Development of a prototype based on an accelerometer sensor for the detection of falls in the elderly. This prototype uses a mobile app to send an alert to emergency services in case of a fall.
STLS: Smart traffic lights system for emergency response vehicles [17]	2019	Android app for monitoring the traffic network of a city to allow emergency vehicles to circulate without conflicts.
State of the Earthquake Field Disaster Investigation Information Service System [18]	2019	Disaster notification system to inform the inhabitants of a country about the arrival of an earthquake to better manage this type of disaster.
Smartphones let surgeons know WhatsApp: an analysis of communication in emergency surgical teams [19]	2015	Study on the improvement of communication and effectiveness in a surgical team thanks to the use of the instant messenger WhatsApp.
MHealth: Blood donation application using android smartphone [20]	2016	App to make blood donations using mobile devices with the Android operating system.
SaveMe: A crime deterrent personal safety android app with a Bluetooth connected hardware switch [21]	2019	Switch connected to a smartphone via Bluetooth that can be pressed for warning about danger to the emergency contact of the victim.
Click away emergency aid scheme by means of intelligent situation assessment [22]	2018	Development of a mobile app to help people access emergency services easily.
Blood donation and life saver app [23]	2018	App for the search of a blood donor with specific characteristics. Obtain possible donors through geolocation and send them an alert message.
Blood bank app using Raspberry PI [24]	2018	Development of an app using Raspberry to shorten blood donation times.
An effective support system of emergency medical services with tablet computers [25]	2015	Study for the evaluation of a system of tablets installed in emergency vehicles to shorten service times and thus improve prehospital medical care.
An Integrated mHealth and Vehicular Sensor Based Alarm System Emergency Alarm Notification System for Long Distance Drivers using Smart Devices and Cloud Networks [26]	2019	App for monitoring the vital signs of professionals who are dedicated to covering long distances on roads to avoid possible traffic accidents.
Intelligent crash detection and emergency communication system for two-wheelers [27]	2019	Accident detection system for two-wheeled vehicles that warn the driver of critical points where an accident can occur through the use of machine learning technology.
Integration of Emergency Web App for Accessing the Emergency Services by Mobile Phones [28]	2013	App for emergency management that can notify emergency services through a smartphone.
Instantaneous feedback pedometer with emergency GPS tracker [29]	2019	App for notifying emergency services in case of sudden heart failure among people who play sports.
Information and communication technologies for enhanced emergency management in Taiwan high-speed rail [30]	2016	System of identification of an environmental hazard integrated into high-speed trains to improve safety.
iEMS1669: An innovative Med Alert App for Thai Emergency Medical System [31]	2017	App to provide a medical alert to emergency services, family, and friends.
Facilitating the collection and dissemination of patient care information for emergency medical personnel [32]	2016	System for sending information from emergency technicians to the hospital to notify about the patient's situation and improve prehospital medical care.
Indoor Localization for Evacuation Management in Emergency Scenarios [33]	2018	System formed by an app and a central monitoring system for the correct evacuation of a building.
An Analysis of WhatsApp Usage for Communication Between Consulting and Emergency Physicians [34]	2016	Study on the use of WhatsApp for communication between the administration of consultations and emergency doctors.

Title	Year	Main contributions
A triggering mechanism for end-to-end IoT eHealth system with connected ambulance vehicles [35]	2018	User monitoring system (generally designed for the elderly) to notify an ambulance in case of emergency.
A Mobile/Cloud Emergency Response Application for Indoor Assisted Living [36]	2014	App for notifying emergency services in case of emergency.
A mobile-based emergency reporting application for the Philippine National Police Emergency Hotline 911: A case for the development of i911 [37]	2018	App for the Philippine National Police to accelerate the response time with the collection of user data.
DETSApp: An App for Disaster Event Tweets Summarization using Images Posted on Twitter [38]	2018	App to summarize a disaster by compiling a Twitter post.
Design and development of a crowdsourcing mobile app for disaster response [39]	2017	App to disseminate geographic information in case of an emergency.
Reducing Traffic Congestion Using Geo-fence Technology: Application for Emergency Car [40]	2014	App for monitoring the traffic network of a city to allow emergency vehicles to circulate without conflict.

Results for the Google Play Store

The results obtained for the Android operating system were analyzed. The categorization of apps provided by the Google Play Store is presented in Table 4.

The results of the apps obtained for Android according to the filter used in the search conducted are presented in Table 5.

Table 4. Categorization of apps for Android.

Categorization	Number of apps
Health and Wellness	56
Medicine	48
Communication	17
Tools	17
Social	15
Lifestyle	10
Travel and Guides	6
Education	5
Business	5
News and Magazines	5
Productivity	4
Maps and Navigation	3
Entertainment	1

Table 5. Categorization of apps by filter for Android.

Categorization	Number of apps
Blood donation OR donación de sangre	73
Emergency OR emergencia	51
SOS	24
112	11
Alert OR alerta	11
Safety OR seguridad	11
eEmergency	8
Disaster OR desastre	3

Results for the Apple App Store

The results obtained for the iOS mobile operating system were analyzed. The categorization of apps provided by the Apple App Store is presented in Table 6.

The results of the apps obtained for iOS according to the filter used in the search conducted are presented in Table 7.

Table 6. Categorization of apps for iOS.

Categorization	Number of apps
Medicine	20
Tools	19
Health and Wellness	17
Travel and Guides	17
Lifestyle	16
Maps and Navigation	10
Business	9
News and Magazines	9
Education	6
Productivity	4
Entertainment	3
Communication	1
Social	1

Table 7. Categorization of apps by filter for iOS.

Categorization	Number of apps
Emergency OR emergencia	38
SOS	32
Alert OR alerta	23
112	13
Safety OR seguridad	12
Disaster OR desastre	7
Blood donation OR donación de sangre	6
eEmergency	1

Comparison of the Results

Comparison of the categorization of the apps on each platform (Google Play Store and Apple App Store) was conducted, and the findings are presented in Table 8.

The results of all apps obtained according to the filter used are presented in Table 9.

In total, 192 selected mobile apps with potential content were obtained from the Google Play Store, while 132 were obtained from the Apple App Store.

On the Android operating system, the most number of results were obtained in the Health and Wellness, and Medicine categories. Out of the 192 apps obtained, 56 apps belonged to the Health and Wellness category. Moreover, 48 apps belonged to the Medicine category. However, on the iOS operating system, in the Apple App Store, we found a more equitable distribution of the number of apps when considering the category. Out of 132 apps, 20 belonged to the category of Medicine, 19 belonged to the category of Tools, 17 belonged to the categories of Health and Wellness, and Travel and Guides, and 16 belonged to the category of Lifestyle.

In total, 73 and 51 apps were obtained from the Google Play Store using the filters “blood donation” OR “donación de sangre” and “emergency” OR “emergencia,” respectively. For the Apple App Store, 38, 32, and 23 mobile apps were obtained with the filters “emergency” OR “emergencia,” “SOS,” and “alert” OR “alerta,” respectively.

Out of 324 apps, 73 and 68 were from the Health and Wellness, and Medicine categories, respectively.

Finally, regarding the categorization according to the filter applied in the search, the three filters through which more apps with potential content were obtained included “emergency” OR “emergencia,” “blood donation” OR “donación de sangre,” and “SOS,” with 89, 79, and 56 results, respectively.

Regarding the price of apps, we found that all identified apps available in the Google Play Store were free. However, 59 mobile apps in the Apple App Store were paid, with the price ranging from US $0.89 to US $5.99. The price of apps is relevant to their use since it can be a critical limitation to their use by people with economic issues.

In total, 79% (22/28) of studies were found in the IEEE Xplore database. Moreover, 17% (5/28) of the studies were selected from Google Scholar and 4% (1/28) from ScienceDirect. Finally, an analysis of the filters used in the search process has been conducted.

According to the filters applied in the search engines, 79% (22/28) of papers resulted from the "emergency" AND "app" filter and 17% (5/28) resulted from the "emergency" AND "mHealth" filter. Moreover, 4% (1/28) of the papers resulted from the "emergency" AND "eHealth" filter. The "eEmergency" filter did not return relevant results.

Table 8. Categorization of apps (Android and iOS).

Categorization	Number of apps
Health and Wellness	73
Medicine	68
Tools	36
Lifestyle	26
Travel and Guides	23
Communication	18
Social	16
Business	14
News and Magazines	14
Maps and Navigation	13
Education	11
Productivity	8
Entertainment	4

Table 9. Categorization of apps by filter (Android and iOS).

Categorization	Number of apps
Emergency OR emergencia	89
Blood donation OR donación de sangre	79
SOS	56
Alert OR alerta	34
112	24
Safety OR seguridad	23
Disaster OR desastre	10
eEmergency	9

Discussion

Principal Findings

The 28 obtained papers were categorized by the authors into the following six different groups: (1) prehospital medical care; (2) apps for disaster management; (3) warning systems for emergency services and medical services; (4) automobile circulation control; (5) communication between medical staff; and (6) apps for blood donation. The classification of the analyzed studies is presented in Table 10.

Warning systems for emergency services and medical services led, with 39% (11/28) of the obtained publications. These systems are usually apps in which medical services are notified with the press of a button. Followed by this, we have the group of apps for disaster management, owing to the problems that exist with different natural disasters in Eastern countries (6/28 [21%] of the obtained publications). These two groups had a higher percentage than the rest of the groups. Communication between medical staff is critical for the success of emergency care services. Effective and efficient methods of communication involving health care staff play major roles in improving global health care services. Apps for blood donation are also critical, since in emergency scenarios, the availability of the correct blood type for a patient is crucial for the recovery process.

Prehospital care should be highlighted to include prehospital medical care methods, as well as automobile circulation control for the circulation of medical vehicles since it directly influences an improvement in prehospital medical care by improving the time it takes for a medical vehicle to care for patients.

mHealth has great potential, as it can provide citizens with the necessary means to manage their health and stay healthy longer. Consequently, people can improve the quality of health care and patient comfort, and help health professionals in their work.

The search for mHealth solutions can contribute to the development of modern, efficient, and sustainable health systems. It is also expected to reduce costly visits to the hospital, help citizens to take charge of their state of health and well-being, and promote health focused on prevention rather than cure. Furthermore, it is an excellent opportunity for the flourishing app sector and entrepreneurs.

This literature review mentioned the relevant usage of mobile apps in the health emergency domain. In addition, this paper stated the need to investigate the realization of more studies on prehospital care. Recent studies are focused on mHealth technology, and they leave aside the different branches that may arise from this technology.

Wearable devices are becoming increasingly relevant not only in applications in the health sector, but also in global mobile telecommunication. The applications can be notified through a mobile device, such as a smartphone or tablet and a wearable device.

Considering the extensive inclusion of 5G technology, mobile communication technology can be assumed to be experiencing a breakthrough, since we can send more information in seconds and with less energy consumption. In addition, the network coverage offered by mobile communication technologies can help to reach remote areas such as mountains. Moreover, at present, mobile networks offer high data transfer rates, which enable remote surgery tasks. Finally, given the increase in the use of mobile devices worldwide, we found a large number of apps.

There was greater availability of apps for Android than for iOS. Furthermore, Android apps were free compared with iOS apps (73/132 [55.5%] were paid). Considering the large number of apps found in the category of Medicine in this study, we can conclude that mobile apps are mainly introduced in this area. Moreover, the category of Health and Wellness involved even more mobile apps. Even though mobile technology has increased, there is much growth in the field of health emergencies and mHealth.

After the completion of this work, three essential aspects are planned as future lines of research. Owing to the tremendous digital transformation experienced by both industry and society, they increasingly require the use of a highly organized infrastructure on the network, that is, on the internet, and the loading or unloading of large amounts of data, which, in the case of health, is of utmost importance. Therefore, the three most relevant technologies that will be increasingly relevant owing to both the demand for cloud services and large volumes of data, and the search for the solvency of vulnerabilities in data management are cybersecurity, big data, and cloud computing. In addition, wearable devices are a critical matter of study in the mHealth app space. Finally, internet of things and smart sensor communication are becoming increasingly widespread and are crucial for enhanced telemedicine.

Table 10. Distribution of the studies per category.

Category	References	Number of studies
I. Prehospital medical care	[14,26,32]	3
II. Apps for disaster management	[15,18,31,33,38,39]	6
III. Warning systems for emergency and medical services	[13,16,17,21,22,28-30,35-37]	11
IV. Automobile circulation control	[27,40]	2
V. Communication between medical staff	[19,25,34]	3
VI. Apps for blood donation	[20,23,24]	3

Conclusion

We conducted an analysis of the current status of mHealth technologies and apps for medical emergencies. The PRISMA methodology was used in this review. First, the available literature of the previous 10 years (2009-2019) was analyzed. Second, a review of mobile apps in the two common virtual storage platforms (Google Play Store and Apple App Store) was carried out. The Google Play Store and Apple App Store are for the Android and iOS operating systems, respectively.

In total, 28 papers in the field of medical emergencies were included. These studies were categorized into six different groups. Overall, 39% (11/28) of the included studies were related to warning systems for emergency services and 21% (6/28) were associated with disaster management apps.

In total, 324 mobile apps were found, with 59.3% (n=192) identified in the Google Play Store and 40.7% (n=132) identified in the Apple App Store. All identified mobile apps in the Google Play Store were free, and in the Apple App Store, 55.5% (73/132) of the identified apps were paid, with the price ranging from US $0.89 to US $5.99.

Acknowledgments

This research has been partially supported by the European Commission and the Ministry of Industry, Energy and Tourism under the project AAL-20125036 named “Wetake Care: ICT-based Solution for (Self-) Management of Daily Living.”

Conflicts of Interest

None declared.

References

1. Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares. Instituto Nacional de Estadística. 2019. URL: https://www.ine.es/prensa/tich_2019.pdf [accessed 2020-01-19]
2. Fernandez-Luque L, Labarta JI, Palmer E, Koledova E. Content Analysis of Apps for Growth Monitoring and Growth Hormone Treatment: Systematic Search in the Android App Store. JMIR Mhealth Uhealth 2020 Feb 18;8(2):e16208 [FREE Full text] [doi: 10.2196/16208] [Medline: 32130162]
3. Gazdecki A. 9 Mobile Technology Trends For 2017 (Infographic). BiznessApps. URL: https://www.biznessapps.com/blog/mobile-technology-trends/ [accessed 2019-11-06]
4. La Sociedad de la Información en España 2016. Fundación Telefónica. URL: https://www.fundaciontelefonica.com/cultura-digital/publicaciones/la-sociedad-de-la-informacion-en-espana-2016/558/ [accessed 2020-01-19]
5. Martinez-Millana A, Jarones E, Fernandez-Llatas C, Hartvigsen G, Traver V. App Features for Type 1 Diabetes Support and Patient Empowerment: Systematic Literature Review and Benchmark Comparison. JMIR Mhealth Uhealth 2018 Nov 21;6(11):e12237 [FREE Full text] [doi: 10.2196/12237] [Medline: 30463839]
6. Thomas L. Global Mobile Health Market Size, Share, Growth, Trends, and Forecast, 2016-2022. America News Hour. URL: https://www.americanewshour.com/2019/09/26/global-mobile-health-market-size-share-growth-trends-and-forecast-2016-2022/63341/ [accessed 2020-01-19]
7. Gaziel-Yablowitz M, Schwartz DG. A Review and Assessment Framework for Mobile-Based Emergency Intervention Apps. ACM Comput. Surv 2018 Apr 14;51(1):1-32. [doi: 10.1145/3145846]
8. Franco-Martín MA, Muñoz-Sánchez JL, Sainz-de-Abajo B, Castillo-Sánchez G, Hamrioui S, de la Torre-Díez I. A Systematic Literature Review of Technologies for Suicidal Behavior Prevention. J Med Syst 2018 Mar 05;42(4):71. [doi: 10.1007/s10916-018-0926-5] [Medline: 29508152]
9. Moher D, Cook D, Eastwood S, Olkin I, Rennie D, Stroup D. Improving the Quality of Reports of Meta-Analyses of Randomised Controlled Trials: The QUOROM Statement. Onkologie 2000 Dec 15;23(6):597-602. [doi: 10.1159/000055014] [Medline: 11441269]
10. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
11. Góngora Alonso S, Hamrioui S, de la Torre Díez I, Motta Cruz E, López-Coronado M, Franco M. Social Robots for People with Aging and Dementia: A Systematic Review of Literature. Telemed J E Health 2019 Jul;25(7):533-540. [doi: 10.1089/tmj.2018.0051] [Medline: 30136901]
12. de la Torre Díez I, Alonso SG, Hamrioui S, López-Coronado M, Cruz EM. Systematic Review about QoS and QoE in Telemedicine and eHealth Services and Applications. J Med Syst 2018 Aug 29;42(10):182. [doi: 10.1007/s10916-018-1040-4] [Medline: 30155565]
13. Huffman A. With the Proliferation of Mobile Medical Apps, Which Ones Work Best in the Emergency Department? Annals of Emergency Medicine 2015 Aug;66(2):A13-A15. [doi: 10.1016/j.annemergmed.2015.06.010]
14. Sarlan A, Xiong FK, Ahmad R, Ahmad WF, Bhattacharyya E. Pre-hospital emergency notification system. 2015 Presented at: 2015 International Symposium on Mathematical Sciences and Computing Research (iSMSC); May 19-20, 2015; Ipon, Malaysia. [doi: 10.1109/ISMSC.2015.7594047]
15. Li P, Miyazaki T, Wang K, Guo S, Zhuang W. Vehicle-assist resilient information and network system for disaster management. IEEE Trans. Emerg. Topics Comput 2017 Jul 1;5(3):438-448. [doi: 10.1109/tetc.2017.2693286]
16. Bhati N. MHealth based ubiquitous fall detection for elderly people. 2017 Presented at: 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT); July 3-5, 2017; Delhi, India. [doi: 10.1109/ICCCNT.2017.8204033]
17. Almuraykhi KM, Akhlaq M. STLS: Smart Traffic Lights System for Emergency Response Vehicles. 2019 Presented at: 2019 International Conference on Computer and Information Sciences (ICCIS); April 3-4, 2019; Sakaka, Saudi Arabia. [doi: 10.1109/ICCISci.2019.8716429]
18. Sun BT, Hu HB, Chen XZ. State of the Earthquake Field Disaster Investigation Information Service System. 2019 Presented at: 13th Symposium on Piezoelectrcity, Acoustic Waves and Device Applications (SPAWDA); January 11-14, 2019; Harbin, China. [doi: 10.1109/SPAWDA.2019.8681791]

19. Johnston MJ, King D, Arora S, Behar N, Athanasiou T, Sevdalis N, et al. Smartphones let surgeons know WhatsApp: an analysis of communication in emergency surgical teams. Am J Surg 2015 Jan;209(1):45-51. [doi: 10.1016/j.amjsurg.2014.08.030] [Medline: 25454952]
20. Fahim M, Cebe HI, Rasheed J, Kiani F. mHealth: Blood donation application using android smartphone. 2016 Presented at: Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP); July 21-23, 2016; Konya, Turkey. [doi: 10.1109/DICTAP.2016.7543997]
21. Tripti NF, Farhad A, Iqbal W, Zaman HU. SaveMe: A Crime Deterrent Personal Safety Android App with a Bluetooth Connected Hardware Switch. 2019 Presented at: 9th IEEE Control and System Graduate Research Colloquium (ICSGRC); August 3-4, 2018; Shah Alam, Malaysia. [doi: 10.1109/ICSGRC.2018.8657489]
22. Pughazendi N, Sathish Kumar R, Karthikeyan A, Seshan PV. Click away emergency aid scheme by means of intelligent situation assessment. 2018 Presented at: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC); March 22-23, 2017; Melmaruvathur, India. [doi: 10.1109/ICCPEIC.2017.8290343]
23. Annish Brislin MR, Albert Mayan J, Aroul Canessane R, Anish Hamlin MR. Blood donation and life saver app. 2017 Presented at: 2nd International Conference on Communication and Electronics Systems (ICCES); October 19-20, 2017; Coimbatore, India. [doi: 10.1109/CESYS.2017.8321318]
24. Pohandulkar SS, Khandelwal CS. Blood Bank App using Raspberry PI. 2018 Presented at: 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS); December 21-22, 2018; Belgaum, India. [doi: 10.1109/CTEMS.2018.8769143]
25. Yamada KC, Inoue S, Sakamoto Y. An effective support system of emergency medical services with tablet computers. JMIR Mhealth Uhealth 2015 Feb 27;3(1):e23 [FREE Full text] [doi: 10.2196/mhealth.3293] [Medline: 25803096]
26. Kang JJ, Venkatraman S. An Integrated mHealth and Vehicular Sensor Based Alarm System Emergency Alarm Notification System for Long Distance Drivers using Smart Devices and Cloud Networks. 2018 Presented at: 28th International Telecommunication Networks and Applications Conference (ITNAC); November 21-23, 2018; Sydney, NSW, Australia. [doi: 10.1109/ATNAC.2018.8615441]
27. Satya RL, Kaviya R, Valarmathi R. Intelligent Crash Detection and Emergency communication system for Two Wheelers. 2018 Presented at: 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT); February 15-17, 2018; Chennai, India. [doi: 10.1109/IC3IoT.2018.8668180]
28. Gómez B, Juiz C. Integration of Emergency Web App for accessing the emergency services by mobile phones. 2013 Presented at: 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013); October 9-12, 2013; Lisbon, Portugal. [doi: 10.1109/HealthCom.2013.6720658]
29. Nair MB, Kumar SR, Kishore NA, Mohan N, Anudev J. Instantaneous feedback pedometer with emergency GPS tracker. 2018 Presented at: 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC); August 30-31, 2018; Palladam, India. [doi: 10.1109/I-SMAC.2018.8653718]
30. Jen YY, Jason Chang SK. Information and communication technologies for enhanced Emergency Management in Taiwan High Speed Rail. 2016 Presented at: 2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT); August 23-25, 2016; Birmingham, UK. [doi: 10.1109/ICIRT.2016.7588552]
31. Jayapravitra Y, Prasertsil W, Limim N, Phatthanacharoensuk I, Songsinviboon T. iEMS1669: An innovative Med Alert App for Thai Emergency Medical System. 2017 Presented at: 2017 International Conference on Digital Arts, Media and Technology (ICDAMT); March 1-4, 2017; Chiang Mai, Thailand. [doi: 10.1109/ICDAMT.2017.7904979]
32. Fleshman MA, Argueta IJ, Austin CA, Lee HH, Moyer EJ, Gerling GJ. Facilitating the collection and dissemination of patient care information for emergency medical personnel. 2016 Presented at: 2016 IEEE Systems and Information Engineering Design Symposium (SIEDS); April 29, 2016; Charlottesville, VA, USA. [doi: 10.1109/SIEDS.2016.7489306]
33. Depari A, Flammini A, Fogli D, Magrino P. Indoor Localization for Evacuation Management in Emergency Scenarios. 2018 Presented at: 2018 Workshop on Metrology for Industry 4.0 and IoT; April 16-18, 2018; Brescia, Italy. [doi: 10.1109/metroi4.2018.8428343]
34. Gulacti U, Lok U, Hatipoglu S, Polat H. An Analysis of WhatsApp Usage for Communication Between Consulting and Emergency Physicians. J Med Syst 2016 Jun;40(6):130. [doi: 10.1007/s10916-016-0483-8] [Medline: 27083574]
35. Elsaadany A, Sedky A, Elkholy N. A triggering mechanism for end-to-end IoT eHealth system with connected ambulance vehicles. 2017 Presented at: 8th International Conference on Information, Intelligence, Systems & Applications (IISA); August 27-30, 2017; Larnaca, Cyprus. [doi: 10.1109/iisa.2017.8316405]
36. Psaila G, Scandurra P, Rovelli S, Mazzucchelli E, Taiocchi M. A Mobile/Cloud Emergency Response Application for Indoor Assisted Living. 2014 Presented at: 2014 Eighth International Conference on Next Generation Mobile Apps, Services and Technologies; September 10-12, 2014; Oxford, UK. [doi: 10.1109/ngmast.2014.39]
37. Edillo SB, Garrote PJ, Domingo LC, Malapit AG, Fabito BS. A mobile based emergency reporting application for the Philippine National Police Emergency Hotline 911: A case for the development of i911. 2018 Presented at: Tenth International Conference on Mobile Computing and Ubiquitous Network (ICMU); October 3-5, 2017; Toyama, Japan. [doi: 10.23919/icmu.2017.8330110]

38. Layek AK, Pal A, Saha R, Mandal S. DETSApp: An App for Disaster Event Tweets Summarization using Images Posted on Twitter. 2018 Presented at: Fifth International Conference on Emerging Applications of Information Technology (EAIT); January 12-13, 2018; Kolkata, India. [doi: 10.1109/eait.2018.8470398]
39. Li L, Ulaganathan MN. Design and development of a crowdsourcing mobile app for disaster response. 2017 Presented at: 25th International Conference on Geoinformatics; August 2-4, 2017; Buffalo, NY, USA. [doi: 10.1109/geoinformatics.2017.8090943]
40. Noei S, Santana H, Sargolzaei A, Noei M. Reducing Traffic Congestion Using Geo-fence Technology: Application for Emergency Car. In: EMASC '14: Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities. 2014 Presented at: 2014 ACM Multimedia Conference; November 2014; Orlando, FL, USA p. 15-20. [doi: 10.1145/2661704.2661709]

Abbreviations

mHealth: mobile health
PDA: personal digital assistant device
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUOROM: Quality of Reporting of Meta-Analysis

Edited by D Schwartz, L Buis; submitted 02.03.20; peer-reviewed by M Yablowitz, I Yoo, A Martinez-Millana, JH Lee; comments to author 09.06.20; revised version received 29.07.20; accepted 17.08.20; published 11.12.20.

Please cite as:
Plaza Roncero A, Marques G, Sainz-De-Abajo B, Martín-Rodríguez F, del Pozo Vegas C, Garcia-Zapirain B, de la Torre-Díez I
Mobile Health Apps for Medical Emergencies: Systematic Review
JMIR Mhealth Uhealth 2020;8(12):e18513
URL: http://mhealth.jmir.org/2020/12/e18513/
doi:10.2196/18513
PMID:33306037

Review

Mobile Health Strategies to Tackle Skin Neglected Tropical Diseases With Recommendations From Innovative Experiences: Systematic Review

Carme Carrion[1,2,3], MSc, PhD; Noemí Robles[1,2,3], PhD; Oriol Sola-Morales[4], MD, PhD; Marta Aymerich[1,2], MD, PhD, MPH; Jose Antonio Ruiz Postigo[5], MD, PhD

[1]eHealth Lab Research Group, School of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

[2]eHealth Center, Universitat Oberta de Catalunya, Barcelona, Spain

[3]Red de Investigación en Servicios Sanitarios en Enfermedades Crónicas (REDISSEC), Barcelona, Spain

[4]Health Innovation Technology Transfer, Barcelona, Catalonia, Spain

[5]Prevention, Treatment and Care Unit, Department of Control of Neglected Tropical Diseases, World Health Organization, Geneve, Switzerland

Corresponding Author:
Carme Carrion, MSc, PhD
eHealth Lab Research Group
School of Health Sciences
Universitat Oberta de Catalunya
Rambla Poble Nou 156
Barcelona, 08035
Spain
Phone: 34 1 416 340 3200
Email: mcarrionr@uoc.edu

Abstract

Background: Neglected tropical diseases (NTDs) represent a diverse group of 20 communicable diseases that occur in tropical and subtropical areas in 149 countries, affecting over 1 billion people and costing developing economies billions of dollars every year. Within these diseases, those that present lesions on the skin surface are classified as skin NTDs (sNTDs). Mobile health interventions are currently being used worldwide to manage skin diseases and can be a good strategy in the epidemiological and clinical management of sNTDs.

Objective: We aimed to analyze existing evidence about mobile health interventions to control and manage sNTDs in low- and middle-income countries (LMICs) and make recommendations for what should be considered in future interventions.

Methods: A systematic review was conducted of the MEDLINE, Embase, and Scopus databases over 10 years up to April 30, 2020. All types of clinical studies were considered. Data were synthesized into evidence tables. Apps were selected through a comprehensive systematic search in the Google Play Store and Apple App Store conducted between March 20 and April 15, 2020.

Results: From 133 potentially relevant publications, 13 studies met our criteria (9.8%). These analyzed eight different interventions (three SMS text messaging interventions and five app interventions). Six of the 13 (46%) studies were community-based cross-sectional studies intended to epidemiologically map a specific disease, mainly lymphatic filariasis, but also cutaneous leishmaniasis, leprosy, and NTDs, as well as sNTDs in general. Most of the studies were considered to have a high (5/13, 39%) or moderate (4/13, 31%) risk of bias. Fifteen apps were identified in the Google Play Store, of which three were also in the Apple App Store. Most of the apps (11/15, 73%) were targeted at health care professionals, with only four targeted at patients. The apps focused on scabies (3/15, 20%), lymphatic filariasis (3/15, 20%), cutaneous leishmaniasis (1/15, 7%), leprosy (1/15, 7%), yaws and Buruli ulcer (1/15, 7%), tropical diseases including more than one sNTDs (3/15, 20%), and NTDs including sNTDs (2/15, 13%). Only 1 (7%) app focused on the clinical management of sNTDs.

Conclusions: All mobile health interventions that were identified face technological, legal, final user, and organizational issues. There was a remarkable heterogeneity among studies, and the majority had methodological limitations that leave considerable room for improvement. Based on existing evidence, eight recommendations have been made for future interventions.

(JMIR Mhealth Uhealth 2020;8(12):e22478) doi:10.2196/22478

KEYWORDS

mHealth; mobile health; neglected tropical diseases; skin neglected tropical diseases; apps; SMS text messaging; low- and middle-income countries

Introduction

According to the World Health Organization (WHO), neglected tropical diseases (NTDs) represent a diverse group of 20 communicable diseases that occur in tropical and subtropical conditions in 149 countries, affecting over 1 billion people and costing developing economies billions of dollars every year [1]. These diseases include protozoal, bacterial, helminth, and viral infections [2] that cause vast suffering, stigma, and disability, and frequently lead to death [3]. As a result, NTDs trap impoverished people in a cycle of poverty and disease. Collectively, NTDs are among the most devastating of communicable diseases in terms of not only the global health burden (26.1 million disability-adjusted life-years) [3], but also the impact on development and overall economic productivity in low- and middle-income countries (LMICs).

The WHO has further categorized NTDs that primarily present as lesions on the skin (lumps or swelling, ulcers, swollen limbs, and patches on the face or body) into the skin NTD (sNTD) group. This includes Buruli ulcers, cutaneous leishmaniasis, post-kala-azar dermal leishmaniasis, leprosy, lymphatic filariasis (lymphoedema and hydrocele), mycetoma, onchocerciasis, fungal infections, scabies, and yaws. sNTDs not only cause considerable disability and stigma, but also exacerbate poverty [4-6]. The integration of mapping, surveillance, clinical diagnosis, and management has only been achieved in a limited range of settings and disease groupings [7]. Unfortunately, there has been relatively little investment in laboratory research, epidemiology, diagnostic tools, and management strategies to control tropical skin diseases. Visual examination provides an opportunity to screen people in communities or children in schools to identify multiple conditions in a single visit. This common approach to skin diseases justifies the integrated delivery of health care interventions to both increase cost-effectiveness and expand coverage [8,9].

Various innovative methods have been used to enhance the clinical management and epidemiological surveillance of, among others, skin and infectious diseases worldwide. These include, but are not limited to, technological methods such as telemedicine, artificial intelligence, and mobile health (mHealth) [10]. mHealth has been proven to be a promising tool to improve the diagnosis and treatment of several diseases such as skin cancer [11]. Telemedicine interventions have been implemented to perform teleconsultation and to improve the diagnosis and treatment of several diseases in countries with poor resources, and they have been proven to be effective in achieving its goal [12,13]. Teledermatological interventions, consisting of real-time videoconferencing or asynchronous transition of images between different dermatologists or between patients and clinicians, have been used for some years. Numerous studies have been published describing the potential of teledermatology in low-income settings, such as the review performed by Médecins Sans Frontières to assess the tele-expertise system it had implemented to improve access to specialized clinical support for its field health workers [14]. These interventions are not yet scaled up, and although trials seem to confirm they are effective, they show some limitations that should be considered, such as the need for both a dermatologist available on demand and a stable and strong internet connection [15]. Artificial intelligence also holds promise in the diagnosis of skin conditions to a good degree of accuracy [16] and selecting the best treatment for a specific infectious disease [17]. However, these techniques are still under development and not yet available for implementation in the clinical management of sNTDs [18].

The WHO Global Observatory for eHealth defines mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [19]. The prevention and management of chronic diseases through mHealth strategies has increased over recent years, mainly in high-income countries. However, there is still a lack of evidence for its efficacy, effectiveness, and safety [20]. Attempts have been made to improve the effective surveillance and control of infectious disease outbreaks [21], but very few interventions have been implemented in LMICs where sNTDs are endemic.

Given the high number of smartphone users worldwide (around 3 billion) [22] and the high penetration of smartphones in groups with low socioeconomic status, health-related mobile apps provide an opportunity to overcome traditional barriers to the control and clinical management of sNTDs in LMICs [23,24]. Nevertheless, the vast increase in low-cost health-related apps that are not regulated by health care policymakers raises important areas of concern, including quality, usability, and the need to educate consumers regarding the potentially beneficial (or harmful) content of apps [25]. Currently, of the over 325,000 health-related apps available [26], over 500 are skin related, with 90 providing self-surveillance and diagnosis [27]. Evaluation of six apps for skin cancer diagnosis has been conducted in high-income countries with discouraging results [28], and content analysis of 123 apps demonstrated that all were in need of improvement [29].

Therefore, it is clear that research on implementing mHealth strategies is required to improve the epidemiological surveillance and clinical management of sNTDs in LMICs. The main objective of this article is to analyze existing evidence about mHealth interventions to control and manage sNTDs in LMICs and make recommendations of what should be considered in future interventions.

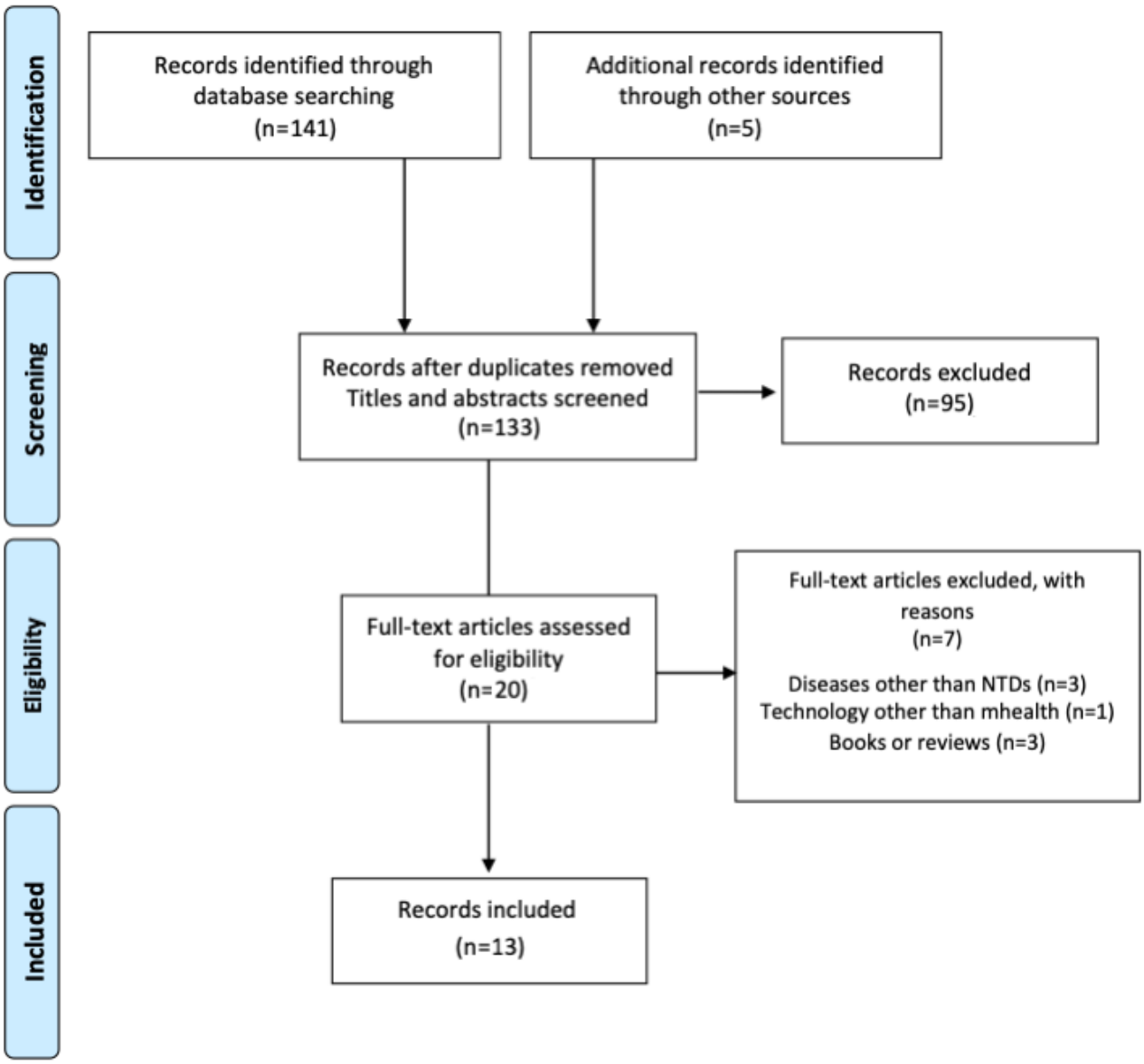
Methods

Systematic Review

Information Sources

Searches were conducted in the following databases: MEDLINE, Embase, and Scopus. This was complemented with a manual search of references in key journal archives. All published articles in the 10 years up to April 30, 2020, were considered with no restrictions on language. The reference lists of all selected studies were cross-checked for additional reports. A flow diagram of papers selected has been reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [30] (Figure 1).

Figure 1. PRISMA flow diagram of the selection of papers for inclusion in the review. NTD: neglected tropical disease.



Search Strategy

The search strategy included both controlled vocabulary and free-text terms. The terms were smartphone, neglected tropical diseases, Buruli ulcer, fungal disease, leishmaniasis, leprosy, lymphatic filariasis, lymphoedema, mycetoma, onchocerciasis, scabies, yaws, mHealth, SMS, and apps (Multimedia Appendix 1).

Eligibility Criteria

The criteria for inclusion were reporting on mHealth diagnoses, treatment, epidemiological surveillance, and prevention of sNTDs. All trial designs, regardless of risk of bias, were considered eligible. Animal studies were excluded. Letters, editorials, protocols, and reviews were not included.

Study Selection and Data Collection Process

All identified references were imported into Mendeley (v1.18), and duplicate references were eliminated. Articles that met the inclusion criteria were full-text reviewed by two independent reviewers. Any disagreements were resolved by a third reviewer. Study features and outcomes were entered into a database specifically designed for this review. Risk of bias was assessed

according to Scottish Intercollegiate Guidelines Network (SIGN) codes for study assessment [31].

App Selection

Search Strategy

Apps were identified through a comprehensive systematic search of the Google Play Store and Apple App Store between March 20 and April 15, 2020, using 14 search terms in English and Spanish across all store categories (Multimedia Appendix 1).

Screening Procedure

Each term was searched separately in both app stores, with two researchers performing identical independent searches. Information in the app store description for each result was reviewed. The inclusion criteria were as follows: (1) the purpose should be diagnosis, treatment, prevention, patient empowerment, health literacy, or epidemiological surveillance; (2) the main disease should be one of the 10 sNTDs; and (3) the app should be available in English or Spanish. Apps were excluded if they did not have a clear focus on the epidemiological or clinical management of sNTDs and focused on other issues, such as conferences, atlases of skin diseases, and a general approach to tropical diseases. Apps that met the eligibility criteria were selected and their URLs were saved. After the independent search, the two researchers compared the lists of apps they identified as eligible and performed analyses. Light and early versions of the apps were excluded in favor of full late versions.

Data Collection

Information and features were extracted from app store descriptions and available screenshots. In cases where not enough information was provided, the apps were downloaded, installed, and tested in smartphones running Android or iOS operating system.

Results

Selection of Studies

A total of 133 potentially relevant publications were identified as eligible. After screening the titles and abstracts, 20 out of 133 (15.0%) were accepted for full-text review. Out of these 20, 7 (35%) were excluded for not meeting the inclusion criteria. They focused on diseases other than sNTDs (n=3), used interventions without any mHealth element (n=1), or were reviews or book chapters (n=3). After peer-review, 13 articles were included in this nonquantitative review (Figure 1).

The main characteristics of the 13 studies included are detailed in Table 1. Studies appear in chronological order and then alphabetical order of the first author. The studies were highly heterogeneous in their design, resulting from their heterogeneous objectives. Six (46%) studies [32-37] were reported to be community-based cross-sectional studies with the main objective of epidemiologically mapping a specific disease based on its clinical symptoms. Four out of 13 (31%) studies had different designs as follows: a mixed-methods approach to define challenges to be considered [38]; different implementation scenarios for a specific mHealth intervention [39]; the sustainability of an intervention [40]; and the process of developing a specific mHealth tool or assessing its usability [10,38]. One trial was a cross-sectional study analyzing the accuracy of a new diagnosis strategy [41], and another was a prospective cohort study testing the efficacy of an mHealth strategy in training health care workers [42]. There was only one controlled nonrandomized study to test the effectiveness of an intervention [43]. Most of the studies had a high (5/13, 39%) or moderate (4/13, 31%) risk of bias according to SIGN quality criteria. Only three of the included studies were considered to have a low risk of bias [33,35,37].

Most of the studies (n=8) were focused on lymphatic filariasis alone or together with podoconiosis or onchocerciasis. These define strategies to estimate the prevalence of the diseases or the most common clinical manifestations (hydrocele or lymphoedema). There was one study about NTDs in general [40] and one about sNTDs [10]. Only two studies focused on sNTDs other than lymphatic filariasis. Of these, one was on cutaneous leishmaniasis [38] and one was on leprosy [43]. All studies were conducted in LMICs where sNTDs are endemic, mainly Africa (n=7). Two study areas were in the American continent, and all others were in Africa and Asia. The studies took place in the following 13 different countries: Tanzania (n=4), Ethiopia (n=3), Bangladesh (n=2), Malawi (n=2), Brazil (n=1), Bolivia (n=1), Colombia (n=1), Ghana (n=1), Indonesia (n=1), Nepal (n=1), Nigeria (n=1), Mozambique (n=1), and Peru (n=1). Seven of the studies analyzed interventions in both rural and urban areas, four in rural areas only, and two in urban areas.

Table 1. Characteristics of the selected studies.

Reference[a]	Study design	Study population	NTD[b]	Tool	Objective	Intervention	Outcomes	Risk of bias
Sime 2014 [32]	Community-based cross-sectional study	Ethiopia (659 districts) n=129,959 Urban and rural areas	LF[c] and PDC[d]	LINKS system	To perform integrated mapping of the two diseases	Immunochromatographic card tests of blood samples to detect circulating *Wuchereria bancrofti*. Data collection through questionnaires on a smartphone (GPS function included), and data sent to a cloud server for analysis.	Real-time data collection. Prevalence of lymphoedema (n=8,110) and *Wuchereria bancrofti* (n=139). Budget halved with smartphone-based data collection.	Moderate Selection bias
Luz 2015 [40]	Mixed-methods qualitative study	Southwestern Amazon region (Brazil, Perú, and Bolivia) n=47 (15 medical doctors, 15 researchers, 13 nurses, and four staff) Rural area	NTD	Nu-case	To explore the sustainability of the intervention	Design of Nu-case prototype. Surveys, questionnaires, sketching, and storyboarding to define actors, tasks, needs, and possible design.	High perceived potential of Nu-case in diagnosis, disease monitoring and surveillance, medical records, case notification, and medical research. Low perceived potential in health care management, assisting diagnosis for HCPs[e] other than medical doctors, and HCP training tools.	High Clustered data analysis with a very small sample size.
Stanton 2015 [33]	Community-based cross-sectional survey	Malawi (Chikwawa district) n=107,331 Ghana (Ahanta West district) n=45,402 Rural area	LF	Measure SMS	To map clinical manifestation of LF (LE[f] and HYC[g]) To pilot the Measure SMS tool	HSAs[h] in Malawi and CHWs[i] in Ghana reported individual LE and HYC case data.	Ghana: prevalence of 17.7 (LE) and 33.0 (HYC) per 10,000. Malawi: prevalence of 76.9 (LE) and 70.5 (HYC) per 10,000. 17% of SMS messages in Malawi and 16% in Ghana contained an error (41% were easy to solve).	Low
Mableson 2017 [39]	Implementation study	Ethiopia, Malawi, Tanzania, Nepal, and Bangladesh n=22 million people Urban and rural areas	LF	Measure SMS-morbidity	To define implementation scenarios for LF mapping	Health workers collect data in a traditional paper-based survey and then send an SMS (two tier) or directly send an SMS (one tier) to a central smartphone connected to a cloud server.	Four potential implementation reporting scenarios: (1) urban, high-endemic setting, two tier (2) rural, high-endemic setting, one tier (3) rural, high-endemic setting, two tier (4) urban and rural, low-endemic setting, one tier	High No methodological details
Mwingira 2017 [34]	Community-based cross-sectional survey	Tanzania (Mtwara Municipal Council) n=108,299 Mainly urban areas	LF	SMS through the GeoPoll platform	To estimate prevalences of HYC and LE	Interactive survey via SMS to randomized selected users.	Response ratio of 15.2%; n=492 (78% male); mean age 20.1 (SD 6.5) years LE signs = 22.2% (95% CI 17.4-24.8) HYC signs = 20.6% (95% CI 16.6-25.0)	High Self-reported data Incentives given to participants Gender and age biased

Reference[a]	Study design	Study population	NTD[b]	Tool	Objective	Intervention	Outcomes	Risk of bias
Mwingira 2017 [35]	Community-based cross-sectional survey	Tanzania (Dar es Salaam) n=5 million Urban area	LF	Measure SMS-morbidity tool	To locate LF patients and estimate prevalence	SMS survey in a three-phase intervention: paper-based data collection checked by a supervisor and sent by SMS to a local server. Data checked and sent to a cloud-based server.	Prevalence of 133.6 per 10,000; n=6.889. Less than 20% of SMS messages had formatting errors. More than 95% of SMS messages were sent to the cloud server.	Low Some implementation challenges
Pedram 2017 [41]	Cross-sectional study	Control (n=77) and infected (n=313) Various countries in Africa, Asia, and America Rural and urban areas	LF and ONC[j]	Smartphone-based microscope	To assess the accuracy of a new diagnostic test	All samples tested with the new test including a portable smartphone-based microscope.	Sensitivity of 71% and specificity around 100% for all tested species.	Moderate Information bias
Karim 2018 [36]	Community-based cross-sectional survey	Bangladesh (high and low endemic districts) n=65 million people at risk Rural and urban areas	LF	Measure SMS	To determine HYC and LE prevalence and develop clinical risk maps for targeted interventions	Paper-based census, active case findings, and summary data per patient sent through the system to a central database. In low endemic districts, case findings via health facility data and confirmed by mHealth[k] trained professionals sending SMS.	Prevalence of 125.4 per 100,000 in high-endemic districts and 2.4 per 100,000 in low-endemic districts.	Moderate Different methods used mHealth strategy not fully described and validated
Martindale 2018 [37]	Community-based cross-sectional study	Ethiopia (two districts: Hawella Tula and Besa) n=460,722 Rural area	LF and PDC	Measure SMS-morbidity tool	To pilot feasibility and utility for an integrated mapping of the two diseases	Survey conducted by health extension workers (n=59). Clinical cases reported by paper-based standard forms and by SMS.	Paper-based cases reported (n=2.377) SMS method cases reported (n=2.372) *P* value of .94 Prevalence of 64 per 10,000. Cost saving (13.7%) with the SMS method.	Low Potential information bias in the real condition
Mieras 2018 [10]	Development and pilot study	HCP in Nigeria and in Mozambique Rural and urban areas	sNTDs	SkinApp	To support HCP in the early diagnosis of sNTDs in low resource settings	Development of SkinApp with 29 diseases included (six sNTDs) and assessment of user friendliness (semistructured interviews and focus groups). Implementation study	SkinApp is considered a good decision support system.	High No methodological details
Navarro 2018 [38]	Consensus report	Colombia (Tumaco) Rural area	CL[l]	Guaral App	To define challenges and requisites of mHealth tools	Development of Guaral App to be adopted by volunteer community workers for diagnosis and mapping of patients with CL.	Key aspects are: -Sociotechnical context -Systems analysis -Human-centered design	Not applicable
Akoko 2019 [42]	Prospective cohort study	Tanzania (Mtwara and Lindi regions) Rural and urban areas	HYC	WhatsApp platform	To test the efficacy of a mobile platform as an adjunct in supervision and support for non-surgical clinicians when practicing hydrocelectomy	Didactic and practical training conducted by two experts. Photographs shared though the WhatsApp platform for group discussion, final approval of surgery, on table and postoperative complications.	Fifteen NPCs[m] trained and able to perform 1337 hydrocelectomies in 1250 patients. Mean procedure duration of 50.2 min (SD 0.24) Complication rate <2.16%	Moderate Small sample of NPCs Selection bias

Reference[a]	Study design	Study population	NTD[b]	Tool	Objective	Intervention	Outcomes	Risk of bias
Rachmani 2019 [43]	Nonrandomized controlled longitudinal observational study	Indonesia (Pekalongan District Java) n=124 control, n=64 intervention group Rural and urban areas	Leprosy	e-leprosy framework	To evaluate the effectiveness of an e-leprosy framework in PHC[n].	Implementation of an e-leprosy framework for primary health care for 19 months.	Abandon rate of 21%. Increase of 21% in on-time completion and 14.6% in attendance rates.	High Nonrandomization Sample selection bias

[a]Studies are in chronological order and then alphabetical order of first author.

[b]NTD: neglected tropical disease.

[c]LF: lymphatic filariasis.

[d]PDC: podoconiosis.

[e]HCP: health care professional.

[f]LE: lymphoedema.

[g]HYC: hydrocele.

[h]HSA: salaried health surveillance assistant.

[i]CHW: community health worker.

[j]ONC: onchocerciasis.

[k]mHealth: mobile health.

[l]CL: cutaneous leishmaniasis.

[m]NPC: nonphysician clinician.

[n]PHC: primary healthcare.

Elements Included in the mHealth Interventions

Only eight different mHealth interventions were identified (Table 2). Three were based on sending SMS text messages between health workers and a central web system [33-35,37,39,41,43]. Text messaging is an easy technology to use and, as these studies show, is useful for motivating patients to complete treatments. Health care services can be equipped with systems providing automated text message reminders for disease control [43]. The Measure SMS and Measure SMS-morbidity strategies have been tested and implemented in several different contexts. The ability to view and assess the quality of patient identification data in real time is a great asset. Enabling information on morbidity burden to be known almost instantaneously, as opposed to having to wait for the collation and digitization of paper forms, also facilitates more efficient provision of necessary care [35].

Table 2. Elements and challenges described in the mobile health interventions of the selected studies.

Tool Name	Elements included				Challenges identified	References
	SMS	Web	App	Cloud server		
e-Leprosy framework	+	+	–	–	• Cost-benefit analysis • Need for a comprehensive strategy • Perceived usefulness of the technology • Perceived ease of use of the technology • No internet connection in remote rural areas	Rachmani 2019 [43]
GeoPoll and SMS	+	+	–	–	No challenges identified in the study	Mwingira 2017 [34]
Guaral App	–	–	+	+	• Sociotechnical context • Human-centered design • Iterative design • Usability • Technical concerns	Navarro 2018 [38]
LINKS system	–	–	+	+	• Data ownership • Lack of technical expertise • Cost of smartphones • HCP[a] perceptions about digital data collection • Batteries running out of charge • Lack of network	Sime 2014 [32]
Measure SMS and Measure SMS-morbidity	+	+	–	+	• Data accuracy • HCP training • Technological barriers • Availability of mobile phones • Cost • Poor network coverage • Battery supply • Feasibility • Health system structure and organizational change	Stanton 2015 [33] Karim 2019 [36] Mableson 2017 [39] Mwingira 2017 [35] Martindale 2018 [37]
Nu-case	–	–	+	+	• Heterogeneity of linguistic cultural backgrounds • Potential legal issues • Network signal coverage • Durability and battery life • Image and data quality • Empowerment of local health workers	Luz 2015 [40]
SkinApp	–	–	+	+	• Technical requirements • Cost • Need of context-specific adaptation (language, culture, and epidemiological situation)	Mieras 2018 [10]
WhatsApp	–	–	+	–	• Data privacy • Highly time consuming • Work required on mentor and mentee relationship • Internet connection requirement	Akoko 2019 [42]

[a]HCP: health care professional.

Five interventions were conducted through apps and cloud servers. One utilized the commercially available WhatsApp app [42], and the others used either very simple ad-hoc developed apps able to collect data and send to a server [32] or apps that included a wider range of elements and took a more comprehensive approach, for example, Guaral App [38], Nu-case [40], and SkinApp [10]. Only Guaral App and SkinApp are available from the app stores (Table 3).

Table 3. Characteristics of apps for skin neglected tropical diseases available from app stores in alphabetical order.

Name (date last updated)	Marketplace (number of downloads)	Objective	Target user	Disease
EndNTDs App (October 2017)	Google Play Store (+100)	To create a community of champions who are at the forefront in the fight against NTDs[a] in Zimbabwe, Africa, and globally	HCP[b]	NTDs (including skin NTDs)
Guaral RPC (February 2020)	Google Play Store (+10)	To improve early diagnosis	HCP	CL[c]
Lepra Reaction Basic management guide (October 2017)	Google Play Store (+100)	To conduct classification and quantification of severity to plan for appropriate clinical management	HCP	Leprosy
LymEX (July 2019 Google Play Store and January 2020 Apple App Store)	Google Play Store (+50); Apple App Store	To improve self-care	Patients	Lymphoedema
LymVol (September 2019)	Google Play Store (+10)	To measure and calculate limb volume of those affected by edema	HCP	Lymphoedema
Recognize Hydrocele disease (September 2019)	Google Play Store (+500)	To improve information about causes, treatment, and complications	Patients	Hydrocele
Recognize scabies (September 2019)	Google Play Store (+100)	To improve knowledge about the disease	Patients	Scabies
Scabies disease (December 2017)	Google Play Store (+500)	To improve clinical management of the disease	HCP	Scabies
Scabies Disease: Treatment (October 2019)	Google Play Store (+100)	To manage symptoms and treatment	Patients	Scabies
SkinApp (October 2019)	Google Play Store (+10); Apple App Store	To act as a diagnostic tool and source of information on signs, symptoms, and therapy	HCP	NTD- and HIV-related skin diseases
Skin NTDs App (July 2020)	Google Play Store (+60); Apple App Store	To diagnose and identify signs and symptoms of sNTDs through their visible characteristics	HCP	Skin NTDs
Task Force Tropical Data (August 2016)	Google Play Store (+1000)	To collect data	Task Force for Global Health members	NTDs (including skin NTDs)
Tropical Diseases (August 2019)	Google Play Store (+1000)	To provide a detailed overview of the etiology, pathophysiology, epidemiology, diagnosis, and treatment of tropical diseases	HCP	TDs[d] including CL, BU[e], and leprosy
Tropical Diseases (December 2019)	Google Play Store (+100)	To provide a detailed overview of the cause, diagnosis, prognosis, risk factors, prevention, and treatment of the stated diseases	HCP	TDs including CL, BU, and leprosy
WIDP (January 2020)	Google Play Store (+10)	To control the epidemiological situation	HCP and WHO[f] staff	Yaws and BU

[a]NTD: neglected tropical disease.

[b]HCP: health care professional.

[c]CL: cutaneous leishmaniasis.

[d]TD: tropical disease.

[e]BU: Buruli ulcer.

[f]WHO: World Health Organization.

Results showed that mHealth strategies do achieve their primary goals, but many challenges require consideration. Both SMS and app interventions face (1) technological, (2) legal, (3) final user, and (4) organizational issues (Table 2) as follows:

- The main technological issues are poor network coverage or no internet connection in specific areas, batteries running out of charge, poor data and image quality, and lack of technical requirements.
- Legal issues, including data privacy and ownership, are important concerns that must be resolved.
- Final users can be health care professionals, patients, and citizens, who all require adequate digital literacy and to feel more empowered.
- Final users need to perceive the technology as possessing a high level of usefulness and ease of use.
- The organizational challenges most often identified relate to resources (cost, time, and availability of smartphones), implementation planning, cultural and language barriers, and the need for user features to be fully completed when implementing comprehensive mHealth strategies.

mHealth interventions using apps also involve extra challenges in the development period, such as consideration of the sociotechnical context, and the need for human-centered and iterative design to improve usability, feasibility, and user experience.

Selection of Apps

Fifteen apps were identified in the Google Play Store, with three of these also available in the Apple App Store. Most of the apps (11/15, 73%) were targeted at health care professionals, with only four targeted at patients. The number of app downloads can be taken as a proxy indicator of use. However, while Google Play Store comprehensively provides this information, the Apple App Store provides it to the app developer only. Google Play Store data revealed that only two of the apps had more than 1000 downloads. These were Task Force Tropical Data, which is mainly used to collect data, and Tropical Diseases, which focuses on various tropical diseases but only provides information on three sNTDs (Buruli ulcer, cutaneous leishmaniasis, and leprosy). Most of the apps have recently been updated or uploaded, with only four not updated over the last 12 months. The apps addressed scabies (3/15, 20%), lymphatic filariasis (3/15, 20%), cutaneous leishmaniasis (1/15, 7%), leprosy (1/15, 7%), yaws and Buruli ulcer (1/15, 7%), tropical diseases including more than one sNTDs (3/15, 20%), and NTDs including sNTDs (2/15, 13%). Only 1 (7%) app was related to the clinical management of sNTDs (Table 3).

Discussion

Principal Findings

We reviewed the literature on the evidence for mobile apps and have providing a set of recommendations for the future development and implementation of such tools. Our review pinpointed 13 studies dealing with mobile health interventions for managing sNTDs in LMICs. A descriptive evidence synthesis showed that most of the studies had a high or moderate risk of bias according to SIGN quality criteria, and only three were considered to have a low risk of bias. Only two studies focused on sNTDs other than lymphatic filariasis, one focused on cutaneous leishmaniasis, and one focused on leprosy.

In addition, eight different mHealth interventions were identified, of which three were based on texting by SMS and five were conducted through apps and cloud servers. Moreover, we acknowledged 15 apps available in app stores. Most of the apps (11/15, 73%) were targeted at health care professionals, with only four targeted at patients. Results showed that mHealth strategies do achieve their primary goals, but many challenges require consideration. All mHealth interventions face technological, legal, final user, and organizational issues.

The disease most targeted in mHealth interventions for sNTDs is lymphatic filariasis. The Global Program to Eliminate Lymphatic Filariasis as a public health problem was launched in 2000, and the global elimination goal of 2020 established in 2012 was probably a key factor in driving the Ministry of Health, the WHO, and implementing partners to scale up surveillance and morbidity management activities for this disease. There is an urgent need for a rapid and adaptable tool to gather patient estimates in order for national programs to appropriately forecast, plan, and deliver a basic package of care to those suffering from the disabling and debilitating clinical manifestations of lymphatic filariasis in an affordable manner [39]. Three of the apps on the market address scabies, but as no trials have been identified, there is a lack of existing evidence about mHealth interventions to manage this disease. Other sNTDs, such as cutaneous leishmaniasis and leprosy, have been managed with pilot interventions through mHealth strategies. It is interesting to note that a number of trials and apps are focused on an integrated approach to the various sNTDs. The most recent app developed by WHO aims to provide support for diagnosis by using an automated algorithm to provide potential conditions from the selection of sNTD signs and symptoms.

The provision of health care services to populations in remote regions, often dispersed over large geographical areas, has long been considered a challenging issue from the perspective of health systems management [44]. Telemedicine approaches have been employed in some places for remote assistance for diagnosis and treatment, with a degree of success [45], but neglected diseases, and therefore sNTDs, require more than the usual functionality of telemedicine systems [40]. mHealth interventions in high-income countries are focused on apps, but as sNTDs mainly affect populations with low resources in LMICs, most of the interventions tested so far were based on SMS text messaging, and only a few app strategies have been attempted. Using text messages to change patient behavior and achieve targeted health outcomes also requires a comprehensive strategy but demands fewer resources. Studies in Indonesia have addressed the apparent difficulties in leprosy control programs and how a health information system could assist, and concluded that SMS text messaging and web-based applications could be good strategies for implementing continuous monitoring and recording of patients [43]. The advantage of SMS is “internet-less” connectivity, and it allows fast transmission of surveillance data, whereas store-and-forward teledermatology represents an ideal approach to assess skin lesions and provide long-distance support to individual diagnosis because images are shared and can be assessed. The cost-effectiveness of store-and-forward teledermatology increases when patients are

required to travel farther distances to access dermatology services. SMS and store-and-forward teledermatology are complementary and contribute to improve the epidemiological and clinical management of sNTDs.

By bringing essential central health services closer to peripheral areas, innovative technological methods like telemedicine, mHealth apps, and SMS text messages can help to bridge the gap between the burden of skin diseases and the lack of capable staff in resource-poor settings [10]. mHealth interventions enable rapid data collection, easy monitoring and supervision of data reporting, better management of sNTDs, and efficient provision of necessary care. They also appear to be lower cost strategies than paper-based forms that require more human and financial resources. The four studies that conducted cost analyses [33,35-37] did not have sufficient robust data, but all concluded that mHealth interventions are less costly than paper-based epidemiological mapping. By managing and taking full ownership of data and the implementation process, local teams will be empowered, and the establishment of valuable data management and health surveillance capacities within the local team and at the country level will empower national health systems [35].

Recommendations Based on the Existing Evidence

Based on evidence presented in the studies identified in this systematic review, the following eight recommendations will enable sNTD mHealth-based interventions to move forward from innovation to implementation:

1. No one should be left behind. Patients from all regions must be targeted to benefit from the proposed interventions. This requires translation of the tools into several languages. At least into Portuguese and Spanish in the Americas, and English, French, and Portuguese in Africa.
2. Users must be empowered. Final users of the interventions (health care professionals and/or patients) should receive sufficient training to improve their digital literacy and make appropriate use of the tools.
3. Complexity must be addressed. The adoption of digital health technology is a complicated process that needs to be thoroughly considered before and during implementation.
4. Utility and simplicity must be perceived. Health care professionals, patients, and healthy citizens should grasp the utility and user friendliness of the proposed technology, in order for these two elements to become enablers and not barriers.
5. Technological requirements must be considered from the very outset. The availability of smartphones and potential problems with electricity or internet networks must be addressed as part of a comprehensive strategy with a specific aim.
6. A long-term mHealth platform must be put in place. The success of an mHealth intervention depends on the existence of an mHealth platform to not only facilitate adoption of the tool, but also guarantee sustained effective use.
7. Two-tiered processes are required for improvement. In the first stages of implementing an mHealth intervention, it is important to have two-tiered processes to refine and optimize the process in an iterative way.
8. The tool must respond to needs. Interventions are embedded in a specific health service; therefore, additional tools must be considered according to the need.

Limitations

One of the main limitations of this review is publication bias. References from other sources, such as conferences and meetings, have not been included. Although the number of scientific journals that publish mHealth-related articles has increased in recent years, there is a lot of grey literature surrounding this field that we may have missed. Moreover, the heterogeneity of interventions and populations has made it difficult to synthesize results, and consequently, findings need to be considered with caution. Most studies in this review were conducted in East Africa, and the total number of countries represents 9% of those with NTDs worldwide. This limits the extrapolation of the results.

Conclusions

The potential for mHealth interventions to improve the epidemiological and clinical management of sNTDs is yet to be reached. There are very few good quality studies and tools, and most of these have methodological limitations, leaving considerable room for improvement. The majority of sNTDs are not addressed by any mHealth tool. This research has enabled us to identify eight recommendations for the future development and implementation of mHealth interventions for managing sNTDs in LMICs.

Acknowledgments

All authors contributed equally. The authors thank GOC Networking and Marta Pulido, MD, PhD, for editing the manuscript and editorial assistance.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 55 KB - mhealth_v8i12e22478_app1.docx]

References

1. Control of Neglected Tropical Diseases. World Health Organization. URL: http://www.who.int/neglected_diseases/diseases/en/# [accessed 2020-12-14]
2. Working to overcome the global impact of neglected tropical diseases: First WHO report on Neglected Tropical Diseases. World Health Organization. URL: https://www.who.int/neglected_diseases/resources/9789241564090/en/ [accessed 2020-12-14]
3. Verrest L, Dorlo TPC. Lack of Clinical Pharmacokinetic Studies to Optimize the Treatment of Neglected Tropical Diseases: A Systematic Review. Clin Pharmacokinet 2017 Jun 15;56(6):583-606 [FREE Full text] [doi: 10.1007/s40262-016-0467-3] [Medline: 27744580]
4. Gómez LJ, van Wijk R, van Selm L, Rivera A, Barbosa M, Parisi S, et al. Stigma, participation restriction and mental distress in patients affected by leprosy, cutaneous leishmaniasis and Chagas disease: a pilot study in two co-endemic regions of eastern Colombia. Trans R Soc Trop Med Hyg 2020 Jul 01;114(7):476-482 [FREE Full text] [doi: 10.1093/trstmh/trz132] [Medline: 32052043]
5. Pires M, Wright B, Kaye PM, da Conceição V, Churchill RC. The impact of leishmaniasis on mental health and psychosocial well-being: A systematic review. PLoS One 2019 Oct 17;14(10):e0223313 [FREE Full text] [doi: 10.1371/journal.pone.0223313] [Medline: 31622369]
6. van 't Noordende AT, Aycheh MW, Schippers A. The impact of leprosy, podoconiosis and lymphatic filariasis on family quality of life: A qualitative study in Northwest Ethiopia. PLoS Negl Trop Dis 2020 Mar;14(3):e0008173 [FREE Full text] [doi: 10.1371/journal.pntd.0008173] [Medline: 32155158]
7. Engelman D, Fuller LC, Solomon AW, McCarthy JS, Hay RJ, Lammie PJ, et al. Opportunities for Integrated Control of Neglected Tropical Diseases That Affect the Skin. Trends Parasitol 2016 Nov;32(11):843-854. [doi: 10.1016/j.pt.2016.08.005] [Medline: 27638231]
8. Mitjà O, Marks M, Bertran L, Kollie K, Argaw D, Fahal AH, et al. Integrated Control and Management of Neglected Tropical Skin Diseases. PLoS Negl Trop Dis 2017 Jan;11(1):e0005136 [FREE Full text] [doi: 10.1371/journal.pntd.0005136] [Medline: 28103250]
9. Yotsu RR, Kouadio K, Vagamon B, N'guessan K, Akpa AJ, Yao A, et al. Skin disease prevalence study in schoolchildren in rural Côte d'Ivoire: Implications for integration of neglected skin diseases (skin NTDs). PLoS Negl Trop Dis 2018 May 17;12(5):e0006489 [FREE Full text] [doi: 10.1371/journal.pntd.0006489] [Medline: 29771976]
10. Mieras L, Taal A, Post E, Ndeve A, van Hees C. The Development of a Mobile Application to Support Peripheral Health Workers to Diagnose and Treat People with Skin Diseases in Resource-Poor Settings. Trop Med Infect Dis 2018 Sep 15;3(3):102 [FREE Full text] [doi: 10.3390/tropicalmed3030102] [Medline: 30274498]
11. Rat C, Hild S, Rault Sérandour J, Gaultier A, Quereux G, Dreno B, et al. Use of Smartphones for Early Detection of Melanoma: Systematic Review. J Med Internet Res 2018 Apr 13;20(4):e135 [FREE Full text] [doi: 10.2196/jmir.9392] [Medline: 29653918]
12. Patterson V. Telemedicine for epilepsy support in resource-poor settings. Front Public Health 2014 Aug 21;2:120 [FREE Full text] [doi: 10.3389/fpubh.2014.00120] [Medline: 25191650]
13. Wootton R, Geissbuhler A, Jethwani K, Kovarik C, Person D, Vladzymyrskyy A, et al. Long-running telemedicine networks delivering humanitarian services: experience, performance and scientific output. Bull World Health Org 2012 May 01;90(5):341-347. [doi: 10.2471/blt.11.099143]
14. Delaigue S, Morand J, Olson D, Wootton R, Bonnardot L. Teledermatology in Low-Resource Settings: The MSF Experience with a Multilingual Tele-Expertise Platform. Front Public Health 2014;2:233 [FREE Full text] [doi: 10.3389/fpubh.2014.00233] [Medline: 25453029]
15. Surovi N, Kiber A, Kashem A, Babi KN. Study and development of algorithm of different skin diseases analysis using image processing method. Asian J Biomed Pharmaceut Sci 2017;7(60) [FREE Full text]
16. Yadav N, Kumar V, Shrivastava U. Skin Diseases Detection Models using Image Processing: A Survey. IJCA 2016 Mar 17;137(12):34-39. [doi: 10.5120/ijca2016909001]
17. Garcia-Vidal C, Sanjuan G, Puerta-Alcalde P, Moreno-García E, Soriano A. Artificial intelligence to support clinical decision-making processes. EBioMedicine 2019 Aug;46:27-29 [FREE Full text] [doi: 10.1016/j.ebiom.2019.07.019] [Medline: 31303500]
18. Peiffer-Smadja N, Dellière S, Rodriguez C, Birgand G, Lescure F, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clin Microbiol Infect 2020 Oct 12;26(10):1300-1309. [doi: 10.1016/j.cmi.2020.02.006] [Medline: 32061795]
19. 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]
20. Puigdomenech Puig E, Robles N, Saigí-Rubió F, Zamora A, Moharra M, Paluzie G, et al. Assessment of the Efficacy, Safety, and Effectiveness of Weight Control and Obesity Management Mobile Health Interventions: Systematic Review. JMIR Mhealth Uhealth 2019 Oct 25;7(10):e12612 [FREE Full text] [doi: 10.2196/12612] [Medline: 31654566]

21. Wood CS, Thomas MR, Budd J, Mashamba-Thompson TP, Herbst K, Pillay D, et al. Taking connected mobile-health diagnostics of infectious diseases to the field. Nature 2019 Feb;566(7745):467-474 [FREE Full text] [doi: 10.1038/s41586-019-0956-2] [Medline: 30814711]
22. 39+ smartphone statistics you should know in 2020. Review 42. 2020. URL: https://review42.com/smartphone-statistics/ [accessed 2020-12-14]
23. Rizvi SMH, Schopf T, Sangha A, Ulvin K, Gjersvik P. Teledermatology in Norway using a mobile phone app. PLoS One 2020 Apr 27;15(4):e0232131 [FREE Full text] [doi: 10.1371/journal.pone.0232131] [Medline: 32339203]
24. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine Learning in Dermatology: Current Applications, Opportunities, and Limitations. Dermatol Ther (Heidelb) 2020 Jun;10(3):365-386 [FREE Full text] [doi: 10.1007/s13555-020-00372-0] [Medline: 32253623]
25. Kho J, Gillespie N, Horsham C, Snoswell C, Vagenas D, Soyer HP, et al. Skin Doctor Consultations Using Mobile Teledermoscopy: Exploring Virtual Care Business Models. Telemed J E Health 2020 Nov 14;26(11):1406-1413. [doi: 10.1089/tmj.2019.0228] [Medline: 32058835]
26. 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 2020-12-14]
27. Flaten HK, St Claire C, Schlager E, Dunnick CA, Dellavalle RP. Growth of mobile applications in dermatology - 2017 update. Dermatol Online J 2018 Feb 15;24(2) [FREE Full text] [Medline: 29630159]
28. Freeman K, Dinnes J, Chuchu N, Takwoingi Y, Bayliss SE, Matin RN, et al. Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies. BMJ 2020 Feb 10;368:m127 [FREE Full text] [doi: 10.1136/bmj.m127] [Medline: 32041693]
29. Charbonneau DH, Hightower S, Katz A, Zhang K, Abrams J, Senft N, et al. Smartphone apps for cancer: A content analysis of the digital health marketplace. Digit Health 2020 Feb 11;6:2055207620905413 [FREE Full text] [doi: 10.1177/2055207620905413] [Medline: 32110428]
30. 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]
31. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network (SIGN); 2008.
32. Sime H, Deribe K, Assefa A, Newport MJ, Enquselassie F, Gebretsadik A, et al. Integrated mapping of lymphatic filariasis and podoconiosis: lessons learnt from Ethiopia. Parasit Vectors 2014 Aug 27;7:397 [FREE Full text] [doi: 10.1186/1756-3305-7-397] [Medline: 25164687]
33. Stanton MC, Mkwanda SZ, Debrah AY, Batsa L, Biritwum N, Hoerauf A, et al. Developing a community-led SMS reporting tool for the rapid assessment of lymphatic filariasis morbidity burden: case studies from Malawi and Ghana. BMC Infect Dis 2015 May 16;15:214 [FREE Full text] [doi: 10.1186/s12879-015-0946-4] [Medline: 25981497]
34. Mwingira UJ, Downs P, Uisso C, Chikawe M, Sauvage-Mar M, Malecela MN, et al. Applying a mobile survey tool for assessing lymphatic filariasis morbidity in Mtwara Municipal Council of Tanzania. Mhealth 2017 Mar 15;3:8-8 [FREE Full text] [doi: 10.21037/mhealth.2017.03.01] [Medline: 28567405]
35. Mwingira U, Chikawe M, Mandara WL, Mableson HE, Uisso C, Mremi I, et al. Lymphatic filariasis patient identification in a large urban area of Tanzania: An application of a community-led mHealth system. PLoS Negl Trop Dis 2017 Jul 14;11(7):e0005748 [FREE Full text] [doi: 10.1371/journal.pntd.0005748] [Medline: 28708825]
36. Karim MJ, Haq R, Mableson HE, Sultan Mahmood ASM, Rahman M, Chowdhury SM, et al. Developing the first national database and map of lymphatic filariasis clinical cases in Bangladesh: Another step closer to the elimination goals. PLoS Negl Trop Dis 2019 Jul;13(7):e0007542 [FREE Full text] [doi: 10.1371/journal.pntd.0007542] [Medline: 31306409]
37. Martindale S, Mableson HE, Kebede B, Kiros FH, Tamiru A, Mengistu B, et al. A comparison between paper-based and m-Health tools for collating and reporting clinical cases of lymphatic filariasis and podoconiosis in Ethiopia. Mhealth 2018 Oct;4:49-49 [FREE Full text] [doi: 10.21037/mhealth.2018.09.12] [Medline: 30505847]
38. Navarro A, Rubiano L, Arango JD, Rojas CA, Alexander N, Saravia NG, et al. Developing mobile health applications for neglected tropical disease research. PLoS Negl Trop Dis 2018 Nov 1;12(11):e0006791 [FREE Full text] [doi: 10.1371/journal.pntd.0006791] [Medline: 30383809]
39. Mableson HE, Martindale S, Stanton MC, Mackenzie C, Kelly-Hope LA. Community-based field implementation scenarios of a short message service reporting tool for lymphatic filariasis case estimates in Africa and Asia. Mhealth 2017 Jul 21;3:28-28 [FREE Full text] [doi: 10.21037/mhealth.2017.06.06] [Medline: 28828375]
40. Luz S, Masoodian M, Cesario M. Disease surveillance and patient care in remote regions: an exploratory study of collaboration among health-care professionals in Amazonia. Behaviour & Information Technology 2014 Jan 13;34(6):548-565. [doi: 10.1080/0144929x.2013.853836]
41. Pedram B, Pasquetto V, Drame PM, Ji Y, Gonzalez-Moa MJ, Baldwin RK, et al. A novel rapid test for detecting antibody responses to Loa loa infections. PLoS Negl Trop Dis 2017 Jul 27;11(7):e0005741 [FREE Full text] [doi: 10.1371/journal.pntd.0005741] [Medline: 28749939]

42. Akoko L, Mwanga A, Chikawe M, Lutainulwa E, Ngoma D, Nshalla A, et al. Supervision and support in surgical practice using mobile platform: a case of mass hydrocele surgeries in remote regions. Mhealth 2019 Sep;5:41-41 [FREE Full text] [doi: 10.21037/mhealth.2019.09.05] [Medline: 31620468]
43. Rachmani E, Hsu C, Chang PWS, Jumanto J, Fuad A, Ningrum DNA, et al. Encouraging On-Time Completion of Leprosy Patients Treatment: Implementing E-Leprosy Framework to Primary Health Care in Indonesia. Asia Pac J Public Health 2019 May;31(4):296-305. [doi: 10.1177/1010539519847355] [Medline: 31104477]
44. Dussault G, Franceschini MC. Not enough there, too many here: understanding geographical imbalances in the distribution of the health workforce. Hum Resour Health 2006 May 27;4:12 [FREE Full text] [doi: 10.1186/1478-4491-4-12] [Medline: 16729892]
45. Hailey D, Roine R, Ohinmaa A. Systematic review of evidence for the benefits of telemedicine. J Telemed Telecare 2002 Dec 02;8 Suppl 1(1_suppl):1-30. [doi: 10.1258/1357633021937604] [Medline: 12020415]

Abbreviations

LMIC: low- and middle-income country
mHealth: mobile health
NTD: neglected tropical disease
SIGN: Scottish Intercollegiate Guidelines Network
sNTD: skin neglected tropical disease
WHO: World Health Organization

Edited by L Buis; submitted 13.07.20; peer-reviewed by C Johnson, M Karim; comments to author 03.08.20; revised version received 28.09.20; accepted 20.11.20; published 31.12.20.

Please cite as:
Carrion C, Robles N, Sola-Morales O, Aymerich M, Ruiz Postigo JA
Mobile Health Strategies to Tackle Skin Neglected Tropical Diseases With Recommendations From Innovative Experiences: Systematic Review
JMIR Mhealth Uhealth 2020;8(12):e22478
URL: http://mhealth.jmir.org/2020/12/e22478/
doi:10.2196/22478
PMID:33382382

Original Paper

Issues Associated With the Management and Governance of Sensor Data and Information to Assist Aging in Place: Focus Group Study With Health Care Professionals

Inga Hunter[1], MB, MA, MPhil; Phoebe Elers[2], BAppSc, GDip, MhM, PhD; Caroline Lockhart[1], RN, MBS; Hans Guesgen[3], Dipl-Inform, Dr rer nat, Dr habil; Amardeep Singh[3], BTech, MTech; Dick Whiddett[1], BSc, MA, PhD

[1]School of Management, Massey University, Palmerston North, New Zealand

[2]School of Communication, Journalism and Marketing, Massey University, Palmerston North, New Zealand

[3]School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Corresponding Author:
Dick Whiddett, BSc, MA, PhD
School of Management
Massey University
Private Bag 11 222
Palmerston North, 4442
New Zealand
Phone: 64 6 350 5701
Email: r.j.whiddett@massey.ac.nz

Abstract

Background: Smart home and telemonitoring technologies have often been suggested to assist health care workers in supporting older people to age in place. However, there is limited research examining diverse information needs of different groups of health care workers and their access to appropriate information technologies.

Objective: The aim of this study was to investigate the issues associated with using technologies that connect older people to their health care providers to support aging in place and enhance older people's health and well-being.

Methods: Seven focus group discussions were conducted comprising 44 health care professionals who provided clinic-based or in-home services to community-dwelling older people. Participants were asked about their information needs and how technology could help them support older people to age in place. The recordings of the sessions were transcribed and thematically analyzed.

Results: The perspectives varied between the respondents who worked in primary care clinics and those who worked in community-based services. Three overarching themes were identified. The first theme was "access to technology and systems," which examined the different levels of technology in use and the problems that various groups of health care professionals had in accessing information about their patients. Primary care professionals had access to good internal information systems but they experienced poor integration with other health care providers. The community-based teams had poor access to technology. The second theme was "collecting and sharing of information," which focused on how technology might be used to provide them with more information about their patients. Primary care teams were interested in telemonitoring for specific clinical indicators but they wanted the information to be preprocessed. Community-based teams were more concerned about gaining information on the patients' social environment. The third theme was that all respondents identified similar "barriers to uptake": cost and funding issues, usability of systems by older people, and information security and privacy concerns.

Conclusions: The participants perceived the potential benefits of technologies, but they were concerned that the information they received should be preprocessed and integrated with current information systems and tailored to the older people's unique and changing situations. Several management and governance issues were identified, which needed to be resolved to enable the widespread integration of these technologies into the health care system. The disconnected nature of the current information architecture means that there is no clear way for sensor data from telemonitoring and smart home devices to be integrated with other patient information. Furthermore, cost, privacy, security, and usability barriers also need to be resolved. This study highlights the importance and the complexity of management and governance of systems to collect and disseminate such information. Further research into the requirements of all stakeholder groups and how the information can be processed and disseminated is required.

(JMIR Mhealth Uhealth 2020;8(12):e24157) doi:10.2196/24157

KEYWORDS

smart home; home monitoring technology; aging in place; information governance; information management; older people; support network; aging; elderly health

Introduction

Like in other developed nations, life expectancy in New Zealand is increasing, which is resulting in an aging population [1]. It is estimated that the proportion of the New Zealand population aged 65 years and older will increase to 21%-26% in 2043 and 24%-33% in 2068 [2]. As older people have higher rates of chronic conditions and disabilities that require regular support [3], this can place an increased demand on health care services. An option favored by policymakers [4] and older people [5] is to support individuals to remain in their own homes for longer and avoid residential care, which is known as "aging in place."

Various technologies have been used to support older people as they age in place, including home-monitoring devices [6], purpose-built smart homes [7], intelligent cognitive assistants [8], and web-based health information resources [9]. However, many technologies have the limitation of only treating 1 condition in isolation, rather than the older person as a whole, who may be dealing with a range of health issues and receiving services from a range of formal health care providers.

Many older people also rely on informal support networks of friends, neighbors, and family members who provide ongoing practical and emotional support such as personal care, household work assistance, company, and emotional assistance. Such informal support networks would also benefit from access to information from these support technologies; Fischer et al [10] reports the following about informal support networks, "Tools for the elderly should consider the whole care network and take into account who will be using the tool, who has access to what information, and how these factors may change over time."

This paper presents some of the findings from an exploratory project that investigates how technologies that connect older people to their informal and formal support networks could assist aging in place and enhance older people's health and well-being. In the initial phase of the project, we explored the requirements of the many stakeholder groups involved; in the second phase, these requirements informed the design of prototype technology [11], which has been used and evaluated by older people and their informal support networks [12].

Stakeholder requirements were identified in 3 ways; in each case, issues of information governance were found to be very important. First, we solicited the perspectives of experts in telemedicine and health informatics by organizing a workshop [13]; this research highlighted the importance of data integration, security, and control. Second, we conducted in-depth interviews with older people and their self-identified informal support networks [14]; these interviews emphasized the importance of information security and access controls. Finally, in this paper, we present the findings related to the needs of the formal support network from 7 focus groups, which consisted of 44 health care professionals working with community-dwelling older people. The focus groups explored the types and sources of patient-related information the various groups needed to support their work, how they accessed the information, and how telemedicine or smart home technologies might provide them with additional useful information about their patients. The findings of this study highlight the different resources and information requirements of the many stakeholder groups who may be involved, thereby showing that information governance issues are important.

The work reported here adds to the scholarly body of knowledge by examining the various and differing information requirements and technological capabilities of different groups of health care professionals who could benefit from using home-monitoring technologies to support the health of older people living in their own homes. The qualitative nature of the research means that care should be taken when generalizing the findings beyond the study locations; however, this study is valuable because it highlights the need to address issues of information access and governance, which have been largely neglected in related health informatics research to date.

Methods

Focus Groups

Seven face-to-face focus groups were conducted with participants employed at health care organizations that provide support for older people. The focus groups were conducted at the respective organizations where the participants were employed during normal working hours. On average, each focus group lasted for 1 hour. The study procedures were approved by the Massey University Human Ethics Committee (SAO 16/65).

Recruitment

Participants were recruited through purposive convenience sampling, which aimed to gather a range of contrasting perspectives. We contacted a range of health care organizations that provide various support services for older people in the Manawatū region in New Zealand, requesting their cooperation and that information about the study be distributed to staff. The focus groups were voluntary, and interested participants chose to attend and participate. They did not receive financial incentives for participation; 44 participants took part in the study. They were predominantly women and their average age was 46.6 years. Further information about the focus groups is outlined in Table 1.

Focus groups 1 to 4 were composed of employees of the local public health system. These different professional teams, physiotherapists, occupational therapists, social workers, and nurses were based at the local hospital but provided services to community-dwelling patients in their homes. Primarily, they provide ongoing services to people who have been discharged from the hospital and to clients who have been referred from primary care practices. Each of the other 3 focus groups consisted of team members from primary care practices, which were based in clinics within the community. These groups

primarily provide services at their clinics with limited community-based services. The sample size was appropriate because while there was some variation in the perspectives between the hospital-based teams and the primary care teams, there were repeated commonalities across the focus groups, whereby researchers deemed that data saturation was met.

Table 1. Composition of the focus groups.

Characteristics	Focus group 1 (n=6)	Focus group 2 (n=6)	Focus group 3 (n=5)	Focus group 4 (n=5)	Focus group 5 (n=8)	Focus group 6 (n=9)	Focus group 7 (n=5)
Organization type	Public health system	Public health system	Public health system	Public health system	Primary care practice	Primary care practice	Primary care practice
Participants' occupations	3 physiotherapists 2 occupational therapists 1 clinical coordinator	6 social workers	5 social workers	5 nurses	1 general practitioner 4 nurses 1 clinical director 1 business manager 1 clerk	2 general practitioner 1 urgent care doctor 2 nurses 1 social worker 1 clerk 2 directors	1 social worker 2 geriatricians 2 nurses

Focus Group Design and Content

Focus groups were used because they are an efficient method for obtaining data from multiple participants [15], and they can stimulate in-depth discussion around key ideas [16]. The flexible nature of the focus groups means that researchers can probe and clarify meanings that may be implied or unclear [17]. Conducting focus groups involves the negotiation of complex power and social dynamics between the researcher and participants and within the participant group in response to the research setting and wider context [18]. For instance, within health organizations, there are likely to be existing hierarchical and professional structures. Although this can never be entirely overcome in a focus group setting, we tried to address this by listening to what was being said, by whom, and under what conditions, and asking and probing for different participants to share their opinions.

Focus group questions were developed to address the project's purpose and increase consistency across sessions. The development of these questions was guided by a workshop conducted with participants attending a health informatics conference [13], which provided another lens of the research procedures to help reduce an unconscious bias in the research design. Each focus group began with the distribution of participant information sheets and consent forms, followed by an introduction about the project and the use of home-monitoring technology and information and communication technologies connecting older people to their informal and formal support networks to assist aging in place. Participants were then asked about their information needs, what information should be collected and transferred, who should receive this information, as well as the potential ethical concerns.

Analysis

The focus groups were transcribed and thematically analyzed [19] inductively in NVivo Version 11.0 (QSR International). Owing to the nature of the focus groups, individual participants were not identified. Thematic analysis is widely used within the social and health sciences [20] as a tool to examine "repeated patterns of meaning" [18] or as a way of identifying and making sense of commonalities within data sets [19]. The coding and theme development were inductive, using an iterative process that involved reading and rereading the data sets to establish initial codes that covered the key ideas discussed and then combining similar codes under the themes. Following this process, the themes were reviewed alongside the original data set as a way of checking that the data set was correctly represented and that important data were not missed. The analysis was undertaken by a primary researcher, which was discussed within the research team throughout the analysis and cross-checked by a second researcher.

Results

Overview of the Findings

The New Zealand public health care system has a bifurcated structure and complex funding structures, which complicates the sharing of information systems and of patient information. The government funds health care on a regional basis through several district health boards, which directly run their local hospital-based services and some community services that are based at the hospital, such as those delivered by the members of focus groups 1-4. The services that are delivered directly by the district health boards are totally tax-payer funded and free at point of care. The delivery of local primary care services is essentially outsourced by the district health boards to several independent primary care practices, such as those observed in focus groups 5-7. Most of the services they provide are heavily subsidized by the district health boards but most patients are required to make some additional payment for the services they receive. The primary care practices each operate their own sophisticated patient management systems.

The analysis of the focus groups identified 3 overarching themes: (1) access to technology and systems, (2) collecting and sharing of information, and (3) barriers to uptake. Each theme has several subthemes, which are summarized in Table 2 and described below. An important observation from our analysis of the different focus groups is the contrast in the facilities and the perspectives between the members of the teams

working in the community within the public health system and those of the primary care teams.

Table 2. Summary of the major themes and perspectives of the respondent groups.

Themes, subthemes	Perspectives	
	Hospital-based teams, focus groups 1-4	Primary care teams, focus groups 5-7
Access to technology and systems		
Technologies used	Multiple fragmented systems, paper-based notes	Good integrated internal system
Major limitations	No access to primary care systems, no mobile access to systems	Poor integration with hospital systems
Collecting and sharing of information		
Additional information required	Information on background and social environment	Less interest in social information
Interest in telemonitoring	Monitoring of falls highly desirable, little desire for clinical monitoring	Monitoring of falls highly desirable, interest in telemonitoring specific clinical indicators, but preprocessing of data required
Barriers to uptake		
Usability	Concerns about dexterity requirements and cognitive decline of patients	Concerns about dexterity requirements and cognitive decline of patients
Cost	Concerns about system funding and cost to patients	Concerns about system funding and cost to patients
Security, privacy, and confidentiality	Concerns about patient privacy and possible breeches of confidentiality	Concerns about patient privacy and possible breeches of confidentiality

Theme 1: Access to Technology and Systems

Despite many years of development and implementation of integrated health information systems, many respondents complained about the way patient information remains fragmented within multiple systems. There are multiple systems within different parts of the hospital, which are not universally accessible to the staff members of the different focus groups who might need the information, and this leads to frustration and inefficiencies.

> *…Probably the easiest way is to say multiple computer systems. And not everybody has access to the same information.* [Focus group 3]

While the hospital-based teams have some access to hospital computer systems, their own on-going notes are paper-based and only a final summary is shared to an electronic system.

> *…The referrals get processed into the computer but then that's printed off…all our records are paper-based.* [Focus group 4]
>
> *…If I have finished seeing Mrs. Smith, we do a discharge summary and type that up and it goes on to a clinical portal. So, discharge summaries or clinical letters can go onto their file but the actual running notes of "I saw someone and provided something" is all written by hand.* [Focus group 1]

This leads to problems coordinating care within and across teams.

> *…I mean there are just lots of people collecting data and then keeping it to themselves essentially and if you want to get as much data on this person as you can, you have to physically run around grabbing files and saying what are your thoughts and what have you noticed, rather than actually having it as a central accessible thing.* [Focus group 1]

Another issue for the hospital-based staff is the lack of access to information from the primary care clinics, which operate their own systems.

> *…One big gap is that we don't have access to the GP records. We need to physically ring, have a chat. We can't just login, you know just to kind of get an idea.* [Focus group 3]

In contrast to the hospital-based teams, the primary care teams run their own sophisticated electronic systems, but they also complained about the lack of integration with other service providers.

> *…Pretty much everything from the hospital in terms of what information you receive, district nursing discharges, physio, everything we get… comes via fax or by post, we have to scan it into the notes.* [Focus group 5]

All groups also raised the issue of the difficulty of obtaining a complete picture of the patients and the presence of gaps in information from other government organizations or from nongovernment service providers. For example, many clients paid for personal alarm systems, which could summon paramedical assistance when needed, for example, in case of a fall, but these service providers tended not to share their patient information with the public health system.

> *…Older patients see so many other providers and nongovernment service providers and social services*

that we don't ever get any feedback about. [Focus group 5]

Theme 2: Collecting and Sharing of Information

As noted in Theme 1, all respondents complained of the difficulties of locating patient information, which could be distributed across multiple disconnected systems, and all felt that better integrated systems would be beneficial for everyone. Focus groups also explored the additional information they would like to have available, which was not currently being collected. Some of this information could be collected by remote monitoring technologies but some information would need to be collected and shared by practitioners after their visits. Again, the perspective sometimes varied depending on whether the focus group participants worked in the community or in the primary care practice. Community practitioners, in particular, wanted more information about patients' background data and social environment before they made a home visit.

…[You] go and see Mrs. Smith, she has no arms and legs, isn't English speaking, has a pressure ulcer, falls over 10 times a day, but nothing is written on the referral. [Focus group 1]

…You don't really know what you're walking into. You could be going into an 80-year-old at home with no fire lit, no food, or you could be going to the opposite. [Focus group 4]

This could also be a matter of personal safety.

…So it would be really nice for us to see, right, alcoholic, maybe we should take somebody else with us on this particular visit. [Focus group 1]

All participants thought it would be useful if technology could retrieve information about patients' social well-being, such as whether they are socializing and leaving the house and their physical well-being such as their diet, home temperature, and physical activity.

…cause they could be isolated and not communicating outside of their home environment, and we may not know that. Similarly, you know, about not accessing food. [Focus group 6]

All participants were particularly concerned about receiving detailed information about older patients' falls; this information could be captured by wearable technologies or video systems. Information about falls is important when deciding if extra resources are needed or if someone should no longer be living on their own. Objective evidence would often be useful rather than relying on reports from clients or their family, who often differ in their perspectives.

…Knowing how often someone had fallen would be evidence. Quite often we have to put quite an argument to the Ministry of Health for funding for things and if you said subjectively, I am sure that this person has bad balance and is at risk of falling, but if you had a concrete "they've fallen 10 times in the last 6 weeks" then that's 10 falls. [Focus group 1]

…Actually a fall is a good event to focus on because even though they wear their alarms, they don't often push them, elderly people, so, you know, it's good. [Focus group 7]

The primary care doctors were also interested in the possibilities of using technology to improve the telemonitoring of patients' parameters such as blood pressure, blood glucose levels, peak flow readings, and weight.

…I suppose that's already happening, I know [xxx] has got quite a few patients that have got the blood pressure monitoring that they email him blood pressure results, the same with blood sugar, glucose results. [Focus group 5]

However, they were concerned about being overwhelmed with data and were enthusiastic about the idea of information being preprocessed and filtered and only passed on when exceptions arise.

…You don't want stuff pushed at you, you just want to see the exceptions, wow, something really strange is happening, yeah. So very much processed and filtered before it gets to you. You don't want a daily update on everything. [Focus group 5]

Furthermore, any new information should be seamlessly integrated into the existing systems.

…That's a really good point, cause we don't want too many databases we have to go into. It's a nightmare… you've got to consolidate in some way. Yeah, no absolutely. [Focus group 6]

However, primary care teams were clear that they were unwilling to assume the additional responsibility of managing such systems.

…Why does it always have to go back to a nurse or a doctor? …we're already busy and overwhelmed with our workloads that we may not be the right people. [Focus group 5]

Theme 3: Barriers to Uptake

Participants were optimistic about technology retrieving information and were clear about the potential barriers to uptake. This theme encompasses 3 subthemes: usability, cost, and security and privacy. Participants were concerned about software complexity and usability issues associated with aging, such as dexterity and cognitive decline. Many participants indicated that the training required for older people would be substantial and that their desire to learn to use technology would be mixed.

…I give them the computer cord (to connect the glucose monitor to their computer) to use download express software, to download their blood glucose monitor and email it to me...I haven't had any success...The program is quite convoluted for them...younger techie people just pick it up. [Focus group 6]

…With ManageMyHealth [patient portal] they're keen to engage. But then it's a struggle. And quite often we'll just take them out the back and actually do it all for them. [Focus group 6]

Participants considered cost to be another barrier and discussed how it is already a barrier for their uptake of health products.

> *…When I talk to patients about personal alarms, it always comes down to the cost and "how much are they? I can't afford that"… so cost, I would say, would be a huge thing, especially for the elderly.* [Focus group 2]

> *…The majority of the elderly have not got that extra $15 a week to pay for a blister pack (medicine sachet packaging)… $60 a month—that's whether you're going to keep the heater on.* [Focus group 5]

Ensuring privacy and data security was considered to be a requirement to protect patients. Some participants felt that clients would have concerns about privacy and data security, although this barrier did not present as strongly compared to usability and cost.

> *…They get quite protective of their privacy because as they get older they feel they are losing more and more.* [Focus group 1]

> *…Any piece of equipment or technology, privacy is a big thing.* [Focus group 1]

It was clear from the participants that sharing information with informal support networks is common, although some participants indicated the need to obtain informed consent.

> *…We try and help to facilitate that everybody is on the same page and know all the information.* [Focus group 3]

> *…Well, we have to gain the patient's consent for information to be shared with family. And we have to be very cautious about what medical information we share.* [Focus group 3]

Respondents had concerns about sharing information with family and how the information might be used as there may be conflicts within the family, for example, about what are the most appropriate arrangements for care.

> *…We might also talk to them if we have concerns about the patient's safety… it's a fine line between breaching confidentiality and what is appropriate to actually divulge to a support person.* [Focus group 4]

> *…Are we monitoring for the right reasons? Is that person being monitored for the health benefits or are they being monitored because... you know there is the safety side of it but there is also the dilemma of what the other dimensions of the family are wanting. Is it exactly what your client wants?* [Focus group 1]

Discussion

Overview

The work reported in this paper is part of the requirements-gathering phase of an exploratory project, which investigates how technologies that connect older people to their informal and formal support networks could assist aging in place and enhance older people's health and well-being. Seven focus groups consisting of 44 health care professionals considered their information needs and the ways that technology could help meet some of those needs. The limited size of the study means that care should be taken not to overgeneralize the conclusions, but several common themes emerged within the findings.

Although the use of telemonitoring and smart home technologies to support older people has been proposed since many years, there has been very limited uptake [21,22], especially within the public health care systems [23,24]. However, many older people and their families are recognizing the benefits that these monitoring technologies can provide in terms of safety and peace of mind for all concerned, and commercial organizations are starting to provide related products and services [25,26]. Such organizations tend to work on a fee-for-service model, where the user pays to have their information collected, monitored, and acted upon as necessary in a limited number of ways. These systems tend to be closed cloud-based ecosystems with the information remaining with the service provider.

For monitoring systems to reach their full potential to improve the well-being of all older people, they must be expanded and integrated into the wider health care system. This raises the information governance issues of how the data and information collected by smart home technologies and other monitoring technologies can be effectively integrated with other systems and made available to the health care workforce.

Access to Technology and Systems

The first major theme to emerge from the focus groups was that there remains a lack of integration between current information systems, especially between primary care and secondary care institutions. Such a lack of integration has been identified as a major problem in previous studies [27-29]. Secondary care also has multiple computer-based systems as well as paper-based notes, which inhibits the easy sharing of appropriate information between health care professionals. Furthermore, many other government departments, nongovernmental organizations, and private service providers often possess useful information about patients, which is not shared with the health service. Participants were frustrated by their inadequate access to information, as practitioners described having little information about patients before visits. This is not only a safety concern, but it is also well-established that information access can impact decision-making and the quality of care and efficiency [30]. This is a particular issue in aged care, as older people are more likely to have chronic conditions and disabilities [3]. Although the New Zealand Ministry of Health is working toward a more integrated information ecosystem [31], the current information system fragmentation does not place the patient at the center of health care delivery. This disconnected information architecture means that currently, there is no clear way for sensor data from telemonitoring and smart home devices to be integrated with other patient information. In some ways, primary care might seem to be the most appropriate site for storing the information since primary care clinics are seen as the site of ongoing, long-term relationships with the patients [32,33]; however, the participants from primary care organizations in this study were unwilling to take on the responsibility. Ultimately, it might be

appropriate to create a new organization to take on the responsibility for the implementation and governance of systems to collect, process, and disseminate this monitoring information to all interested parties in the health care sector and to informal support networks.

Collecting and Sharing of Information

Research into telemonitoring and smart homes has identified many ways to collect information that can be useful for various health care providers. However, focus group participants emphasized that they were concerned that they should not be overloaded with information. They wanted to be able to tailor the delivery of information for each individual patient and to have it preprocessed so that only exceptional situations were identified and flagged. The feasibility of processing information on exceptions with blood pressure recordings by telemedicine systems has been demonstrated in previous research [34,35] and similar techniques should be integrated into future developments. Participants also wanted the information to be integrated with their existing patient management system so that they only needed to access 1 system.

Barriers to Uptake

The findings also reaffirm several user requirements deduced from interviews previously undertaken with older people and informal support networks [14]. Some issues addressed by these requirements correspond to barriers already identified in health informatics research. For example, cost is an issue for many older people [21,36], and governments should undertake detailed cost-benefit analyses of these systems to determine what level of public funding is appropriate. Usability is another barrier [21,36]; therefore, the technology should be designed for ease-of-use. A direct user interface may not be suitable for some older people, particularly for those living with cognitive decline. A number of systems have been developed that use motion sensor technologies to passively monitor the home environment and send alerts when certain out-of-the-ordinary events occur [37,38], and future developments should consider the integration of such systems.

Concerns about privacy and the desire to control the distribution of information were also seen as barriers to uptake, which have also been widely recognized in other studies [23,39,40]. The information needs of informal support networks have been largely neglected in previous research [21], even though they are often essential for enabling aging in place [21], and inadequate information access has been shown to negatively impact their ability to assist older people [10]. An important finding was that participants indicated that older people's support networks are frequently a considerable part of information exchange, both for retrieving and providing information. However, this is generally undertaken through an informal means, rather than through a specific technological solution. This is an example of technology failing to address the realities of supporting aging in place as a collaborative and information-critical activity. However, providing access to information from home-monitoring systems to an older person's informal support network will require much stricter controls to ensure privacy and confidentiality than is necessary for a system that is restricted to health care personnel.

Limitations

The findings of this study are limited to the 44 participants recruited and the organizational settings in which they work. Reflective of the nature of qualitative enquiry [41], the participant recruitment was not drawn from a random sample of subjects, but it rather comprised individuals and organizations in the Manawatū region in New Zealand that actively volunteered to be involved in the study. This resulted in the participant number varying across the focus group sessions and the participants only being employed at general health practices and hospital departments, although they represented a range of health professionals. There are also limitations that are inherent to the nature of conducting focus groups, which can involve complex social dynamics, especially within organizations; therefore, the findings should be considered as a product of the groups and not as individuals [42]. Care should be taken when generalizing the findings drawn from the study beyond the study locations, and future research examining this topic should certainly take a wider scope.

Conclusion

The 44 participants working with community-dwelling older people in this study want more information about their patients' well-being within their homes, thereby demonstrating a potential for home monitoring and information communication technologies to connect older people to their formal support networks. This would not only assist health practitioners but would also support aging in place, which is socially and economically beneficial, as life expectancy continues to rise [3]. Our analysis of the discussions within the focus groups highlights the importance and the complexity of management and governance of systems to collect and disseminate such information. Issues that will need to be considered in the future are as follows.

1. Where will the data be stored and who will be responsible for its governance?
2. Who should pay for the technology and the ongoing running costs?
3. How will data be processed before it is presented?
4. How will the information be integrated into other systems used by health care providers?
5. To what extent should the information be released to informal support networks?

Further research into the requirements of all stakeholder groups is required to address these issues, which need to be resolved before the true potential of these technologies can be realized in practice.

Acknowledgments

This study was supported by the Health Research Council of New Zealand [reference number 16/679]. We would like to thank all the participants in this study for their time and cooperation.

Authors' Contributions

All authors made a direct substantial intellectual contribution to this study. All authors approved the final version.

Conflicts of Interest

None declared.

References

1. OECD. Country statistical profiles. 2016. URL: http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT [accessed 2020-01-23]
2. Statistics NZ. National population projections: 2016 (base) -2068. URL: http://archive.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationProjections_HOTP2016.aspx [accessed 2020-11-23]
3. Ministry of Health. Healthy ageing strategy. 2016. URL: https://www.health.govt.nz/system/files/documents/publications/healthy-ageing-strategy_june_2017.pdf [accessed 2020-11-23]
4. Ministry of social Development. Positive ageing goals and key actions. 2001. URL: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/planning-strategy/positive-ageing/goals-and-actions.html [accessed 2020-11-23]
5. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]
6. Canally C, Doherty S, Doran DM, Goubran RA. Using integrated bio-physiotherapy informatics in home health-care settings: A qualitative analysis of a point-of-care decision support system. Health Informatics J 2015 Jun;21(2):149-158 [FREE Full text] [doi: 10.1177/1460458213511346] [Medline: 24835146]
7. Loane J, O'Mullane B, Bortz B, Knapp RB. Looking for similarities in movement between and within homes using cluster analysis. Health Informatics J 2012 Sep;18(3):202-211 [FREE Full text] [doi: 10.1177/1460458212445501] [Medline: 23011815]
8. Wolters MK, Kelly F, Kilgour J. Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. Health Informatics J 2016 Dec;22(4):854-866 [FREE Full text] [doi: 10.1177/1460458215593329] [Medline: 26276794]
9. Nyman SR, Yardley L. Usability and acceptability of a website that provides tailored advice on falls prevention activities for older people. Health Informatics J 2009 Mar;15(1):27-39 [FREE Full text] [doi: 10.1177/1460458208099866] [Medline: 19218310]
10. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. Int J Med Inform 2014 Sep;83(9):624-635 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.005] [Medline: 24996581]
11. Singh A, Guesgen H, Whiddett D, Hunter I, Elers P, Lockhart C. Intelligent Digital Environment to Assist Elderly People: System Design and Evaluation. Wellington: Health Informatics New Zealand; 2018 Presented at: Proceedings of the 2018 HiNZ Conference; 21-23 November; Wellington, New Zealand p. 22 URL: https://cdn.ymaws.com/www.hinz.org.nz/resource/resmgr/conference_2018/HiNZ18_proceedings_19.12.19.pdf
12. Hunter I, Whiddett D, Elers P, Lockhart C, Guesgen H, Singh A. Home is where the health is: Findings from the prototyping of a system that connects older adults to their support networks to assist ageing in place in Proceedings of the 2018 HiNZ Conference; 2018. URL: https://cdn.ymaws.com/www.hinz.org.nz/resource/resmgr/conference_2018/HiNZ18_proceedings_19.12.19.pdf [accessed 2020-11-23]
13. Elers P, Whiddett D, Hunter I, Lockhart C, Guesgen H. Supporting older adults to age in place using technology: The shifting boundaries between formal and informal support groups. Proceedings of the 2017 Health Informatics New Zealand Conference (HiNZ 2017). 2017. URL: https://c.ymcdn.com/sites/www.hinz.org.nz/resource/resmgr/conference_2017/HiNZ17_proceedings.pdf [accessed 2020-11-23]
14. Elers P, Hunter I, Whiddett D, Lockhart C, Guesgen H, Singh A. User Requirements for Technology to Assist Aging in Place: Qualitative Study of Older People and Their Informal Support Networks. JMIR Mhealth Uhealth 2018 Jun 06;6(6):e10741 [FREE Full text] [doi: 10.2196/10741] [Medline: 29875083]
15. Krueger R, Casey M. Focus groups: A practical guide for applied research (3rd ed). Thousand Oaks, CA: Sage; 2014.
16. Stewart D, Shamdasani P. Focus groups: Theory and practice (Vol 20). Thousand Oaks, CA: Sage; 2014.
17. Balch G, Mertens D. Focus Group Design and Group Dynamics: Lessons from Deaf and Hard of Hearing Participants. American Journal of Evaluation 1999;20(2):265-277. [doi: 10.1177/109821409902000208]
18. Dutta M, Elers P. Interpretive approaches to listening. In: Worthington D, editor. Handbook of listening. London: Wiley; 2020:41-54.

19. BraunV, ClarkeV, TerryG. Thematic analysis. In: Rohleder P, Lyons A, editors. Qualitative research in clinical and health psychology. London: Palgrave MacMillan; 2015:95-113.
20. 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]
21. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014 Apr;83(4):235-248. [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]
22. Ienca M, Wangmo T, Jotterand F, Kressig RW, Elger B. Ethical Design of Intelligent Assistive Technologies for Dementia: A Descriptive Review. Sci Eng Ethics 2018 Aug;24(4):1035-1055. [doi: 10.1007/s11948-017-9976-1] [Medline: 28940133]
23. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: A systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
24. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Hinder S, Rouncefield M. What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services. BMC Med 2015 Apr 23;13(1). [doi: 10.1186/s12916-015-0279-6]
25. St John Medical Alarm. URL: https://www.stjohn.org.nz/medical-alarms/ [accessed 2020-11-23]
26. Medical Alarms. Securely. URL: https://securely.nz/medical-alarms/ [accessed 2020-11-23]
27. Auschra C. Barriers to the Integration of Care in Inter-Organisational Settings: A Literature Review. International Journal of Integrated Care 2018;18(1):5 [FREE Full text] [doi: 10.5334/ijic.3068]
28. Salomi MJA, Claro PB. Adopting Healthcare Information Exchange among Organizations, Regions, and Hospital Systems toward Quality, Sustainability, and Effectiveness. TI 2020;11(03):58-97. [doi: 10.4236/ti.2020.113005]
29. Eden KB, Totten AM, Kassakian SZ, Gorman PN, McDonagh MS, Devine B, et al. Barriers and facilitators to exchanging health information: a systematic review. Int J Med Inform 2016 Apr;88:44-51 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.01.004] [Medline: 26878761]
30. Heart T, Ben-Assuli O, Shabtai I. A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. Health Policy and Technology 2017 Mar;6(1):20-25. [doi: 10.1016/j.hlpt.2016.08.002]
31. Ministry of Health. Digital Health Strategic Framework. 2020. URL: https://www.health.govt.nz/our-work/digital-health/digital-health-strategic-framework [accessed 2020-11-23]
32. Atmore C. General practice evolution in New Zealand: hybridisation in action. J Prim Health Care 2017;9(3):220. [doi: 10.1071/hc17043]
33. Downs A. From Theory to Practice: the promise of primary care in New Zealand. Wellington, New Zealand: Fulbright; 2017. URL: https://www.fulbright.org.nz/wp-content/uploads/2017/09/DOWNS-From-Theory-to-Practice-The-Promise-of-Primary-Care-in-New-Zealand-.pdf [accessed 2020-03-01]
34. Kayıkçıoğlu İ, Akdeniz F, Kayıkçıoğlu T, Kaya İ. Real-Time Monitoring of ST Change for Telemedicine. CMBEBIH 2017. IFMBE Proceedings 2017;62 [FREE Full text] [doi: 10.1007/978-981-10-4166-2_101]
35. Massaro A, Maritati V, Savino N, Galiano A, Convertini D, De Fonte E, et al. A Study of a Health Resources Management Platform Integrating Neural Networks and DSS Telemedicine for Homecare Assistance. Information 2018 Jul 19;9(7):176. [doi: 10.3390/info9070176]
36. Grace A, Gleasure R. Supporting the critical role of family carers in wellness management. Health Policy and Technology 2017 Jun;6(2):242-249. [doi: 10.1016/j.hlpt.2017.04.002]
37. Patel PA, Gunnarsson C. A Passive Monitoring System in Assisted Living Facilities: 12-Month Comparative Study. Phys Occup Ther Geriatr 2012 Mar;30(1):45-52 [FREE Full text] [doi: 10.3109/02703181.2011.650298] [Medline: 22448082]
38. Alexander G, Wakefield B, Rantz M, Skubic M, Aud MA, Erdelez S, et al. Passive Sensor Technology Interface to Assess Elder Activity in Independent Living. Nursing Research 2011;60(5):318-325 [FREE Full text] [doi: 10.1097/nnr.0b013e318225f3e1]
39. Peek STM, Luijkx KG, Rijnaard MD, Nieboer ME, van der Voort CS, Aarts S, et al. Older Adults' Reasons for Using Technology while Aging in Place. Gerontology 2016;62(2):226-237 [FREE Full text] [doi: 10.1159/000430949] [Medline: 26044243]
40. Kapadia V, Ariani A, Li J, Ray PK. Emerging ICT implementation issues in aged care. Int J Med Inform 2015 Nov;84(11):892-900. [doi: 10.1016/j.ijmedinf.2015.07.002] [Medline: 26250987]
41. Morse JM. Critical Analysis of Strategies for Determining Rigor in Qualitative Inquiry. Qual Health Res 2015 Sep;25(9):1212-1222. [doi: 10.1177/1049732315588501] [Medline: 26184336]
42. Short S. Focus group interviews. In: Perecman S, Curran S, editors. A handbook for social science field research: Essays & bibliographic sources on research design and methods. Thousand Oaks, CA: Sage; 2006:103-115.

Edited by L Buis; submitted 06.09.20; peer-reviewed by K Day, J Soar; comments to author 07.10.20; revised version received 19.10.20; accepted 09.11.20; published 02.12.20.

Please cite as:
Hunter I, Elers P, Lockhart C, Guesgen H, Singh A, Whiddett D
Issues Associated With the Management and Governance of Sensor Data and Information to Assist Aging in Place: Focus Group Study With Health Care Professionals
JMIR Mhealth Uhealth 2020;8(12):e24157
URL: https://mhealth.jmir.org/2020/12/e24157
doi:10.2196/24157
PMID:33263551

Original Paper

Ambulatory Phonation Monitoring With Wireless Microphones Based on the Speech Energy Envelope: Algorithm Development and Validation

Chi-Te Wang[1,2,3], MD, PhD; Ji-Yan Han[4], BSc; Shih-Hau Fang[2,5], PhD; Ying-Hui Lai[4], PhD

[1]Department of Otolaryngology Head and Neck Surgery, Far Eastern Memorial Hospital, New Taipei City, Taiwan

[2]Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan

[3]Department of Special Education, University of Taipei, Taipei, Taiwan

[4]Department of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan

[5]Ministry of Science and Technology Joint Research Center for Artificial Intelligence Technology and All Vista Healthcare, Taoyuan, Taiwan

Corresponding Author:
Ying-Hui Lai, PhD
Department of Biomedical Engineering
National Yang-Ming University
No155, Sec 2, Linong Street
Taipei, 112
Taiwan
Phone: 886 228267021
Fax: 886 228210847
Email: yh.lai@gm.ym.edu.tw

Abstract

Background: Voice disorders mainly result from chronic overuse or abuse, particularly in occupational voice users such as teachers. Previous studies proposed a contact microphone attached to the anterior neck for ambulatory voice monitoring; however, the inconvenience associated with taping and wiring, along with the lack of real-time processing, has limited its clinical application.

Objective: This study aims to (1) propose an automatic speech detection system using wireless microphones for real-time ambulatory voice monitoring, (2) examine the detection accuracy under controlled environment and noisy conditions, and (3) report the results of the phonation ratio in practical scenarios.

Methods: We designed an adaptive threshold function to detect the presence of speech based on the energy envelope. We invited 10 teachers to participate in this study and tested the performance of the proposed automatic speech detection system regarding detection accuracy and phonation ratio. Moreover, we investigated whether the unsupervised noise reduction algorithm (ie, log minimum mean square error) can overcome the influence of environmental noise in the proposed system.

Results: The proposed system exhibited an average accuracy of speech detection of 89.9%, ranging from 81.0% (67,357/83,157 frames) to 95.0% (199,201/209,685 frames). Subsequent analyses revealed a phonation ratio between 44.0% (33,019/75,044 frames) and 78.0% (68,785/88,186 frames) during teaching sessions of 40-60 minutes; the durations of most of the phonation segments were less than 10 seconds. The presence of background noise reduced the accuracy of the automatic speech detection system, and an adjuvant noise reduction function could effectively improve the accuracy, especially under stable noise conditions.

Conclusions: This study demonstrated an average detection accuracy of 89.9% in the proposed automatic speech detection system with wireless microphones. The preliminary results for the phonation ratio were comparable to those of previous studies. Although the wireless microphones are susceptible to background noise, an additional noise reduction function can alleviate this limitation. These results indicate that the proposed system can be applied for ambulatory voice monitoring in occupational voice users.

(JMIR Mhealth Uhealth 2020;8(12):e16746) doi:10.2196/16746

KEYWORDS

voice disorder; speech envelope; phonation habits; background noise; noise reduction; adaptive threshold; dosimetry; phonotrauma

Introduction

Human voice is produced via the periodic vibrations of vocal folds, driven by expiratory airflow. Cumulative voice loads and excessive vocal fold vibrations result in phonotraumatic injuries, such as vocal nodules and polyps [1]. The common symptoms of dysphonia (ie, phonation discomfort) include hoarseness, vocal fatigue, increased effort, and throat pain, which may limit the performance and long-term careers of occupational voice users [2]. Dysphonia also results in considerable financial losses for individuals and society [3,4]; the estimated annual cost associated with dysphonia is US $2.5 billion [5]. In addition, voice-related disorders significantly lower the quality of life in terms of physical functioning, general health, bodily pain, fatigue, and role limitation [6].

The most recognized risk for voice disorders is occupational voice overuse, commonly found in salespeople, industrial/factory workers, teachers, clergy, lecturers, and singers [6,7]. Among these occupations, the teaching profession has been significantly investigated by academic researchers [8-10]. In comparison to other occupations, teachers are more likely to report voice problems and the negative effects of dysphonia on their work performance [11]. Roy et al [2] reported that the prevalence of voice disorders was significantly higher in teachers (137/1243, 11.0%) in comparison to nonteachers (80/1288, 6.2%). The lifetime prevalence of dysphonia for teachers (717/1243, 57.7%) was also significantly higher than that for nonteachers (371/1288, 28.8%).

Voice therapy has been widely applied as the first-line treatment for voice disorders related to voice overuse or abuse [12,13]. By implementing multiple treatment strategies, voice therapy can effectively ameliorate the dysphonic symptoms, lower the phonation effort, and improve the voice quality [13]. However, one of the major challenges for voice therapy is the carryover of voicing techniques and habits taught during treatment sessions into their daily lives. To facilitate the maintenance of adequate phonation behavior, serial studies proposed the concept of ambulatory voice monitoring with promising results [14-16]. Most of these studies used a contact microphone or accelerometer attached to the anterior neck [17,18]. Subsequent studies demonstrated that this technology could significantly aid patients in controlling and tracking their vocal hyperfunction [19,20]. Another device for ambulatory voice monitoring was designed as a neck collar embedded with a contact microphone [21]. Although contact microphones and accelerometers can accurately detect phonation via the vibration of neck skin, the wiring and taping involved with these devices may cause discomfort in users. Furthermore, voice usage was mostly analyzed over a certain period with post hoc feedback [22], whereas real-time monitoring of the phonation ratio has not yet been reported.

To overcome the limitations in current devices, we propose a novel automatic speech detection system using a wireless microphone to capture acoustic signals from users, which can eliminate the discomfort associated with the wiring and taping of contact microphones. Our study hypothesizes that the speech energy envelope received via a wireless microphone can be used for ambulatory phonation monitoring. To examine this hypothesis, we designed this research with the following objectives: (1) to investigate the detection accuracy of speech, (2) to compare the measured phonation ratio and length of speech segments with those in existing literature, and (3) to examine the robustness of the noise reduction algorithm in simulated noisy conditions.

Methods

Overall Study Design

We proposed an automatic speech detection system using a wireless microphone for real-time ambulatory voice monitoring. We invited 10 teachers to participate in the pilot study. We designed an adaptive threshold (AT) function to detect the presence of speech based on the energy envelope. All participants were equipped with a wireless microphone during a teaching session (around 40-60 minutes) in a quiet classroom (background noise <55 dB sound pressure level [SPL]). We developed software for manually labeling the speech segments according to the time and frequency domains. We randomly selected 25 utterances (10 seconds each) from the recorded audio files to acquire the coefficients required for the AT function using a genetic algorithm (GA). Another 5 random utterances were used to test the accuracy of the automatic speech detection system using manually labeled data as the ground truth. We also mimicked scenarios of noisy backgrounds by mixing 4 different types of noise (at a signal-to-noise ratio [SNR] of 0, 3, and 5 dB) into the original recordings. An adjuvant noise reduction function using a log minimum mean square error (logMMSE) [23] algorithm was applied to counteract the influence on detection accuracy.

Participants

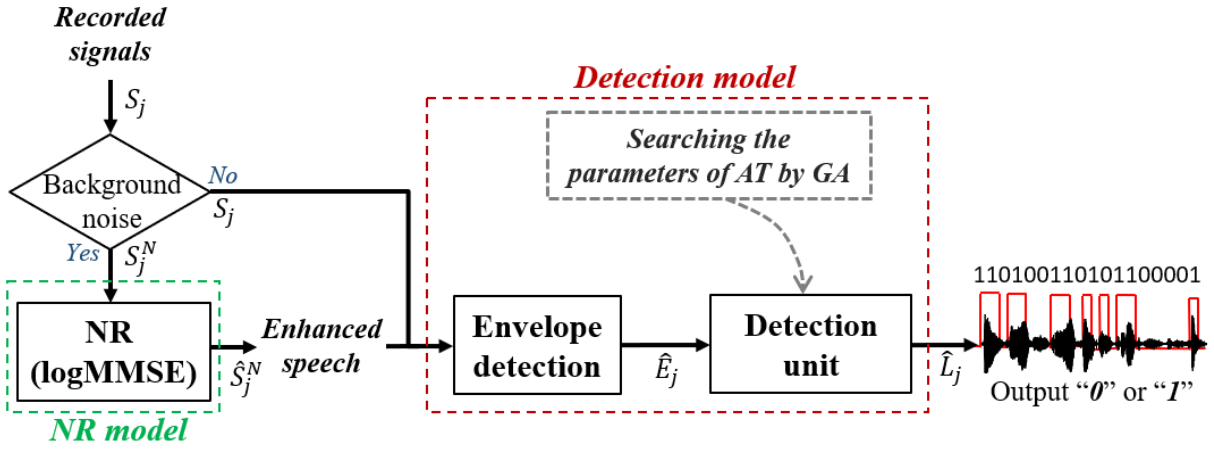
We invited 10 teachers to participate in this study. This study was conducted at Far Eastern Memorial Hospital and National Yang-Ming University. The study protocol was approved by the Research Ethics Review Committee of Far Eastern Memorial Hospital (FEMH 108019-E). For the first period of study, we recruited 5 teachers from April to June 2019; for the second period, we recruited another 5 teachers from February to April 2020. Each teacher was provided with a wireless microphone during a regular teaching session of 40-60 minutes. The average background noise level was controlled under 55 dB SPL, established as the controlled environment test condition.

Automatic Speech Detection System

Overview

Figure 1 illustrates the automatic speech detection system proposed in this study. The main corpus of this system is the detection model, which automatically divides acoustic signals into speech and nonspeech segments based on the energy envelope. We used a frame size of 32 milliseconds with a sampling rate of 16 kHz. Under simulated noisy conditions, the noise reduction model can be turned on to alleviate the effects from the background noise.

Figure 1. Proposed automatic speech detection system. AT: adaptive threshold; GA: genetic algorithm; logMMSE: logarithm minimum mean square error; NR: noise reduction.



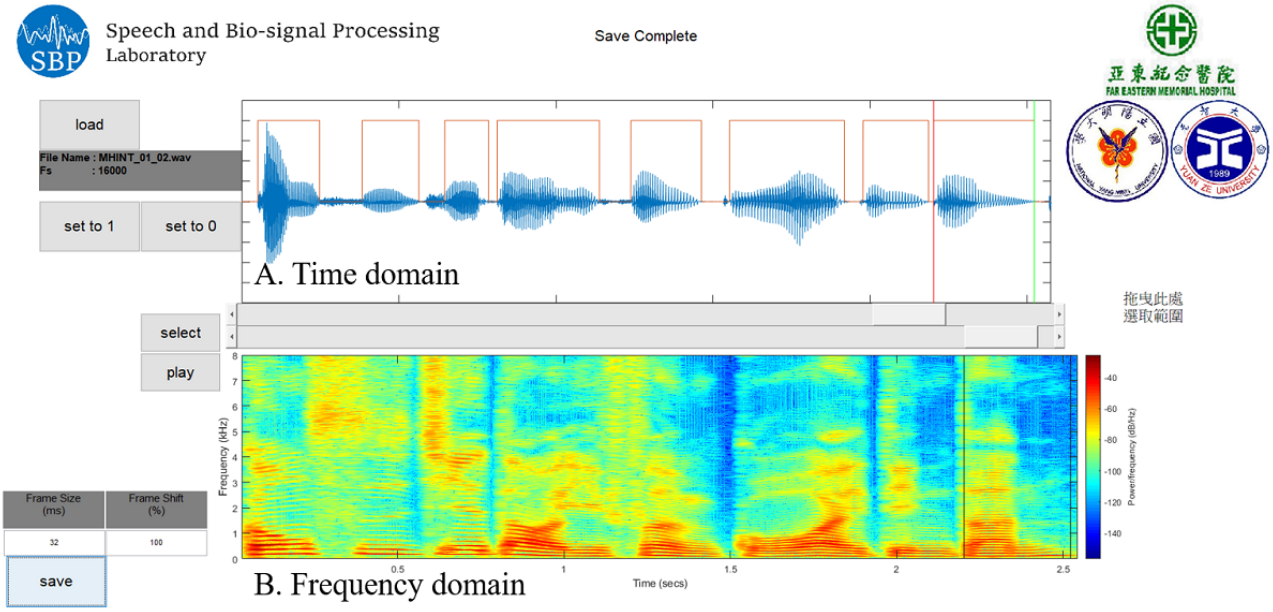
Detection Model

The signals (S_j) recorded from the wireless microphone were converted to envelope ($\hat{E}_j$) by an "envelope detection" unit; the power energy can be used in this unit. Then, the "detection unit" predicted whether the input frames were speech or nonspeech by comparing the value of $\hat{E}_j$ with that of the AT. The AT can be calculated using Equation (1); it is based on the energy of the input frame and 3 consecutive preceding frames.

Here, a_i represents the coefficients of the energy envelope of the current frame (i=0) and 3 successive preceding frames (i=1 to 3). It should be noted that the 3 successive preceding frames are included in this equation according to the best performance observed in our pilot study. $\hat{E}_{j-i}$ represents the input acoustic energy features at the j-i frame index and b is the bias. When the value of $\hat{E}_j$ exceeded the threshold derived from the AT in Equation (1), the system generated "*1*" as the output, indicating that this frame was recognized as speech. However, if the value was lower than the threshold derived from the AT in Equation (1), the system generated "*0*" (nonspeech) as the output.

The 5 coefficients required to calculate the AT were defined by the following two steps: (1) manually labeling speech segments of the recoded audio files and (2) using a GA to search for these 5 coefficients. In the first step, we developed software (Figure 2) to manually label the speech segments according to their time and frequency domains. We applied the GA [24] to search for these 5 coefficients for the AT function based on 25 randomly selected utterances of 10 seconds each (details are provided in Multimedia Appendix 1). After acquiring the coefficients required for the AT function, another 5 random utterances were used to test the accuracy of the automatic speech detection system. Similarly, the manual labeling of the speech segments was considered as the ground truth. The overall accuracy of each subject was calculated based on each frame (ie, by dividing the predicted number of speech frames by the total number of speech frames labeled manually).

Figure 2. Overview of user interfaces in the proposed labeling tool. The selected speech segments are displayed as red brackets.



Noise Reduction Model

Because a wireless microphone is an air-conducted device that is susceptible to background noise, we performed additional experiments to examine the performance of the noise reduction model. We mimicked the presence of background noise by mixing the recorded speech signal (S_i) with 4 different common background noises (crowd cheering noise, sharp speech noise, street noise, and white noise, shown in Multimedia Appendix 2) at 3 SNR levels (0, 3, and 5 dB), denoted by [redacted]). The noisy signals ([redacted]) were then processed by the logMMSE algorithm to obtain enhanced signals ([redacted]). Next, the [redacted] were sent into the detection model for energy and speech detection. We evaluated the performance of the noise reduction model by comparing the accuracy of speech detection under simulated noisy conditions with or without the noise reduction function.

Measuring Phonation Ratio and Duration of Speech Segments

To examine the applicability of the automatic speech detection system, we calculated the phonation ratio of the participants as shown in Equation (2), which is a common approach used to analyze phonation habits and usage [19,25].

[redacted]

In addition, we calculated the duration and distribution of the phonation and nonphonation segments in a similar manner as in previous literature [22].

Results

Figure 3 displays the average recognition accuracy of the automatic speech detection system, which was 89.9% (frame-based) in the controlled environment, ranging from 81.0% (67,357/83,157) to 95.0% (199,201/209,685). Figures 4 and 5 illustrate the phonation ratio for the 10 teachers evaluated during the teaching session. On average, the phonation ratio ranged from 44.0% (33,019/75,044) to 78.0% (68,785/88,186). We also noted a drastic decrease in the phonation ratio in subject 5 at approximately 20 minutes (asterisk, Figure 4). After reviewing the recorded audio file, we observed that this teacher did not speak for a while because he left the podium to fetch chalk; this example further demonstrated the excellent sensitivity of the proposed automatic speech detection system in practical scenarios. Figures 6 and 7 illustrate the distribution of the speech and nonspeech segments in logarithmic scales. Analytical results showed that the durations of most of the speech and nonspeech segments were less than 10 seconds.

Figure 3. Average accuracy of the automatic speech detection system with respect to 10 teachers in a controlled environment.

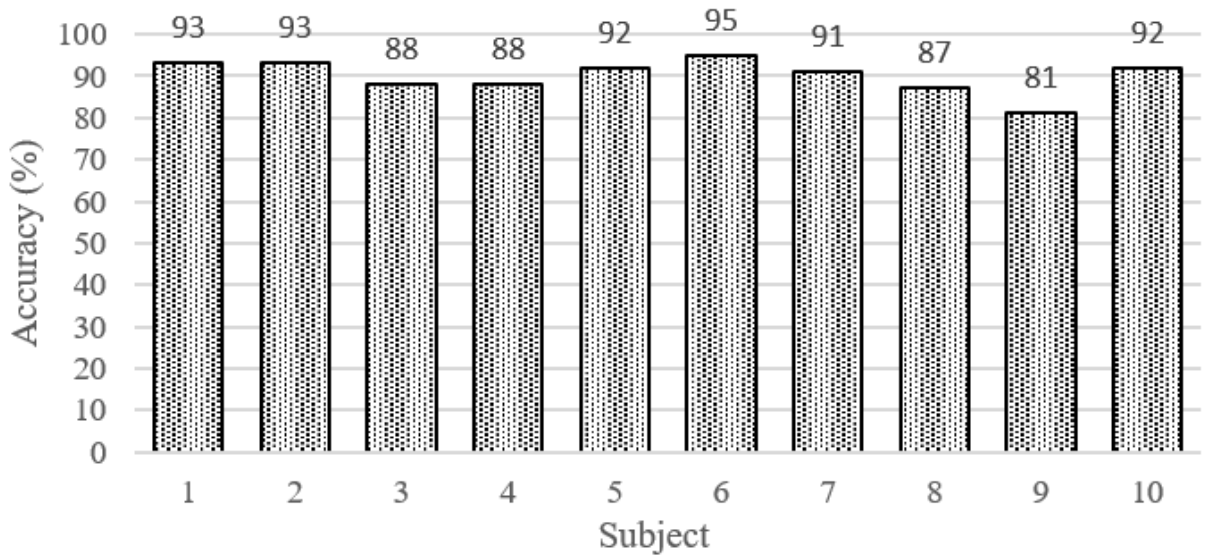


Figure 4. Phonation ratio over time for the 10 teachers (Subjects 1-6).

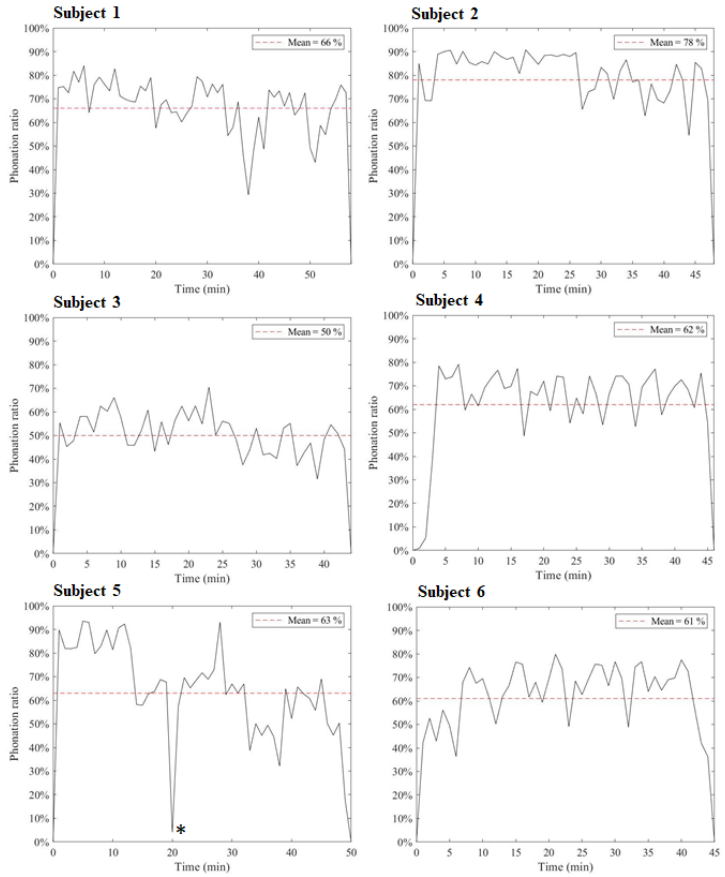


Figure 5. Phonation ratio over time for the 10 teachers (Subjects 7-10).

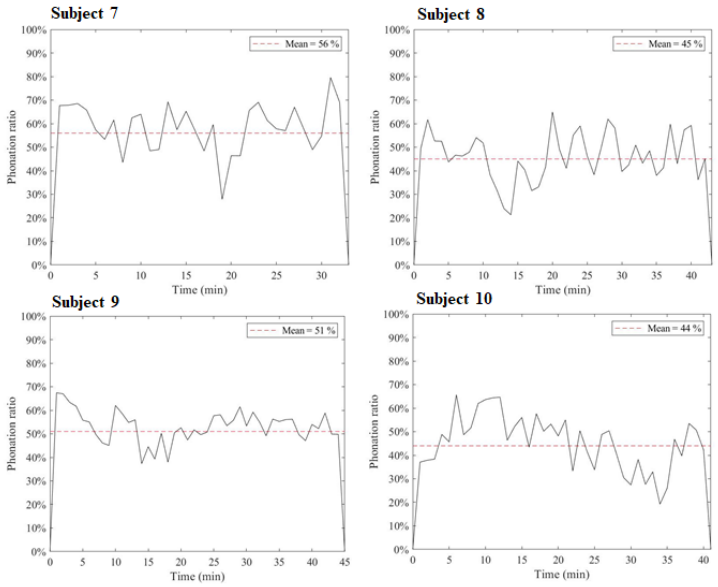


Figure 6. Speech and nonspeech segments measured by the automatic speech detection system. The x-axis indicates the length of the speech segments in a logarithmic scale, while the y-axis represents the occurrences during the recording period in a logarithmic scale (Subjects 1-6).

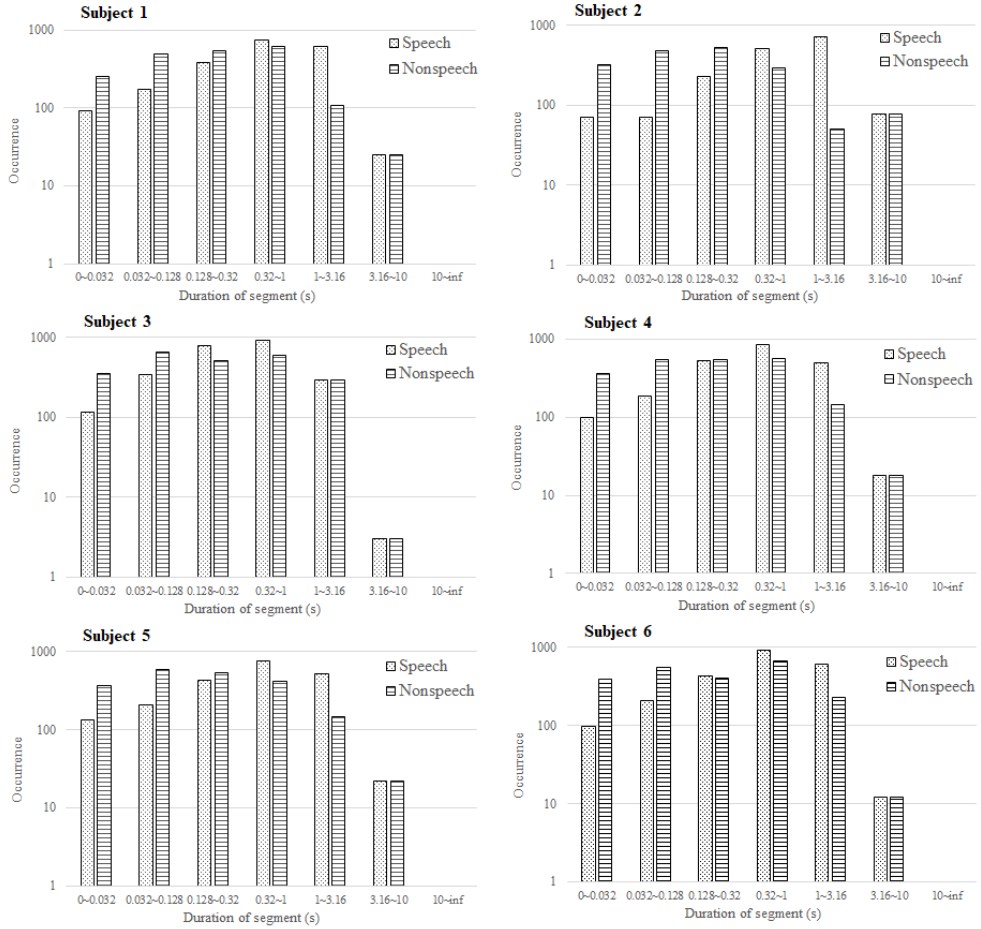
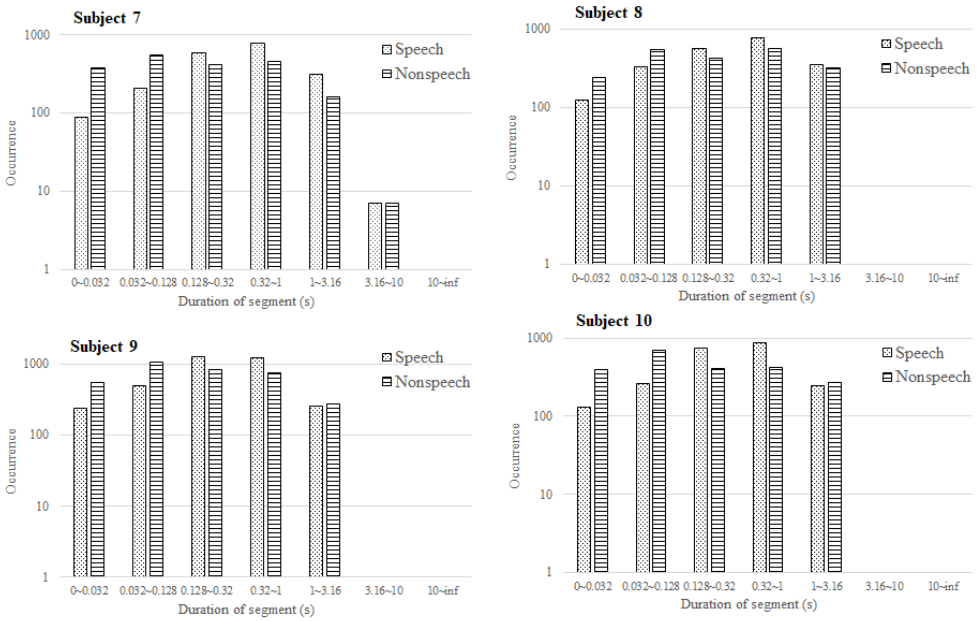


Figure 7. Speech and nonspeech segments measured by the automatic speech detection system. The x-axis indicates the length of the speech segments in a logarithmic scale, while the y-axis represents the occurrences during the recording period in a logarithmic scale (Subjects 7-10).



Figures 8 and 9 present a comparison of the same recordings under the controlled environment and simulated noisy conditions. We noticed that the detection accuracy dropped significantly under noisy conditions, indicating that the performance of the automatic speech detection system can be easily affected by the presence of noise without a noise reduction function. On average, the additional noise decreased the accuracy by approximately 33%, 36%, 34%, and 36% for 4 different types of noise: crowd cheering noise, sharp speech noise, street noise, and white noise, respectively.

Figure 10 presents an example of the relationship between the speech envelope and AT with and without the noise reduction function. Under the controlled environment, the AT was higher than the speech signal during nonspeech segments; in contrast, the energy envelope exceeded the AT in the presence of speech (Figure 10A). However, when the speech signal was contaminated by background noise, the overall energy exceeded the AT in both the speech and nonspeech segments (Figure 10B); thereby the proposed system may not be as effective in differentiating between speech and nonspeech signals. After enabling the logMMSE noise reduction function (Figure 10C), the AT could accurately detect the segments of speech versus nonspeech. On average, the additional noise reduction function yielded an average improvement of 4.9%, 27.5%, 19.3%, and 29.8% under the conditions of crowd cheering noise, sharp speech noise, street noise, and white noise, respectively. The detailed improvements are provided in Multimedia Appendix 3.

Figure 8. Accuracy of the automatic speech detection system with the presence of 2 different types of background noise (crowd cheer noise and speech sharp noise) at 3 SNR levels. The first bar of each graph indicates the accuracy of the original recording under the controlled environment. SNR: signal-to-noise ratio.

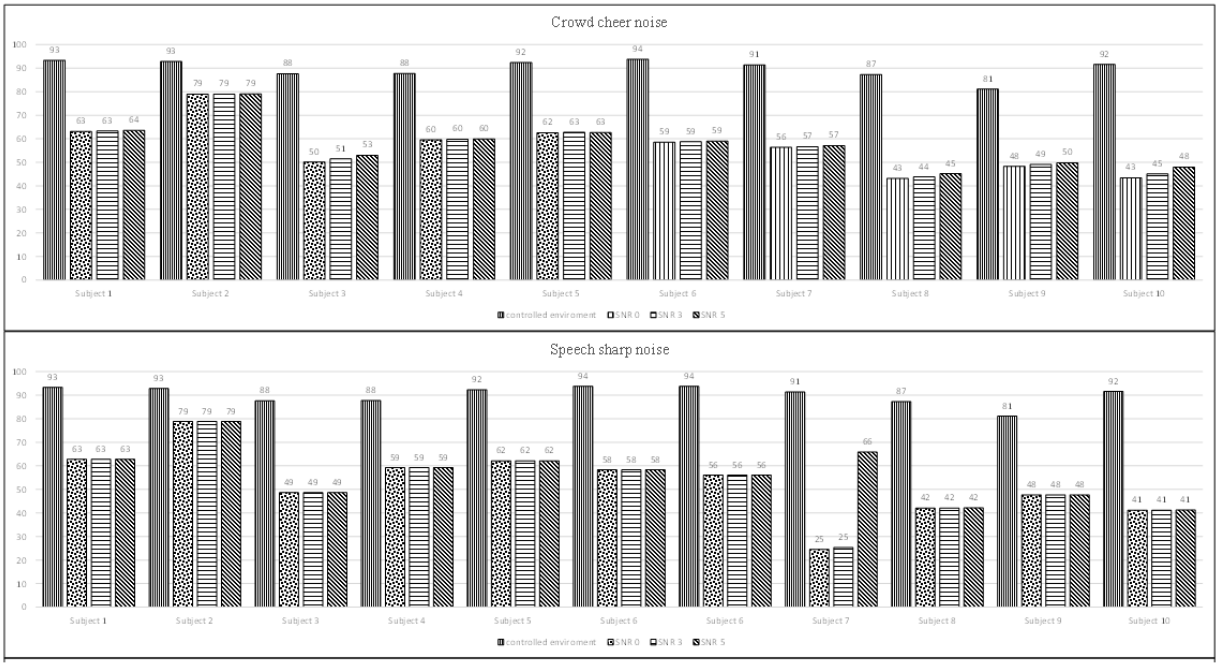


Figure 9. Accuracy of the automatic speech detection system with the presence of 2 different types of background noise (street noise and white noise) at 3 SNR levels. The first bar of each graph indicates the accuracy of the original recording under the controlled environment. SNR: signal-to-noise ratio.

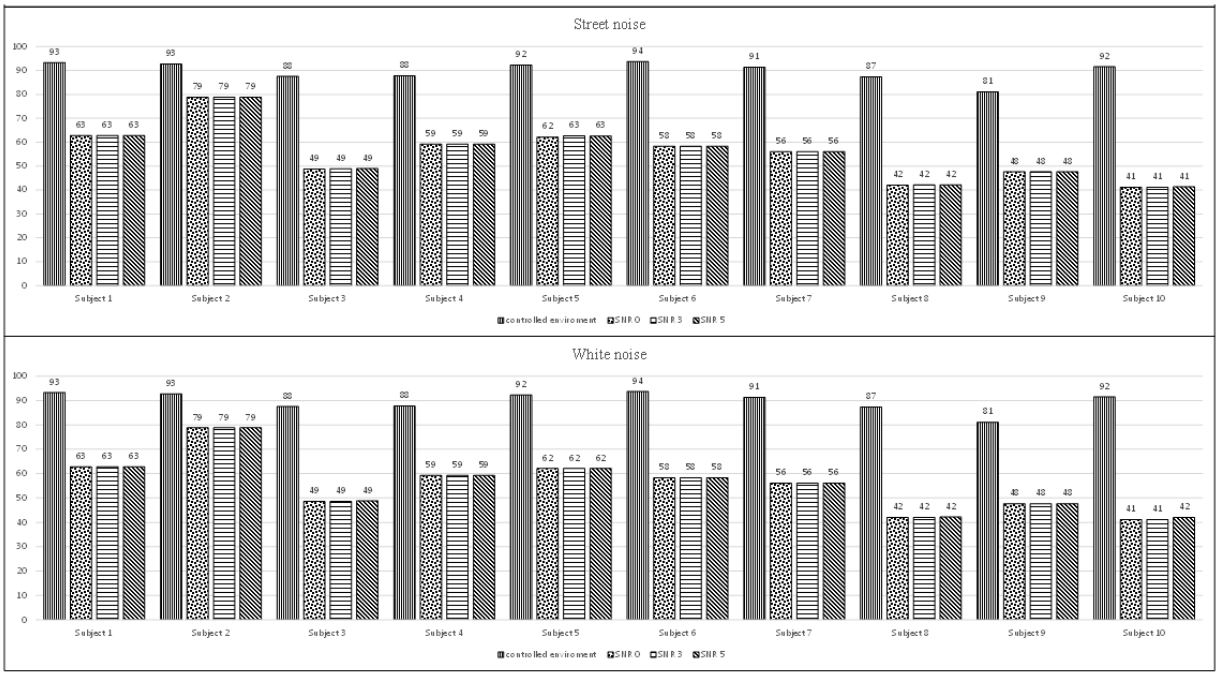
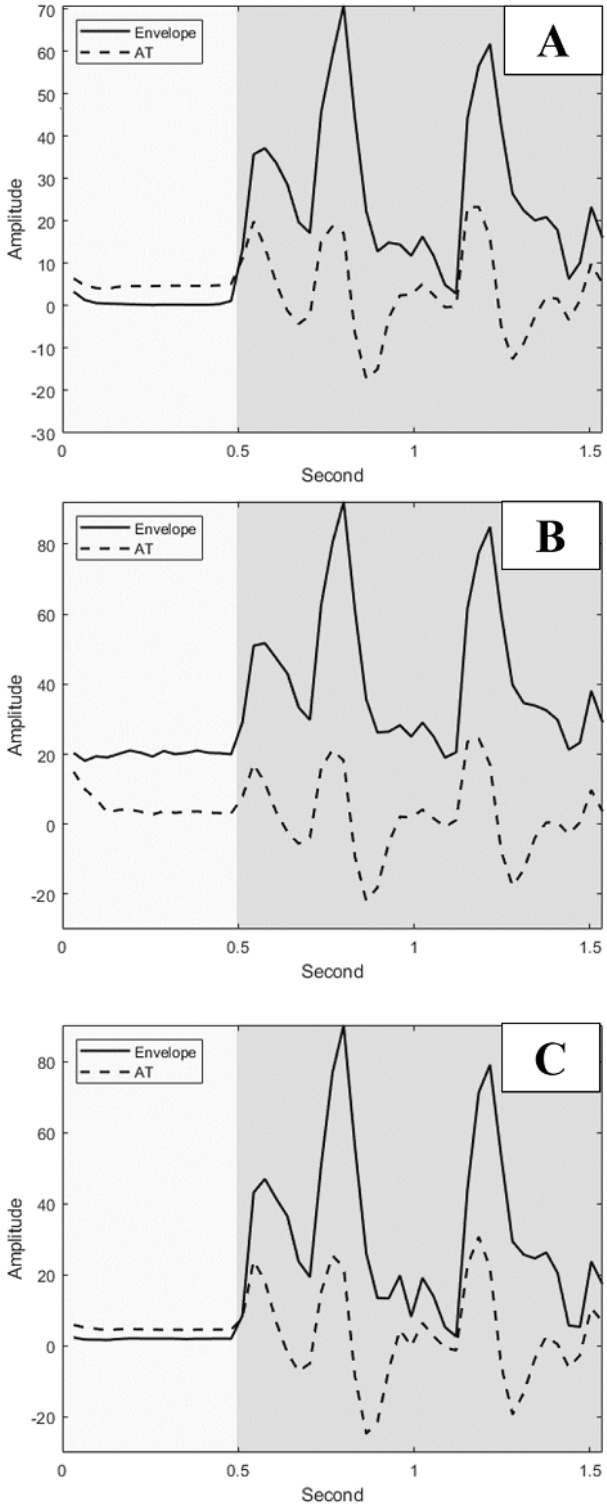


Figure 10. Example of the relationship between the speech envelope and AT in (A) the controlled environment, (B) noisy conditions without the noise reduction function, and (C) noisy conditions with the noise reduction function. Each part of the figure presents the same speech; the gray background denotes the speech segments, while the other areas indicate the nonspeech segments. AT: adaptive threshold.



Discussion

Principal Findings

In this study, we proposed an ambulatory phonation monitoring system with a wireless microphone. The results demonstrated that the proposed system can accurately differentiate between speech and nonspeech segments based on the energy envelope in a controlled environment. The implementation of an additional noise reduction function using a logMMSE algorithm can effectively reduce the impact of background noise. Preliminary results of the phonation ratio and the distribution of speech segments in 10 teachers were compatible with those in previous literature [18,19].

Applicability and Accuracy of Automatic Speech Detection System

Most studies in the existing literature used a neck accelerometer to detect the vibrations of vocal folds via skin [18,25]. Although contact microphones can effectively suppress the effects of background noise [16], they may not always be convenient for the users, owing to the cumbersome wiring and taping. In contrast, the wireless microphone used in this study eliminated the discomfort associated with the wiring and taping of contact microphones. All the participants reported good tolerance using wireless microphone, without any physical discomfort during the teaching session.

Previous studies [22] applied predefined criteria to detect voice activity, such as the fundamental frequency during normal speaking (ie, 70 to 1000 Hz), SPL greater than 30 dB, and a low/high ratio of at least 22 dB. In this study, we specifically designed software to manually label the speech segments (Figure 2), which served as the ground truth for examining the detection accuracy of this novel system. Figure 3 demonstrates an average detection accuracy of 89.9% in the controlled environment, which established the applicability and reliability of the proposed system.

In comparison to previous studies [18,25], our results demonstrated a higher phonation ratio (range: 44.0%-78.0%; Figure 3) owing to the continual lecturing of the teachers in the classroom. Similarly, the durations of most of the speech segments were less than 10 seconds. We did not observe long durations of silence (nonspeech) in this study (Figures 6 and 7). In contrast, previous studies recorded the phonation ratio throughout the day (except sleeping) [18,25]; thus, longer silence periods were more likely to be documented.

Benefits of Noise Reduction Function

Because wireless microphones are more susceptible to background noise, we examined the effectiveness of the additional noise reduction function by mixing 4 different types of background noise to simulate noisy conditions. Our results showed that the noise reduction function using the logMMSE algorithm can improve the detection accuracy by up to 45.8% (maximum) in stable noise conditions (eg, sharp speech noise and white noise) (Multimedia Appendix 3); however, logMMSE works less efficiently in competing voice signals (eg, crowd cheering noise), resulting in an improvement of approximately 5%, similar to previous literature [26]. Accordingly, other noise reduction approaches, such as deep learning [27], may be more robust for enhancing the automatic speech detection system in the future. Additionally, automatic gain control [28] can also be integrated into the system to normalize the input volume and improve the accuracy in cases where sudden changes are observed in the input volume.

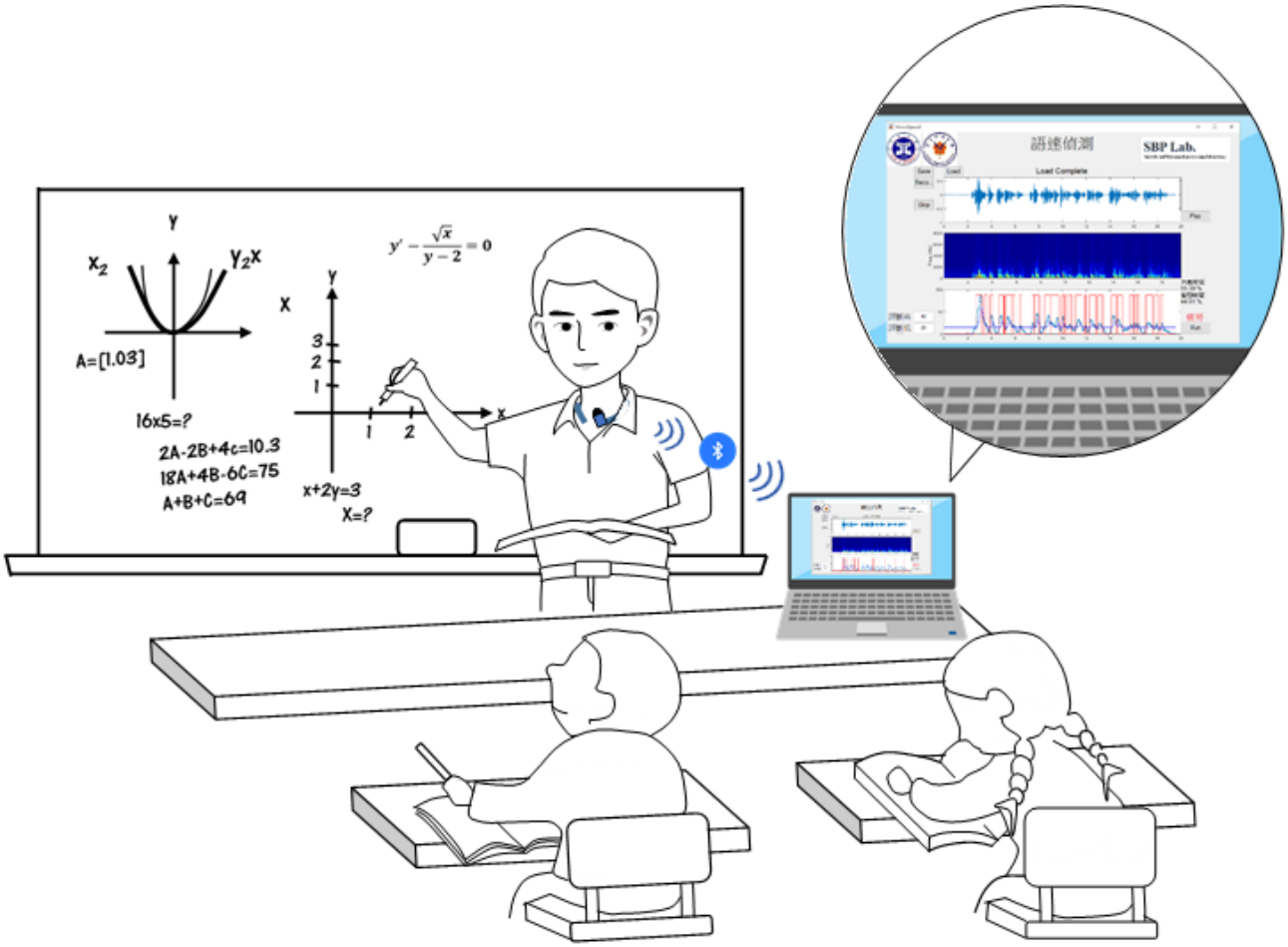
Speaker Identification

The proposed system yielded comparable accuracy in most of the test conditions and an additional noise reduction function further improved the performance of the proposed system in noisy conditions. However, there is still room for improvement in some challenging conditions (eg, video sound or sudden increase in volume). We observed that subject 3 played a video clip with speech context during the class, and the loud speech from the video was misidentified as the speech of subject 3. For subject 4, several conversations took place between the teacher and students, which also caused the voice signals from the students to be misidentified as the speech of subject 4 and reduced the accuracy. One way to alleviate this inherent limitation of wireless microphones (ie, susceptibility to noise and competitive speakers) is the use of a microphone array with a beamforming algorithm that can fix (or adapt to adjust) the recorded position to distinguish between the speech of the speaker and background noise or other speakers. Another option to improve our system is implementing the speaker identification algorithm [29]; however, it requires significantly higher computing power to handle complex features (such as i-vector or x-vector [30,31]) using deep learning–based technology.

Future Perspective

The study results suggest that the proposed automatic speech detection system with wireless microphone can be applied in practical scenarios to overcome the limitations of contact microphone for ambulatory phonation monitoring. The proposed system can be further implemented on personal laptops (or mobile phone devices) for daily use and timely feedback, as illustrated in Figure 11. By monitoring the baseline phonation ratio, doctors and speech language pathologists can prescribe a certain threshold of phonation ratio based on individual conditions. Upon exceeding this limit, an alarm signal (flash or sound) could be sent to the user to ensure that they take enough breaks; promising results are available with respect to this concept [32] but it requires further evidential support from ongoing studies.

Figure 11. Exemplary use of the automatic speech detection system by a teacher.



Although the proposed automatic speech detection system achieved 89.9% accuracy in this study for the proposed ambulatory phonation monitoring, it still has room for improvement. More recently, deep learning–based automatic speech recognition (ASR) [33] and natural language processing (NLP) [34,35] systems were proven to achieve higher speech recognition efficiency for conventional communication between human-machine applications (eg, Amazon Alexa, Google Home, and Apple Siri). These deep learning–based ASR and NLP systems could be applied in ambulatory phonation monitoring; however, some critical issues need to be addressed. For example, ASR and NLP technologies might violate the user's privacy because they recognize the context of the user's speech. In contrast, the automatic speech detection system of this study is energy-based; it will not directly access the content of speech and might be more acceptable to the users. In addition, ASR and NLP technologies require high computing power, especially when a deeper structure of the neural network is implanted to achieve higher speech recognition accuracy. A cloud-based ASR and NLP system could be effective in alleviating this limitation; however, the recorded speech data still needs to be uploaded to the server, which may lead to additional privacy and security issues. More recently, phonetic posteriorgram features obtained from the acoustic model of the ASR system was introduced for speech processing applications, and it has proven to achieve benefits in many tasks [36-38]. Following the success of phonetic posteriorgram, our future study could apply its features and deep learning technology to improve the performance of the current model.

Furthermore, this system can also be extended for detecting speech and communication disorders [39] (eg, Parkinson disease [40] and depression [41]). However, such work may require more sophisticated features of voice signals and computation techniques, such as the combination of the Mel frequency cepstral coefficients and deep neural networks, which was used in a previous study [42]. With the significant advancements in smartphones and smart home devices, the proposed automatic speech detection system can potentially be implemented in these devices to further decrease the clumsiness of any additional devices [43].

Limitations

The first limitation of this study is the small number of participants (N=10). A larger cohort is required to obtain more robust evidence for the clinical use of automatic speech detection for ambulatory phonation monitoring. In addition, only teachers were recruited owing to the approved IRB protocol. Other occupations with high vocal demands (eg, salespeople and customer service representatives) will be included in the future to expand the potential use of the proposed system. Second, the proposed automatic speech detection system cannot precisely identify the speech of the speaker in the presence of loud competing background noise or other speakers. To overcome this issue, algorithms that require higher computing power, such as speaker identification or microphone array algorithms, could

be used in future studies. Lastly, this automatic speech detection system requires manually labeling the recorded speech for model training. Considering the high accuracy achieved in this study, future research does not need to record the original voice content, so the confidentiality of the participants can be better protected.

Conclusions

This study proposed an automatic speech detection system comprising a wireless microphone to receive the acoustic signals and an adaptive threshold for speech detection based on the energy envelope. The proposed system demonstrated a speech detection accuracy of 89.9%, and the analytical results for the phonation ratio and speech segments were comparable to those of previous research. Moreover, the use of an unsupervised noise reduction function (logMMSE) can improve the robustness of the proposed system in noisy conditions. These results imply that the proposed system can be a potential tool for ambulatory voice monitoring in occupational voice users.

Acknowledgments

This study was supported by the Ministry of Science and Technology of Taiwan under the 108-2314-B-418-013 and 109-2218-E-010-004 projects and Far Eastern Memorial Hospital-National Yang-Ming University Joint Research (FEMH-NYMU108DN34).

Conflicts of Interest

The authors CTW and YHL are the inventors of “Real-time monitor system of phonation,” (Taiwan Patent No. TW I626647).

Multimedia Appendix 1
Using genetic algorithm to determine the parameters of the adaptive threshold.
[DOCX File , 395 KB - mhealth_v8i12e16746_app1.docx]

Multimedia Appendix 2
Spectrogram of the background noises used in this study.
[DOCX File , 1301 KB - mhealth_v8i12e16746_app2.docx]

Multimedia Appendix 3
Mean improvements in recognition accuracies when using the noise reduction function for the proposed auto speech detection system under simulated noisy conditions.
[DOCX File , 26 KB - mhealth_v8i12e16746_app3.docx]

References

1. Titze IR, Martin DW. Principles of voice production. Saddle River City, NJ: Prentice Hall; 1994:169-190.
2. Roy N, Merrill RM, Thibeault S, Parsa RA, Gray SD, Smith EM. Prevalence of voice disorders in teachers and the general population. J Speech Lang Hear Res 2004 Apr;47(2):281-293. [doi: 10.1044/1092-4388(2004/023)] [Medline: 15157130]
3. Van Houtte E, Claeys S, Wuyts F, Van Lierde K. The impact of voice disorders among teachers: vocal complaints, treatment-seeking behavior, knowledge of vocal care, and voice-related absenteeism. J Voice 2011 Sep;25(5):570-575. [doi: 10.1016/j.jvoice.2010.04.008] [Medline: 20634042]
4. Cohen SM, Kim J, Roy N, Asche C, Courey M. The impact of laryngeal disorders on work-related dysfunction. Laryngoscope 2012 Jul;122(7):1589-1594. [doi: 10.1002/lary.23197] [Medline: 22549455]
5. Verdolini K, Ramig LO. Review: occupational risks for voice problems. Logoped Phoniatr Vocol 2001;26(1):37-46. [Medline: 11432413]
6. 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]
7. Williams NR. Occupational groups at risk of voice disorders: a review of the literature. Occup Med (Lond) 2003 Oct;53(7):456-460. [doi: 10.1093/occmed/kqg113] [Medline: 14581643]
8. Simberg S, Laine A, Sala E, Rönnemaa AM. Prevalence of voice disorders among future teachers. J Voice 2000 Jun;14(2):231-235. [doi: 10.1016/s0892-1997(00)80030-2] [Medline: 10875574]
9. Chen SH, Chiang SC, Chung YM, Hsiao LC, Hsiao TY. Risk factors and effects of voice problems for teachers. J Voice 2010 Mar;24(2):183-90, quiz 191. [doi: 10.1016/j.jvoice.2008.07.008] [Medline: 19481416]
10. Bermúdez de Alvear RM, Barón FJ, Martínez-Arquero AG. School teachers' vocal use, risk factors, and voice disorder prevalence: guidelines to detect teachers with current voice problems. Folia Phoniatr Logop 2011;63(4):209-215. [doi: 10.1159/000316310] [Medline: 20938203]
11. Smith E, Lemke J, Taylor M, Kirchner HL, Hoffman H. Frequency of voice problems among teachers and other occupations. J Voice 1998 Dec;12(4):480-488. [doi: 10.1016/s0892-1997(98)80057-x] [Medline: 9988035]

12. Stemple JC, Roy N, Klaben BK. Clinical voice pathology: Theory and management. San Diego, CA: Plural Publishing; 2018.
13. Chen SH, Hsiao T, Hsiao L, Chung Y, Chiang S. Outcome of resonant voice therapy for female teachers with voice disorders: perceptual, physiological, acoustic, aerodynamic, and functional measurements. J Voice 2007 Jul;21(4):415-425. [doi: 10.1016/j.jvoice.2006.02.001] [Medline: 16581227]
14. Ryu S, Komiyama S, Kannae S, Watanabe H. A newly devised speech accumulator. ORL J Otorhinolaryngol Relat Spec 1983;45(2):108-114. [doi: 10.1159/000275632] [Medline: 6341919]
15. Cheyne HA, Hanson HM, Genereux RP, Stevens KN, Hillman RE. Development and testing of a portable vocal accumulator. J Speech Lang Hear Res 2003 Dec;46(6):1457-1467. [doi: 10.1044/1092-4388(2003/113)] [Medline: 14700368]
16. Popolo PS, Svec JG, Titze IR. Adaptation of a Pocket PC for use as a wearable voice dosimeter. J Speech Lang Hear Res 2005 Aug;48(4):780-791. [doi: 10.1044/1092-4388(2005/054)] [Medline: 16378473]
17. Carroll T, Nix J, Hunter E, Emerich K, Titze I, Abaza M. Objective measurement of vocal fatigue in classical singers: a vocal dosimetry pilot study. Otolaryngol Head Neck Surg 2006 Oct;135(4):595-602 [FREE Full text] [doi: 10.1016/j.otohns.2006.06.1268] [Medline: 17011424]
18. Titze I, Hunter E, Svec JG. Voicing and silence periods in daily and weekly vocalizations of teachers. J Acoust Soc Am 2007 Jan;121(1):469-478 [FREE Full text] [doi: 10.1121/1.2390676] [Medline: 17297801]
19. Mehta DD, Zañartu M, Feng SW, Cheyne HA, Hillman RE. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. IEEE Trans Biomed Eng 2012 Nov;59(11):3090-3096 [FREE Full text] [doi: 10.1109/TBME.2012.2207896] [Medline: 22875236]
20. Remacle A, Morsomme D, Finck C. Comparison of vocal loading parameters in kindergarten and elementary school teachers. J Speech Lang Hear Res 2014 Apr 01;57(2):406-415. [doi: 10.1044/2013_JSLHR-S-12-0351] [Medline: 24129011]
21. Searl J, Dietsch AM. Tolerance of the VocaLog™ Vocal Monitor by Healthy Persons and Individuals With Parkinson Disease. J Voice 2015 Jul;29(4):518.e13-518.e20. [doi: 10.1016/j.jvoice.2014.09.011] [Medline: 25726068]
22. Van Stan JH, Mehta DD, Sternad D, Petit R, Hillman RE. Ambulatory Voice Biofeedback: Relative Frequency and Summary Feedback Effects on Performance and Retention of Reduced Vocal Intensity in the Daily Lives of Participants With Normal Voices. J Speech Lang Hear Res 2017 Apr 14;60(4):853-864 [FREE Full text] [doi: 10.1044/2016_JSLHR-S-16-0164] [Medline: 28329366]
23. Scalart P, Filho JV. Speech enhancement based on a priori signal to noise estimation. In: IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings. 1996 Presented at: IEEE International Conference on Acoustics, Speech, and Signal Processing; 1996; Atlanta, GA p. 629-632. [doi: 10.1109/icassp.1996.543199]
24. John H H. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. In: Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. Boston: A Bradford Book; 1992.
25. Van Stan JH, Mehta DD, Zeitels SM, Burns JA, Barbu AM, Hillman RE. Average Ambulatory Measures of Sound Pressure Level, Fundamental Frequency, and Vocal Dose Do Not Differ Between Adult Females With Phonotraumatic Lesions and Matched Control Subjects. Ann Otol Rhinol Laryngol 2015 Nov;124(11):864-874 [FREE Full text] [doi: 10.1177/0003489415589363] [Medline: 26024911]
26. Wang SS, Tsao Y, Wang HLS, Lai YH, Li LPH. A deep learning-based noise reduction approach to improve speech intelligibility for cochlear implant recipients in the presence of competing speech noise. In: Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). 2017 Presented at: Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC); 2017; Kuala Lumpur p. 808-812. [doi: 10.1109/apsipa.2017.8282144]
27. Xu Y, Du J, Huang Z, Dai LR, Lee CH. Multi-objective learning and mask-based post-processing for deep neural network-based speech enhancement. arXiv. 2017. URL: https://arxiv.org/abs/1703.07172 [accessed 2020-10-14]
28. Shan T, Kailath T. Adaptive algorithms with an automatic gain control feature. IEEE Trans. Circuits Syst 1988 Jan;35(1):122-127. [doi: 10.1109/31.1709]
29. Reynolds D, Rose R. Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans. Speech Audio Process 1995 Jan;3(1):72-83. [doi: 10.1109/89.365379]
30. Kanagasundaram A, Vogt R, Dean DB, Sridharan S, Mason MW. I-vector based speaker recognition on short utterances. 2011 Presented at: 12th Annual Conference of the International Speech Communication Association (ISCA); 2011; Florence, Italy p. 2341-2344 URL: https://eprints.qut.edu.au/46313/1/IS110023.PDF
31. David S, Garcia-Romero D, Sell G, Povey D, Khudanpur S. X-Vectors: Robust DNN Embeddings for Speaker Recognition. NY, US: IEEE; 2018 Presented at: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 15-20 April 2018; Calgary, AB, Canada. [doi: 10.1109/ICASSP.2018.8461375]
32. Van Stan JH, Mehta DD, Petit RJ, Sternad D, Muise J, Burns JA, et al. Integration of Motor Learning Principles Into Real-Time Ambulatory Voice Biofeedback and Example Implementation Via a Clinical Case Study With Vocal Fold Nodules. Am J Speech Lang Pathol 2017 Feb 01;26(1):1-10 [FREE Full text] [doi: 10.1044/2016_AJSLP-15-0187] [Medline: 28124070]

33. Zhang W, Cui X, Finkler U, Kingsbury B, Saon G, Kung D. Distributed deep learning strategies for automatic speech recognition. 2019 Presented at: ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2019; Brighton, United Kingdom p. 5706-5710. [doi: 10.1109/icassp.2019.8682888]
34. Kaufman DR, Sheehan B, Stetson P, Bhatt AR, Field AI, Patel C, et al. Natural Language Processing-Enabled and Conventional Data Capture Methods for Input to Electronic Health Records: A Comparative Usability Study. JMIR Med Inform 2016 Oct 28;4(4):e35 [FREE Full text] [doi: 10.2196/medinform.5544] [Medline: 27793791]
35. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review. JMIR Med Inform 2019 Apr 27;7(2):e12239 [FREE Full text] [doi: 10.2196/12239] [Medline: 31066697]
36. Tian X, Chng ES, Li H. A Speaker-Dependent WaveNet for Voice Conversion with Non-Parallel Dat. 2019 Presented at: Interspeech 2019; 2019; Graz, Austria p. 201-205. [doi: 10.21437/interspeech.2019-1514]
37. Sun L, Li K, Wang H, Kang S, Meng H. Phonetic posteriorgrams for many-to-one voice conversion without parallel data training. NY, US: IEEE; 2016 Presented at: IEEE International Conference on Multimedia and Expo (ICME); 2016; Seattle, WA. [doi: 10.1109/ICME.2016.7552917]
38. Chen CY, Zheng WZ, Wang SS, Tsao Y, Li PC, Li YH. Enhancing Intelligibility of Dysarthric Speech Using Gated Convolutional-based Voice Conversion System. In: IEEE Interspeech. 2020 Presented at: IEEE Interspeech; 2020; Shanghai.
39. Furlong LM, Morris ME, Erickson S, Serry TA. Quality of Mobile Phone and Tablet Mobile Apps for Speech Sound Disorders: Protocol for an Evidence-Based Appraisal. JMIR Res Protoc 2016 Nov 29;5(4):e233 [FREE Full text] [doi: 10.2196/resprot.6505] [Medline: 27899341]
40. Erdogdu Sakar B, Serbes G, Sakar CO. Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. PLoS One 2017;12(8):e0182428 [FREE Full text] [doi: 10.1371/journal.pone.0182428] [Medline: 28792979]
41. Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R, et al. Major depressive disorder discrimination using vocal acoustic features. J Affect Disord 2018 Jan 01;225:214-220. [doi: 10.1016/j.jad.2017.08.038] [Medline: 28841483]
42. Fang S, Tsao Y, Hsiao M, Chen J, Lai Y, Lin F, et al. Detection of Pathological Voice Using Cepstrum Vectors: A Deep Learning Approach. J Voice 2019 Sep;33(5):634-641. [doi: 10.1016/j.jvoice.2018.02.003] [Medline: 29567049]
43. Chung AE, Griffin AC, Selezneva D, Gotz D. Health and Fitness Apps for Hands-Free Voice-Activated Assistants: Content Analysis. JMIR Mhealth Uhealth 2018 Sep 24;6(9):e174 [FREE Full text] [doi: 10.2196/mhealth.9705] [Medline: 30249581]

Abbreviations

ASR: automatic speech recognition
AT: adaptive threshold
GA: genetic algorithm
logMMSE: log minimum mean square error
NLP: natural language processing
SPL: sound pressure level

Edited by G Eysenbach; submitted 23.10.19; peer-reviewed by B Johnson, C He; comments to author 20.01.20; revised version received 03.06.20; accepted 03.10.20; published 03.12.20.

Please cite as:
Wang CT, Han JY, Fang SH, Lai YH
Ambulatory Phonation Monitoring With Wireless Microphones Based on the Speech Energy Envelope: Algorithm Development and Validation
JMIR Mhealth Uhealth 2020;8(12):e16746
URL: https://mhealth.jmir.org/2020/12/e16746
doi:10.2196/16746
PMID:33270033

Original Paper

COVID-19–Related Disruptions and Increased mHealth Emergency Use Intention: Experience Sampling Method Study

Zhenduo Zhang[1*], PhD; Li Zhang[1*], PhD; Junwei Zheng[2*], PhD; Huan Xiao[1], PhD; Zhigang Li[3], PhD

[1]School of Management, Harbin Institute of Technology, Harbin, China

[2]Department of Construction Management, Kunming University of Science and Technology, Kunming, China

[3]School of Economics and Management, Beijing Polytechnic, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Zhigang Li, PhD
School of Economics and Management
Beijing Polytechnic
No 9, Liangshuihe 1st Street
Beijing Economic and Technological Development Zone
Beijing
China
Phone: 86 13651205747
Email: lizhigang@bpi.edu.cn

Abstract

Background: The COVID-19 pandemic has become a global public health event, which has raised concerns regarding individuals' health. Individuals need to cope with the COVID-19 pandemic with guidelines on symptom recognition, home isolation, and maintain mental health. Besides routine use of mobile health (mHealth) such as accessing information to keep healthy, individuals can use mHealth services in situations requiring urgent medical care, which is defined as mHealth emergency use. It is not known whether individuals have increased their daily mHealth services emergency use as a result of disruptions caused by the COVID-19 pandemic.

Objective: The purpose of this diary analysis study is to assess the influences of daily disruptions related to the COVID-19 pandemic on individuals' mHealth emergency use. The secondary purpose of this study is to explore the mediating role of COVID-19–induced strain and the moderating role of promotion regulatory focus in the relationship between daily disruptions of COVID-19 and mHealth emergency use. Drawing from the cognitive activation theory of stress, we investigated the underlying mechanism and boundary condition of the influence of COVID-19–related disruptions on daily mHealth emergency use.

Methods: To test the proposed model, this study adopts the experience sampling method to collect daily data. The experience sampling method helps researchers to capture participants' fluctuations in emotions, mental engagement in an activity, and experienced stress. This study collected 550 cases nested in 110 samples in mainland China to test the conceptual model. In addition, we employed hierarchical linear modeling analysis to test the effect of COVID-19–related disruptions on mHealth emergency use.

Results: We found that COVID-19–related disruptions increased COVID-19–induced strain (γ=0.24, P<.001) and mHealth emergency use on a daily basis (γ=0.28, P<.001). COVID-19–induced daily strain mediated this relationship (effect=0.09, 95% CI 0.05-0.14). Promotion regulatory focus moderated the relationship between COVID-19–induced strain and mHealth emergency use (γ=0.35, P=.02). In addition, the indirect relationship between disruptions and mHealth emergency use intentions through COVID-19–induced strain is contingent upon promotion regulatory focus: this relationship was stronger in those with high promotion regulatory focus (effect=0.12, 95% CI 0.06-0.19) than in those with low promotion regulatory focus (effect=0.06, 95% CI 0.02-0.11).

Conclusions: Event disruption of the COVID-19 pandemic induced mHealth emergency use intention through increased psychological strain. Furthermore, individuals' promotion regulatory focus amplified this indirect relationship. Our findings extend our understanding of the factors underlying mHealth emergency use intention and illustrate the potential contingent role of promotion regulatory focus in the cognitive activation theory of stress. This study also opens avenues for future research on mHealth emergency use intention in other countries and cultural settings.

(JMIR Mhealth Uhealth 2020;8(12):e20642) doi:10.2196/20642

KEYWORDS

mobile health services; emergency use intention; event disruption; COVID-19–induced strain; promotion regulatory focus; mHealth; COVID-19

Introduction

Background

The COVID-19 pandemic has become a worldwide public health event. This has resulted in greater concerns regarding one's health and well-being [1]. Similar to the research of Morgeson et al [2], this study considers event disruption as a negative influence on behaviors of health information system use. Event disruption is defined as a discontinuity in the environment where the external situation has somehow changed [3]. The COVID-19 pandemic has likely transformed people's routines and, thus, requires an additional investment of resources to adapt effectively to cope with increased health concerns [4].

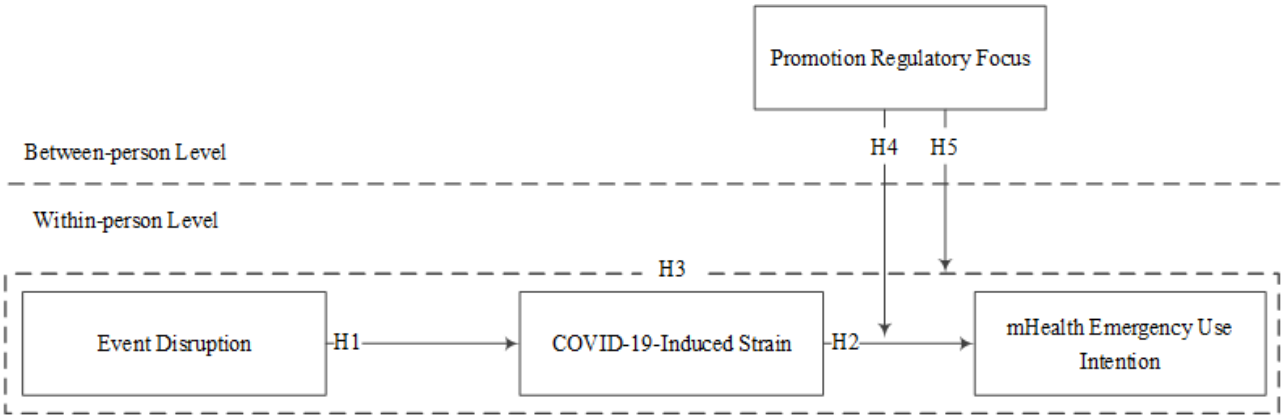
Mobile health (mHealth) service is defined as health care practice supported by mobile devices. Given that our research focuses on the mHealth service in the COVID-19 pandemic, the mHealth service in this study includes apps that health care professionals use to treat clinical disease, reinforce treatment adherence, provide consultation to the users, and educate users on self-monitoring of the disease COVID-19 [5]. mHealth service is an essential component of health information technology, which has the potential to enhance health care quality and promote good health [6-8]. Besides routine use, individuals use mHealth services in situations requiring urgent medical care, which is defined as mHealth emergency use [8]. The aims of mHealth emergency use are to help users to acquire appropriate care in urgent situations and improve the efficiency of treatment toward specific disease [9]. mHealth services can facilitate the fast delivery of health care information to the users, assisting users to make medical decisions in emergencies [8]. The outbreak of the COVID-19 pandemic is a serious public health event and threatens everyone's health. However, it is not yet known whether the COVID-19 pandemic influences mHealth emergency use. Furthermore, the pandemic situation changes on a daily basis, resulting in corresponding daily changes in disruptions and their effects. Therefore, our first research question asks whether event disruption of the COVID-19 pandemic increases daily mHealth emergency use.

The cognitive activation theory of stress (CATS) [10,11] addresses the role of cognitive appraisal and interpretation in shaping the way an individual responds to stressful events [12]. When the event is regarded as threatening and challenging, the individual may experience strain, which is defined as an unpleasant subjective experience toward a specific event [13]. Considering the disruptions caused by the COVID-19 pandemic, the strain will likely be induced and consequently impact an individual's attitudes and preferred coping strategies [14]. Our second research question is whether psychological strain mediates the relationship between event disruption of the COVID-19 pandemic and mHealth emergency use on a daily basis.

Whether an individual decides to use mHealth to cope with COVID-19–induced strain is contingent upon their preferred method to deal with problems [15]. Prior research has demonstrated the critical role of promotion regulatory focus in facilitating individuals' adoption of new technologies [16,17]. Promotion regulatory focus denotes a sensitivity to gains and a desire for advancement and growth [18]. Individuals with strong promotion regulatory focus have greater intentions to use mHealth to help cope with COVID-19–induced strain. The third research question of this study is to investigate whether the indirect relationship between event disruption and mHealth emergency use intention through COVID-19–induced strain is contingent upon individuals' promotion regulatory focus.

To address our three research questions, we used the experience sampling method to test the conceptual model (see Figure 1). Our study has three potential contributions to mHealth literature. First, this study examines the temporal relationship between the event disruption of the COVID-19 pandemic and mHealth emergency use intention. Second, this study explores the underlying mechanism through which the event disruption impacts mHealth emergency use intention by examining the mediating role of COVID-19–induced strain. Third, this study depicts the boundary condition under which event disruption is more or less influential on mHealth emergency use intention through COVID-19–induced strain by exploring the moderating role of promotion regulatory focus.

Figure 1. Conceptual model. H: hypothesis; mHealth: mobile health.



Literature Review of CATS

CATS proposes that stress occurs with a discrepancy between desired outcomes and reality [19]. Individuals will feel stress when the future is unpredictable or their resources are not sufficient to cope with the demands [11], whereas if individuals have control over the situations to achieve favorable outcomes, the stress will not occur [20]. Cognitive appraisal is the central role of cognitive activation [11]. Primary and secondary appraisals are two sequential stages in the cognitive appraisal process [21]. In the primary appraisal stage, individuals will evaluate the target events as harmful or beneficial. If such events are regarded as harmful, individuals will then evaluate the extent that they can overcome such harmful events in the secondary appraisal [22]. When they cannot handle the stressful events, the strain would arise. Individuals are motivated to take the necessary strategies and change their attitudes to get rid of strain [23]. Especially when individuals anticipate that their chosen responses to stressful events will be associated with a favorable outcome, they are coping [11].

CATS offers us a framework to elaborate on the influences of the COVID-19 pandemic on individuals' mHealth emergency use. The unpredictable and detrimental characteristics of the COVID-19 pandemic change individuals' life and work. Confronted with such changes, the strain will arise in individuals and further shape their attitudes and coping behavior. mHealth is an effective instrument to realize disease prevention and health promotion [24], which helps individuals to successfully cope with the COVID-19 pandemic. Therefore, this study adopts CATS as the overarching theory to explain the indirect relationship between the COVID-19 pandemic and mHealth emergency use intention through COVID-19–induced strain.

Hypothesis Development

Event Disruption, mHealth Emergency Use Intention, and COVID-19–Induced Strain

When events are urgent, unpredictable, unexpected, and threatening, they are regarded as stressful and may result in negative psychological, physical, and physiological outcomes [25]. CATS posits that strain or stress experience arises from challenging stressful events [26,27].

This corresponds with the event disruptions of the COVID-19 pandemic, which reflect change and discontinuity of the external situation [2]. The COVID-19 pandemic has significantly changed the way individuals live and work. For instance, Chinese citizens are required to quarantine and work from home [28]. The COVID-19 pandemic has transformed daily routines, requiring individuals to invest considerable resources and energies to adapt effectively. Risk of exposure to the virus, uncertainty about income, shortages of protective equipment, and irregular work hours have all contributed to the increase in stress experience or strain [29,30]. Additionally, the COVID-19 pandemic situation changes on a daily basis, and therefore, the effects on stress levels also change on a daily basis. Taken together, we hypothesize that:

- Hypothesis (H)1: Event disruption is positively associated with COVID-19–induced strain on a daily basis.

CATS links stressful events with coping behavior [11]. Stress coping refers to the constant adaptation of cognitive and behavioral efforts to deal with specific demands deriving from external or internal situations [31]. Two determinants of coping behavior are the specific context of the stressful situation and the individual's expectations about the outcomes [11,31].

Regarding the context of stress, an individual's coping response depends on their appraisal of the demands and resources available to handle the stressful event [32]. COVID-19–induced strain mainly arises from the uncertainty of infection and the risk of exposure to the virus [29]. To maintain their health status and prevent unexpected infection, individuals are driven to seek relevant information and help from professional medical personnel [33], which can be accommodated by mHealth. Through integrating advanced technology, mHealth can be accessed using portable and wireless communication equipment such as a tablet, mobile phone, or wearable device [8]. The remote and instant availability of mHealth provides convenience for recipients, especially during the COVID-19 pandemic [34].

In terms of expectations, the choice of coping behavior is determined by an individual's anticipated outcome [8]. If desirable outcomes are expected, individuals are more likely to choose positive coping strategies in response to stressful experiences [35]. The COVID-19–induced strain has increased public health concerns. The uncertainty and unpredictability of COVID-19 require that individuals access instant and accurate information. Appropriate consultation and treatment can be provided via mHealth in situations of urgency [36]. Research

has demonstrated the vital role that mHealth plays in the implementation of prehospital measures in response to specific diseases [37]. It effectively facilitates the delivery of health care services and accurate health-related information [38]. Based on these findings, we infer that mHealth is a preferred coping tool for daily COVID-19–induced strain. Therefore, we hypothesize the following:

- H2: COVID-19–induced strain is positively correlated with mHealth emergency use intention on a daily basis.

The event disruption of the COVID-19 pandemic makes the external situations unpredictable and uncertain. The changes in life induce stress experience in individuals. To cope with event disruption and COVID-19–induced strain on a daily basis, individuals are more likely to use mHealth in urgent situations to promote health status and prevent a specific disease. In this vein, we further hypothesize that:

- H3: COVID-19–induced strain mediates the relationship between event disruption and daily mHealth emergency use intention.

Moderating Role of Promotion Regulatory Focus

CATS suggests that the choice of coping behavior is determined by the interaction of personal and contextual variables [31]. Regulatory focus is regarded as one of the most important personality variables impacting coping behavior [15] and explains the motivations underlying goal setting [16]. The literature divides regulatory focus into two categories: promotion and prevention focus [39]. An individual with high prevention regulatory focus tries to ensure that they meet the minimum requirements, whereas those with high promotion regulatory focus strive to optimize the situation to achieve ideals and nurturance [40]. Research has shown that promotion regulatory focus plays a critical role in facilitating coping with stress positively [41]. Given that the use of mHealth is an active effort toward resolving COVID-19–induced strain, we adopted promotion rather than prevention regulatory focus to represent a personal preference of coping strategies.

In the context of COVID-19–induced strain, individuals with high promotion regulatory focus may regulate their actions and attitudes to achieve favorable outcomes [42] by generating potential approaches and strategies. When experiencing COVID-19–induced strain, they may seek immediate access to accurate medical information and health care services to reduce their health concerns as well as those of people who are close to them [29]. mHealth may be particularly suited to these individuals because of its timeliness and accessibility. In contrast, individuals with low promotion regulatory focus will not prioritize an optimal outcome and therefore will not consider access to health care services as urgent. These individuals are less likely to use mHealth to help cope with COVID-19–induced strain. Therefore, we hypothesize the following:

- H4: Promotion regulatory focus will moderate the relationship between daily COVID-19–induced strain and mHealth emergency use intention, such that the relationship is stronger in the condition of high promotion regulatory focus than in the condition of low promotion regulatory focus.

As previously mentioned, event disruptions of the COVID-19 pandemic cause unpredictable and unfavorable changes in personal and work life, which elevates stress experience. This induced strain may drive individuals to use mHealth in urgent situations, especially those with high promotion regulatory focus, as this will allow them to promote good health and prevent disease. We hypothesize that:

- H5: Promotion regulatory focus will moderate the indirect relationship between event disruption and daily mHealth emergency use intention through COVID-19–induced strain, such that the indirect relationship is stronger in the condition of high promotion regulatory focus than in the low promotion regulatory focus.

Methods

Data Collection

Based on the research of Du et al [43], we used convenience sampling to recruit our participants. We sought help from the administrative staff at the university. With their help, we contacted the alumni who updated their contact information within 2 years. We recruited the participants through the alumni networks of the public university in China. Through this convenience sampling method, we invited the participation of the alumni living or working in diverse areas in China, thereby bolstering the external validity of the research findings. The inclusion criteria included having a certain degree of smartphone use experience (≥1 year), living or working in mainland China, and having a mobile phone or tablet connected to the internet. The WeChat smartphone app was adopted for this study because of its popularity in China. A research group was created on WeChat, with an initial invitation to 150 potential participants to join the group.

The data collection contained two stages. On February 23, 2020, participants were asked to complete a baseline questionnaire regarding demographic information (gender, age, education) and promotion regulatory focus. From February 24 to 28, 2020, participants were sent a website link at 11 AM that assessed event disruptions and at 5 PM that assessed COVID-19–induced strain and mHealth emergency use intention on each day. Participants were asked to complete the questionnaires within 2 hours. Of the 150 individuals invited, we collected 550 matched responses from a total of 110 participants, yielding an effective response rate of 73.3%. The 110 participants received a ¥25 (about US $3.53) inconvenience allowance.

Measurement Development

All of the measures of the constructs were developed based on previous research. We adapted each item to fit the daily gathering of data. For instance, one item of the original work strain scale is “I often feel too tense due to my work.” We adapted it as “Due to COVID-19 Pandemic, I lived and worked under a great deal of tension today” to fit the COVID-19 pandemic and the daily research context. Specifically, in accordance with suggestions from Donald et al [44], short scales were adopted to minimize the nonresponse rate. A 7-point Likert scale was used, ranging from 1 (*strongly disagree*) to 7 (*strongly*

agree). The complete scale is shown in Multimedia Appendix 1.

Daily Measurement

Event Disruptions

Measures for event disruptions were adapted from four items developed by Morgeson et al [2]. The sample item was "Today, COVID-19 pandemic disrupted my ability to get its work done." The average score of this scale ranged from 1 to 7. The scale yielded a Cronbach alpha of .93.

COVID-19–Induced Strain

The COVID-19–induced strain was measured by three items adapted from the scale developed by House and Rizzo [45]. The sample item was "Due to COVID-19 Pandemic, I lived and worked under a great deal of tension today." The range of the average score was from 1 to 7. The reliability of this scale was .76.

Emergency Use Intention

mHealth emergency use intention was measured by three items developed by Liu et al [8]. The sample item was "Today, I intended to use mHealth services under urgent medical requirements." The range of the score was from 1 to 7. The Cronbach alpha of this scale was .92.

Baseline Measurement

Promotion Regulatory Focus

The regulatory focus has been regarded as a personality trait, which is stable and not probable to change in a short time. Thus, this study put promotion regulatory focus at the baseline measurement [46-48]. Promotion regulatory focus was measured by nine items developed by Lockwood et al [49]. The sample item of this scale was "I frequently imagine how I will achieve my hopes and aspirations." The average score of this scale ranged from 1 to 7. The alpha of this scale was .93.

Control Variables

We also collected demographic data including gender, education, age, and chronic disease, as they may influence mHealth use intention [8,50].

Analytical Strategy

The data was nested, as the data were collected using the experience sampling method. The data had a two-level hierarchical structure, where daily level or within-person level data was positioned at level one and individual level or between-person level data was positioned at level two [51]. We, therefore, used hierarchical linear modeling (HLM) for our analyses [52].

The analysis contained two stages. First, we investigated the within-person level variance in the daily variables. The results showed about a 71%-85% variance for the within-person level for event disruption, COVID-19–induced strain, and mHealth emergency use intention, justifying the use of HLM. Second, we performed HLM (version 6.08) using a restricted maximum likelihood estimation for the parameter analyses. We conducted a moderated mediation model analysis with a random slope and used robust estimators in level one to indicate the daily or within-person effect. The daily variables (event disruption, COVID-19–induced strain, and mHealth emergency use intention) were group-centered.

Results

Participants

Table 1 shows the results of demographic information. Among the 110 participants, 54.5% (n=60) were males. Of the participants, 0.9% (n=1) only received primary school education, 1.8% (n=2) graduated from junior school, 27.3% (n=30) graduated from senior high school, 23.6% (n=26) held college certificates, and 46.4% (n=51) held Bachelor's degrees or above. Regarding chronic disease, 88.4% (n=97) of the participants did not have a chronic disease. For the distribution of age, 1.8% (n=2) of the participants were younger than 26 years, 45.5% (n=50) ranged from 26 to 35 years, 32.7% (n=36) ranged from 36 to 45 years, and 20% (n=22) were older than 45 years.

Table 1. Participants' demographic data (N=110).

Characteristic	Participants, n (%)
Gender	
Male	60 (54.5)
Female	50 (45.5)
Chronic disease	
No	97 (88.4)
Yes	13 (11.6)
Education	
Primary school	1 (0.9)
Senior school	2 (1.8)
High school	30 (27.3)
College	26 (23.6)
Bachelor's and above	51 (46.4)
Age (years)	
<26	2 (1.8)
26-35	50 (45.5)
36-45	36 (32.7)
≥46	22 (20.0)

Multilevel Confirmatory Analysis

Given that our daily data were nested, we adopted multilevel confirmatory factor analysis rather than exploratory factor analysis to test the validity of the measurements and the common method variance [43,53]. The results showed that the proposed four-factor model yielded a better model fit (χ^2_{57}=151.22; root mean square error of approximation 0.06; comparative fit index 0.96; Tucker–Lewis index 0.95; standardized root mean square residual 0.03) than any other model. The results are shown in Table 2.

Table 2. Results of multilevel confirmatory factor analysis.

Models	Chi-square (*df*)	△ chi-square	*P* value	RMSEA[a]	CFI[b]	TLI[c]	SRMR[d] (within)
EU[e], ED[f], LS[g], PF[h]	151.22 (57)	N/A[i]	N/A	0.06	0.96	0.95	0.03
EU+ED, LS, PF	451.15 (59)	299.93	<.001	0.11	0.84	0.78	0.19
EU+LS, ED, PF	420.60 (59)	269.38	<.001	0.11	0.85	0.80	0.12
EU, LS+ED, PF	440.35 (59)	289.13	<.001	0.11	0.85	0.79	0.16
EU+LS+ED, PF	658.05 (60)	506.83	<.001	0.14	0.76	0.67	0.21

[a]RMSEA: root mean square error of approximation.

[b]CFI: comparative fit index.

[c]TLI: Tucker–Lewis index.

[d]SRMR: standardized root mean square residual.

[e]EU: mobile health emergency use intention.

[f]ED: event disruption.

[g]LS: COVID-19–induced strain.

[h]PF: promotion regulatory focus.

[i]N/A: not applicable.

Descriptive Statistics

Tables 3 and 4 show the means, SDs, reliabilities, and correlations of the variables both at the within- and between-person levels. The results showed that event disruption was significantly correlated with daily mHealth emergency use (r=0.20, P<.001) and daily COVID-19–induced strain (r=0.29, P<.001). Event disruption was associated with daily COVID-19–induced strain (r=0.27, P<.001).

Table 3. Within-person level (N=550) means, SDs, and correlations.

Variables	Mean (SD)	1	2	3
1. mHealth[a] emergency use intention[b]	4.84 (0.99)			
r		1		
P value		N/A[c]		
2. Event disruption[d]	3.81 (0.69)			
r		0.20	1	
P value		<.001	N/A	
3. COVID-19–induced strain[e]	4.20 (0.77)			
r		0.29	0.27	1
P value		<.001	<.001	N/A

[a]mHealth: mobile health.
[b]Cronbach alpha=.92.
[c]N/A: Not applicable.
[d]Cronbach alpha=.93.
[e]Cronbach alpha=.76.

Table 4. Between-person level (N=110) means, SDs, and correlations.

Variables	Mean (SD)	1	2	3	4	5
1. Gender	1.45 (0.50)					
r		1				
P value		N/A[a]				
2. Education	4.13 (0.94)					
r		0.68	1			
P value		<.001	N/A			
3. Chronic disease	1.12 (0.32)					
r		0.06	0.07	1		
P value		.19	.12	N/A		
4. Age	2.71 (0.80)					
r		–0.49	–0.24	–0.16	1	
P value		<.001	<.001	<.001	N/A	
5. Promotion regulatory focus[b]	4.08 (0.35)					
r		0.03	–0.15	–0.02	–0.15	1
P value		.69	<.001	.57	<.001	N/A

[a]N/A: Not applicable.
[b]Cronbach alpha=.93.

Hierarchical Linear Modeling

Daily event disruption had a significant positive relationship with both COVID-19–induced strain (Model 1: γ=0.24, *P*<.001) and mHealth emergency use intention (Model 2: γ=0.28, *P*<.001). Results of model 3 showed that COVID-19–induced strain was significantly positively correlated with mHealth emergency use intention (γ=0.36, *P*<.001) when mHealth emergency use intention was simultaneously regressed on strain and event disruption. H1 and H2 were supported. The results are shown in Table 5.

To further explore the mediating role of COVID-19–induced strain on the temporal relationship between event disruption and mHealth emergency use intention, a Monte Carlo bootstrapping test was performed using R (version 3.5.3; R Foundation for Statistical Computing). Both the direct relationship (effect=0.18, 95% CI 0.06-0.30) and indirect relationship (effect=0.09, 95% CI 0.05-0.14) were significant. The results are summarized in Table 6. H3 was supported.

The results for promotion regulatory focus are shown in model 4. The interaction of promotion regulatory focus with COVID-19–induced strain was positively associated with mHealth emergency use intention (γ=0.35, *P*=.02). To further explore the moderating role of promotion regulatory focus, we performed a Monte Carlo bootstrapping test. The moderating effect of promotion regulatory focus on the relationship between COVID-19–induced strain and mHealth emergency use intention was significantly stronger in the condition of high promotion regulatory focus (effect=0.49, 95% CI 0.34-0.65) than in the condition of low promotion regulatory focus (effect=0.25, 95% CI 0.09-0.50). The difference was also significant (effect=0.24, 95% CI 0.04-0.45), supporting H4. The moderating role of promotion regulatory focus is shown in Figure 2.

Finally, we used a Monte Carlo bootstrapping test to examine the moderated mediation model. The results showed that the indirect relationship between daily event disruption and mHealth emergency use intention through COVID-19–induced strain was significantly stronger when promotion regulatory focus was high (effect=0.12, 95% CI 0.06-0.19) than when it was low (effect=0.06, 95% CI 0.02-0.11). The difference between the two effects was significant (effect=0.06, 95% CI 0.001-0.12), supporting H5.

Table 5. Results of hierarchical linear model analysis.

Variables	COVID-19–induced strain			mHealth[a] emergency use intention								
	Model 1[b]			Model 2[c]			Model 3[d]			Model 4[e]		
	γ	SE	*P* value	γ	SE	*P* value	γ	SE	*P* value	γ	SE	*P* value
Intercepts	3.43	0.21	<.001	4.23	0.30	<.001	4.22	0.31	<.001	4.20	0.31	<.001
Between-person (N=110)												
Gender	–0.14	0.09	.115	–0.13	0.20	.51	–0.11	0.19	.55	–0.12	0.19	.55
Education	0.09	0.04	<.001	0.08	0.09	.38	0.08	0.09	.38	0.08	0.09	.36
Chronic disease	0.13	0.10	.20	0.31	0.14	.03	0.29	0.15	.06	0.29	0.15	.06
Age	0.02	0.05	.59	0.04	0.08	.61	0.05	0.08	.52	0.05	0.08	.50
Promotion regulatory focus	0.30	0.09	<.001	0.37	0.16	.02	0.43	0.16	.009	0.39	0.16	.02
Within-person (N=550)												
Event disruption	0.24	0.05	<.001	0.28	0.06	<.001	0.19	0.06	<.001	0.18	0.06	.004
COVID-19–induced strain	N/A[f]	N/A	N/A	N/A	N/A	N/A	0.36	0.06	<.001	0.37	0.06	<.001
Interactive item												
COVID-19–induced strain × promotion regulatory focus	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.35	0.15	.02

[a]mHealth: mobile health.

[b]Pseudo R^2=0.11.

[c]Pseudo R^2=0.07.

[d]Pseudo R^2=0.10.

[e]Pseudo R^2=0.12.

[f]N/A: not applicable.

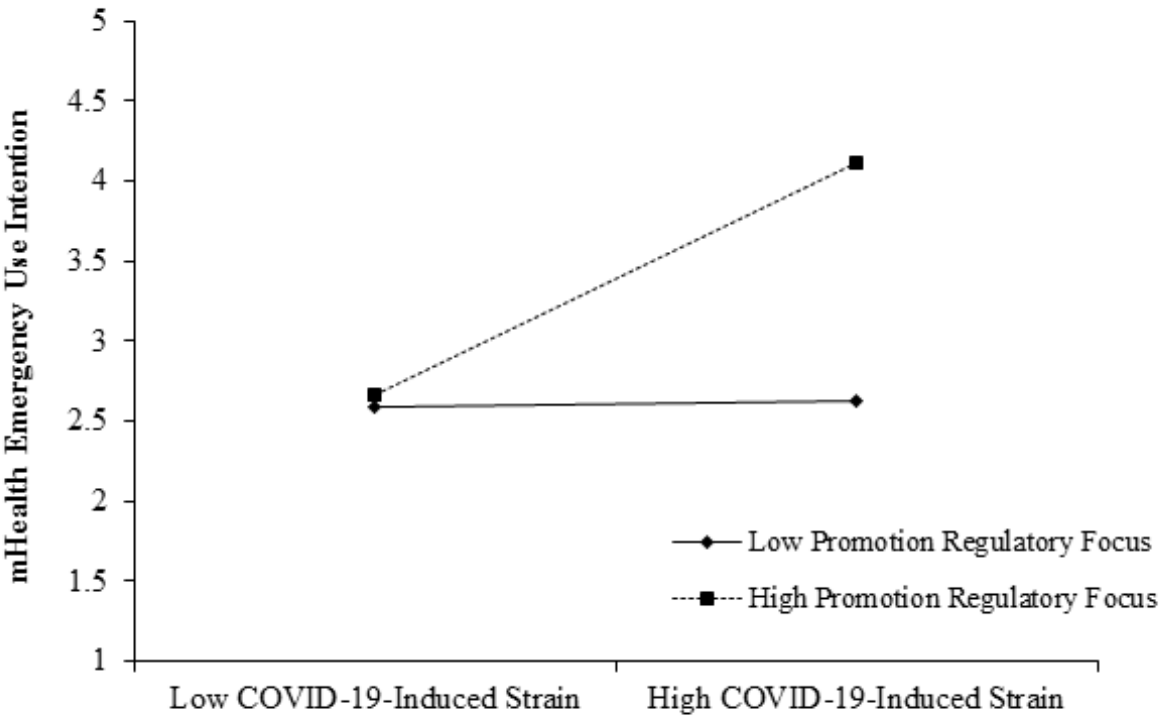
Table 6. Results of the Monte Carlo bootstrapping test.

Effect	Estimator	SE	95% CI[a]
Moderating effect of promotion regulatory focus			
Low (M[b] – SD)	0.25	0.08	0.09-0.40
High (M + SD)	0.49	0.08	0.34-0.65
Difference	0.24	0.11	0.04-0.45
Mediating model of COVID-19–induced strain			
Direct effect	0.18	0.06	0.06-0.30
Indirect effect	0.09	0.02	0.05-0.14
Moderated mediation model			
Low (M – SD)	0.06	0.02	0.02-0.11
High (M + SD)	0.12	0.03	0.06-0.19
Difference	0.06	0.03	0.01-0.12

[a]Bootstrapping=20,000.

[b]M: mean.

Figure 2. Moderating role of promotion regulatory focus. mHealth: mobile health.



Discussion

We developed a conceptual model to explore how daily event disruptions of the COVID-19 pandemic predict mHealth emergency use intention. This study provides key findings and contributions to both mHealth research and practitioners.

Principal Findings

This study presents three significant key findings. First, event disruptions of the COVID-19 pandemic are associated with increased daily mHealth emergency use intentions. Event disruptions of the COVID-19 pandemic represent the discontinuity of daily routines [2], resulting in the need for individuals to adapt their behaviors and attitudes [4]. During the COVID-19 pandemic, a major change has been an increased concern for disease prevention [54]. However, the uncertainty of the SARS-CoV-2 virus has driven people to find a means to access immediate medical information and health care services in urgent situations [55], for which mHealth may be a solution. Given the fluctuations of event disruptions caused by the COVID-19 pandemic, this study also explores the temporal

influences of event disruption on mHealth emergency use intention. The results found that individuals increased their intention to use mHealth to help deal with the event disruptions of the COVID-19 pandemic.

Second, this study found that COVID-19–induced strain mediated the relationship between event disruption and mHealth emergency use intention on a daily basis. According to CATS, stressful experiences or strain arises from the lack of resources to effectively deal with the demands of stressful events [56]. The SARS-CoV-2 virus is novel and highly infectious, putting high levels of stress on the public in general, which also threatens their mental health [57]. Consequently, individuals have a need for finding ways to alleviate this strain. Our research found that COVID-19–induced strain strengthens the mechanism through which event disruption facilitates mHealth emergency use intention.

Third, our research findings suggest that promotion regulatory focus amplifies the indirect effect of event disruption on mHealth emergency use intention through daily COVID-19–induced strain. The interaction of personality traits and contextual variables determines the choice of coping behavior [31]. When confronted with a stressful experience, individuals who are striving to maintain their health are likely to seek tools to acquire relevant medical information [58]. This is consistent with our finding that promotion regulatory focus plays a contingent role in the association between COVID-19–induced strain and mHealth emergency use intention.

Theoretical Implications

This study provides several theoretical contributions to mHealth literature. First, this study contributes to the mHealth literature by identifying the temporal influences of event disruption and mHealth emergency use intention. The COVID-19 pandemic was used as an example situation to explore the influence of event disruption caused by an emergent health crisis on the use of mHealth. This study extended this line of research by not only incorporating event disruption as an influencing factor for mHealth emergency use intention but also by examining mHealth emergency use intention on a daily basis. In doing so, this study contributes to mHealth literature by identifying a new type of mHealth use intention and examining its proximal antecedent.

Second, our study uncovered an underlying mechanism by examining the mediating role of COVID-19–induced strain. Previous studies investigating mHealth use intention mainly focused on the influences of technological and psychological factors [8,59]. For instance, Hoque [60] found a positive relationship between perceived usefulness and mHealth use intention, while Liu et al [8] found that perceived autonomy increased the use of mHealth. This study extends this line of research concerning the influences of psychological characteristics by incorporating COVID-19–induced strain. We identified COVID-19–induced strain as a contributor to mHealth emergency use intention. Furthermore, our exploration of the impact of event disruption on strain showed that changes in the external environment increased individuals' health concerns, associating with elevated stress levels. Individuals were likely to use mHealth for health self-management and to reduce strain. We have extended prior research by identifying the role of strain in shaping mHealth emergency use intention and gaining a better understanding of how a public health crisis impacts personal strain and mHealth emergency use intention.

Third, we have also enriched the understanding of CATS by incorporating promotion regulatory focus into our model. Previous research using CATS primarily focused on the role of expectations in shaping an individual's response to stressful events [11]. The expected outcome is what motivates an individual to adopt certain coping strategies [61]. However, little is known about the criteria for evaluating specific expected outcomes. We propose that promotion regulatory focus could be a possible explanation. For individuals with high promotion regulatory focus, the expected outcome would be to maintain personal health when confronted with psychological strain caused by the COVID-19 pandemic. We found that promotion regulatory focus prompts individuals to rely more on mHealth to cope with psychological strain in a health emergency. This provides useful insight into the formation process of expectations.

Practical Implications

This study has practical implications for mHealth providers during a public health crisis. When the public experiences a health crisis, many people use mHealth services, which helps deal with psychological strain. We recommend that service providers develop specific services to cater to the needs of the public. For instance, remote primary diagnosis and health monitoring for a specific disease can be integrated into mHealth. This would enable individuals to incorporate mHealth into their daily lives and allow effective self-monitoring, even in urgent situations.

In addition, it would be useful for service providers to consider the role of regulatory focus, as individuals with high promotion regulatory focus are more likely to use mHealth when confronted with a health emergency. Service providers may adopt the regulatory focus scale as a primary screening method to select potential users and provide them with specific functions and services. This would offer providers with opportunities to increase user compliance.

Limitations and Future Research

This study has several limitations and points out directions for future research. First, we did not establish a causal relationship between event disruption and mHealth emergency use intention. Moreover, although we collected two-wave data on a daily basis, we cannot conclude that daily event disruption predicts psychological strain and mHealth emergency use intention because we did not manipulate the event disruptions of the COVID-19 pandemic. Future research could use a cross-lagged panel design to infer the causal relationship between event disruption and mHealth use intention.

Second, it is also not possible to rule out common method variance. The experience sampling method controls for common method variance to a certain degree, as confirmed by the multilevel confirmatory factor analysis. However, our data were collected through self-report questionnaires, and therefore, our

results may still have been impacted by common method variance. Future research could acquire objective data to minimize the potential effects of common method variance. This could be implemented through gathering mHealth app browsing history during a public health emergency.

Another limitation is that our study was conducted in China. Further research is needed to test the generalizability of our findings in other countries. The development of the mHealth industry differs across the world and mHealth use will depend on the stage of development of this technology. Therefore, future research in other countries will need to additionally consider these factors.

Conclusions

mHealth provides individuals with a platform to access health care services. The results showed that event disruption of the COVID-19 pandemic induced mHealth emergency use intention through increased psychological strain. Furthermore, individuals' promotion regulatory focus amplified this indirect relationship. Our findings extend our understanding of the factors underlying mHealth emergency use intention and illustrate the potential contingent role of promotion regulatory focus in CATS. This study also opens avenues for future research on mHealth emergency use intention in other countries and cultural settings.

Acknowledgments

Our research is supported by the National Natural Science Foundation of China (71701083), Yunnan Province Basic Research Planning Project (Grant No. 2019FB084), and the 12th Five-years Plan of Beijing science of education (CAA15006).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire: measurement of the major constructs.
[DOCX File , 13 KB - mhealth_v8i12e20642_app1.docx]

References

1. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
2. Morgeson FP, Mitchell TR, Liu D. Event system theory: an event-oriented approach to the organizational sciences. Acad Manage Rev 2015 Oct;40(4):515-537. [doi: 10.5465/amr.2012.0099]
3. Hoffman AJ, Ocasio W. Not all events are attended equally: toward a middle-range theory of industry attention to external events. Organ Sci 2001 Aug;12(4):414-434. [doi: 10.1287/orsc.12.4.414.10639]
4. Zellmer-Bruhn ME. Interruptive events and team knowledge acquisition. Manage Sci 2003 Apr;49(4):514-528. [doi: 10.1287/mnsc.49.4.514.14423]
5. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile health apps on COVID-19 launched in the early days of the pandemic: content analysis and review. JMIR mHealth uHealth 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: 10.2196/19796] [Medline: 32609622]
6. 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]
7. Bauer AM, Thielke SM, Katon W, Unützer J, Areán P. Aligning health information technologies with effective service delivery models to improve chronic disease care. Prev Med 2014 Sep;66:167-172 [FREE Full text] [doi: 10.1016/j.ypmed.2014.06.017] [Medline: 24963895]
8. Liu F, Ngai E, Ju X. Understanding mobile health service use: An investigation of routine and emergency use intentions. Int J Inf Manage 2019 Apr;45:107-117. [doi: 10.1016/j.ijinfomgt.2018.09.004]
9. Yang L, Su G, Yuan H. Design principles of integrated information platform for emergency responses: the case of 2008 Beijing Olympic games. Inf Syst Res 2012 Sep;23(3-part-1):761-786. [doi: 10.1287/isre.1110.0387]
10. Ursin H, Eriksen HR. The cognitive activation theory of stress. Psychoneuroendocrinology 2004 Jun;29(5):567-592. [doi: 10.1016/S0306-4530(03)00091-X] [Medline: 15041082]
11. Meurs JA, Perrewé PL. Cognitive activation theory of stress: an integrative theoretical approach to work stress. J Manage 2010 Nov 05;37(4):1043-1068. [doi: 10.1177/0149206310387303]
12. Mikolajczak M, Luminet O. Trait emotional intelligence and the cognitive appraisal of stressful events: an exploratory study. Pers Individual Differences 2008 May;44(7):1445-1453. [doi: 10.1016/j.paid.2007.12.012]
13. Lambert EG, Qureshi H, Frank J, Klahm C, Smith B. Job stress, job involvement, job satisfaction, and organizational commitment and their associations with job burnout among Indian police officers: a research note. J Police Crim Psych 2017 Apr 17;33(2):85-99. [doi: 10.1007/s11896-017-9236-y]

14. Alexander D, Monk JS, Jonas AP. Occupational stress, personal strain, and coping among residents and faculty members. J Med Educ 1985 Nov;60(11):830-839. [doi: 10.1097/00001888-198511000-00002] [Medline: 4057224]
15. Bergvik S, Sørlie T, Wynn R. Approach and avoidance coping and regulatory focus in patients having coronary artery bypass graft surgery. J Health Psychol 2010 Sep;15(6):915-924. [doi: 10.1177/1359105309359542] [Medline: 20453051]
16. Lai C, Hsu JS, Li Y. Leadership, regulatory focus and information systems development project team performance. Int J Project Manage 2018 Apr;36(3):566-582. [doi: 10.1016/j.ijproman.2017.11.001]
17. Ahmadi S, Khanagha S, Berchicci L, Jansen JJP. Are managers motivated to explore in the face of a new technological change? The role of regulatory focus, fit, and complexity of decision-making. J Manage Stud 2017 Feb 06;54(2):209-237. [doi: 10.1111/joms.12257]
18. Shah J, Higgins ET. Expectancy x value effects: regulatory focus as determinant of magnitude and direction. J Pers Soc Psychol 1997 Sep;73(3):447-458. [doi: 10.1037//0022-3514.73.3.447] [Medline: 9294896]
19. Ursin H, Eriksen HR. Cognitive activation theory of stress (CATS). Neurosci Biobehav Rev 2010 May;34(6):877-881. [doi: 10.1016/j.neubiorev.2009.03.001] [Medline: 20359586]
20. Karaman MA, Watson JC. Examining associations among achievement motivation, locus of control, academic stress, and life satisfaction: a comparison of U.S. and international undergraduate students. Pers Individual Differences 2017 Jun;111:106-110. [doi: 10.1016/j.paid.2017.02.006]
21. Gomes AR, Teixeira PM. Stress, cognitive appraisal and psychological health: testing instruments for health professionals. Stress Health 2016 Apr;32(2):167-172. [doi: 10.1002/smi.2583] [Medline: 24891033]
22. Youn S, Shin W. Teens' responses to Facebook newsfeed advertising: the effects of cognitive appraisal and social influence on privacy concerns and coping strategies. Telematics Inform 2019 May;38:30-45. [doi: 10.1016/j.tele.2019.02.001]
23. Hewett R, Liefooghe A, Visockaite G, Roongrerngsuke S. Bullying at work: cognitive appraisal of negative acts, coping, wellbeing, and performance. J Occup Health Psychol 2018 Jan;23(1):71-84. [doi: 10.1037/ocp0000064] [Medline: 27936829]
24. 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. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
25. Hale D, Ployhart RE, Shepherd W. A two-phase longitudinal model of a turnover event: disruption, recovery rates, and moderators of collective performance. Acad Manage J 2016 Jun;59(3):906-929. [doi: 10.5465/amj.2013.0546]
26. Hurrell JJ, Nelson DL, Simmons BL. Measuring job stressors and strains: where we have been, where we are, and where we need to go. J Occup Health Psychol 1998 Oct;3(4):368-389. [doi: 10.1037//1076-8998.3.4.368] [Medline: 9805282]
27. Schmitt A, Den Hartog DN, Belschak FD. Transformational leadership and proactive work behaviour: a moderated mediation model including work engagement and job strain. J Occup Organ Psychol 2016 Jan 20;89(3):588-610. [doi: 10.1111/joop.12143]
28. Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y. A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. Gen Psychiatr 2020;33(2):e100213 [FREE Full text] [doi: 10.1136/gpsych-2020-100213] [Medline: 32215365]
29. Pfefferbaum B, North CS. Mental health and the Covid-19 pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
30. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. Nat Hum Behav 2020 May;4(5):460-471. [doi: 10.1038/s41562-020-0884-z] [Medline: 32355299]
31. Folkman S, Lazarus RS, Dunkel-Schetter C, DeLongis A, Gruen RJ. Dynamics of a stressful encounter: cognitive appraisal, coping, and encounter outcomes. J Pers Soc Psychol 1986 May;50(5):992-1003. [doi: 10.1037//0022-3514.50.5.992] [Medline: 3712234]
32. Zhao X, Xia Q, Huang W. Impact of technostress on productivity from the theoretical perspective of appraisal and coping processes. Inf Manage 2020 Dec;57(8):103265. [doi: 10.1016/j.im.2020.103265]
33. James DCS, Harville C. eHealth literacy, online help-seeking behavior, and willingness to participate in mHealth chronic disease research among African Americans, Florida, 2014-2015. Prev Chronic Dis 2016 Nov 17;13:E156 [FREE Full text] [doi: 10.5888/pcd13.160210] [Medline: 27854421]
34. Grasselli G, Pesenti A, Cecconi M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. JAMA 2020 Apr 28;323(16):1545-1546. [doi: 10.1001/jama.2020.4031] [Medline: 32167538]
35. Friedman-Wheeler DG, Litovsky AR, Prince KR, Webbert J, Werkheiser A, Carlson E, et al. Do mood-regulation expectancies for coping strategies predict their use? A daily diary study. Int J Stress Manage 2019 Aug;26(3):287-296. [doi: 10.1037/str0000115]
36. Choi BY, Blumberg C, Williams K. Mobile integrated health care and community paramedicine: an emerging emergency medical services concept. Ann Emerg Med 2016 Mar;67(3):361-366. [doi: 10.1016/j.annemergmed.2015.06.005] [Medline: 26169927]
37. Eapen ZJ, Turakhia MP, McConnell MV, Graham G, Dunn P, Tiner C, et al. Defining a mobile health roadmap for cardiovascular health and disease. J Am Heart Assoc 2016 Jul 12;5(7) [FREE Full text] [doi: 10.1161/JAHA.115.003119] [Medline: 27405809]

38. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future Mobile-Health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. J Med Internet Res 2016 Jun 10;18(6):e135 [FREE Full text] [doi: 10.2196/jmir.5066] [Medline: 27287668]
39. Higgins ET. Promotion and prevention: regulatory focus as a motivational principle. In: Zanna MP, editor. Advances in Experimental Social Psychology. San Diego, CA: Academic Press; 1998:1-46.
40. Cho H, Roh S, Park B. Of promoting networking and protecting privacy: effects of defaults and regulatory focus on social media users' preference settings. Comput Hum Behav 2019 Dec;101:1-13. [doi: 10.1016/j.chb.2019.07.001]
41. Woltin K, Sassenberg K, Albayrak N. Regulatory focus, coping strategies and symptoms of anxiety and depression: a comparison between Syrian refugees in Turkey and Germany. PLoS One 2018;13(10):e0206522 [FREE Full text] [doi: 10.1371/journal.pone.0206522] [Medline: 30359455]
42. Kim J, Chen K, Davis WE, Hicks JA, Schlegel RJ. Approaching the true self: promotion focus predicts the experience of authenticity. J Res Pers 2019 Feb;78:165-176. [doi: 10.1016/j.jrp.2018.12.001]
43. Du D, Derks D, Bakker AB. Daily spillover from family to work: a test of the work-home resources model. J Occup Health Psychol 2018 Apr;23(2):237-247. [doi: 10.1037/ocp0000073] [Medline: 28150994]
44. Donald JN, Atkins PW, Parker PD, Christie AM, Ryan RM. Daily stress and the benefits of mindfulness: examining the daily and longitudinal relations between present-moment awareness and stress responses. J Res Pers 2016 Dec;65:30-37. [doi: 10.1016/j.jrp.2016.09.002]
45. House RJ, Rizzo JR. Toward the measurement of organizational practices: scale development and validation. J App Psychol 1972;56(5):388-396. [doi: 10.1037/h0033444]
46. Johnson RE, King DD, (Joanna) Lin S, Scott BA, Jackson Walker EM, Wang M. Regulatory focus trickle-down: how leader regulatory focus and behavior shape follower regulatory focus. Organizational Behav Hum Decision Processes 2017 May;140:29-45. [doi: 10.1016/j.obhdp.2017.03.002]
47. Manczak EM, Zapata-Gietl C, McAdams DP. Regulatory focus in the life story: prevention and promotion as expressed in three layers of personality. J Pers Soc Psychol 2014 Jan;106(1):169-181. [doi: 10.1037/a0034951] [Medline: 24377362]
48. Lim S, Ilies R, Koopman J, Christoforou P, Arvey RD. Emotional mechanisms linking incivility at work to aggression and withdrawal at home: an experience-sampling study. J Manage 2016 Jun 23;44(7):2888-2908. [doi: 10.1177/0149206316654544]
49. Lockwood P, Jordan CH, Kunda Z. Motivation by positive or negative role models: regulatory focus determines who will best inspire us. J Pers Soc Psychol 2002 Oct;83(4):854-864. [Medline: 12374440]
50. 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]
51. Liu W, Song Z, Li X, Liao Z. Why and when leaders' affective states influence employee upward voice. Acad Manage J 2017 Feb;60(1):238-263. [doi: 10.5465/amj.2013.1082]
52. Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd ed. Thousand Oaks, CA: Sage; 2002.
53. Margola D, Fenaroli V, Sorgente A, Lanz M, Costa G. The Family Relationships Index (FRI): multilevel confirmatory factor analysis in an Italian community sample. Eur J Psychological Assess 2019 May;35(3):335-345. [doi: 10.1027/1015-5759/a000427]
54. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]
55. Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing. JAMA 2020 Apr 14;323(14):1341-1342. [doi: 10.1001/jama.2020.3151] [Medline: 32125371]
56. Useche SA, Montoro L, Alonso F, Pastor JC. Psychosocial work factors, job stress and strain at the wheel: validation of the copenhagen psychosocial questionnaire (COPSOQ) in professional drivers. Front Psychol 2019;10:1531. [doi: 10.3389/fpsyg.2019.01531] [Medline: 31312166]
57. Bo H, Li W, Yang Y, Wang Y, Zhang Q, Cheung T, et al. Posttraumatic stress symptoms and attitude toward crisis mental health services among clinically stable patients with COVID-19 in China. Psychol Med 2020 Mar 27:1-2 [FREE Full text] [doi: 10.1017/S0033291720000999] [Medline: 32216863]
58. Chen L, Baird A, Rai A. Mobile health (mHealth) channel preference: an integrated perspective of approach - avoidance beliefs and regulatory focus. JAIS 2019;20(12):1743-1773. [doi: 10.17705/1jais.00584]
59. Meng F, Guo X, Peng Z, Zhang X, Vogel D. The routine use of mobile health services in the presence of health consciousness. Electronic Commerce Res Applications 2019 May;35:100847. [doi: 10.1016/j.elerap.2019.100847]
60. Hoque MR. An empirical study of mHealth adoption in a developing country: the moderating effect of gender concern. BMC Med Inform Decis Mak 2016 May 03;16:51 [FREE Full text] [doi: 10.1186/s12911-016-0289-0] [Medline: 27142844]
61. House RJ, Shapiro HJ, Wahba MA. Expectancy theory as a predictor of work behavior and attitude: a re‐evaluation of empirical evidence. Decision Sci 1974 Jul;5(3):481-506. [doi: 10.1111/j.1540-5915.1974.tb00632.x]

Abbreviations

CATS: cognitive activation theory of stress
H: hypothesis
HLM: hierarchical linear modeling
mHealth: mobile health

Edited by L Buis; submitted 06.06.20; peer-reviewed by W Zhang; comments to author 17.07.20; revised version received 03.08.20; accepted 11.12.20; published 30.12.20.

Please cite as:
Zhang Z, Zhang L, Zheng J, Xiao H, Li Z
COVID-19–Related Disruptions and Increased mHealth Emergency Use Intention: Experience Sampling Method Study
JMIR Mhealth Uhealth 2020;8(12):e20642
URL: http://mhealth.jmir.org/2020/12/e20642/
doi:10.2196/20642
PMID:33315579

Original Paper

Effectiveness of a Mobile-Based Influenza-Like Illness Surveillance System (FluMob) Among Health Care Workers: Longitudinal Study

May Oo Lwin[1], PhD; Jiahui Lu[2], PhD; Anita Sheldenkar[1], MSc; Chitra Panchapakesan[1], PhD; Yi-Roe Tan[3], MPH; Peiling Yap[4], PhD; Mark I Chen[3], MPH, PhD; Vincent TK Chow[5], PhD; Koh Cheng Thoon[6], MBBS; Chee Fu Yung[6], MFPHM, FFPHM; Li Wei Ang[3], MSc; Brenda SP Ang[3], MBBS, MPH

[1]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

[2]School of New Media and Communication, Tianjin University, Tianjin, China

[3]National Centre for Infectious Diseases, Singapore, Singapore

[4]The International Digital Health & Artificial Intelligence Research Collaborative (I-DAIR), Graduate Institute of Geneva, Geneva, Switzerland

[5]National University of Singapore, Singapore, Singapore

[6]KK Women's and Children's Hospital, Singapore, Singapore

Corresponding Author:
Jiahui Lu, PhD
School of New Media and Communication
Tianjin University
No. 92 Weijin Road
Tianjin, 300072
China
Phone: 86 18222418810
Email: lujiahui@tju.edu.cn

Abstract

Background: Existing studies have suggested that internet-based participatory surveillance systems are a valid sentinel for influenza-like illness (ILI) surveillance. However, there is limited scientific knowledge on the effectiveness of mobile-based ILI surveillance systems. Previous studies also adopted a passive surveillance approach and have not fully investigated the effectiveness of the systems and their determinants.

Objective: The aim of this study was to assess the efficiency of a mobile-based surveillance system of ILI, termed FluMob, among health care workers using a targeted surveillance approach. Specifically, this study evaluated the effectiveness of the system for ILI surveillance pertaining to its participation engagement and surveillance power. In addition, we aimed to identify the factors that can moderate the effectiveness of the system.

Methods: The FluMob system was launched in two large hospitals in Singapore from April 2016 to March 2018. A total of 690 clinical and nonclinical hospital staff participated in the study for 18 months and were prompted via app notifications to submit a survey listing 18 acute respiratory symptoms (eg, fever, cough, sore throat) on a weekly basis. There was a period of study disruption due to maintenance of the system and the end of the participation incentive between May and July of 2017.

Results: On average, the individual submission rate was 41.4% (SD 24.3%), with a rate of 51.8% (SD 26.4%) before the study disruption and of 21.5% (SD 30.6%) after the disruption. Multivariable regression analysis showed that the adjusted individual submission rates were higher for participants who were older (<30 years, 31.4% vs 31-40 years, 40.2% [*P*<.001]; 41-50 years, 46.0% [*P*<.001]; >50 years, 39.9% [*P*=.01]), ethnic Chinese (Chinese, 44.4% vs non-Chinese, 34.7%; *P*<.001), and vaccinated against flu in the past year (vaccinated, 44.6% vs nonvaccinated, 34.4%; *P*<.001). In addition, the weekly ILI incidence was 1.07% on average. The Pearson correlation coefficient between ILI incidence estimated by FluMob and that reported by Singapore Ministry of Health was 0.04 (*P*=.75) with all data and was 0.38 (*P*=.006) including only data collected before the study disruption. Health care workers with higher risks of ILI and influenza such as women, non-Chinese, allied health staff, those who had children in their households, not vaccinated against influenza, and reported allergy demonstrated higher surveillance correlations.

Conclusions: Mobile-based ILI surveillance systems among health care workers can be effective. However, proper operation of the mobile system without major disruptions is vital for the engagement of participants and the persistence of surveillance power. Moreover, the effectiveness of the mobile surveillance system can be moderated by participants' characteristics, which highlights the importance of targeted disease surveillance that can reduce the cost of recruitment and engagement.

(JMIR Mhealth Uhealth 2020;8(12):e19712) doi:10.2196/19712

KEYWORDS

participatory surveillance; syndromic surveillance; mobile phone; influenza-like illness; health care workers

Introduction

Seasonal influenza is an acute respiratory infection that causes an estimated 290,000-650,000 deaths worldwide [1]. As a major travel hub in a tropical area of southeast Asia, Singapore is at high risk of both imported southern and northern strains of seasonal influenza, exposing the country to the virus throughout the year without discrete peaks [2,3]. The economic burden of the influenza threat is estimated to produce 2.5 million lost days of work and US $470 million costs per year in Singapore [4,5].

Although there are existing influenza surveillance systems in Singapore, the timeliness and effectiveness of these systems can be improved. Since the influenza A (H1N1) outbreak in 2009, virological and clinical surveillance systems have been set up nationwide based on outpatients with influenza-like illness (ILI) visiting public health care institutions [2,6]. However, these methods can be less efficient due to the delay of visiting doctors or the culture of self-medication. This may especially be the case in Singapore because acute respiratory syndrome is often perceived as mild symptoms that can be self-cured or cured via alternative medicines such as traditional Chinese medicine [7-9]. In addition, despite the strength of existing surveillance systems, the maintenance cost can be high [10]. These limitations highlight the importance of a complementary surveillance system that is robust, in real-time, and cost-effective.

The recent idea of crowdsourcing has been promoting the evolution of participatory infectious disease surveillance, which has greatly improved the timeliness and effectiveness of influenza surveillance. Crowdsourcing in health surveillance encourages the general public to contribute their disease information via the internet. The United States, Australia, and Europe have been using internet-based influenza surveillance systems for over a decade [11-13]. The general public is invited via email to submit weekly reports about ILI and provide information on health care–seeking activities. The surveillance power of these internet-based systems is generally good. For example, ILI incidence estimated by Influenzanet (a European influenza surveillance system) in the Netherlands was found to have moderate to high correlations (ie, ranging from 0.42 to 0.89) with those estimated by traditional reporting systems during a 6-year period from 2003 to 2008 [14]. Similar findings were reported in other countries such as Belgium and Portugal using data collected across 10 years [13]. This suggests that internet-based participatory surveillance systems are a valid sentinel for ILI surveillance.

However, we recognize at least three significant limitations in extant work on participatory surveillance of influenza. First, current internet-based surveillance systems may have methodological limitations as they adopt a passive recruitment approach, for example via publicity campaigns involving television, radio, and newspaper, which engage several thousands of participants per season [13,14]. Although large samples can often yield better surveillance power, participant recruitment and system maintenance can be very costly and challenging using these methods [15]. Recruitment via mass advertisements has also been found to underrepresent important populations who are at high risk of ILI and influenza, such as children and the elderly [14]. Recent virological studies have demonstrated that targeted disease surveillance, which is based on specific populations such as university students and military recruits, is not only cost-effective but also complementary to existing systems [16,17]. However, the effectiveness of internet-based surveillance systems using targeted samples has not been thoroughly explored.

Second, although existing studies have demonstrated that current internet-based systems have good surveillance power [13], they have not examined how the power may vary between different populations from a targeted surveillance approach. It is reasonable to assume that people with different occupations, age groups, or lifestyles may demonstrate distinct effects on surveillance power. For example, the elderly, children, and those who have not been vaccinated against influenza are likely to be more vulnerable to the seasonal epidemic. These populations may be more suitable for the early detection of influenza epidemics. Therefore, it is important to understand what populations will demonstrate the highest surveillance power, which can strengthen the targeted surveillance approach and reduce recruitment costs.

Third, although previous studies also investigated determinants of participation in influenza surveillance [18], many studies have not considered determinants that are specific to mobile-based app systems, which are increasingly being used by the public and should be utilized. Unlike websites that can be rigid in their use, mobile apps can be utilized at any location within the user's own timeframe [19]. This could assist in increasing participant engagement. However, mobile apps often require updates corresponding to the updating operating systems of mobile devices. Understanding these systematic pros and cons of different participatory approaches can highlight strategies to overcome the impact, increase engagement, and reduce maintenance costs.

To address the above limitations, the aim of this study was to present a mobile-based integrated surveillance system of ILI among health care workers, called FluMob, as a complementary system to the existing influenza surveillance in Singapore. FluMob is a digitally integrated syndromic surveillance system exclusively designed for health care workers in clinical settings (see detailed descriptions of the system in the Methods section) [20]. Health care workers are considered as a priority group for influenza preventive measures because they are at high risk of exposure to influenza. Health care workers account for 20%-30% of all influenza cases during regular influenza seasons worldwide [21,22]. In Singapore, this number can grow up to 40% in particular pandemics such as during the spread of severe

acute respiratory syndrome (SARS) in 2003 [23]. In addition, health care workers can cause nosocomial outbreaks [24,25], because they can act as potential vectors for influenza transmission to high-risk patients. Lack of preventive measures of influenza for health care workers can lead to higher mortality and morbidity of older and immunocompromised patients [26,27]. Therefore, real-time surveillance systems targeted at health care workers are imperative to prevent health care costs due to additionally infected patients and the loss of workforce among health care workers. Such surveillance systems are invaluable as they can also be easily adapted to emerging respiratory infections such as COVID-19.

In this study, we tested the FluMob system over 2 years from 2016 to 2018 in health care worker populations. Data were collected on demographics, lifestyle, and weekly reports of ILI symptoms and health care–seeking behaviors. This study had two main objectives: (1) to evaluate the effectiveness of FluMob regarding the participation rate among health care workers and its surveillance power for ILI, and (2) to identify factors that can moderate the effectiveness of the app. The study was disrupted for maintenance of the system and the end of the incentive. Although we tried to minimize the potential impact of this disruption by notifying participants about the app relaunch, it is important to investigate whether and how these disruptions would affect the level of participation and surveillance power of the participatory surveillance system. This knowledge will offer invaluable lessons for future research. Thus, this study examined the following research questions: (1) What is the level of participation of FluMob among health care workers, and what factors would moderate the participation rate? (2) What is the level of ILI surveillance power of FluMob among health care workers, and what factors would moderate the ILI surveillance power?

Methods

Design and Setting

This study was approved by the National Health Group Domain Specific Review Board (DSRB Ref: 2014/01009) and the SingHealth Centralised Institutional Review Board. From April 2016 to March 2018, the research team administered the FluMob system in two large hospitals in Singapore: a designated hospital for screening and treatment during communicable disease outbreaks (Tan Tock Seng Hospital) and another that deals with pediatric patients (KK Women's and Children's Hospital). Between May 7 and July 15 in 2017, the system was shut down for maintenance. Participants were prompted with notifications of the app relaunch. More information on the development of the app and participant engagement can be found in Lwin et al [20].

Intervention: FluMob System

FluMob is a digitally integrated syndromic surveillance system that integrates the ubiquitous access to the internet and the simple portability of smartphones. FluMob was developed using a mobile-based participatory epidemiological approach that relies on crowd involvement for disease surveillance and control through mobile technologies. Particularly, the system comprises a responsive web portal and mobile operating systems (ie, Android and iOS), allowing access from various types of mobile devices and web browsers. The FluMob app requires participants to register in the system upon their first use and to subsequently log in with a user identification and password. Health care workers can use the app easily and conveniently to provide reports of ILI on a weekly basis.

The system also has an analytic component for instantaneous surveillance. As soon as a report is submitted, the data input will be stored in a central database securely and confidentially. The research administrators (ie, the research team and other stakeholders) can access and view reports via an analytical module that is integrated into central servers. This allows researchers and clinicians to gather real-time surveillance of ILI, which can inform prevention and management strategies. A more detailed description of technical specifications and operating environments is available elsewhere [20].

Participants and Procedures

A convenience sampling method was employed to recruit participants. Clinical and nonclinical hospital staff from all departments were invited via mass emails to participate in the study in April 2016. Respondents who were above 21 years old and owned mobile devices in Android or iOS systems qualified and were asked to download the FluMob app from the corresponding software store for free. A total of 200 health care workers from KK Women's and Children's Hospital and 500 health care workers from Tan Tock Seng Hospital were recruited and given a unique user ID in the system for data tracking.

Participants were informed that they would participate in the study for 18 months, although the study continued for a further 6 months. No notification was sent to participants when their official study end dates approached. Participants were provided 20 Singapore dollars (US $15) as an incentive if they submitted 10 or more reports for each half of the first year. To evaluate whether the participants would continue to use the app without the external motivator, we stopped the token of appreciation after 1 year of app use (ie, they were not paid after the first year). Notably, as the time of the system maintenance and that of the end of the incentive for most participants were well aligned, their effects could not be distinguished. Therefore, we termed this period as the "study disruption period" to indicate disruptions from both the system maintenance and the incentive end.

Upon registration, participants completed a background survey capturing demographics and influenza vaccination status. The data collected included age, gender, ethnicity, number of children in a household, job category and department in the hospital, and influenza vaccination records in the previous year. Eight participants did not complete the registration and thus were excluded from the sample. Two participants were excluded from the final sample because they did not submit a full background survey, with missing demographic and lifestyle information. Finally, 690 participants were included in the analysis.

On a weekly basis, participants were prompted via app notifications to complete a short questionnaire of acute respiratory symptoms through the FluMob app. They reported

whether they had experienced any symptoms from a list since their last report, including fever, chills, runny or blocked nose, sneezing, cough, sore throat, shortness of breath, muscle or joint pain, headache, malaise, loss of appetite, colored phlegm, watery and bloodshot eyes, nausea, vomiting, diarrhea, stomach ache, and chest pain [28].

Analysis Strategies

Individual Submission Rate and Weekly Reporting Rate

Although participants could submit weekly reports even after their study end dates, we excluded entries for estimations of individual participation rates and weekly submission rates. The individual submission rate was calculated at the individual level as the percentage of submitted weekly surveys over the number of weeks in the 18-month period excluding the period of system maintenance. The weekly reporting rate was calculated as the percentage of reporting participants over the cumulative number of participants at the week level.

To investigate the effect of the study disruption, we compared the individual submission rates before and after the disruption with a paired-sample *t* test. To understand the effectiveness of the implementation of FluMob in health care workers, we conducted a multivariable analysis, with the individual submission rate as the outcome and individual factors as predictors in a linear regression model. This analysis was designed to compare individual submission rates across participants by demographics, hospital departments, and vaccination status. To better understand the characteristics of participants who committed to submit weekly reports throughout the study, regardless of the study disruption, we defined participants who submitted at least two reports after the system maintenance as committed users. A multivariable logistic regression analysis was conducted using participant characteristics as predictors. All analyses were conducted with SPSS version 25.0 (IBM Corp, Armonk, NY, USA).

ILI Incidence and Surveillance Power

ILI was defined as fever (≥38.0°C) accompanied by a cough or a sore throat following the definition of the Ministry of Health (MOH) in Singapore [2]. As participants could submit weekly reports after their end dates of the study, we utilized all entries for the ILI incidence estimations. ILI incidence in week *i* was calculated by dividing the number of participants reporting ILI in week *i* by the total number of reports in week *i*.

On less than 1% of occasions in which participants submitted multiple surveys within a week, the survey that reported respiratory symptoms, or the latest report if none of the surveys reported symptoms, was retained in the analysis. If participants reported ILI symptoms in consecutive weeks, the symptoms were considered to be from the same episode [14,28]. Thus, the reporting of ILI in the later consecutive weeks was removed from the calculation of weekly ILI incidence.

To evaluate the surveillance power of the FluMob system, we compared the weekly ILI incidence based on FluMob (ILI%_FluMob) with the weekly ILI incidence reported by the Singapore MOH (ILI%_MOH) [29]. The ILI%_MOH was estimated as the ILI incidence among attendances for acute upper respiratory infection in the government-funded primary care clinics. The Pearson correlation coefficient between the 4-week moving proportion of ILI%_FluMob and the 4-week moving proportion of ILI%_MOH was used to measure the surveillance power of the FluMob system.

To examine the effect of the study disruption on the surveillance power of the FluMob system, we compared the surveillance correlation between ILI%_FluMob and ILI%_MOH estimated by all data with that estimated by only data collected before the disruption. Furthermore, to investigate what factors can determine the surveillance power of the system, we compared the surveillance correlations across different individual characteristics.

Results

Participant Demographics and Characteristics

The participant flow diagram is shown in Figure 1. Table 1 shows the demographic and other characteristics of our sample. The majority of participants were women, aged under 40 years (531/690, 77.0%), and with no children in their households. Most of the participants reported receiving an influenza vaccination in the past year and had no allergic conditions.

Figure 1. Participant flow chart. TTSH: Tan Tock Seng Hospital; KKH: KK Women's and Children's Hospital.

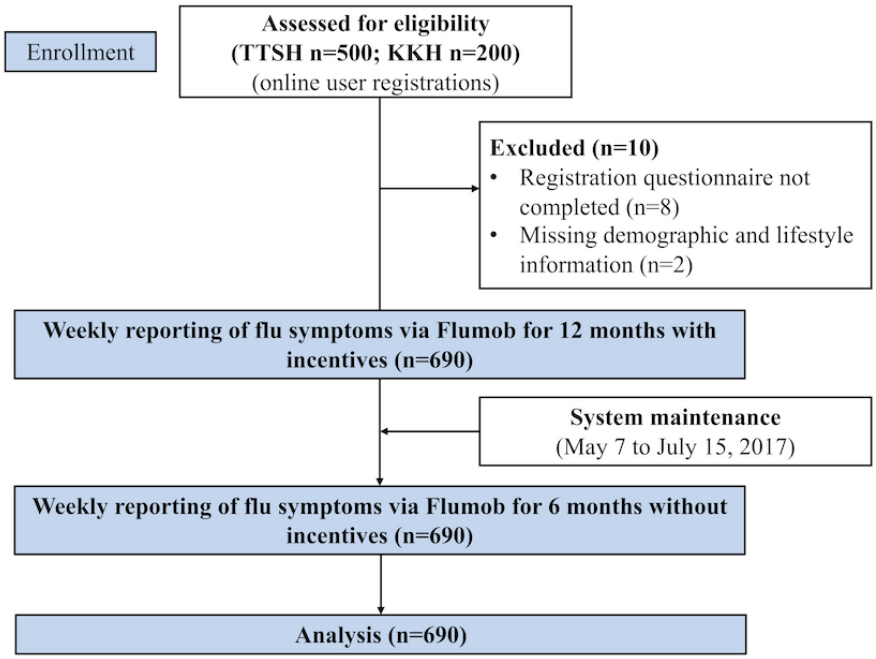


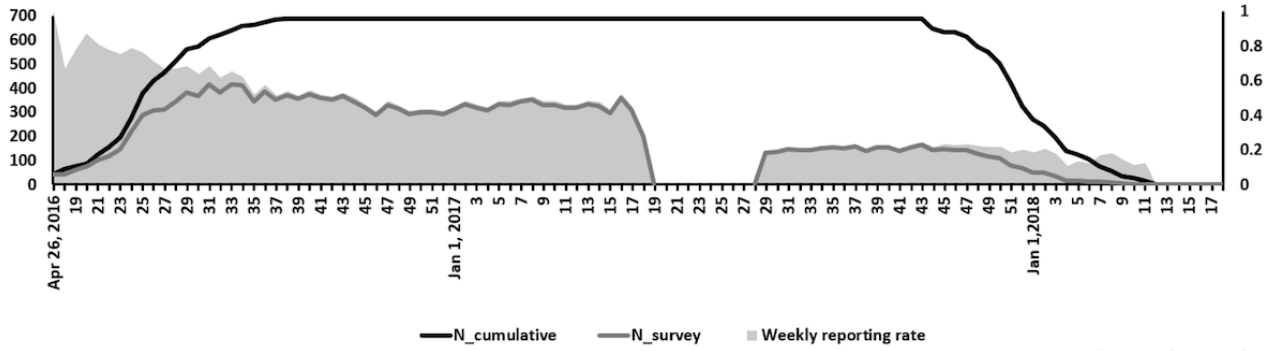
Table 1. Demographic and other characteristics of the health care workers recruited for this study (N=690).

Characteristics	Participants, n (%)
Gender	
Male	110 (15.9)
Female	580 (84.1)
Age (years)	
21-30	258 (37.4)
31-40	273 (39.6)
41-50	92 (13.3)
>50	67 (9.7)
Job category	
Administration and others	114 (16.5)
Ancillary	105 (15.2)
Allied health	186 (27.0)
Medical	58 (8.4)
Nursing	285 (41.3)
Ethnicity	
Chinese	389 (56.4)
Others	301 (43.6)
Children in household	
Yes	127 (18.4)
No	563 (81.6)
Influenza vaccine in the past year	
Yes	564 (81.7)
No	126 (18.3)
Allergy	
Yes	178 (25.8)
No	512 (74.2)

Level of Participation

Among the 690 health care workers recruited for this study, 671 (97.2%) submitted at least one weekly symptoms survey after registration. The weekly reporting rate stabilized at about 50% before the study disruption but decreased to around 20% after the disruption (Figure 2). On average, the individual submission rate was 41.4% (SD 24.3%), with a rate of 51.8% (SD 26.4%) before the disruption and of 21.5% (SD 30.6%) after the disruption. The paired *t* test revealed that the average individual submission rate before the disruption was significantly higher than that after the disruption (t_{689}=26.9, *P*<.001).

Figure 2. Weekly reporting rates among health care workers in FluMob. N_cumulative: cumulative sample size across time; N_survey: number of submitted surveys.



As shown in Table 2, the multivariate regression analysis indicated that participants aged below 30 years had lower individual submission rates than those aged above 30 years. Participants of Chinese ethnicity had a higher individual submission rate than participants of other ethnicities. Participants who had received an influenza vaccine in the year before the survey submitted more weekly reports than those who had not. No other significant results were found. In addition, to avoid the effect of the study disruption on the findings, we conducted the analysis for the period before the disruption, and similar results were observed (data not shown).

Table 2. Results of multivariable linear regression on individual submission rates by participant characteristics (N=690).

Characteristic	Adjusted submission rate	Unstandardized coefficient, estimate (SE)	*P* value
Gender			
Female (Reference)	0.39	N/A[a]	N/A
Male	0.40	0.009 (0.026)	.74
Age (years)			
21-30 (Reference)	0.31	N/A	N/A
31-40	0.40	0.087 (0.021)	<.001
41-50	0.46	0.149 (0.029)	<.001
>50	0.40	0.084 (0.033)	.01
Job category			
Nursing (Reference)	0.39	N/A	N/A
Administration and others	0.42	0.030 (0.035)	.38
Ancillary	0.44	0.043 (0.027)	.12
Allied health	0.37	–0.019 (0.023)	.41
Medical	0.35	–0.045 (0.036)	.22
Ethnicity			
Chinese (Reference)	0.44	N/A	N/A
Others	0.35	–0.096 (0.019)	<.001
Children in household			
No (Reference)	0.40	N/A	N/A
Yes	0.38	–0.008 (0.024)	.74
Influenza vaccine in the past 1 year			
No (Reference)	0.34	N/A	N/A
Yes	0.45	0.102 (0.023)	<.001
Allergy			
No (Reference)	0.40	N/A	N/A
Yes	0.39	–0.002 (0.020)	.92

[a]N/A: not applicable.

In total, 298 participants who submitted at least two reports after the system maintenance were identified as committed users. Multivariable logistic regression analysis (Table 3) showed that committed users were more likely to be above 30 years of age, ethnic Chinese, and those who were vaccinated in the previous year.

Table 3. Results of multivariable logistic regression of committed users by characteristics (N=298).

Characteristic	Committed users, n (%)	Adjusted odds ratio	95% CI	*P* value
Gender				
Female (Reference)	252 (43.5)	1.00	N/A[a]	N/A
Male	46 (41.8)	0.79	0.50-1.24	.31
Age (years)				
21-30 (Reference)	95 (36.8)	1.00	N/A	N/A
31-40	125 (45.8)	1.73	1.19-2.51	.004
41-50	49 (53.3)	2.22	1.34-3.69	.002
>50	29 (43.3)	1.36	0.77-2.40	.29
Job category				
Nursing (Reference)	113 (39.7)	1.00	N/A	N/A
Administration and others	31 (55.4)	1.79	0.98-3.26	.06
Ancillary	53 (50.5)	1.48	0.95-2.42	.08
Allied health	75 (40.3)	0.94	0.62-1.40	.74
Medical	26 (44.8)	1.12	0.60-2.09	.71
Ethnicity				
Chinese (Reference)	182 (47.0)	1.00	N/A	N/A
Others	115 (38.2)	0.59	0.42-0.83	.002
Children in household				
No (Reference)	248 (44.1)	1.00	N/A	N/A
Yes	50 (39.4)	0.73	0.48-1.10	.13
Flu vaccine in the past year				
No (Reference)	44 (34.9)	1.00	N/A	N/A
Yes	254 (45.0)	1.59	1.05-2.40	.03
Allergy				
No (Reference)	220 (43.0)	1.00	N/A	N/A
Yes	78 (43.8)	1.04	0.73-1.48	.83

[a]N/A: not applicable.

Level of ILI Surveillance Power

Overall, 25.1% of all participants reported at least one episode of ILI during the investigation period. On average, the weekly ILI incidence was 1.07%. Figure 3 shows the 4-week moving proportions of ILI%_FluMob and ILI%_MOH. Figure 4 is a scatter plot of these two 4-week moving proportions before and after the system maintenance. Table 4 shows that when including all data, there was no significant linear relationship between the 4-week moving proportion of ILI%_FluMob and ILI%_MOH. However, when only data before the study disruption were used for analysis, the 4-week moving proportion of ILI%_FluMob was significantly correlated with that of ILI%_MOH.

Figure 3. Comparisons between influenza-like illness (ILI) incidence estimated by FluMob (ILI%_FluMob) and that reported by the Singapore Ministry of Health (ILI%_MOH). Plotted values are 4-week moving proportions.

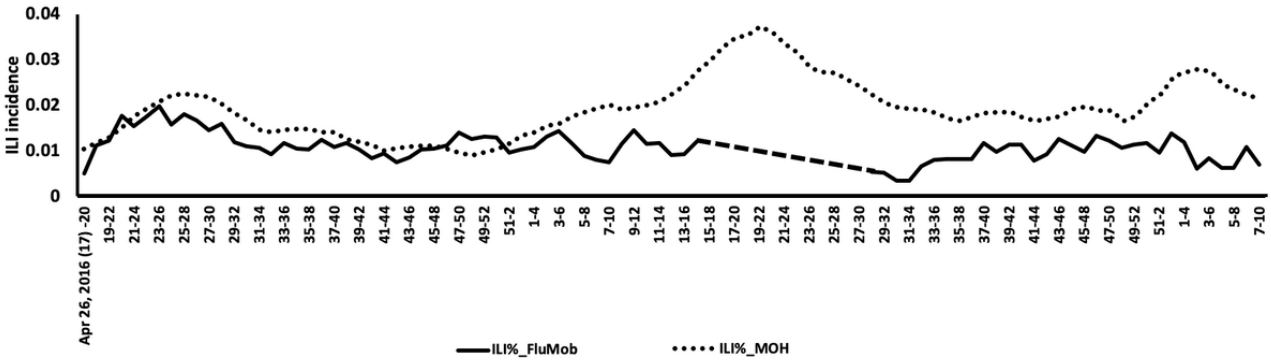


Figure 4. Scatter plot of percentage of influenza-like illness (ILI) estimated by FluMob and the Singapore Ministry of Health (MOH) using data collected before and after the system maintenance disruption.

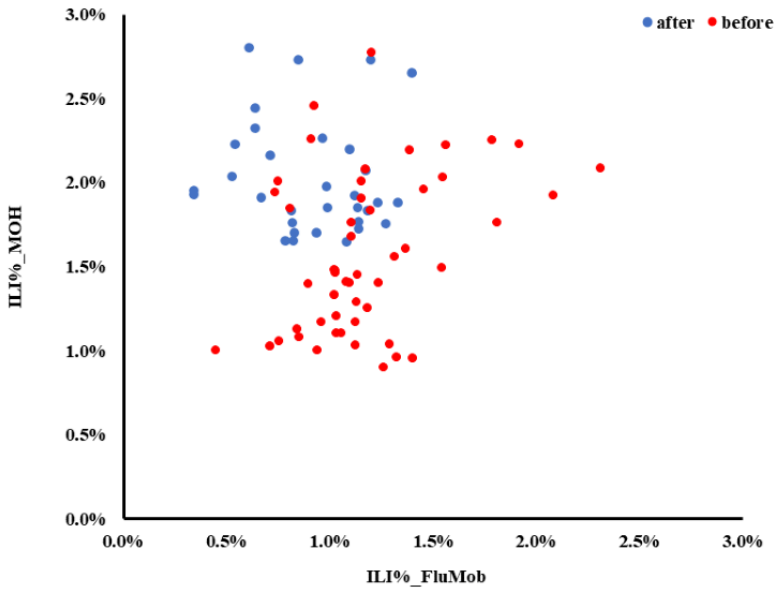


Table 4 also shows the surveillance correlations by individual characteristics. With data before the study disruption, participants who were female, aged 41-50 years, non-Chinese, and with children in their households had higher surveillance correlations between ILI%_FluMob and ILI%_MOH. In addition, participants who worked in allied health, had not received an influenza vaccine in the past year, and reported allergy demonstrated higher surveillance correlations.

Table 4. Pearson correlations coefficients (r) between 4-week moving proportions of influenza-like illness incidence estimated by FluMob and those of Singapore Ministry of Health.

Variables	All data		Data before maintenance	
	r	*P* value	*r*	*P* value
Overall	0.04	.75	0.38	.006
Gender				
Female	0.13	.25	0.46	.001
Male	–0.29	.007	–0.07	.62
Age (years)				
21-30	0.08	.47	0.31	.03
30-40	–0.16	.14	0.00	.99
41-50	0.35	.001	0.64	<.001
>50	–0.13	.24	–0.05	.73
Job category				
Administration and others	–0.11	.34	–0.23	.11
Ancillary	–0.16	.16	0.22	.12
Allied health	0.47	<.001	0.48	<.001
Medical	–0.21	.06	–0.19	.18
Nursing	–0.15	.17	0.19	.18
Ethnicity				
Chinese	–0.07	.53	0.10	.50
Others	0.12	.29	0.39	.004
Children in household				
Yes	0.12	.29	0.44	.001
No	–0.03	.79	0.25	.08
Influenza vaccine in the past year				
Yes	–0.14	.22	0.26	.06
No	0.52	<.001	0.45	.001
Allergy				
Yes	0.09	.42	0.44	.001
No	–0.03	.80	0.18	.20

Discussion

Principal Findings

The primary aim of participatory disease surveillance systems is to provide a cost-effective and timely method for controlling and preventing epidemics. Internet-based platforms have enabled the gathering of disease information from the general public. Mobile-based apps can strengthen these surveillance platforms by integrating multifunctional and interactive components, and by accessing hard-to-reach populations. Targeted surveillance and knowledge of determinants of the system effectiveness are important, as they facilitate the planning and implementation of recruitment and reduce management costs. Toward this end, the aim of this study was to evaluate the effectiveness of a mobile-based ILI surveillance system for health care workers (ie, FluMob) and to examine the determinants of the effectiveness.

Our study suggests that mobile-based systems are feasible for respiratory infections surveillance within hospitals. The findings showed that health care workers were generally committed to participating in the ILI surveillance. They submitted one report every other week on average within the first year of surveillance, and approximately half of the participants regularly submitted weekly reports. Demonstrating this feasibility of targeted samples such as health care workers in hospitals is of practical significance, as it implies that our mobile-based system can be translated into regular surveillance to prevent outbreaks of influenza and other emerging respiratory infections in health care settings. Importantly, as the current COVID-19 outbreak has recurred after the initial pandemic wave, it is imperative to

understand the effectiveness of our system to facilitate management of COVID-19 within the health care setting.

However, smooth operation of the system is essential to maintain its feasibility. In this study, only one-fifth of participants submitted weekly reports after the study disruption that included both a 10-week downtime of the system and the end of incentives. Unfortunately, we were not able to disentangle the effects of system maintenance and incentive end on participation due to the two events occurring consecutively. Nevertheless, these findings suggest that future research should consider the possible impacts of system maintenance and incentives on participation in surveillance programs.

The study findings highlight the potential of a targeted approach in participatory surveillance in several ways. First, our analyses show that individual characteristics are important determinants of participation. Greater participation was associated with older age and being vaccinated against influenza, which is consistent with past evidence suggesting that individuals with these characteristics are more concerned about their personal health and thus more likely to participate [12,18]. One novel finding was that being non-Chinese was associated with lower participation of ILI surveillance. This could be because of ethnic disparities in socioeconomic status and in turn health care–seeking behaviors in Singapore that favor the major Chinese group over the minority groups [30,31].

Second, these results demonstrate that mobile-based ILI surveillance of the targeted health care worker sample can be complementary to traditional surveillance based on the general public. The ILI incidence estimated by FluMob had a moderate correlation with that of the Singapore MOH within the first year of surveillance. However, the study disruption significantly weakened the surveillance correlation. We speculate that those who continued to submit weekly reports after the study disruption might be more health-conscious, which thus biased the ILI incidence estimations.

Third, and importantly, further analyses demonstrated that surveillance effectiveness could be distinct across different populations, thereby highlighting the potential of using a targeted sample for early detection of disease epidemics. Specifically, women had a higher power of ILI surveillance than men, possibly because women are more vulnerable to influenza due to physiological differences in sex [30,32]. Older age, not being vaccinated, and having allergic conditions were associated with higher surveillance power of ILI, which is likely because these are risk factors for influenza and other types of acute respiratory infections [28,33]. Thus, participants in those groups are more vulnerable to influenza and ILI epidemics. Furthermore, the ILI incidence estimated from health care workers who had children in their household demonstrated a higher predictive correlation with the ILI incidence reported by the MOH than those who did not, suggesting that they might be more exposed to illness via their young children [28]. Overall, the above findings suggest that demographic groups with higher risks of ILI and influenza would be more effective to target in ILI surveillance. This finding may be critical in future participatory surveillance research, as it facilitates the planning and implementation of sample recruitments using a targeted approach.

Future studies should include medical or laboratory confirmation of ILI to further boost accurate surveillance. However, and interestingly, the ILI incidence estimated by health care workers within FluMob was at a similar level to that reported by the MOH. This finding differs from previous studies showing that ILI incidence estimated by internet-based systems was often 3 to 10 times higher than that of the government reports [13,14,34]. One possible reason for this finding is that health care workers, especially those in the FluMob sample who work in hospitals for communicable disease and pediatric patients, are knowledgeable about acute respiratory symptoms.

Although our findings have significant theoretical and practical implications in ILI surveillance, the study has some limitations. The surveillance correlations found in this study (ie, r=0.0-0.4) were relatively low compared with those in previous internet-based surveillance studies [13,14]. Owing to the targeted approach, a smaller sample size was recruited in comparison to past research, which would inevitably lower the surveillance power. Furthermore, the system had a major disruption during the investigation. The downtime caused a loss of data during a big seasonal influenza outbreak in 2017, along with discouragement of participation and ultimately the loss of surveillance power. If such data were available, stronger surveillance correlations would have been obtained. Hence, such disruptions should be avoided with better study planning, funding support, and technological support. However, this disruption also offered us a valuable opportunity to investigate the impact of such breaks in accessibility on the effectiveness of mobile-based participatory ILI surveillance. Overall, this study was able to demonstrate a significant surveillance correlation, suggesting that the targeted surveillance approach can be an advantageous and complementary approach for participatory infectious disease surveillance.

Conclusion

This study is among the first to evaluate the effectiveness of a mobile-based participatory ILI surveillance system and its determinants regarding participation engagement and surveillance power from a targeted health surveillance approach. The findings suggest that mobile-based systems can be effective for participatory surveillance. However, smooth operation of the mobile app without major disruptions is vital for the engagement of participants and the persistence of surveillance power. In addition, the effectiveness of the mobile-based system can be moderated by participants' characteristics, which highlights the importance of targeted disease surveillance that can reduce the cost of recruitment and engagement. Our findings have significant theoretical and practical implications on participatory health surveillance based on mobile phones. Future research should identify factors that can improve the effectiveness of mobile-based participatory apps for infectious disease surveillance.

Acknowledgments

This research was supported by the Singapore MOH's National Medical Research Council under its Communicable Diseases—Public Health Research Grant (CDPHRG13NOV020). The authors would like to acknowledge the contribution of broader team members, Huarong Xu and Jie Chen, who helped with data collection in the hospitals.

Authors' Contributions

ML, JL, AS, CP, YT, PY, VC, BA, and MC were involved in conceptualization of the study. JL and AS wrote the main sections of the paper. ML, JL, AS, YT, VC, MC, and LWA were involved in the overall editing. CY and KT were involved in data collection. BA is the principal investigator and overall coordinator of the project.

Conflicts of Interest

None declared.

References

1. Influenza (Seasonal). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) [accessed 2020-11-14]
2. Ang LW, Tien WS, Lin RT, Cui L, Cutter J, James L, et al. Characterization of influenza activity based on virological surveillance of influenza-like illness in tropical Singapore, 2010-2014. J Med Virol 2016 Dec 10;88(12):2069-2077. [doi: 10.1002/jmv.24566] [Medline: 27152935]
3. Saha S, Chadha M, Al Mamun A, Rahman M, Sturm-Ramirez K, Chittaganpitch M, et al. Influenza seasonality and vaccination timing in tropical and subtropical areas of southern and south-eastern Asia. Bull World Health Organ 2014 May 01;92(5):318-330 [FREE Full text] [doi: 10.2471/BLT.13.124412] [Medline: 24839321]
4. Lee VJ, Tok MY, Chow VT, Phua KH, Ooi EE, Tambyah PA, et al. Economic analysis of pandemic influenza vaccination strategies in Singapore. PLoS One 2009 Sep 22;4(9):e7108 [FREE Full text] [doi: 10.1371/journal.pone.0007108] [Medline: 19771173]
5. Ng TP, Pwee KH, Niti M, Goh LG. Influenza in Singapore: assessing the burden of illness in the community. Ann Acad Med Singap 2002 Mar;31(2):182-188. [Medline: 11957555]
6. Ang LW, Lim C, Lee VJM, Ma S, Tiong WW, Ooi PL, et al. Influenza-associated hospitalizations, Singapore, 2004-2008 and 2010-2012. Emerg Infect Dis 2014 Oct;20(10):1652-1660. [doi: 10.3201/eid2010.131768] [Medline: 25275710]
7. Tan JPT, Freathy P. Consumer decision making and store patronage behaviour in Traditional Chinese Medicine (TCM) halls in Singapore. J Retail Consum Serv 2011 Jul;18(4):285-292. [doi: 10.1016/j.jretconser.2011.02.002]
8. Chui WK, Li SC. Advice-giving on self-medication: perspectives of community pharmacists and consumers in Singapore. J Clin Pharm Ther 2005 Jun;30(3):225-231. [doi: 10.1111/j.1365-2710.2005.00637.x] [Medline: 15896239]
9. Quah SR. Self-medication in Singapore. Singapore Med J 1985 Apr;26(2):123-129. [Medline: 4035394]
10. Blaya JA, Fraser HSF, Holt B. E-health technologies show promise in developing countries. Health Aff (Millwood) 2010 Feb;29(2):244-251. [doi: 10.1377/hlthaff.2009.0894] [Medline: 20348068]
11. Dalton CB, Carlson SJ, Butler MT, Elvidge E, Durrheim DN. Building influenza surveillance pyramids in near real time, Australia. Emerg Infect Dis 2013 Nov;19(11):1863-1865. [doi: 10.3201/eid1911.121878] [Medline: 24207165]
12. Debin M, Turbelin C, Blanchon T, Bonmarin I, Falchi A, Hanslik T, et al. Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. PLoS One 2013 Sep 11;8(9):e73675 [FREE Full text] [doi: 10.1371/journal.pone.0073675] [Medline: 24040020]
13. van Noort SP, Codeço CT, Koppeschaar CE, van Ranst M, Paolotti D, Gomes MGM. Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. Epidemics 2015 Dec;13:28-36 [FREE Full text] [doi: 10.1016/j.epidem.2015.05.001] [Medline: 26616039]
14. Friesema IHM, Koppeschaar CE, Donker GA, Dijkstra F, van Noort SP, Smallenburg R, et al. Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in The Netherlands. Vaccine 2009 Oct 23;27(45):6353-6357. [doi: 10.1016/j.vaccine.2009.05.042] [Medline: 19840672]
15. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, et al. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clin Microbiol Infect 2014 Jan;20(1):17-21 [FREE Full text] [doi: 10.1111/1469-0691.12477] [Medline: 24350723]
16. Seah SG, Lim EA, Kok-Yong S, Liaw JC, Lee V, Kammerer P, et al. Viral agents responsible for febrile respiratory illnesses among military recruits training in tropical Singapore. J Clin Virol 2010 Mar;47(3):289-292 [FREE Full text] [doi: 10.1016/j.jcv.2009.12.011] [Medline: 20079688]
17. Tan AL, Virk RK, Tambyah PA, Inoue M, Lim EA, Chan K, et al. Surveillance and clinical characterization of influenza in a university cohort in Singapore. PLoS One 2015 Mar 19;10(3):e0119485 [FREE Full text] [doi: 10.1371/journal.pone.0119485] [Medline: 25790305]

18. Bajardi P, Vespignani A, Funk S, Eames KT, Edmunds WJ, Turbelin C, et al. Determinants of follow-up participation in the Internet-based European influenza surveillance platform Influenzanet. J Med Internet Res 2014 Mar 10;16(3):e78 [FREE Full text] [doi: 10.2196/jmir.3010] [Medline: 24613818]
19. Dimmick J, Feaster JC, Hoplamazian GJ. News in the interstices: The niches of mobile media in space and time. New Media Soc 2010 May 18;13(1):23-39. [doi: 10.1177/1461444810363452]
20. Lwin MO, Yung CF, Yap P, Jayasundar K, Sheldenkar A, Subasinghe K, et al. FluMob: Enabling Surveillance of Acute Respiratory Infections in Health-care Workers Mobile Phones. Front Public Health 2017;5:49. [doi: 10.3389/fpubh.2017.00049] [Medline: 28367433]
21. Kawana A, Teruya K, Kirikae T, Sekiguchi J, Kato Y, Kuroda E, et al. "Syndromic surveillance within a hospital" for the early detection of a nosocomial outbreak of acute respiratory infection. Jpn J Infect Dis 2006 Dec;59(6):377-379 [FREE Full text] [Medline: 17186956]
22. Taylor G, Mitchell R, McGeer A, Frenette C, Suh KN, Wong A, Canadian Nosocomial Infection Surveillance Program. Healthcare-associated influenza in Canadian hospitals from 2006 to 2012. Infect Control Hosp Epidemiol 2014 Feb 10;35(2):169-175. [doi: 10.1086/674858] [Medline: 24442080]
23. Koh D, Lim MK, Chia SE, Ko SM, Qian F, Ng V, et al. Risk perception and impact of Severe Acute Respiratory Syndrome (SARS) on work and personal lives of healthcare workers in Singapore: what can we learn? Med Care 2005 Jul;43(7):676-682. [doi: 10.1097/01.mlr.0000167181.36730.cc] [Medline: 15970782]
24. Pollara C, Piccinelli G, Rossi G, Cattaneo C, Perandin F, Corbellini S, et al. Nosocomial outbreak of the pandemic Influenza A (H1N1) 2009 in critical hematologic patients during seasonal influenza 2010-2011: detection of oseltamivir resistant variant viruses. BMC Infect Dis 2013 Mar 07;13:127 [FREE Full text] [doi: 10.1186/1471-2334-13-127] [Medline: 23496867]
25. Carnicer-Pont D, White D, Pike C, Lyons M. Influenza A outbreak in a community hospital in south east Wales, February 2005. Euro Surveill 2005 Feb 17;10(2):E050217.2. [doi: 10.2807/esw.10.07.02645-en] [Medline: 16685099]
26. Potter J, Stott DJ, Roberts MA, Elder AG, O'Donnell B, Knight PV, et al. Influenza vaccination of health care workers in long-term-care hospitals reduces the mortality of elderly patients. J Infect Dis 1997 Jan 01;175(1):1-6 [FREE Full text] [doi: 10.1093/infdis/175.1.1] [Medline: 8985189]
27. Gruneir A, Kwong JC, Campitelli MA, Newman A, Anderson GM, Rochon PA, et al. Influenza and seasonal patterns of hospital use by older adults in long-term care and community settings in Ontario, Canada. Am J Public Health 2014 Feb;104(2):e141-e147. [doi: 10.2105/AJPH.2013.301519] [Medline: 24328631]
28. Adler AJ, Eames KT, Funk S, Edmunds WJ. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey. BMC Infect Dis 2014 May 01;14(1):232 [FREE Full text] [doi: 10.1186/1471-2334-14-232] [Medline: 24885043]
29. Weekly infectious disease bulletin. Ministry of Health Singapore. URL: https://www.moh.gov.sg/resources-statistics/infectious-disease-statistics/2020/weekly-infectious-diseases-bulletin [accessed 2020-11-14]
30. Lim RBT, Zheng H, Yang Q, Cook AR, Chia KS, Lim WY. Ethnic and gender specific life expectancies of the Singapore population, 1965 to 2009 - converging, or diverging? BMC Public Health 2013 Oct 26;13(1):1012 [FREE Full text] [doi: 10.1186/1471-2458-13-1012] [Medline: 24160733]
31. Chen T, Tan PJ, Chan A. Ethnicity predicts falls among community-dwelling older adults in Singapore. Geriatr Gerontol Int 2018 Jan 30;18(1):72-79. [doi: 10.1111/ggi.13143] [Medline: 28857421]
32. Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. J Leukoc Biol 2012 Jul 01;92(1):67-73 [FREE Full text] [doi: 10.1189/jlb.0811427] [Medline: 22131346]
33. Eames KTD, Brooks-Pollock E, Paolotti D, Perosa M, Gioannini C, Edmunds WJ. Rapid assessment of influenza vaccine effectiveness: analysis of an internet-based cohort. Epidemiol Infect 2012 Jul;140(7):1309-1315. [doi: 10.1017/S0950268811001804] [Medline: 21906412]
34. Tilston NL, Eames KTD, Paolotti D, Ealden T, Edmunds WJ. Internet-based surveillance of Influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic. BMC Public Health 2010 Oct 27;10:650 [FREE Full text] [doi: 10.1186/1471-2458-10-650] [Medline: 20979640]

Abbreviations

ILI: influenza-like illness
MOH: Ministry of Health
SARS: severe acute respiratory syndrome

Edited by G Eysenbach; submitted 29.04.20; peer-reviewed by W Evans, J Chan; comments to author 22.06.20; revised version received 29.07.20; accepted 30.10.20; published 07.12.20.

Please cite as:
Lwin MO, Lu J, Sheldenkar A, Panchapakesan C, Tan YR, Yap P, Chen MI, Chow VTK, Thoon KC, Yung CF, Ang LW, Ang BSP
Effectiveness of a Mobile-Based Influenza-Like Illness Surveillance System (FluMob) Among Health Care Workers: Longitudinal Study
JMIR Mhealth Uhealth 2020;8(12):e19712
URL: https://mhealth.jmir.org/2020/12/e19712
doi:10.2196/19712
PMID:33284126

Original Paper

Effective Use of Mobile Electronic Medical Records by Medical Interns in Real Clinical Settings: Mixed Methods Study

SuJin Kim[1,2], MS; Seulji Ku[2], BA; Taerim Kim[3], MD, PhD; Won Chul Cha[1,3,4], MD; Kwang Yul Jung[1,5], MD

[1]Department of Digital Health, Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[2]Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea

[3]Department of Emergency Medicine, Samsung Medical Center, Seoul, Republic of Korea

[4]Health Information and Strategy Center, Samsung Medical Center, Seoul, Republic of Korea

[5]Department of Emergency Medicine, Ewha Womans University Seoul Hospital, Seoul, Republic of Korea

Corresponding Author:
Kwang Yul Jung, MD
Department of Emergency Medicine
Ewha Womans University Seoul Hospital
260 Gonghang-daero
Gangseo-gu
Seoul, 07804
Republic of Korea
Phone: 82 10 9183 0491
Email: mdyul414@gmail.com

Abstract

Background: In South Korea, most graduated medical students undertake a 1-year internship before beginning residency and specialization. Interns usually work in a tertiary hospital and rotate between different, randomly assigned departments to be exposed to different medical specialties. Their jobs are mostly simple and repetitive but are still essential for the patient care process. However, owing to the lack of experience and overwhelming workload, interns at tertiary hospitals in South Korea are usually inefficient, often delaying the entire clinical process. Health care providers have widely adopted mobile electronic medical records (mEMRs) as they have been shown to improve workflow efficiency.

Objective: This study investigates the association between the frequency of mEMR usage and the clinical task completion interval time among interns in a tertiary hospital.

Methods: This mixed methods study was conducted at the Samsung Medical Center, Seoul, South Korea. Interns who worked at the Samsung Medical Center from March 2018 to February 2019 were included. The hospital electronic medical record (EMR) system known as DARWIN (Data Analysis and Research Window for Integrated kNowledge) was launched with PC and mobile. Both versions are actively used in hospitals by personnel in various positions. We collected the log data from the mEMR server and the intern clinical task time-series data from the EMR server. Interns can manage the process of identifying patients, assigning the clinical task, finishing the requested clinical intern tasks, etc, through the use of the mEMR system. We compared the clinical task completion interval among 4 groups of interns divided by the mEMR frequency quantile. Then, System Usability Score (SUS) questionnaires and semistructured interviews were conducted.

Results: The regular mEMR users were defined as those who logged in more than once a day on average and used the mEMR until the level after login. Among a total of 87 interns, 84 used the mEMR to verify the requested clinical tasks. The most frequently used item was “Intern task list.” Analysis of the 4 intern groups revealed an inverse relationship between the median time of the task completion interval and the frequency of mEMR use. Correlation analysis showed that the intern task completion time interval had a significant inverse relationship with the individual frequency of mEMR usage (coefficient=-0.27; 95% CI -0.46 to -0.04; *P*=.02). In the additional survey, the mean SUS value was 81.67, which supported the results of the data analysis.

Conclusions: Our findings suggest that frequent mEMR use is associated with improved work efficiency in hospital interns with good usability of the mEMR. Such finding supports the idea that the use of mEMR improves the effectiveness and workflow efficiency of interns working in hospitals and, more generally, in the context of health care.

(JMIR Mhealth Uhealth 2020;8(12):e23622) doi:10.2196/23622

KEYWORDS

mobile health; mobile EMR; intern; workflow; mHealth; electronic medical record; electronic health record; EHR; EMR; efficiency

Introduction

Background and Significance

Professionals of various occupations, such as doctors, nurse pharmacists, and other supporters, provide patient care in hospitals. In most tertiary hospitals in South Korea, prescribers, such as specialists and residents, determine the appropriate care plan and use computerized provider order entry (CPOE) to order prescriptions. Then a nurse executes the order or passes it on to interns or other supporters. This computerized linear workflow benefits workflow and patient safety [1,2]; however, when overloaded, it inevitably results in inefficiency and delay [3].

Internship is the transition period between being a medical student and becoming a specialty doctor [4]. In South Korea, most graduated medical students undertake a 1-year internship with their physician's license before beginning their specialty resident course [5]. Interns usually work in a tertiary hospital, where they rotate monthly among different, randomly assigned departments to be exposed to different medical specialties. Their jobs are mostly simple and repetitive, but they are essential for the patient care process. These jobs include simple procedures (eg, catheterization, biopsies, monitoring, and sampling), documentation (eg, getting consent forms for radiology or procedures), and prescriptions that do not affect patient care plans (eg, meal changes, simple dressing) [6].

The lack of experience and the workload of interns at tertiary hospitals, where patient needs are substantial and often overwhelming, make interns inevitably susceptible to inefficiency and fatigue [7]. Such inefficiency can halt the entire clinical process and expose patients to the risk of errors. This can ultimately have an adverse influence on patient care and safety [8].

Mobile Electronic Medical Records for Health Care Providers

Before the use of smartphones, interns usually received notifications pertaining to their jobs through various mediums (depending on the policies of each hospital), such as pagers, phone calls, or beepers for SMS texting exclusively used in the hospital communication system [9]. With the widespread use of smartphones, health care professionals have widely adopted mobile electronic medical records (mEMRs) [10-12]. The mEMR has been shown to improve the efficiency and effectiveness of hospital workflow in previous studies [13,14]. However, none of the previous studies evaluated the time efficiency of interns' job achievement in a clinical setting using mEMRs.

Study Objective

This study aims to determine the association between interns' clinical task completion time interval and the frequency of mEMR usage.

Methods

Study Setting

This mixed methods study was conducted at the Samsung Medical Center, Seoul, South Korea. We targeted and analyzed interns who worked at the Samsung Medical Center from March 2018 to February 2019. In South Korea, from the month of March to the following February, interns rotate between various medical departments. To examine the association between mEMR usage and intern performance, we collected the log data from the mEMR server and the intern clinical task time-series data from Samsung Medical Center's Electronic Medical Record (EMR). The study protocol was reviewed and approved by the institutional review board (IRB) of Samsung Medical Center (IRB No. SMC 2019-09-122-001).

Mobile Electronic Medical Records

In July 2016, the hospital EMR system known as DARWIN (Data Analysis and Research Window for Integrated kNowledge) was launched. DARWIN has both PC and mobile versions. DARWIN is actively used in hospitals, and its mobile version is used by hospital personnel in various positions. Mobile DARWIN (mDARWIN) includes a main menu, list-level features, and patient-level features. After login into the mDARWIN, users can select a list-level feature on the first screen from 8-9 main menus.

Interns' Clinical Task Implementation Process

There are 3 types of prescriptions that prescribers such as specialists or residents issue: (1) basic prescription (eg, vital-sign check term, input and output check term, meal, or simple daily care service for postoperation patients, (2) medication prescription, and (3) examination prescription. These prescriptions have associated tasks that are performed by health care providers. With the exception of the tagging of prescriptions (which is performed by nurses), most clinical tasks are performed by interns. When interns receive an alarm about a new task on their mobile device, they verify the clinical task and self-assign the prescriptions to themselves. Then, they conduct the clinical task according to the instruction (Figure 1). Interns can manage this process (ie, identify, assign, and mark the task as complete after finishing the requested clinical task) through the mDARWIN mEMR (Figure 2).

Figure 1. Flowchart of an intern's clinical task process.

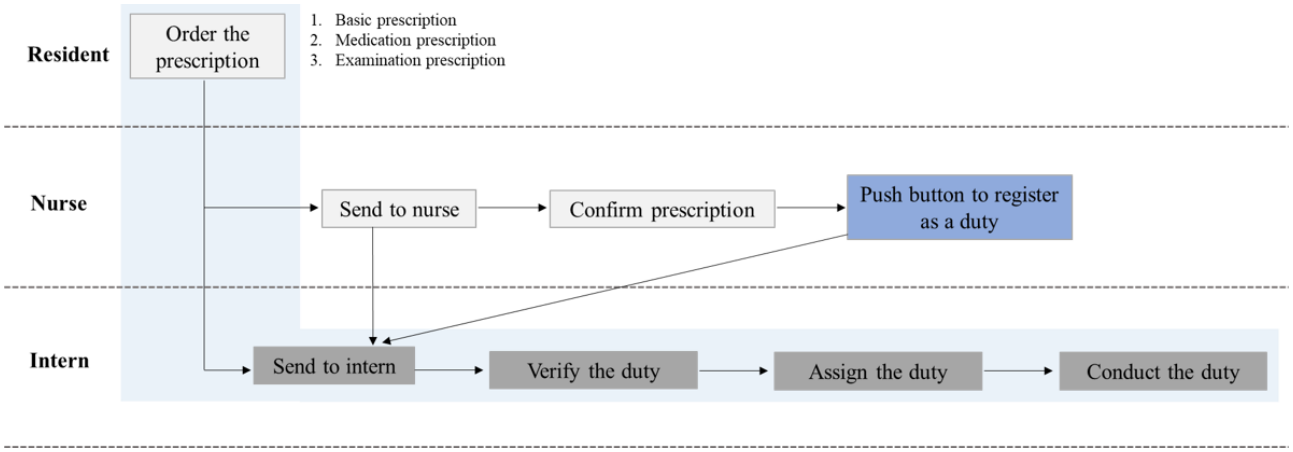
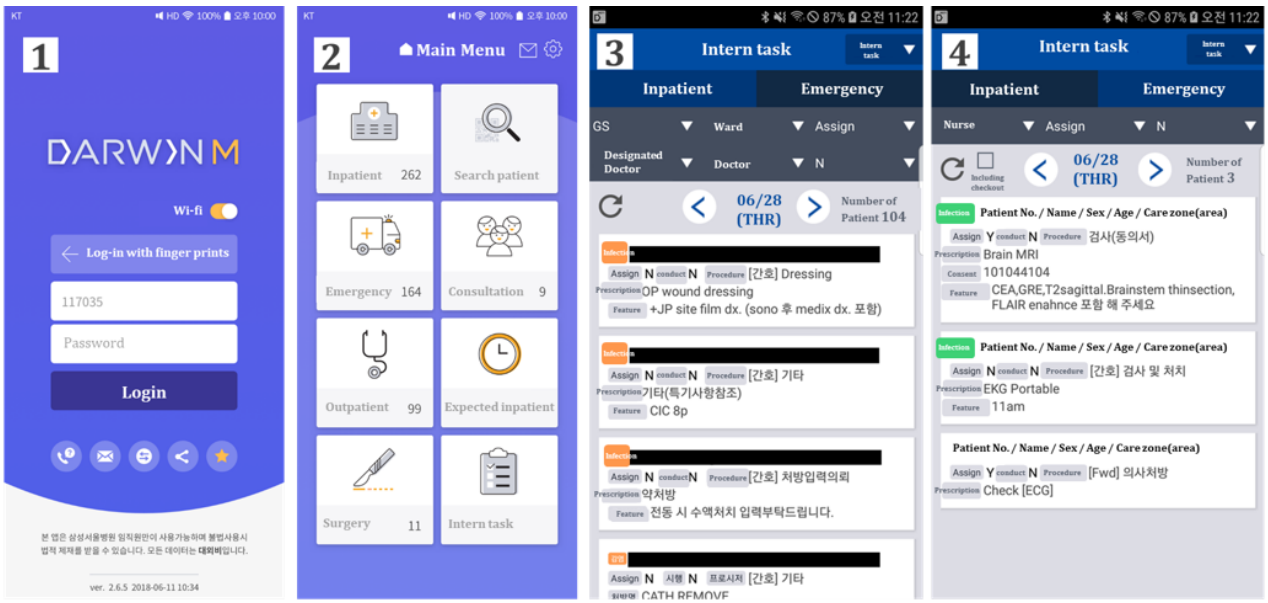


Figure 2. Screenshot of intern tasks from Mobile DARWIN (mDARWIN).



*Screenshots 3 and 4 do not show real patient information. These are prototype screenshots used as examples. These pages show the list of intern tasks for inpatients and emergency patients.

Outcome Measures and Sensitivity Analysis

The primary outcome was the comparison of the time interval to complete the intern tasks after dividing the interns into 4 groups based on the quantile of the frequency of mEMR usage. The definition of the task completion interval time was set from the time the task was requested to the task completion check time. For sensitivity analysis, we verified the correlation between the frequency of mEMR usage and the median time of interval to complete the intern's tasks individually. Subsequently, System Usability Score (SUS) questionnaires were administered and analyzed.

Survey

To investigate the feasibility of the mEMR in a clinical environment, SUS questionnaires were administered to interns [15]. Interns were recruited through a notice posted on the bulletin board in the hospital for 2 months. In addition to the survey, interns were also interviewed at the same time. The survey included 10 questions, scored using a 5-point Likert scale (1=strongly disagree, 5=strongly agree). The SUS calculation formula is as follows:

During the interview, interns were asked questions such as when they mostly used the mEMR, which list they searched the most, where they mostly used the mEMR, and whether mDARWIN helps with their tasks. The interviews were semistructured.

Statistical Analysis

We investigated the log data of interns' mEMRs during the study period. We compared the task completion time interval among the 4 groups using statistical analysis. We compared the task completion interval's median time and 95% confidence interval between the 4 groups.

For sensitivity analysis, we evaluated the correlation between the frequency of log data of interns' mEMR and the individual task completion time interval of interns using the Pearson product-moment correlation coefficient test. *P* values of <.05 were considered statistically significant. All data analyses were

performed using R software (version 3.4.2; R Project for Statistical Computing).

Results

Characteristics of the Subjects

In total, 87 interns performed intern tasks during the study period. A total of 1,081,413 tasks were performed by these interns. Of the 87 interns, 84 regularly used the mEMR and were included in the analysis. However, 3 interns were excluded because 2 had not used mEMR at all and 1 had a total of only 4 log records during the study period; thus, they were considered nonusers. In this context, regular mEMR users were defined as those who logged in more than once a day on average and used the mEMR until the next level after login. Table 1 shows the intern information included in the study and the clinical tasks they received.

Table 1. Information about the study subjects during the study period (n=87).

Participant characteristics	Values, n (%)
Medical interns (n=87)	
mEMR[a] users	84 (97)
Non-mEMR[a] users	3 (3)
Total intern clinical tasks performed (n=1,081,413)	
By location	
Inpatient	940,338 (87.00)
Outpatient	1336 (0.10)
Emergency	139,739 (12.90)
By department	
Medical part	462,018 (42.70)
Surgical part	478,242 (44.20)
Other hospital-based part	141,153 (13.10)
By procedure category	
Request order transcription (from nurse)[b]	348,805 (32.30)
Request order transcription (from doctor) [b]	163,886 (15.20)
Diagnostic test consent form	170,542 (15.80)
Wound dressing	134,503 (12.40)
Diagnostic test	94,596 (8.70)
Diagnostic test and treatment	30,521 (2.80)
Catheter tube insertion	16,058 (1.50)
Administrative paperwork	14,533 (1.30)
Irrigation	14,305 (1.30)
Influenza exam	5938 (0.50)
Enema	3245 (0.30)
Writing slip	2733 (0.30)
Inject medicine	959 (0.10)
Other	80,789 (7.50)

[a]mEMR: mobile electronic medical records.

[b]South Korea's medical system adopts a fee-for-service model for medical service. As the prescription order can only be authorized by a doctor, this category is in relation to the prescription after the act of the nurse or doctor.

Log Data Analysis

During the study period, 489,444 mEMR logs were created by interns. Interns used a total of 43 items within the mEMR, as shown in Multimedia Appendix 1. From the 489,444 logs, 67,147 logs were made in a list-level feature. Among these records, “Intern task list” topped the list with 39,506 tasks. This was followed by “My patient list” and “Surgery history list,” as shown in Table 2.

Table 2. Total number of logs with a list-level feature (n=67,147).

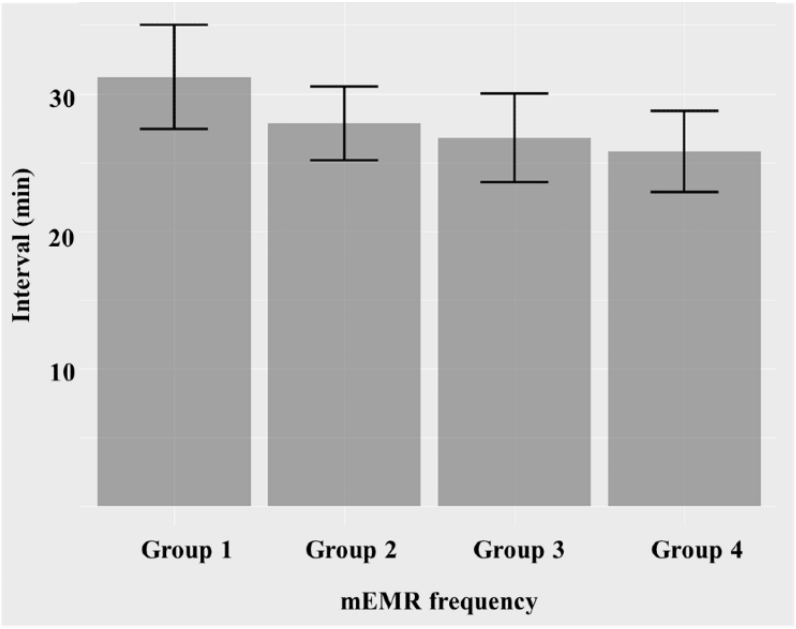
No.	List	Frequency, n (%)
1	Intern task list	39,506 (58.80)
2	My patient list	13,685 (20.40)
3	Surgery history list	8,545 (12.70)
4	Inpatient list	3,963 (5.90)
5	Emergency patient list	663 (1.00)
6	Integrated-view EMR[a] list	241 (0.40)
7	Outpatient list	137 (0.20)
8	Consultation list	110 (0.20)
9	Expected inpatient list	78 (0.10)
10	Scheduled surgery list	46 (0.10)
11	Patient search (through patient ID)	15 (0.00)
12	Other	158 (0.20)

[a]EMR: electronic medical record.

Statistical Outcomes

The comparison of clinical task completion interval consistently showed an inverse relationship between the median time of completion interval and the frequency of mEMR usage (Figure 3). The frequent mEMR user group took a shorter time to complete the requested tasks. Sensitivity analysis with the Pearson product-moment correlation coefficient showed that the intern task completion interval time had a significant inverse relationship with individual frequency of mEMR usage (coefficient -0.27; 95% CI -0.46 to -0.04; *P*=.02). Using the mEMR once reduced the task completion time by approximately 16 seconds (*P*=.02).

Figure 3. Task completion time interval and frequency of mobile electronic medical record (mEMR) usage among 4 intern groups divided by quantile of mEMR usage.



(Group 1: under 2,100 usage; Group 2: 2,101 to 4,125 usage;

Group 3: 4126 to 6,029 usage; Group 4: 6,030 to 13,290 usage)

SUS Survey Outcome

A total of 15 interns completed the SUS survey from December 2019 to January 2020. The mean SUS value for the intern clinical task item in the mEMR was 81.67 (Table 3). The interview and survey were conducted at the same time. Figure 4 shows the key points in the interview that may be useful for future research.

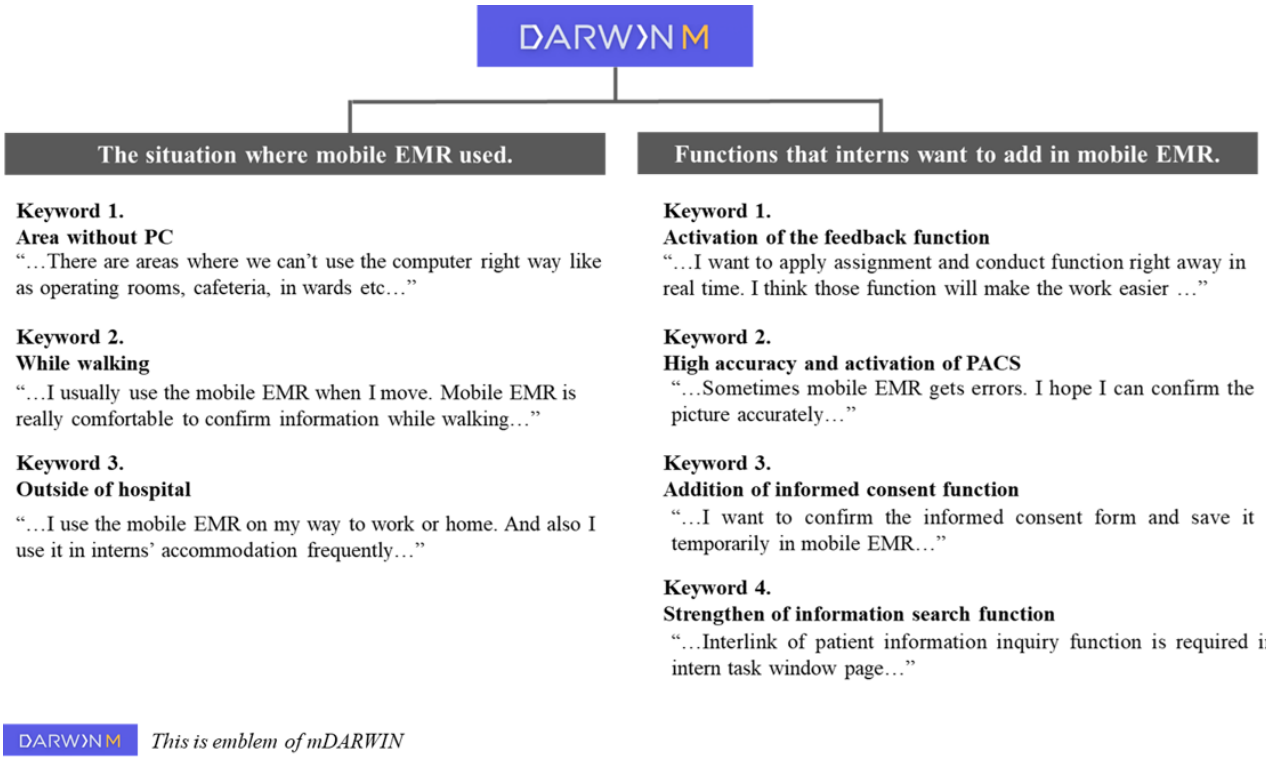
Table 3. System Usability Score (SUS) survey assessing an intern task item in the mobile electronic medical record (mEMR; n=15). The mean value of the 5-point Likert-scale responses was 3.1 (SD 1.6), and the mean SUS value was 81.67 (SD 9.4).

No.	Question	Response, mean (SD)[a]
1	I think I use (intern task) frequently through mDARWIN[b].	4.7 (0.6)
2	I found that using (intern task) through mDARWIN[b] is unnecessarily complex.	1.4 (0.5)
3	I thought that using (intern task) in mDARWIN[b] was easy.	4.5 (0.5)
4	I think technical support is needed to use (intern task) in mDARWIN[b].	2.7 (1.4)
5	I found that (intern task) in mDARWIN[b] was well integrated.	3.9 (1.0)
6	I thought there was too much inconsistency with (intern task) in mDARWIN[b].	1.7 (0.7)
7	I would imagine that most people would learn to use (intern task) through mDARWIN[b] very quickly and easily.	4.7 (0.5)
8	I found that using (intern task) in mDARWIN[b] is very cumbersome.	1.7 (0.9)
9	I felt very confident using (intern task) through mDARWIN[b].	4.0 (0.9)
10	I needed to learn many things before I could get going with (intern task) through mDARWIN[b].	1.5 (0.6)

[a]Responses were scored using a 5-point Likert scale (1=strongly disagree, 5=strongly agree).

[b]mDARWIN: mobile Data Analysis and Research Window for Integrated kNowledge.

Figure 4. Summary of the intern interviews about mobile electronic medical records (EMR).



Discussion

Principal Findings

This study investigated the correlation between the frequency of mEMR usage and the intern task time interval based on mEMR log data and EMR timestamp data. Most interns use mEMR and the task completion time was shorter for interns who used mEMR more frequently. This suggests that mEMR use could effectively enhance hospital workflow time, leading to a faster response in real practice. This result supports the findings of previous studies that indicate that the mEMR is linked to improved workflow efficiency in hospitals by enabling faster responses [13,16,17].

In addition to log data analysis, we also interviewed interns to assess the use of the mEMR for job execution. All the interns who participated in the survey and interview were actively using the mEMR. The mean SUS value was >80, implying that the system is well utilized by the user [18]. Doctors tend to underestimate the various positive workflow effects of mEMR usage [14]; as such, our results are interesting and valuable

enough to analyze motivation. We assume that the obvious and dominant benefit of mEMRs in terms of convenience and time efficiency would make all interns maximize the use of the mEMR compared to other systems such as computers and telephones. Further in-depth surveying and analysis can help increase mEMR usage among hospital health care providers. Our study shows that mEMR use offers both quantitative and qualitative strengths for intern job performance.

Comparison with Prior Work

Studies aiming to investigate the effects of mobile device use among health care providers in hospitals have proven their efficiency via surveys [14,19,20] and in simulations [21,22], and they have shown to be effective in limited spaces such as surgical rooms [23] and emergency departments [24]. However, there are limited quantitative studies assessing the efficiency of mEMR use in clinical practice. Our study results provide further evidence of the efficiency of mEMRs and suggest extending their use to other professionals with relatively similar daily tasks, such as physician assistants (PAs) who are responsible for clinical prescriptions in tertiary hospitals or nurses who are similarly overloaded with work. Further, the use of mEMRs by PAs or nurses would improve workflow efficiency, and ultimately, patient care and patient safety [25].

In recent times, quick response code technology reduces time and errors in patient identification during patient care and procedures [26-28]. Further, the closed-loop medication system, which integrates the barcode medication system and CPOE technology with automated dispensing technology (robots/units), prevents the adverse effects of medication due to administration errors [29]. Future efforts should be directed at combining mEMR use with these technologies to simultaneously achieve efficient workflows and patient safety.

Limitations

This study has some limitations. First, as this study was performed at a single center, the results have limited generalizability. Further, given that there are different job allocations for each occupation depending on the hospital, its feasibility and usability need to be validated in other institutions and environments.

Second, we could not identify the causality of log data as we analyzed the entire log dataset. There is no consideration for context or order between log data and interns' jobs. Although the entire log was sufficient to achieve the study aim, a further observational study using the small cut log of mEMR is needed to analyze the association between behavior and mEMR usage.

Third, we did not consider the priority of specific jobs when assessing performance. Jobs related to emergency situations need to be prioritized over others that can be completed after the emergency situation. However, this study aimed to investigate the general trend of frequent mEMR users and not to compare nonfrequent and frequent users to assess the efficiency of the mEMR. Furthermore, the log data was large enough to distinguish between situations.

Conclusions

By retrospectively analyzing the mEMR log data of hospital interns, this study revealed that more frequent use of the mEMR led to quicker completion of intern jobs. This finding implies the effectiveness of mEMR use for the workflow of interns in hospitals. We used a SUS survey to examine the usability of mEMR, and the survey concluded that the mEMR has a good usability.

Acknowledgments

The authors would like to acknowledge the support of Samsung Medical Center.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Items on the mEMR menu.
[PDF File (Adobe PDF File), 242 KB - mhealth_v8i12e23622_app1.pdf]

References

1. Househ M, Ahmad A, Alshaikh A, Alsuweed F. Patient safety perspectives: the impact of CPOE on nursing workflow. Stud Health Technol Inform 2013;183:367-371. [Medline: 23388315]
2. Page N, Baysari MT, Westbrook JI. A systematic review of the effectiveness of interruptive medication prescribing alerts in hospital CPOE systems to change prescriber behavior and improve patient safety. Int J Med Inform 2017 Dec;105:22-30. [doi: 10.1016/j.ijmedinf.2017.05.011] [Medline: 28750908]
3. Zhivan NA, Diana ML. U.S. hospital efficiency and adoption of health information technology. Health Care Manag Sci 2011 Sep 16;15(1):37-47. [doi: 10.1007/s10729-011-9179-2]
4. Brennan N, Corrigan O, Allard J, Archer J, Barnes R, Bleakley A, et al. The transition from medical student to junior doctor: today's experiences of Tomorrow's Doctors. Med Educ 2010 May;44(5):449-458. [doi: 10.1111/j.1365-2923.2009.03604.x] [Medline: 20518984]

5. Lim H, Ahn D, Ahn S. A Comparative Study on Medical Education Systems in Korea, China and Japan. Korean J Med Educ 2007 Dec 31;19(4):271-278. [doi: 10.3946/kjme.2007.19.4.271]
6. Yeo I, Kim J, Lee M, Shin S, Ryoo H, Park J. Work-related fatigue, stress and depression of the emergency department interns. Journal of The Korean Society of Emergency Medicine 2013;24(4):428-438.
7. Kim K, Lee S, Choi YH. Relationship between occupational stress and depressive mood among interns and residents in a tertiary hospital, Seoul, Korea. Clin Exp Emerg Med 2015 Jun 30;2(2):117-122. [doi: 10.15441/ceem.15.002]
8. Garcia, Abreu, Ramos, Castro, Smiderle, Santos, Bezerra. Influence of Burnout on Patient Safety: Systematic Review and Meta-Analysis. Medicina 2019 Aug 30;55(9):553. [doi: 10.3390/medicina55090553]
9. 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 Aug 20;5(6):323-328. [doi: 10.1002/jhm.790]
10. Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
11. Dimond R, Bullock A, Lovatt J, Stacey M. Mobile learning devices in the workplace: 'as much a part of the junior doctors' kit as a stethoscope'? BMC Med Educ 2016 Aug 17;16(1). [doi: 10.1186/s12909-016-0732-z]
12. Schimpf B, Deanda K, Severenuk DA, Montgomery TM, Cooley GD, Kowalski RG, et al. Integration of Real-Time Electronic Health Records and Wireless Technology in a Mobile Stroke Unit. Journal of Stroke and Cerebrovascular Diseases 2019 Sep;28(9):2530-2536. [doi: 10.1016/j.jstrokecerebrovasdis.2019.06.008]
13. Jung KY, Kim S, Kim K, Lee EJ, Kim K, Lee J, et al. Frequent Mobile Electronic Medical Records Users Respond More Quickly to Emergency Department Consultation Requests: Retrospective Quantitative Study. JMIR Mhealth Uhealth 2020 Feb 14;8(2):e14487. [doi: 10.2196/14487]
14. Duhm J, Fleischmann R, Schmidt S, Hupperts H, Brandt SA. Mobile Electronic Medical Records Promote Workflow: Physicians' Perspective From a Survey. JMIR Mhealth Uhealth 2016 Jun 06;4(2):e70 [FREE Full text] [doi: 10.2196/mhealth.5464] [Medline: 27268720]
15. Brooke J. SUS-A quick and dirty usability scale. Usability evaluation in industry 1996;189(194):4-7. [doi: 10.1201/9781498710411-35]
16. Power JM, Spina SP, Forbes DA, Harder CK, Lalli SL, Loewen PS, et al. Integration of Smartphones into clinical pharmacy practice: An evaluation of the impact on pharmacists efficiency. Health Policy and Technology 2014 Dec;3(4):296-305. [doi: 10.1016/j.hlpt.2014.08.002]
17. Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, et al. An Evaluation of the Use of Smartphones to Communicate Between Clinicians: A Mixed-Methods Study. J Med Internet Res 2011 Aug 29;13(3):e59. [doi: 10.2196/jmir.1655]
18. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies 2009;4(3):A.
19. Panesar P, Jones A, Aldous A, Kranzer K, Halpin E, Fifer H, et al. Attitudes and Behaviours to Antimicrobial Prescribing following Introduction of a Smartphone App. PLoS ONE 2016 Apr 25;11(4):e0154202. [doi: 10.1371/journal.pone.0154202]
20. 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(1). [doi: 10.1186/s12911-016-0249-8]
21. 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. In: AMIA Annu Symp Proc. 2012 Presented at: AMIA Annu Symp Proc ; Medline; 2012; Chicago, US p. 744-752 URL: http://europepmc.org/abstract/MED/23304348
22. Jung KY, Kim T, Jung J, Lee J, Choi JS, Mira K, et al. The Effectiveness of Near-Field Communication Integrated with a Mobile Electronic Medical Record System: Emergency Department Simulation Study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11187. [doi: 10.2196/11187]
23. Khanna V, Sambandam SN, Gul A, Mounasamy V. "WhatsApp"ening in orthopedic care: a concise report from a 300-bedded tertiary care teaching center. Eur J Orthop Surg Traumatol 2015 Jul;25(5):821-826. [doi: 10.1007/s00590-015-1600-y] [Medline: 25633127]
24. Gulacti U, Lok U, Hatipoglu S, Polat H. An Analysis of WhatsApp Usage for Communication Between Consulting and Emergency Physicians. J Med Syst 2016 Jun;40(6):130. [doi: 10.1007/s10916-016-0483-8] [Medline: 27083574]
25. Moore S, Jayewardene D. The use of smartphones in clinical practice. Nursing Management 2014 Jun 26;21(4):18-22. [doi: 10.7748/nm.21.4.18.e1225]
26. Mathivanan P, Edward Jero S, Ramu P, Balaji Ganesh A. QR code based patient data protection in ECG steganography. Australas Phys Eng Sci Med 2018 Nov 5;41(4):1057-1068. [doi: 10.1007/s13246-018-0695-y]
27. Upton J, Olsson-Brown A, Marshall E, Sacco J. Using QR codes to enable quick access to information in acute cancer care. British Journal of Nursing 2017 May 25;26(10):S4-S12. [doi: 10.12968/bjon.2017.26.10.s4]
28. Dixon JL, Smythe WR, Momsen LS, Jupiter D, Papaconstantinou HT. Quick Response codes for surgical safety: a prospective pilot study. Journal of Surgical Research 2013 Sep;184(1):157-163. [doi: 10.1016/j.jss.2013.06.036]
29. Burkoski V, Yoon J, Solomon S, Hall T, Karas A, Jarrett S, et al. Closed-Loop Medication System: Leveraging Technology to Elevate Safety. cjnl 2019 May 07;32(SP):16-28. [doi: 10.12927/cjnl.2019.25817]

Abbreviations

CPOE: computerized provider order entry
DARWIN: Data Analysis and Research Window for Integrated kNowledge
EMR: electronic medical record
mDARWIN: mobile Data Analysis and Research Window for Integrated kNowledge
mEMR: mobile electronic medical record
SUS: System Usability Score

Edited by L Buis; submitted 18.08.20; peer-reviewed by R Krukowski, J Li; comments to author 22.08.20; revised version received 16.10.20; accepted 30.11.20; published 15.12.20.

Please cite as:
Kim S, Ku S, Kim T, Cha WC, Jung KY
Effective Use of Mobile Electronic Medical Records by Medical Interns in Real Clinical Settings: Mixed Methods Study
JMIR Mhealth Uhealth 2020;8(12):e23622
URL: http://mhealth.jmir.org/2020/12/e23622/
doi:10.2196/23622
PMID:33320098

Original Paper

Patterns and Perceptions of Smartphone Use Among Academic Neurologists in the United States: Questionnaire Survey

William Zeiger[1], MD, PhD; Scott DeBoer[2,3], MD, PhD; John Probasco[4], MD

[1]Department of Neurology, University of California, Los Angeles School of Medicine, Los Angeles, CA, United States

[2]Medstar Franklin Square Medical Center, Baltimore, MD, United States

[3]Department of Neurology, Georgetown University, Washington, DC, United States

[4]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
John Probasco, MD
Department of Neurology
Johns Hopkins University School of Medicine
600 N Wolfe St
Meyer 6-113
Baltimore, MD, 21287
United States
Phone: 1 410 955 8174
Fax: 1 410 955 0672
Email: jprobas1@jhmi.edu

Abstract

Background: Smartphone technology is ubiquitous throughout neurologic practices, and numerous apps relevant to a neurologist's clinical practice are now available. Data from other medical specialties suggest high utilization of smartphones in routine clinical care. However, the ways in which these devices are used by neurologists for patient care–related activities are not well defined.

Objective: This paper aims to characterize current patterns of smartphone use and perceptions of the utility of smartphones for patient care–related activities among academic neurology trainees and attending physicians. We also seek to characterize areas of need for future app development.

Methods: We developed a 31-item electronic questionnaire to address these questions and invited neurology trainees and attendings of all residency programs based in the United States to participate. We summarized descriptive statistics for respondents and specifically compared responses between trainees and attending physicians.

Results: We received 213 responses, including 112 trainee and 87 attending neurologist responses. Neurology trainees reported more frequent use of their smartphone for patient care–related activities than attending neurologists (several times per day: 84/112, 75.0% of trainees; 52/87, 59.8% of attendings; P=.03). The most frequently reported activities were internet use, calendar use, communication with other physicians, personal education, and health care–specific app use. Both groups also reported regular smartphone use for the physical examination, with trainees again reporting more frequent usage compared with attendings (more than once per week: 35/96, 36.5% of trainees; 8/58, 13.8% of attendings; P=.03). Respondents used their devices most commonly for the vision, cranial nerve, and language portions of the neurologic examination. The majority of respondents in both groups reported their smartphones as "very useful" or "essential" for the completion of patient care–related activities (81/108, 75.0% of trainees; 50/83, 60.2% of attendings; P=.12). Neurology trainees reported a greater likelihood of using their smartphones in the future than attending neurologists ("very likely": 73/102, 71.6% of trainees; 40/82, 48.8% of attendings; P=.005). The groups differed in their frequencies of device usage for specific patient care–related activities, with trainees reporting higher usage for most activities. Despite high levels of use, only 12 of 184 (6.5%) respondents reported ever having had any training on how to use their device for clinical care. Regarding future app development, respondents rated vision, language, mental status, and cranial nerve testing as potentially being the most useful to aid in the performance of the neurologic examination.

Conclusions: Smartphones are used frequently and are subjectively perceived to be highly useful by academic neurologists. Trainees tended to use their devices more frequently than attendings. Our results suggest specific avenues for future technological development to improve smartphone use for patient care–related activities. They also suggest an unmet need for education on effectively using smartphone technology for clinical care.

(JMIR Mhealth Uhealth 2020;8(12):e22792) doi:10.2196/22792

KEYWORDS

smartphone; mobile phone; mobile apps; mobile health; neurology; neurologic exam; physical exam

Introduction

Smartphones are a ubiquitous presence on hospital wards. Ownership among physicians is nearly universal [1-8], with most regularly using their devices for phone calls, texting, email, and internet [4,6]. Rising rates of usage among physicians have been paralleled by a proliferation of health care–specific smartphone apps. These are generally rated as useful for clinical practice, especially for providing quick access to references and clinical score calculators [1,3-6]. As smartphone technology improves, clinical uses for these devices continue to expand. An increasing number of these devices are being used in creative ways to directly augment and improve bedside patient diagnosis and care. Smartphones can be used for vital signs [9], telemetry monitoring [10], ambulatory electroencephalography (EEG) [11], and even as portable ultrasounds [12], otoscopes [13], and ophthalmoscopes [14].

There are now many neurology-specific smartphone apps, with an emphasis on everything from anatomy to localization, reference materials, education, and documentation [15]. Mobile photo and video capture capabilities can help characterize intermittent symptoms such as seizures [16,17] or allow for remote telemedicine evaluation of acute stroke [18]. For patient monitoring in the home environment, there are symptom trackers (eg, headache diaries [19]) and apps to track functional impairments related to Parkinson disease [20], multiple sclerosis [21], and dementia [22,23]. To assist at the bedside, smartphone apps can now be found to evaluate everything from visual function [24,25] to tremors [26], gait speed [27], joint range of motion [28-30], and spinal deformities [31]. Despite the promise of smartphone technology, little is known about the current use of smartphones by neurologists in patient care or about areas of need to guide future app development. Therefore, we designed a survey to characterize current practice patterns of smartphone use among attending academic neurologists and neurology trainees. We also sought to identify parts of the neurologic examination that neurologists find to be most in need of adjunctive technological innovations.

Methods

The study was approved by the Johns Hopkins University School of Medicine Institutional Review Board. An initial draft questionnaire was developed by the authors and was subsequently refined and validated through 2 focus groups consisting of a total of 4 residents and 3 attendings from the Department of Neurology at the Johns Hopkins University School of Medicine. Focus group participants provided direct oral and written feedback regarding the questionnaire length and subject areas, as well as the clarity, response options, and relevance of items in the questionnaire about their experience using smartphones. Patient care–related activities were clarified to include "communication with or about patients, clinical documentation, physical examination, accessing clinical or reference information, and healthcare specific mobile applications." The final questionnaire was distributed electronically using Qualtrics software.

A letter with an anonymous link to the final 31-item questionnaire was emailed to all program directors and coordinators of academic neurology residency training programs (154 programs in total) in the United States in the spring of 2018. Follow-up reminder emails were sent 1 and 2 months later, and data collection was closed 3 months after the initial invitation. We did not solicit or receive feedback from programs about whether they had distributed the survey to physicians at their program, and as a result, we were unable to calculate a complete response rate for the questionnaire. On the first page of the questionnaire, participants were told the purpose of the survey and the estimated length of time to complete the survey (10 minutes) and were informed that participation was completely voluntary and that participation in the survey would serve as consent to have responses included in the study. Respondents could leave questionnaire items incomplete. No personal or identifying data were collected or stored about respondents. We did not collect information about the institutions to which respondents belonged nor did we attempt to validate self-reported usage data with data logs from respondents' smartphones.

Data were analyzed using MATLAB (MathWorks) and R (R Foundation for Statistical Computing). Questionnaires with incomplete data were included in the analysis. Results are presented with the total number of respondents for each questionnaire item. Primary analysis was done with chi-square tests. When expected counts were low (<5), response categories were binned. When response categories could not be logically binned, a Fisher exact test was used. A threshold for statistical significance of 0.05 was used. Follow-up 2 × 2 contingency tables were created for post hoc testing of individual response categories with Bonferroni correction. For matrix table items with Likert-type scales, data were compared using the Wilcoxon rank sum test with Bonferroni correction.

Results

A total of 213 neurologists responded to the questionnaire, all of whom owned smartphones. We estimate our response rate was about 4% for trainees, based on 112 trainee responses and a total of 2797 neurology residents and fellows in 2018 [32]. Demographics are presented in Table 1. Overall, smartphone use for patient care–related activities was high. The majority of respondents reported using their smartphone several times per day, with trainees reporting more frequent usage (84/112, 75.0% of trainees and 52/87, 59.8% of attending physicians; P=.03) and longer duration of use per day (median of 31-50 minutes for trainees and 11-30 minutes for attending physicians; P=.02) (Table 2). A variety of specific patient care–related activities for which respondents used their devices were surveyed, with

the most frequently reported activities being internet use, calendar use, communication with other physicians, personal education, and health care–specific app use (Figure 1). Trainees reported greater smartphone usage patterns for most activities. The specific mobile apps used by each group are summarized in Figure 1. The majority of respondents in both groups reported that their smartphones were "very useful" or "essential" for the completion of patient care activities (Table 2).

Table 1. Demographics of survey respondents.

Characteristics	All	Attending	Trainee	*P* value
Sex (female), n/N (%)	93/199 (46.7)	37/87 (42.5)	56/112 (50.0)	.36
Age (years), median	30-34	40-49	30-34	<.001
<30, n/N (%)[a]	36/198 (18.2)	1/87 (1.1)	35/111 (31.5)	
30-34, n/N (%)[a]	66/198 (33.3)	6/87 (6.9)	60/111 (54.1)	
35-39, n/N (%)	32/198 (16.2)	17/87 (19.5)	15/111 (13.5)	
40-49, n/N (%)[a]	29/198 (14.6)	29/87 (33.3)	0/111 (0.0)	
50-59, n/N (%)[a]	15/198 (7.6)	14/87 (16.1)	1/111 (0.9)	
>60, n/N (%)[a]	20/198 (10.1)	20/87 (23.0)	0/111 (0.0)	
PGY[b], median	N/A[c]	N/A	3	N/A
Years in practice, median	N/A	14	N/A	N/A

[a]Individual response category was found to be significant upon post hoc testing with Bonferroni correction.

[b]PGY: postgraduate year.

[c]N/A: not applicable.

Respondents were also surveyed regarding current usage of their devices as an aid to the performance of the neurologic examination. Most respondents said they had used their smartphone as an aid to the examination, with more trainees having done so compared to attending physicians (97/108, 89.8% trainees vs 58/83, 69.9% attending physicians; *P*<.001) (Table 2). Frequency of smartphone usage as an aid to the neurologic examination was lower than for overall use, with a median response in both groups of "2-3 times a month." Respondents used their devices most commonly for the vision, cranial nerve, and language portions of the neurologic examination, with trainees reporting more frequent usage compared with attending physicians (Figure 2). The specific smartphone functions most frequently used are summarized in Figure 2. Very few respondents reported having ever received any instruction in the use of a smartphone as an aid to the neurologic examination (Table 2).

Finally, respondents were asked about their expectations regarding future smartphone use. The majority of respondents reported a high likelihood ("likely" or "very likely") of using their devices for patient care–related activities in the future, with trainees reporting higher likelihood (*P*=.005) (Table 2). Subjective likelihood of future device use as an aid to the neurologic examination was also high, with trainees reporting greater likelihoods (median response for trainees was "likely" vs "somewhat likely" for attending physicians; *P*=.05) (Table 2). When asked to imagine that a new mobile app was developed to aid in the performance of the neurologic examination, respondents reported the greatest potential utility for apps enhancing vision, language, and mental status testing (Figure 3). Respondents almost universally expected future use of their devices to be at similar or greater levels than current usage (Table 2).

Given that we found several differences between attending physicians and trainees, we wondered how much of this effect could have been driven by age rather than training status. Therefore, we conducted a subgroup analysis for respondents in the age range with the greatest overlap between attending physicians and trainees (35-39 years). In this age range, we did not find any significant differences between groups for any of the items reported in Table 2. While our study is not sufficiently powered for this type of subgroup analysis, these results do suggest that differences between trainees and attending physicians may be attributable to age rather than training status per se.

Table 2. Patterns of current smartphone usage and predicted future usage.

Usage	All	Attending	Trainee	*P* value
Frequency of use, median	Several times a day	Several times a day	Several times a day	.03
Once a week or less, n/N (%)	18/199 (9.0)	13/87 (14.9)	5/112 (4.5)	
2-3 times a week, n/N (%)	19/199 (9.5)	11/87 (12.6)	8/112 (7.1)	
Once or twice a day, n/N (%)	26/199 (13.1)	11/87 (12.6)	15/112 (13.4)	
Several times a day, n/N (%)	136/199 (68.3)	52/87 (59.8)	84/112 (75.0)	
Duration of use (min), median	11-30 min	11-30 min	31-50 min	.02
<10, n/N (%)[a]	34/199 (17.1)	22/87 (25.3)	12/112 (10.7)	
11-30, n/N (%)	69/199 (34.7)	32/87 (36.8)	37/112 (33.0)	
31-50, n/N (%)	29/199 (14.6)	11/87 (12.6)	18/112 (16.1)	
>50, n/N (%)	67/199 (33.7)	22/87 (25.3)	45/112 (40.2)	
Number of apps, median	<5	<5	<5	.03
<5, n/N (%)[a]	121/198 (61.1)	62/86 (72.1)	59/112 (52.7)	
6-10, n/N (%)[a]	50/198 (25.3)	13/86 (15.1)	37/112 (33.0)	
11-15, n/N (%)	15/198 (7.6)	6/86 (7.0)	9/112 (8.0)	
>15, n/N (%)	12/198 (6.1)	5/86 (5.8)	7/112 (6.3)	
Ever used for examination (yes), n/N (%)	155/191 (81.2)	58/83 (69.9)	97/108 (89.8)	<.001
Frequency of use for examination, median	2-3 times a month	2-3 times a month	2-3 times a month	.03
Less than once a month, n/N (%)	40/154 (26.0)	18/58 (31.0)	22/96 (22.9)	
Once a month, n/N (%)	19/154 (12.3)	8/58 (13.8)	11/96 (11.5)	
2-3 times a month, n/N (%)	32/154 (20.8)	13/58 (22.4)	19/96 (19.8)	
Once a week, n/N (%)	20/154 (13.0)	11/58 (19.0)	9/96 (9.4)	
2-3 times a week, n/N (%)	27/154 (17.5)	7/58 (12.1)	20/96 (20.8)	
Daily, n/N (%)	16/154 (10.4)	1/58 (1.7)	15/96 (15.6)	
Usefulness, median	Very useful	Very useful	Very useful	.12
Not useful at all, n/N (%)	12/191 (6.3)	8/83 (9.6)	4/108 (3.7)	
Minimally useful, n/N (%)	48/191 (25.1)	25/83 (30.1)	23/108 (21.3)	
Very useful, n/N (%)	85/191 (44.5)	34/83 (41.0)	51/108 (47.2)	
Essential, n/N (%)	46/191 (24.1)	16/83 (19.3)	30/108 (27.8)	
Ever received instruction (yes), n/N (%)	12/184 (6.5)	2/82 (2.4)	10/102 (9.8)	.09
Likelihood of future use, median	Very likely	Likely	Very likely	.005
Very unlikely, n/N (%)	2/184 (1.1)	2/82 (2.4)	0/102 (0.0)	
Unlikely, n/N (%)	1/184 (0.5)	1/82 (1.2)	0/102 (0.0)	
Somewhat unlikely, n/N (%)	3/184 (1.6)	3/82 (3.7)	0/102 (0.0)	
Undecided, n/N (%)	9/184 (4.9)	7/82 (8.5)	2/102 (2.0)	
Somewhat likely, n/N (%)	14/184 (7.6)	8/82 (9.8)	6/102 (5.9)	
Likely, n/N (%)	42/184 (22.8)	21/82 (25.6)	21/102 (20.6)	
Very likely, n/N (%)	113/184 (61.4)	40/82 (48.8)	73/102 (71.6)	
Likelihood of future use for examination, median	Likely	Somewhat likely	Likely	.05
Very unlikely, n/N (%)	8/184 (4.3)	6/82 (7.3)	2/102 (2.0)	
Unlikely, n/N (%)	6/184 (3.3)	4/82 (4.9)	2/102 (2.0)	
Somewhat unlikely, n/N (%)	9/184 (4.9)	7/82 (8.5)	2/102 (2.0)	

Usage	All	Attending	Trainee	*P* value
Undecided, n/N (%)	25/184 (13.6)	14/82 (17.1)	11/102 (10.8)	
Somewhat likely, n/N (%)	27/184 (14.7)	11/82 (13.4)	16/102 (15.7)	
Likely, n/N (%)	44/184 (23.9)	17/82 (20.7)	27/102 (26.5)	
Very likely, n/N (%)	65/184 (35.3)	23/82 (28.0)	42/102 (41.2)	
Frequency of future use, median	More	More	Same or more	.37
Never, n/N (%)	2/184 (1.1)	1/82 (1.2)	1/102 (1.0)	
Less than current usage, n/N (%)	1/184 (0.5)	0/82 (0.0)	1/102 (1.0)	
The same as current usage, n/N (%)	80/184 (43.5)	31/82 (37.8)	49/102 (48.0)	
More than current usage, n/N (%)	101/184 (54.9)	50/82 (61.0)	51/102 (50.0)	

[a]Individual response category was found to be significant upon post hoc testing with Bonferroni correction.

Figure 1. Frequency of smartphone and mobile app use. (A) Respondents were asked, "How frequently do you use your smartphone and/or tablet for the following patient care related activities?" (B) Respondents were asked, "How frequently do you use the following types of mobile applications?" A: attending physicians; T: trainees. *Significantly greater usage for the indicated group compared with the other, with Bonferroni adjusted *P*<.05.

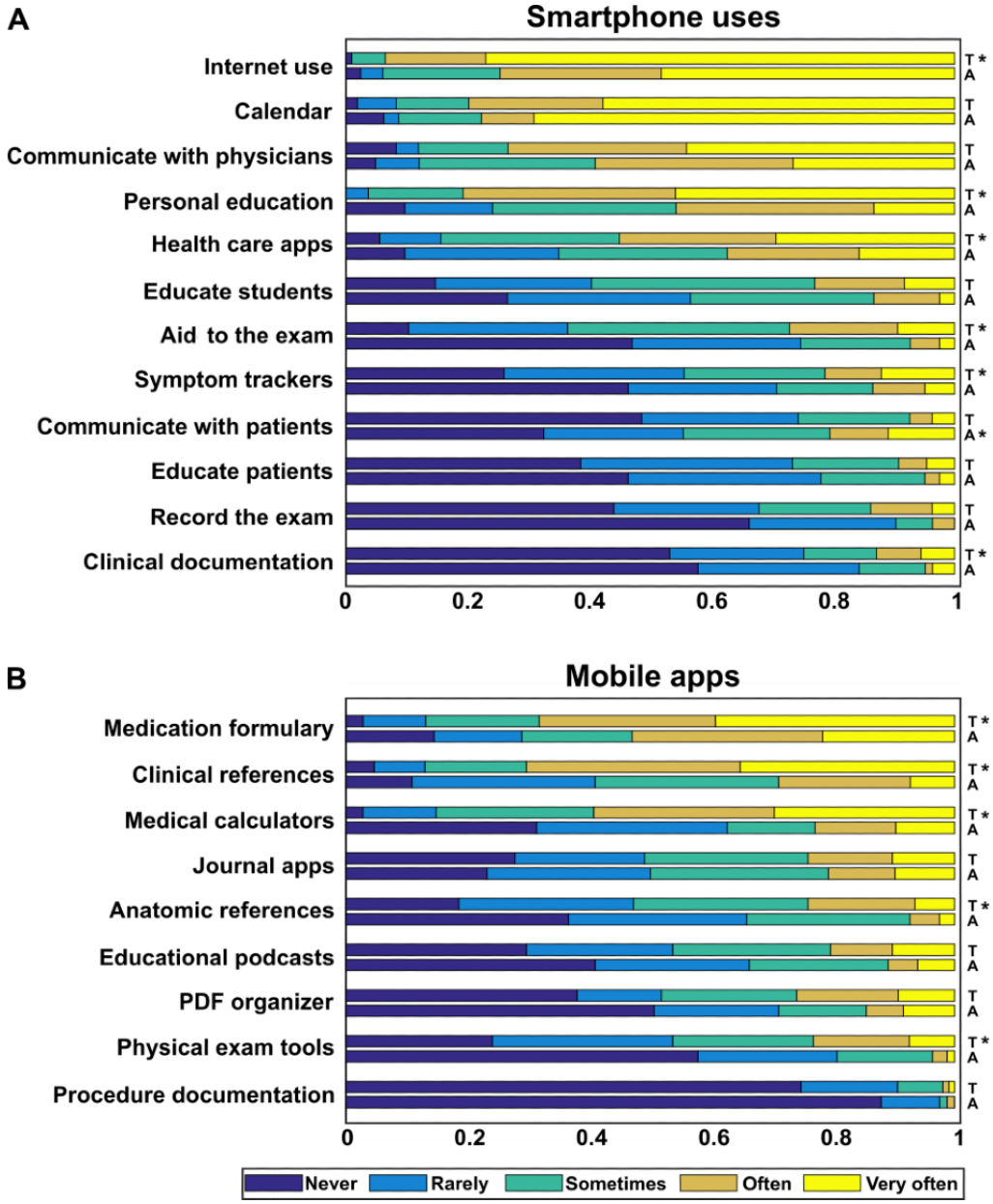


Figure 2. Frequency of smartphone use for the examination. (A) Respondents were asked, "How frequently do you utilize your smartphone and/or tablet for the following parts of the neurologic examination?" (B) Respondents were asked, "How frequently do you use your smartphone and/or tablet for each of the following functions when performing the physical examination?" A: attending physicians; OKN: optokinetic nystagmus; T: trainees. *Significantly greater usage for the indicated group compared with the other, with Bonferroni adjusted $P<.05$.

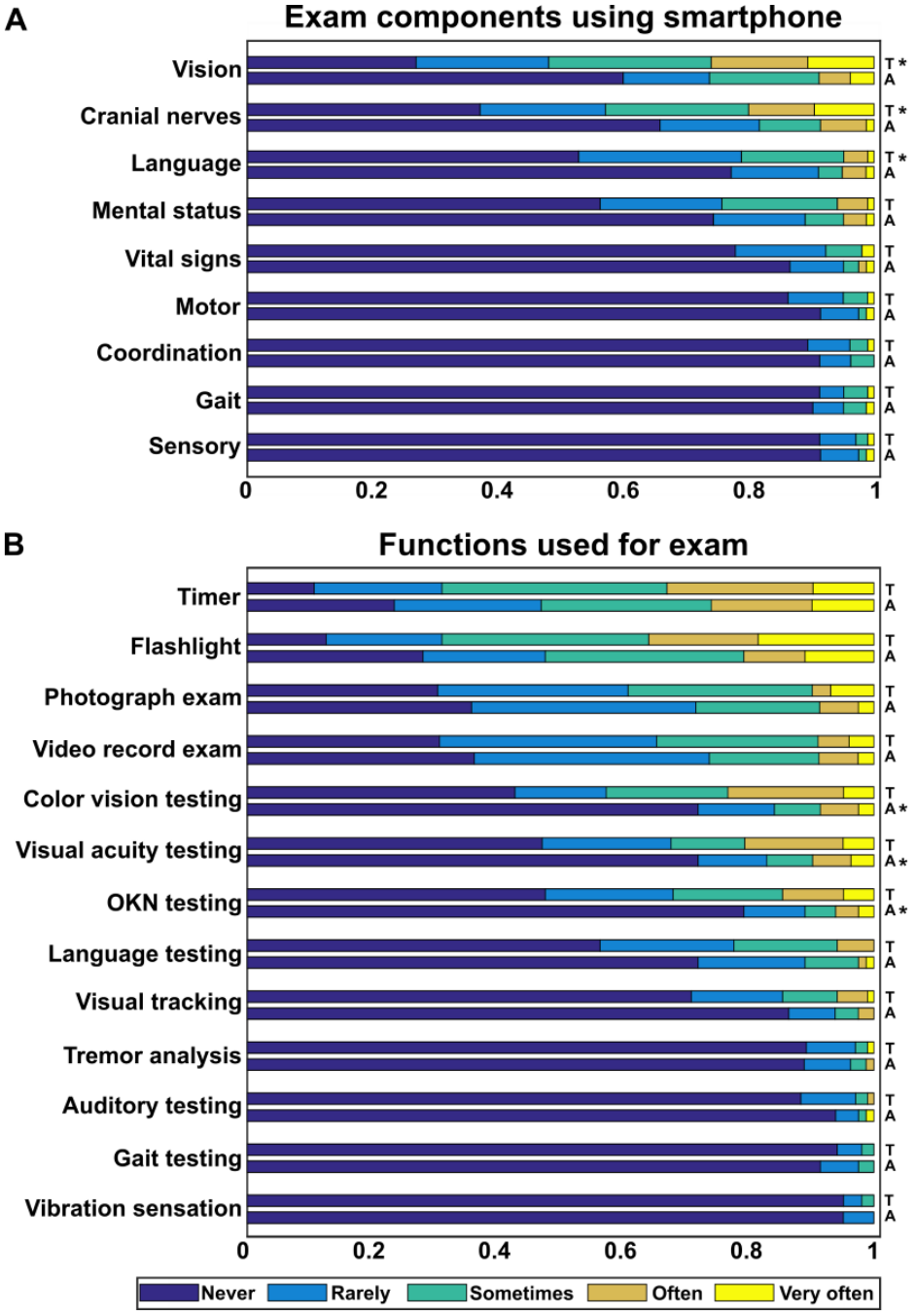
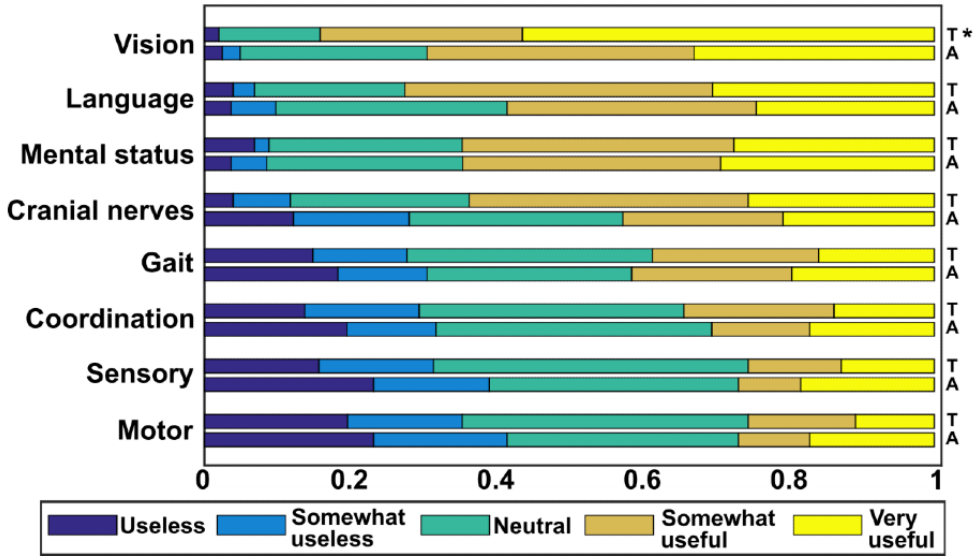


Figure 3. Perceived utility of potential new smartphone apps. Respondents were asked, "Imagine that a new mobile application was developed to aid in the performance of the neurologic examination. Please rank how useful it would be to have an application that could enhance the performance of each area of testing listed below." A: attending physicians; T: trainees. *Significantly greater usage for the indicated group compared with the other, with Bonferroni adjusted $P<.05$.



Discussion

As smartphone technologies improve, neurologists frequently use their mobile devices for patient care–related activities. Standard items from the neurologist's tool kit, such as a wristwatch and penlight, can easily be replaced with basic smartphone functionalities. Apps supplanting more advanced testing are rapidly being incorporated as well, including apps for visual acuity and color vision testing [24,25], tremor analysis [26], ophthalmoscopy [14], cognitive testing [23], and EEG [11]. Our results, which demonstrate ubiquitous and frequent usage of smartphones by neurologists, are in broad agreement with data across other specialties. For example, a 2011 survey distributed to all residents, fellows, and attending physicians participating in Accreditation Council for Graduate Medical Education training programs found that 85% of respondents owned smartphones and 56% used apps in their clinical practice [1]. Similar or greater usage has been found among residents in internal medicine [4,6], pediatrics [2], neurosurgery [7], obstetrics [8], urology [5], and radiation oncology [3] training programs, both in the United States and abroad. We found that the majority of neurologists use their devices for patient care–related activities several times per day, including for communication, educational activities, and health care–specific apps, as well as at the bedside as an aid in the performance of the physical examination. Smartphones were used for multiple portions of the neurologic examination, most commonly for vision, cranial nerve, language, and mental status testing.

These practice patterns are unlikely to be transient, as most respondents in this study predicted high likelihood of future smartphone use, including as an aid to the neurologic examination. We anticipate that the use of smartphone apps in neurologic practice will continue to grow, as trainees use their devices more frequently than attending physicians across a range of smartphone apps and functions. Indeed, neurology trainees tended to use their devices more frequently both for general patient care–related activities and as an aid to the performance of the physical examination. These trends held true when examining specific smartphone apps and functions, with trainees tending to report higher usage for most categories, with the exception of communication with patients. Trainees also reported higher likelihood of future use, though subjective usefulness was similar between trainees and attending physicians. Although not powered for a subgroup analysis, responses were similar between trainees and attending physicians aged 35 to 39 years. This suggests that age may be a significant factor in the overall differences between these groups, with younger neurologists using their devices more, which further emphasizes the likelihood that smartphone use in neurologic practice will continue to grow.

In addition to changes driven by the demographics of neurologists entering the workforce, we expect other factors may increase reliance on smartphone technologies for patient care. The SARS-CoV-2 pandemic has led to a dramatically increased reliance by neurologists on telehealth technologies for remote care delivery [33]. As telehealth grows as a method of care delivery, there will be an increasing need for apps that neurologists and their patients can use to augment their telehealth encounter. Although conducted prior to the pandemic, this survey provides end user insight into areas of need that may guide smartphone app development for neurologic telehealth care. Smartphone apps that augment in-person vision, language, and mental status testing could also be designed for patients to use on their own devices during a telehealth visit, bringing the examination tool kit from the clinic to the patient.

Although a large majority of neurologists use their devices, almost none have had any education on how to do so effectively for clinical practice. The development of such a curriculum could have several benefits, including greater use, increased efficiency, expanded access, improved subjective utility, and potentially, encouragement to spur the next generation of app development. On the other hand, such a curriculum could address mitigation of the negative effects of smartphones, such as impaired sleep [34], distractibility [35,36], burnout [37], and

confidentiality issues [38]. Education should also help physicians vet mobile technologies for incorporation into their practice, as the validity of smartphone apps for patient care is not always well established. Highlighting this fact, a recent review of apps related to emergency medicine found only a small percentage of apps to be clinically relevant [39]. Indeed, most apps and functionalities available today have not yet been well studied. The US Food and Drug Administration has recently recognized the unique challenges posed by mobile medical apps and has begun issuing policy guidance for manufacturers and distributors of these apps [40,41]. However, a comprehensive framework allowing patients and providers to easily evaluate mobile medical apps remains lacking [42].

This study was limited in several ways. The questionnaire was distributed to academic neurology training programs, so these findings may not be generalizable to private practice or nonacademic hospital settings. Participation was voluntary and our sample may have been biased toward neurologists with an interest in technology, who might have been more likely to respond to the questionnaire. All respondents were active smartphone users, and this might have resulted in an overestimation of the frequency of use or subjective usefulness, though based on our own experience, this seems unlikely. In addition, all smartphone use data were self-reported, and we did not validate these with objective data use logs. Finally, although our total number of respondents was large, our overall response rate was likely low. Given a total of 2797 neurology residents and fellows in 2018 [32] and 112 trainee respondents, we estimate that our response rate was approximately 4% for trainees.

In summary, smartphones are a valuable tool in academic neurology not only for communication but also for education and practice. These devices now feature in the neurologist's equipment bag alongside the reflex hammer and tuning fork. Smartphone-owning neurologists expect to continue using their devices in the future. There is opportunity for further refinement of these devices for neurologic practice, limited only by our creativity in the use of features and the development of associated tools, scales, and apps. We anticipate that these ubiquitous handheld devices will in time prove invaluable to the diagnosis and treatment of patients with neurologic disease.

Acknowledgments

The authors received no funding for this study and have no relevant financial disclosures to report.

Authors' Contributions

WZ contributed to the study conception, design, and data acquisition and analysis and drafted and revised the manuscript. SD contributed to the study design and data acquisition and revised the manuscript. JP contributed to the study conception, design, data acquisition, and analysis and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]
2. Katz-Sidlow RJ, Ludwig A, Miller S, Sidlow R. Smartphone use during inpatient attending rounds: prevalence, patterns and potential for distraction. J Hosp Med 2012 Oct;7(8):595-599. [doi: 10.1002/jhm.1950] [Medline: 22744793]
3. Bibault J, Leroy T, Blanchard P, Biau J, Cervellera M, Diaz O, et al. Mobile technology and social media in the clinical practice of young radiation oncologists: results of a comprehensive nationwide cross-sectional study. Int J Radiat Oncol Biol Phys 2014 Sep 01;90(1):231-237. [doi: 10.1016/j.ijrobp.2014.05.012] [Medline: 24986747]
4. Johnson AC, El Hajj SC, Perret JN, Caffery TS, Jones GN, Musso MW. Smartphones in medicine: emerging practices in an academic medical center. J Med Syst 2015 Jan;39(1):164. [doi: 10.1007/s10916-014-0164-4] [Medline: 25526706]
5. Nason GJ, Burke MJ, Aslam A, Kelly ME, Akram CM, Giri SK, et al. The use of smartphone applications by urology trainees. Surgeon 2015 Oct;13(5):263-266. [doi: 10.1016/j.surge.2014.06.008] [Medline: 25199700]
6. O'Connor P, Byrne D, Butt M, Offiah G, Lydon S, Mc Inerney K, et al. Interns and their smartphones: use for clinical practice. Postgrad Med J 2014 Feb;90(1060):75-79. [doi: 10.1136/postgradmedj-2013-131930] [Medline: 24243966]
7. Bergeron D, Iorio-Morin C, Bigder M, Dakson A, Eagles ME, Elliott CA, Canadian Neurosurgery Research Collaborative. Mobile Applications in Neurosurgery: A Systematic Review, Quality Audit, and Survey of Canadian Neurosurgery Residents. World Neurosurg 2019 Jul;127:e1026-e1038. [doi: 10.1016/j.wneu.2019.04.035] [Medline: 30980978]
8. Perry R, Burns RM, Simon R, Youm J. Mobile Application Use Among Obstetrics and Gynecology Residents. J Grad Med Educ 2017 Oct;9(5):611-615 [FREE Full text] [doi: 10.4300/JGME-D-17-00163.1] [Medline: 29075382]
9. Scully CG, Lee J, Meyer J, Gorbach AM, Granquist-Fraser D, Mendelson Y, et al. Physiological parameter monitoring from optical recordings with a mobile phone. IEEE Trans Biomed Eng 2012 Feb;59(2):303-306 [FREE Full text] [doi: 10.1109/TBME.2011.2163157] [Medline: 21803676]

10. Treskes RW, Gielen W, Wermer MJ, Grauss RW, van Alem AP, Dehnavi RA, et al. Mobile phones in cryptogenic strOke patients Bringing sIngle Lead ECGs for Atrial Fibrillation detection (MOBILE-AF): study protocol for a randomised controlled trial. Trials 2017 Aug 29;18(1):402 [FREE Full text] [doi: 10.1186/s13063-017-2131-0] [Medline: 28851409]
11. McKenzie ED, Lim ASP, Leung ECW, Cole AJ, Lam AD, Eloyan A, et al. Validation of a smartphone-based EEG among people with epilepsy: A prospective study. Sci Rep 2017 Apr 03;7:45567 [FREE Full text] [doi: 10.1038/srep45567] [Medline: 28367974]
12. Mertz L. Ultrasound? Fetal monitoring? Spectrometer? There's an app for that!: biomedical smart phone apps are taking healthcare by storm. IEEE Pulse 2012 Mar;3(2):16-21. [doi: 10.1109/MPUL.2011.2181019] [Medline: 22481741]
13. Moshtaghi O, Sahyouni R, Haidar YM, Huang M, Moshtaghi A, Ghavami Y, et al. Smartphone-Enabled Otoscopy in Neurotology/Otology. Otolaryngol Head Neck Surg 2017 Mar;156(3):554-558. [doi: 10.1177/0194599816687740] [Medline: 28118550]
14. Russo A, Morescalchi F, Costagliola C, Delcassi L, Semeraro F. A Novel Device to Exploit the Smartphone Camera for Fundus Photography. J Ophthalmol 2015;2015:823139 [FREE Full text] [doi: 10.1155/2015/823139] [Medline: 26137320]
15. Cohen AB, Nahed BV, Sheth KN. Mobile medical applications in neurology. Neurol Clin Pract 2013 Feb;3(1):52-60 [FREE Full text] [doi: 10.1212/CPJ.0b013e318283ff4f] [Medline: 29406529]
16. Zeiler SR, Kaplan PW. Our digital world: camera phones and the diagnosis of a seizure. Lancet 2009 Jun 20;373(9681):2136. [doi: 10.1016/S0140-6736(09)60304-4] [Medline: 19541039]
17. Tatum WO, Hirsch LJ, Gelfand MA, Acton EK, LaFrance WC, Duckrow RB, OSmartViE Investigators. Assessment of the Predictive Value of Outpatient Smartphone Videos for Diagnosis of Epileptic Seizures. JAMA Neurol 2020 May 01;77(5):593-600. [doi: 10.1001/jamaneurol.2019.4785] [Medline: 31961382]
18. Demaerschalk BM, Vegunta S, Vargas BB, Wu Q, Channer DD, Hentz JG. Reliability of real-time video smartphone for assessing National Institutes of Health Stroke Scale scores in acute stroke patients. Stroke 2012 Dec;43(12):3271-3277 [FREE Full text] [doi: 10.1161/STROKEAHA.112.669150] [Medline: 23160878]
19. Hundert AS, Huguet A, McGrath PJ, Stinson JN, Wheaton M. Commercially available mobile phone headache diary apps: a systematic review. JMIR Mhealth Uhealth 2014;2(3):e36 [FREE Full text] [doi: 10.2196/mhealth.3452] [Medline: 25138438]
20. Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity: The Mobile Parkinson Disease Score. JAMA Neurol 2018 Jul 01;75(7):876-880 [FREE Full text] [doi: 10.1001/jamaneurol.2018.0809] [Medline: 29582075]
21. Maillart E, Labauge P, Cohen M, Maarouf A, Vukusic S, Donzé C, et al. MSCopilot, a new multiple sclerosis self-assessment digital solution: results of a comparative study versus standard tests. Eur J Neurol 2020 Mar;27(3):429-436. [doi: 10.1111/ene.14091] [Medline: 31538396]
22. Tung JY, Rose RV, Gammada E, Lam I, Roy EA, Black SE, et al. Measuring life space in older adults with mild-to-moderate Alzheimer's disease using mobile phone GPS. Gerontology 2014;60(2):154-162. [doi: 10.1159/000355669] [Medline: 24356464]
23. Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E. Cognitive Testing in People at Increased Risk of Dementia Using a Smartphone App: The iVitality Proof-of-Principle Study. JMIR Mhealth Uhealth 2017 May 25;5(5):e68 [FREE Full text] [doi: 10.2196/mhealth.6939] [Medline: 28546139]
24. Ozgur OK, Emborgo TS, Vieyra MB, Huselid RF, Banik R. Validity and Acceptance of Color Vision Testing on Smartphones. J Neuroophthalmol 2018 Mar;38(1):13-16. [doi: 10.1097/WNO.0000000000000637] [Medline: 29438263]
25. O'Neill S, McAndrew DJ. The validity of visual acuity assessment using mobile technology devices in the primary care setting. Aust Fam Physician 2016 Apr;45(4):212-215 [FREE Full text] [Medline: 27052139]
26. Daneault J, Carignan B, Codère C, Sadikot AF, Duval C. Using a smart phone as a standalone platform for detection and monitoring of pathological tremors. Front Hum Neurosci 2012;6:357 [FREE Full text] [doi: 10.3389/fnhum.2012.00357] [Medline: 23346053]
27. Carlson RH, Huebner DR, Hoarty CA, Whittington J, Haynatzki G, Balas MC, et al. Treadmill gait speeds correlate with physical activity counts measured by cell phone accelerometers. Gait Posture 2012 Jun;36(2):241-248 [FREE Full text] [doi: 10.1016/j.gaitpost.2012.02.025] [Medline: 22475727]
28. Quek J, Brauer SG, Treleaven J, Pua Y, Mentiplay B, Clark RA. Validity and intra-rater reliability of an android phone application to measure cervical range-of-motion. J Neuroeng Rehabil 2014 Apr 17;11:65 [FREE Full text] [doi: 10.1186/1743-0003-11-65] [Medline: 24742001]
29. Kurillo G, Han JJ, Obdržálek S, Yan P, Abresch RT, Nicorici A, et al. Upper extremity reachable workspace evaluation with Kinect. Stud Health Technol Inform 2013;184:247-253. [Medline: 23400165]
30. Williams CM, Caserta AJ, Haines TP. The TiltMeter app is a novel and accurate measurement tool for the weight bearing lunge test. J Sci Med Sport 2013 Sep;16(5):392-395. [doi: 10.1016/j.jsams.2013.02.001] [Medline: 23491138]
31. Franko OI, Bray C, Newton PO. Validation of a scoliometer smartphone app to assess scoliosis. J Pediatr Orthop 2012 Dec;32(8):e72-e75. [doi: 10.1097/BPO.0b013e31826bb109] [Medline: 23147635]
32. FREIDA online. American Medical Association. URL: https://freida.ama-assn.org/Freida/user/specStatisticsSearch.do?method=viewDetail&pageNumber=2&spcCd=180 [accessed 2020-02-29]

33. Waldman G, Mayeux R, Claassen J, Agarwal S, Willey J, Anderson E, et al. Preparing a neurology department for SARS-CoV-2 (COVID-19): Early experiences at Columbia University Irving Medical Center and the New York Presbyterian Hospital in New York City. Neurology 2020 May 19;94(20):886-891. [doi: 10.1212/WNL.0000000000009519] [Medline: 32253352]
34. Christensen MA, Bettencourt L, Kaye L, Moturu ST, Nguyen KT, Olgin JE, et al. Direct Measurements of Smartphone Screen-Time: Relationships with Demographics and Sleep. PLoS One 2016;11(11):e0165331 [FREE Full text] [doi: 10.1371/journal.pone.0165331] [Medline: 27829040]
35. Vearrier L, Rosenberger K, Weber V. Use of Personal Devices in Healthcare: Guidelines From A Roundtable Discussion. JournalMTM 2018 Oct;7(2):27-34. [doi: 10.7309/jmtm.7.2.4]
36. Wu RC, Tzanetos K, Morra D, Quan S, Lo V, Wong BM. Educational impact of using smartphones for clinical communication on general medicine: more global, less local. J Hosp Med 2013 Jul;8(7):365-372. [doi: 10.1002/jhm.2037] [Medline: 23713054]
37. Brubaker JR, Beverly EA. Burnout, Perceived Stress, Sleep Quality, and Smartphone Use: A Survey of Osteopathic Medical Students. J Am Osteopath Assoc 2020 Jan 01;120(1):6-17. [doi: 10.7556/jaoa.2020.004] [Medline: 31904778]
38. McKnight R, Franko O. HIPAA Compliance with Mobile Devices Among ACGME Programs. J Med Syst 2016 May;40(5):129. [doi: 10.1007/s10916-016-0489-2] [Medline: 27079578]
39. 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]
40. Policy for Device Software Functions and Mobile Medical Applications. US Food Drug Administration. 2019. URL: http://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications [accessed 2020-02-29]
41. Shuren J, Patel B, Gottlieb S. FDA Regulation of Mobile Medical Apps. JAMA 2018 Jul 24;320(4):337-338. [doi: 10.1001/jama.2018.8832] [Medline: 29971339]
42. Bates DW, Landman A, Levine DM. Health Apps and Health Policy: What Is Needed? JAMA 2018 Nov 20;320(19):1975-1976. [doi: 10.1001/jama.2018.14378] [Medline: 30326025]

Abbreviations

EEG: electroencephalography

Edited by L Buis; submitted 23.07.20; peer-reviewed by R Treskes, T Tollefson; comments to author 15.09.20; revised version received 08.10.20; accepted 20.11.20; published 24.12.20.

Please cite as:
Zeiger W, DeBoer S, Probasco J
Patterns and Perceptions of Smartphone Use Among Academic Neurologists in the United States: Questionnaire Survey
JMIR Mhealth Uhealth 2020;8(12):e22792
URL: http://mhealth.jmir.org/2020/12/e22792/
doi:10.2196/22792
PMID:33361053

Original Paper

A Mobile Phone–Based App for Use During Cognitive Behavioral Therapy for Adolescents With Anxiety (MindClimb): User-Centered Design and Usability Study

Amanda Newton[1*], PhD; Alexa Bagnell[2,3*], MA, MD; Rhonda Rosychuk[1], PhD; Janelle Duguay[3], BA, BSW; Lori Wozney[2], PhD; Anna Huguet[2], PhD; Joanna Henderson[4], PhD; Janet Curran[5], PhD

[1]Department of Pediatrics, University of Alberta, Edmonton, AB, Canada

[2]IWK Health Centre, Halifax, NS, Canada

[3]Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

[4]Centre for Addiction and Mental Health, University of Toronto, Toronto, ON, Canada

[5]School of Nursing, Dalhousie University, Halifax, NS, Canada

[*]these authors contributed equally

Corresponding Author:
Alexa Bagnell, MA, MD
Department of Psychiatry
Dalhousie University
PO Box 9700
5850/5980 University Avenue
Halifax, NS, B3K 6R8
Canada
Phone: 1 902 470 8087
Email: Alexa.Bagnell@iwk.nshealth.ca

Abstract

Background: Mobile device–based tools to help adolescents practice skills outside of cognitive behavioral therapy (CBT) sessions for treating an anxiety disorder may lead to greater treatment gains.

Objective: This study aimed to develop, design, and test the acceptability, learnability, heuristics, and usability of MindClimb, a smartphone-based app for adolescents with anxiety to use between CBT sessions to plan and complete exposure activities using skills (cognitive, relaxation, exposure practice, and reward) learned in treatment.

Methods: This 3-phase study took place from August 2015 to December 2018. In phase 1, the app was designed and developed in consultation with young people and CBT therapists to identify desired functions and content. Feedback was subjected to thematic analysis using a general inductive approach. In phase 2, we conducted 2 high-fidelity testing sessions using the think-aloud approach (acceptability, learnability, usability) and 10-item System Usability Scale with 10 adolescents receiving CBT. The high-fidelity MindClimb app was evaluated by 5 app developers based on Nielsen's usability heuristics and 5-point severity ranking scale. In phase 3, a total of 8 adolescents and 3 therapists assessed the usability of MindClimb during CBT sessions by recording the frequency of skills practice, use of MindClimb features, satisfaction with the app, and barriers and facilitators to app use during treatment.

Results: Feedback from phase 1 consultations indicated that the app should (1) be responsive to user needs and preferences, (2) be easy to use and navigate, (3) have relevant content to the practice of CBT for anxiety, and (4) be aesthetically appealing. Using this feedback as a guide, a fully functional app prototype for usability testing and heuristic evaluation was developed. In phase 2, think-aloud and usability data resulted in minor revisions to the app, including refinement of exposure activities. The average system usability score was 77 in both testing cycles, indicating acceptable usability. The heuristic evaluation by app developers identified only minor errors (eg, loading speed of app content, with a score of 1 on the severity ranking scale). In phase 3, adolescents considered app features for completing exposure (6.2/10) and relaxation (6.4/10) modestly helpful. Both adolescents (average score 11.3/15, SD 1.6) and therapists (average score 10.0/12, 2.6 SD) reported being satisfied with the app.

Conclusions: The user-centered approach to developing and testing MindClimb resulted in a mobile health app that can be used by adolescents during CBT for anxiety. Evaluation of the use of this app in a clinical practice setting demonstrated that adolescents

and therapists generally felt it was helpful for CBT practice outside of therapy sessions. Implementation studies with larger youth samples are necessary to evaluate how to optimize the use of technology in clinical care and examine the impact of the app plus CBT on clinical care processes and patient outcomes.

(JMIR Mhealth Uhealth 2020;8(12):e18439) doi:10.2196/18439

KEYWORDS

anxiety disorders; mobile apps; adolescents; usability testing; development; design; anxiety

Introduction

In recent years, mobile device–based apps have been regarded as an opportunistic and potentially effective adjunct to cognitive behavioral therapy (CBT) for adolescent anxiety disorders. In the treatment context, an app is considered an ecological momentary intervention (EMI) [1] for in vivo skills coaching in an adolescent's natural settings (ie, at home, school, etc instead of a therapist's office) when it is most needed during daily life (ie, in real time). Conventionally, skills practice between treatment sessions has been supported through paper-based workbook activities and assignments. However, this approach to skills practice can be limited if the workbook and assignments are not an integral or natural part of an adolescent's daily life, unlike the use of mobile devices, such as smartphones.

Adolescents can realize greater treatment gains when they practice CBT skills outside of formal treatment (eg, at home or school) [2-4], yet adherence to and persistence with skills practice are common challenges for adolescents during CBT. Skills take a taxing effort, as purposely and repeatedly being in feared situations is mentally and physically uncomfortable. The use of an EMI app during CBT offers several solutions to these challenges. An app can provide structure to skills practice to promote consistency between skills practiced in a treatment session and those practiced in real-life practice situations, which can be less predictable than treatment contexts. Push notifications from smartphone calendars can remind adolescents to practice skills and engage in relaxation and self-rewarding, pleasurable activities to offset discomfort, encourage self-care, and reinforce hard work.

While recent systematic reviews of mobile health (mHealth) apps have identified several CBT-based apps that adolescents with anxiety can use independently to practice exposure, relaxation, and coping strategies [5,6], one app, SmartCAT, has been specifically designed as an EMI app for patient and therapist use during CBT [7]. The SmartCAT app prompts patients to use CBT skills learned in treatment and, via the web, allows therapists to monitor and reward skill use and communicate with patients using an integrated clinician portal for secure 2-way communication. In a feasibility study of SmartCAT, 9 children and adolescents aged 9 to 14 years rated the app as easy to use and used it regularly during treatment [7]. In a follow-up open trial with a larger sample of same-aged patients (n=34), users were highly satisfied with the app and used it an average of 12 times between each in-person CBT session. Other findings included improvements in CBT skills targeted by the app [8].

Our team was interested in developing an EMI app for adolescent (aged 13-18 years) use that would interface with native smartphone features, information, and hardware. As a native app, it would be installed directly from an app store (eg, Google Play or Apple's App Store) onto the smartphone and would not require access to Wi-Fi or a data plan. A native app would also allow our team to take advantage of cross-functional device abilities (eg, push notifications, interaction with calendar) [9,10] and multimedia opportunities (eg, use of camera, voice recordings, videos) to personalize planning and skills practice. We believed that adolescent use of a personal smartphone for exposure activities, relaxation strategies, and self-reward would help promote independent skills practice in real time and provide adolescents with a sense of ownership of their treatment goals, since they would be integrated into their personal device. We also wanted the app to be simple in design, as an app that is not easy to use during stressful skills practice would likely not be used by adolescents. The app design process incorporated academia, end users, and industry to create an engaging and useful intervention tool [11]. The objectives of this study were to develop, design, and test the acceptability, learnability, heuristics, and usability of this native app, which we called MindClimb.

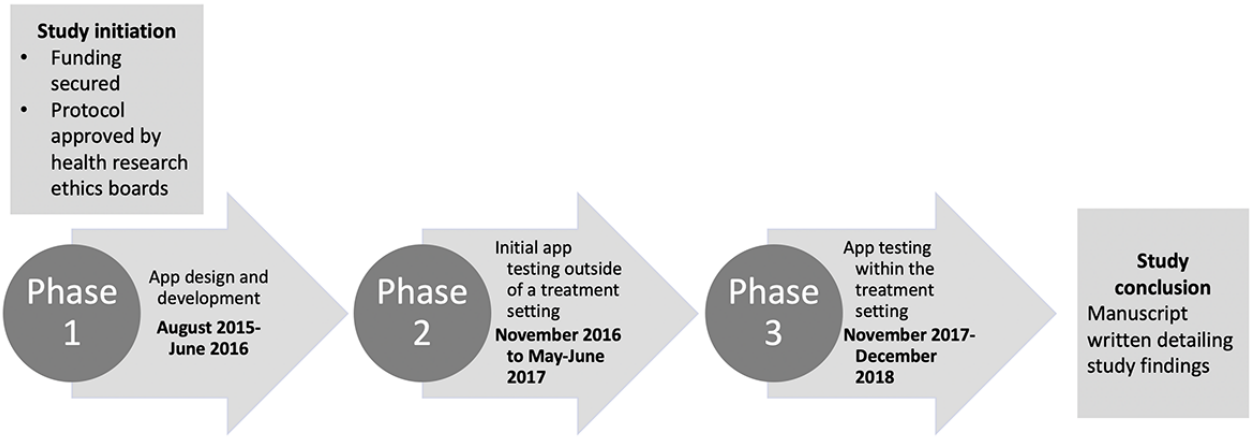
Methods

Study Overview

The study took place from August 2015 to December 2018 (Figure 1) and was conducted in 3 phases. We worked with an app designer and developer to create the app using a user-centered design approach during each study phase [12]. The study design is consistent with recommendations on mHealth technology development and evaluation, including user experience design, development, and alpha and beta testing [13]. Informed assent or consent was obtained from all participants prior to participation. In phase 1, the MindClimb app was designed and developed in consultation with youth and CBT-trained clinicians. In phase 2, adolescents in treatment for an anxiety disorder evaluated the app for acceptability, learnability, and usability in 2 testing cycles that took place outside of the formal treatment setting. In this phase, we also asked app developers to evaluate the app using a set of rules (heuristics) to measure the usability of the user interface. For this phase, we used Nielsen's heuristics [14]. In phase 3, app usability was assessed by adolescents and therapists during CBT treatment for anxiety. This involved reviewing the frequency of app use for skills practice between CBT sessions, satisfaction with app features, and barriers and facilitators to app use. The study was led out of the University of Alberta (Edmonton, Alberta, Canada) and Dalhousie University (Halifax, Nova

Scotia, Canada). The University of Alberta and the IWK Health Centre's Research Ethics Boards approved the study.

Figure 1. Study timeline.



Participants

We recruited a convenience sample of adolescents, young adults, and CBT therapists to provide initial, varied perspectives on app design. Our goal was to recruit 7 to 10 participants. We invited individuals aged 12 to 24 years from across Canada who were members of the Centre for Addiction and Mental Health's National Youth Advisory Committee (NYAC) and CBT therapists specialized in outpatient anxiety treatment at the IWK Health Centre in Halifax, Nova Scotia, Canada, to participate. CBT therapists from the same setting were also invited to participate in phase 3 if one of their patients was recruited into this study phase.

For all phases, we recruited purposeful samples of adolescents aged 13 to 18 years from an outpatient program at the IWK Health Centre in Halifax, Nova Scotia, Canada. At the IWK Health Centre, adolescents are treated until their 19th birthday. Adolescents receiving outpatient treatment at this setting scored ≥25 on the Screen for Child Anxiety Related Emotional Disorders (SCARED) [15], indicating a clinical level of anxiety symptoms. To participate, adolescents needed to be fluent in English. As sex differences are noted in mobile phone use [16], we hoped both adolescent boys and girls would participate. Youth identifying with any gender were welcome to participate, as we were not testing sex- or gender-based differences in app use in this study. Honorariums were provided to participants (phase 1 and 2: Can $25 gift card [US $19.14]; phase 3: Can $40 gift card [US $30.61]). Consistent with recommendations in the literature, our recruitment goals for phases 1 and 2 were 7 to 10 adolescents (phase 1) and 10 adolescents (phase 2) [14,17]. In phase 3, we aimed to enroll 21 adolescents and 7 CBT therapists (with each therapist treating 3 of the enrolled adolescents) to comprise a total sample size of 28 participants. This sample size was chosen assuming a 30% dropout rate for clinicians and patients (5 clinicians with 3 patients per clinician, for a total of 15 participants), which would still allow detection of an effect size of at least 1.1 in the mean change in frequency of app use (between last and first weeks of use; 2-sided, 1-sample *t* test, 80% power, type I error rate of 0.05). PASS (NCSS Statistical Software) was used to determine the detectable effect size, and the sample size was adjusted for potential intracluster correlation (clinicians were considered a cluster; ρ=0.1) using an inflation factor.

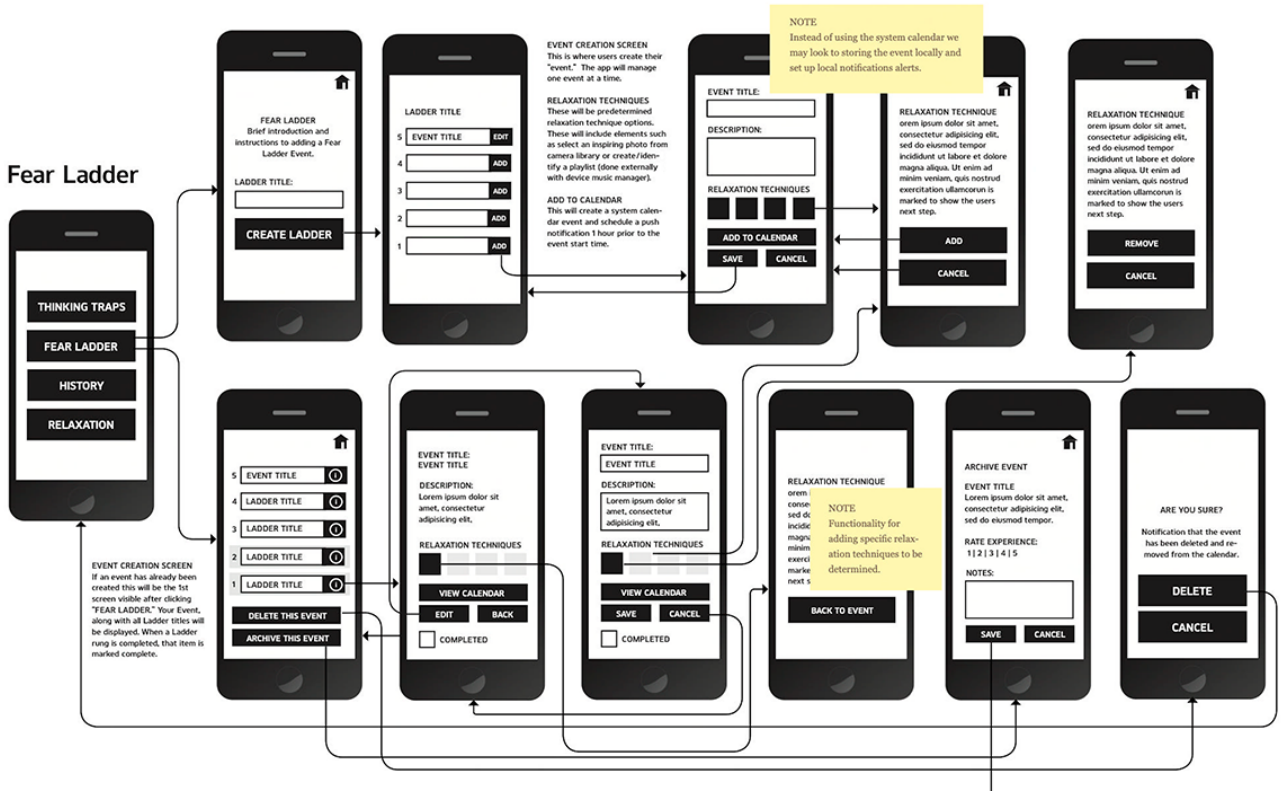
We also recruited a convenience sample of app developers in phase 2 via email using professional contacts identified by our app developer. To be eligible to participate in the heuristic evaluation, developers needed to have postsecondary training in human-computer interaction and experience in designing and evaluating health apps for mobile devices. We considered 5 developer participants to be an adequate sample size for the heuristic-based evaluation [14,17].

Procedures

Phase 1

App development began with wireframing, a process in which our team created black-and-white pictures of MindClimb app screens and arranged them to determine how users would navigate and use the app (Figure 2). We organized the MindClimb wireframes around guiding adolescent users through the development of 3 main components: (1) fear ladders, (2) relaxation, and (3) thinking traps.

Figure 2. Screenshot of an initial MindClimb wireframe.



First, fear ladders are a tool designed to help adolescents work on their exposure goals. Our goal was to allow adolescents to create, select, and manage multiple ladders. Events for each ladder would be synced with the phone's calendar app, and phone notifications would be used to remind and cue adolescents of upcoming events. After completion of the exposure, adolescents could rate the activity in terms of its difficulty. Ladders could also be archived ("history" tab on the wireframe).

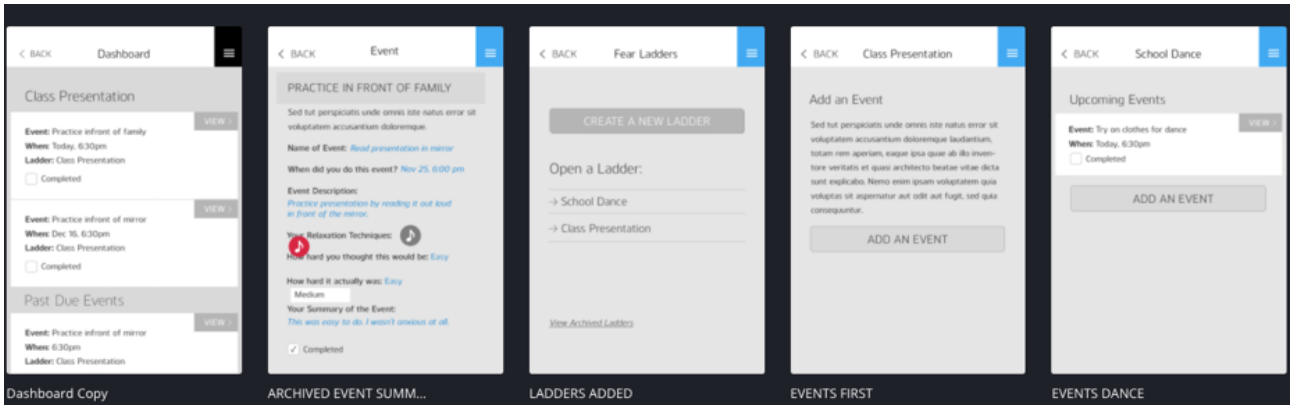
Second, the relaxation section would offer information about relaxation techniques that could be used as strategies before and after exposure events and for general self-care.

Third, the thinking traps section would contain information on and examples of common cognitive errors that could occur during planned exposure activities.

All user actions were mapped out and the details of each interaction were defined. Content (text) was later created to support app functions and tasks.

We used the wireframes to create a predesign prototype of the MindClimb app using InVision (InVisionApp Inc), a digital product design platform [18] used by our app developer. The predesign prototype was visually similar to the black-and-white line art used in the wireframes (Figure 3).

Figure 3. A screenshot of several phone screens for the predesign prototype of the MindClimb app.



The predesign prototype was shared with NYAC members and CBT clinicians during separate 1-hour brainstorming sessions. The youth session was held via a web conference and led by 2 NYAC youth engagement facilitators who were collaborators on the study [19]. Youth were prompted to generate ideas on app aesthetics (ie, look and feel) and desired features of a user experience (eg, what app features would ease use) (see Multimedia Appendix 1). The clinician session was held in person at the IWK Health Centre and facilitated by a research team lead and coordinator and the app developer. As therapists were oriented to the predesign prototype, they were prompted to discuss the flow and approach of the app's features and

whether any essential treatment components were missing from the app (section 1 in Multimedia Appendix 1). We summarized the ideas generated from both sessions by themes and incorporated them into the development and design plans for the app.

We initiated app design after the predesign consultations concluded. During app design, our team completed 3 steps: (1) review of the style board, a collage of different design styles and elements, to establish the design direction of the app, (2) approval of a creative brief to establish the aesthetic approach to be used for the app, and (3) approval of the interface design by reviewing several visual mock-ups of wireframe screens that were based on the direction defined in the creative brief. The result of these steps was a design plan that was used by our app developer to create a low-fidelity MindClimb prototype in the InVision platform [18].

The low-fidelity prototype underwent evaluation by adolescents receiving outpatient CBT treatment. Adolescents participated in a digitally recorded focus group facilitated by a research team lead and coordinator and the app developer. As adolescents interacted with the prototype, they were prompted to discuss app features and use, including difficulty in creating exposure events and navigating the app (verbal ratings on a scale of 1 to 5: 5=hard, 1=easy) (section 2 of Multimedia Appendix 1). We transcribed the recorded sessions and thematically analyzed the discussions using a general inductive approach [20]. We used the themes and verbal ratings to finalize the programming on an iOS version of the MindClimb app that was fully functional (a high-fidelity prototype).

Phase 2

We used a mixed-method approach in 2 iterative testing cycles to refine the high-fidelity MindClimb prototype. In both cycles, we asked adolescents to evaluate the usability and acceptability of the prototype, and in cycle 2 we also evaluated its learnability. In cycle 2 we asked app developers to conduct a heuristic evaluation of the prototype to identify usability problems with the user interface design [14,21]. During testing, MindClimb was supported by TestFlight (Apple Inc) [22], an app designed to support beta testing that was used by our app developer.

Adolescent testing sessions took 1 hour to complete and were led by a research team lead and coordinator and the app developer. The sessions were structured according to a think-aloud testing guide (Multimedia Appendix 2) and digitally recorded to facilitate data analysis. Adolescents who did not own or have access to an iOS operating system smartphone that was required to download MindClimb were provided with a smartphone or tablet on which to use MindClimb during the testing cycle.

In cycle 1, adolescents first participated in a walk-through exercise on how to create an exposure activity (fear hierarchy) using the app's ladder function. Adolescents were then asked to complete their own ladder, and facilitators engaged them during this task by posing questions from the testing guide to document experiences with task completion. Field notes documented by the facilitators were used to inform results interpretation. Adolescents were encouraged to use the app on their own before cycle 2 testing (one week after the first cycle). We chose this timeframe because we wanted to understand learnability in the timeframe in which adolescents would use the app during CBT (ie, between weekly therapy sessions). Cycle 2 proceeded similarly to cycle 1 and included gathering feedback regarding adolescents' use of the app during the previous week. At the end of both cycles, adolescents rated their experience with the app by completing the 10-item System Usability Scale (SUS) [14].

Data from the think-aloud activity were examined using a template approach to analysis [23]. We reviewed the findings from cycle 1 after the session to allow the developer to make necessary modifications to MindClimb prior to cycle 2. We interpreted mean SUS scores using published recommendations: a mean score >70 indicated acceptable usability, 50 to 70 indicated marginal usability, and <50 indicated that the app was not acceptable (score range was 0-100). Given the user-centered process that we undertook for app development, we hypothesized that adolescents would report high acceptability and learnability, as supported by the qualitative think-aloud data and an acceptable usability score (≥70).

Following adolescent testing, we asked mobile app developers to evaluate MindClimb based on Nielsen's usability heuristics using a 5-point severity ranking scale (SRS) [17]. Median scores from the SRS were interpreted as follows: 0=none, 1=cosmetic, 2=minor, 3=major, and 4=catastrophic [24]. We hoped that developers would report a low level of errors (ratings of 1 or 0) using the SRS. Any median SRS scores ≥1 were addressed in a final iOS version of the app, and an Android version was then built.

Phase 3

In the final phase, we conducted a case series study to assess the usability of MindClimb among adolescents and therapists during group CBT for anxiety. The group therapy setting was a more convenient approach to recruit from and to study app use during real-world clinical care. Specific usability objectives were to (1) track frequency of skills practice between CBT sessions through self-report of ladder use (use of exposure hierarchies) and other MindClimb features as part of CBT and (2) identify satisfaction as well as barriers and facilitators to MindClimb use during CBT.

Following enrollment, a research coordinator collected adolescent demographic data (age and gender) and contact information (email address and phone number) and recorded the adolescent's preference for future contact (phone, text, or email). Therapists were asked to provide their names and contact information (phone and email).

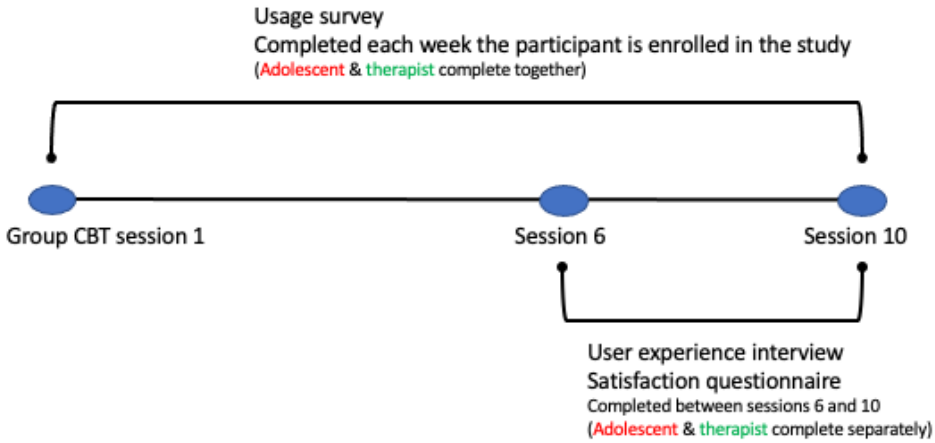
A research team lead, research coordinator, and app developer provided training sessions for adolescents and clinicians on downloading MindClimb to their smartphones and using the app in conjunction with weekly CBT sessions; therapists were not required to download the app, as they would not be actively using it but rather, supporting their patient during use. MindClimb was available for download in both the iOS App Store and Google Play Store. Training was guided by a review of a MindClimb quick start user guide provided to participants.

During training, therapists could also download the app for the purposes of understanding the app's functions so that discussion of app use could be incorporated into treatment sessions, and they were provided the same review of the MindClimb quick start user guide.

Data were collected throughout the group therapy period (Figure 4). Adolescents began using the app with sessions 3 to 9 to coincide with treatment content on exposure activities. The use of key CBT features in MindClimb—ladders (exposure hierarchies), thinking traps, and relaxation strategies—was assessed weekly, with therapists completing the survey with each adolescent to report how frequently the adolescent used MindClimb between treatment sessions and rating the helpfulness of features they used over the same period on a scale of 1 (least helpful) to 10 (most helpful) (Multimedia Appendix 3). After 6 to 7 treatment sessions, adolescents and therapists were contacted to report satisfaction and experience with using MindClimb during treatment. Satisfaction with app use was measured using a study-created questionnaire based on the Client Satisfaction Questionnaire [25] for adolescents (5 questions asked; score range of 0-15, with higher scores indicating higher satisfaction) and therapists (4 questions asked; score range of 0-12, with higher scores indicating higher satisfaction) (Multimedia Appendix 4). User experience interviews were conducted by a research team lead (clinician interviews) or research coordinator (adolescent interviews). Therapists completed the interview in person; adolescents had the option to complete the interview over the phone or through online chat (eg, using iMessage). If an adolescent was not comfortable completing the interview over the phone or using online chat, they had the option to complete an online version of the interview in SurveyMonkey that used open fields for text entry. None of the adolescents in phase 3 chose this online option. Interviews were based on a semistructured guide featuring the 14 domains of the Theoretical Domains Framework (TDF) [26], with versions for therapists (Multimedia Appendix 5) and adolescents (Multimedia Appendix 6). The TDF provided a useful approach to identifying experiences, barriers, and facilitators to MindClimb use. Interviews were digitally recorded and transcribed.

We used descriptive statistics (eg, means with standard deviations, proportions and percentages, medians with range) to summarize demography and MindClimb satisfaction and use. We also used box plots to visually display MindClimb use. We used directed content analysis to synthesize interview data [27]. We were unable to recruit our intended sample size of therapists and patients, so intracluster correlation was not determined for continuous outcomes, and adjusted mean change scores were not calculated.

Figure 4. Phase 3 data collection timeline. CBT: cognitive behavioral therapy.



Results

Phase 1

Predesign consultations on MindClimb were held in November 2015 with 6 NYAC youths via a web conference and in December 2015 with 3 therapists (9 participants in total). Feedback provided on the predesign prototype is summarized in Table 1, alongside the impact of this feedback on MindClimb development and design. Feedback centered on 4 areas for app development and design: (1) the app should be responsive to user needs and preferences, (2) the app should contain features relevant to the practice of CBT for anxiety, (3) the app should be easy to use and navigate, and (4) app design should prioritize aesthetics.

Table 1. Summary of the feedback provided during the predesign consultations.

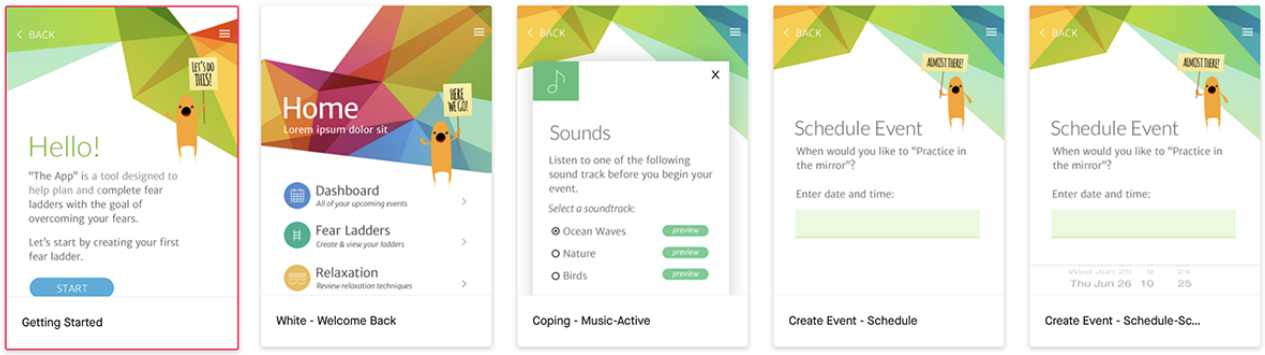
Development area and feedback	Impact on MindClimb development and design
The app should be responsive to user needs and preferences	
Ensure app connectivity with phone calendar (youth)	No change made, as this feature was already intended.
Do not require Wi-Fi connection (youth)	No change made, as this feature was already intended.
Personalization should be possible (youth)	Users will have the ability personalize exposure activities and coping strategies.
Make the app interactive (youth)	The app will contain interactive features, such as rewards and points for app use and pop-ups with positive affirmations.
Build in rewards; completion of exposure activities unlocks new reward content (therapist)	Completion of exposure activities will unlock new reward content.
The app should contain features relevant to the practice of CBT[a] for anxiety	
More information on self-care strategies (therapist)	The app will contain text and videos on self-care strategies.
Consider a live chat with counselors (youth)	No change made, as it is beyond the scope of the purpose of the app.
Include a “tools” section that contains static information along with interactive components (therapist)	The app will contain text and videos on self-care strategies. Interactivity will be limited to videos.
Include an “other” option under suggested self-care and coping strategies to allow patients to identify a personal strategy not listed	The app will include an “other” option under suggested self-care and coping strategies.
Ensure users can repeat scheduled exposure events (therapist)	Users will have the ability to retry and reschedule events.
The app dashboard should list overdue events before upcoming events (therapist)	Events created in the app will appear in a single stream, with overdue events flagged as being overdue. Users will have the option to view and reschedule these events from the dashboard.
The app should be easy to navigate	
Ensure the app is not cluttered and has clean lines (youth)	Content for each app section will be brief. Visual clutter will be avoided.
Make sure there is less content on screens (ie, more pages or screens for content) (youth)	Each screen will support a single action for the user.
Tabs to facilitate use (youth)	No change made; the main dashboard will be used.
Images for guiding steps or how to use (youth)	Videos will be used to provide information on how to practice specific relaxation techniques; images will not be used for guiding the creation of exposure activities, but sample fear hierarchies (ladders) will be provided.
Ensure that it is easy to move between functions (youth)	App functions will be sequential and logical.
Self-explanatory icons (youth)	Each app icon will reflect the section content that it represents.
Ensure that it is easy to scroll through content (youth)	Content for each app section will be brief so that scrolling down for information or functions will not occur.
Ensure transitions between functions are quick (no lagging) (youth)	Push notifications will be limited to reduce the likelihood of noticeable lag in the running of the app.
App design should prioritize aesthetics	
Big lettering for titles, large fonts for important content (youth)	All feedback incorporated into the creative direction and design options (storyboard) developed for the app.
No advertisements (youth)	Incorporated into design options.
“Calm” aesthetic (youth)	Incorporated into design options.
Bright colors, neutral colors, not too many colors (youth)	Incorporated into design options.
Twitter-style text, brief and concise (youth)	Incorporated into design options.
Have a good graphic artist/designer (youth)	Incorporated into design options.
Use sans serif font (youth)	Incorporated into design options.

[a]CBT: cognitive behavioral therapy.

The app was further developed based on youth and clinician feedback. A low-fidelity MindClimb prototype was then created. Screens from the prototype are presented in Figure 5. The prototype native app had full functionality and could be used by youth and therapists from an iOS device without requiring an internet connection.

Feedback on the low-fidelity prototype was collected in June 2016 from 4 adolescents receiving outpatient CBT. We were unable to recruit our intended minimum of 7 participants. Participating adolescents reported preferring a colorful palette with white accents. They understood what the MindClimb app was for by looking at it and found it easy to create exposure events in the app (4 ratings of "1"). Ease of navigation was also rated as easy (ratings of "1" and "2") and app sections were not confusing. Adolescents indicated that preferred terms in the app were "step ladders" rather than "fear ladders" for exposure activities and "helpful thoughts" versus "thinking traps" in order to shift to more positive or neutral wording. The app was modified to use "step ladders," but we chose not to use "helpful thoughts," as this change in wording did not accurately convey the content in the thinking traps section of the app. However, we did review the relaxation section of the app to ensure helpful thoughts were conveyed.

Figure 5. Screens from the low-fidelity MindClimb prototype.



Phase 2

Testing the high-fidelity MindClimb prototype took place in November 2016. A total of 10 adolescents participated in cycle 1, and 7 of these adolescents also participated in cycle 2. Of the 10 participants, 5 were female (50%) and 3 were male (30%); gender was not documented for 2 adolescents. The mean age across 7 adolescents was 15.1 (SD 0.90) years; age was not documented for 3 adolescents. The mean SUS score from both cycles was 77, indicating acceptable usability. Results from the think-aloud activities revealed only minor issues with the app. Issues identified by the adolescents in the testing cycles and changes made to the app are outlined in Table 2. Changes to the app were made prior to the heuristic-based evaluation.

Table 2. Summary of adolescent feedback generated during usability testing cycles that was incorporated into MindClimb.

App feature and feedback	Changes applied
Security and privacy	
Cycle 1: password protection is important	Incorporated an option for the user to create a PIN[a] that must be entered each time the app is opened
Cycle 2: "a password to get into the app or ladders"	Syncing of events with main calendar is optional
Notifications	
Cycle 1 generated varied feedback on the frequency of user notifications, including wanting daily reminders, reminders 1 to 2 times per week, and notifications after a period of 1 to 2 weeks of inactivity	Added single "We've missed you!" notification to pop up after 1 week of having not used the app
Cycle 2: alerts were described as sometimes overwhelming and stressful. Adolescents wished for an option to turn off notifications	Users will receive 1 notification alert when their event is scheduled to occur
Step ladder creation and management	
Cycles 1 and 2: a scale of 1 to 10 to rate the difficulty of exposure activities is more appropriate than 1 to 5	Difficulty scale for exposure activities changed to 1 to 10
Cycle 2: adolescents wanted the option to complete an exposure activity several times before proceeding to the next step in their ladder	Added the ability to repeat an event
Cycle 2: scheduling specific dates and times for each exposure activity might be difficult to plan; app should allow users to "opt out of dates" and dates should not be mandatory	Selection of dates for exposure activities is optional
Cycle 2: adolescents wanted ability to change the order of exposure activities and an easy way to "move the steps around"	Exposure activities, regardless of difficulty rating, can be rearranged by dragging and dropping
Cycle 2: adolescents wanted the ability to not have a coping strategy attached to an exposure activity	Activities can be created without selecting a coping strategy
Positive reinforcement and rewards	
Cycle 1: "Quotes would be good. Ones that make you smile more or positivity"; creatures were a popular reward, especially the ability to upgrade characters; "increase the content" that the monster character says	Incorporated random short positive messages provided by the monster characters at various points within app
Cycle 2: adolescents wanted monster characters to relate quotes to them; "it's fun to use, but the point system doesn't do anything right now. If the points system unlocked another guy (character)"	Added 3 additional monster characters (total of 4) that can be unlocked for completing activities within the app
Positive self-talk	
Cycle 2: adolescents wanted to be able to record "own memo (voice), positive self-talk"	Allow users to record their own positive self-talk messages and replay them during exposure activities
Relaxation techniques and thinking traps	
Cycle 1: adolescents wanted a voice that could walk them through the relaxation techniques	Added voice narration and video as a way for users to access selected relaxation techniques during events
Cycles 1 and 2: too much text on deep breathing and thinking traps	Text edited for length and clarity
Cycle 2: techniques learned in CBT[b] are not in the app	Added "acting as if" relaxation techniques

[a]PIN: personal identification number.

[b]CBT: cognitive behavioral therapy.

The heuristic evaluation took place between May and June 2017. A total of 5 developers, all male, participated. For the majority of the app's usability heuristics, the median score indicated no issue. A cosmetic issue regarding load time was identified for the error prevention heuristic and was corrected. Results from the evaluation are presented in Table 3.

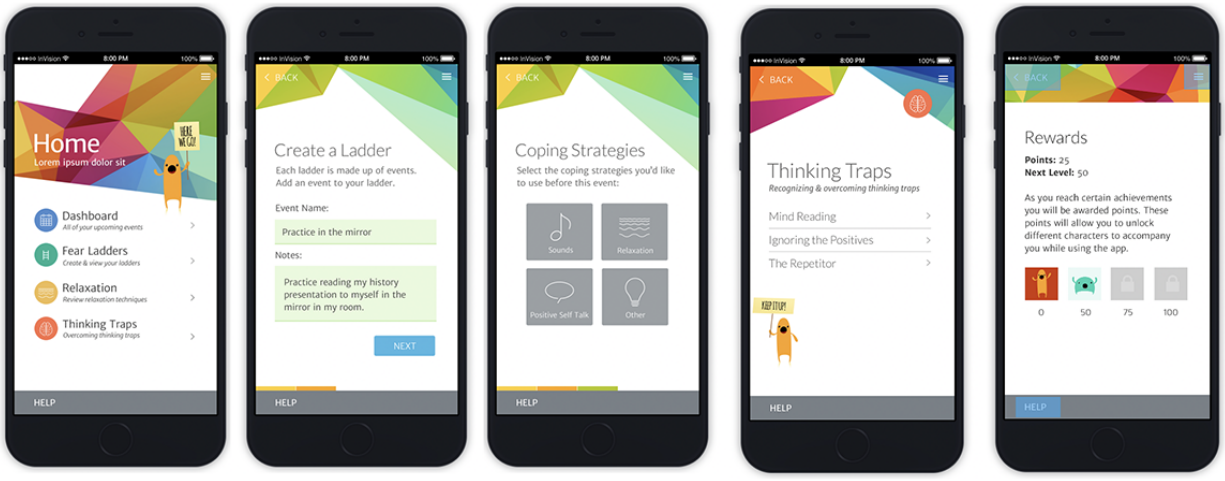
Table 3. Severity ranking scale results from mobile app developers.

Usability heuristic	Median score[a]	Developer comments
1. Visibility of system status	0	"It would be nice to have to see your points or progress somewhere all the time" [Developer 5]
2. Match between system and the real world	0	"Instructions on 'Create an Event' page not very descriptive. Could include example of an event. (e.g., Practice in front of mirror)" [Developer 1]
3. User control and freedom	1	"No immediate 'undo' option for event completion choices, but events can be edited anytime to change completion choice" [Developer 2]
4. Consistency and standards	1	"If different terminology is used in treatment it may take a moment for the terms used in MindClimb to register with users" [Developer 3]
5. Error prevention	0	"The app sometimes takes a long time (~20secs) to move from the splash screen to the home screen on iPhone 5S" [Developer 2]
6. Recognition rather than recall	0	"When adding a relaxation technique, I selected a category and chose a specific technique. When you tap Save and return to the category selection screen, it would be nice to see the specific technique you chose and not only the category selected before hitting next" [Developer 5]
7. Flexibility and efficiency of use	0	"Not a lot of system tailoring present, but given the target audience and system scope, likely not required or useful" [Developer 2]
8. Aesthetic and minimalist design	0	"Some subheading may not be needed (e.g., 'select a ladder to view its events')" [Developer 1]
9. Help users recognize, diagnose, and recover from errors	0	"Did not encounter any error conditions" [Developer 5]
10. Help and documentation	0	"Help documentation was readily available on all screens" [Developer 5]

[a]Heuristic scoring range: 0=none, 1=cosmetic, 2=minor, 3=major, and 4=catastrophic.

Screenshots of the final version of MindClimb are provided in Figure 6. Based on the high-fidelity evaluations from adolescents and app developers, the final version of the app has the following key content and functions: (1) step ladders, (2) dashboard, (3) relaxation, (4) thinking traps, (5) rewards, (6) settings, and (7) help sessions.

Figure 6. Screenshots of key content domains for MindClimb.



First, the app contains step ladders, formerly termed fear ladders. Up to four coping strategies built into the app (eg, audio clips of soothing sounds, text instruction and videos of relaxation techniques) can be identified for each event. When creating ladders, the adolescent indicates how difficult they think it will be to complete the exposure (scale of 1 [easy] to 10 [hard]); after completion of the exposure, adolescents can rate the activity again in terms of its difficulty to actually carry out. Users are able to repeat exposures as part of the step ladder features; ladders can be archived for review of progress and activities.

Second, the dashboard contains all of the upcoming events created for each step ladder. The dashboard provides a quick overview of scheduled events sorted by date.

Third, for the relaxation feature, text and audiovisual clips are available in conjunction with information about relaxation

techniques that can be used as strategies before and after exposure events and for general self-care.

Fourth, the thinking traps section contains brief information on and examples of common cognitive errors.

Fifth, a rewards section tracks adolescent progress with step ladder events. Points are awarded for event and ladder completion; different point levels allow the adolescent to unlock different MindClimb characters that appear with encouraging messages while the app is being used.

Sixth, a settings screen allows the adolescent to protect MindClimb with a 4-digit personal identification number (PIN), set a security question and answer should the PIN be forgotten, and turn on calendar syncing so that MindClimb events sync with the phone's calendar app.

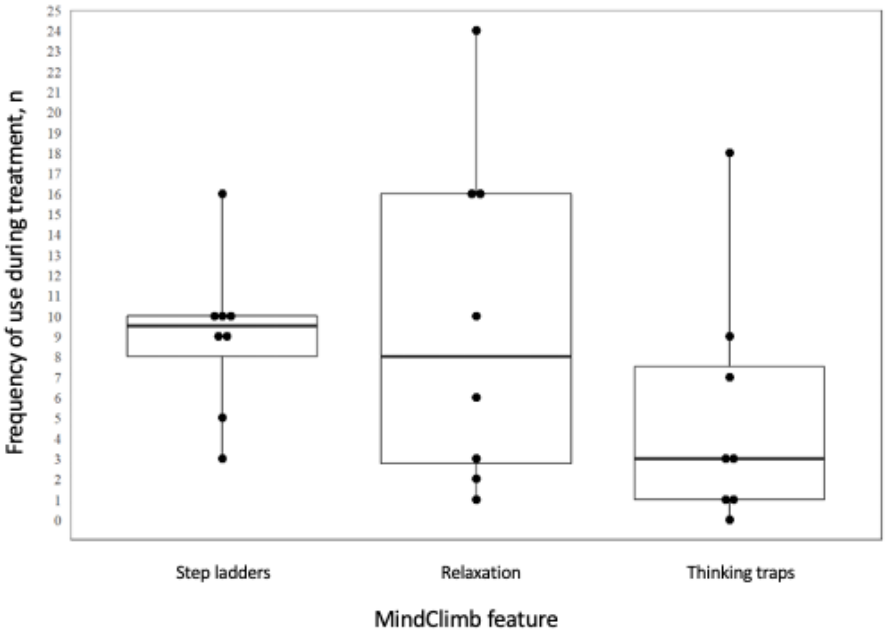
Last, a help session supports use in each content domain (eg, sample ladders are provided in the step ladders section).

Phase 3

A total of 8 adolescents (all girls) and 3 therapists (all women) used MindClimb during group therapy between November 2017 and December 2018. Another 3 adolescents consented to participate, but 1 withdrew from the study and 2 did not download the app for study participation. The 2 adolescents who did not download the app for use did not provide information on why they chose not to use the app. The mean age across the participating adolescents was 14.0 (SD 1.5) years. The median SCARED score reported by adolescents was 56 (range of 46-62); the median SCARED score reported by parents regarding their children was 46 (range of 23-60).

Over the group treatment sessions, MindClimb use began during session 3 (with 2 adolescents), session 4 (with 1 adolescent), session 5 (with 3 adolescents), or session 6 (with 2 adolescents). All the adolescents reported using the app between sessions, and use ranged between features and adolescents (Figure 7); none of the adolescents reported use in the final week of treatment. The largest range of use occurred with the relaxation feature of the app, with some users using it infrequently (once) and some using it regularly (24 uses over a 4-week period). Use of the thinking traps feature was also varied (range of 0-18 uses). The step ladders feature was the most regularly used, with most adolescents using it 9 to 10 times during group treatment.

Figure 7. Median use of MindClimb features over 7 sessions of group cognitive behavioral therapy.



The average satisfaction score among adolescents was 11.3 (SD 1.4). Adolescents rated the step ladder (median rating of 7) and relaxation (median rating of 7) features of MindClimb as slightly more helpful in completing weekly exposure activities than the thinking traps feature (median rating of 6) (Table 4). The average satisfaction score among therapists was 10.0 (SD 2.0). Therapists reported slightly higher overall satisfaction with the app compared with adolescents (Table 4).

Table 4. Adolescent and therapist satisfaction with MindClimb.

Question	Responses, n (%)
Adolescents	
Q1. Satisfaction with the amount of help they received from MindClimb	
Quite dissatisfied	0 (0)
Indifferent or mildly dissatisfied	0 (0)
Mostly satisfied	6 (75)
Very satisfied	2 (25)
Q2. Satisfaction with the kind of support they wanted from MindClimb	
No, definitely not	0 (0)
No, not really	1 (13)
Yes, generally	6 (75)
Yes, definitely	1 (13)
Q3. Whether they would recommend MindClimb to a friend if they needed similar help	
No, definitely not	0 (0)
No, I don’t think so	0 (0)
Yes, I think so	5 (63)
Yes, definitely	3 (38)
Q4. Satisfaction with the amount of support they received from MindClimb to help them deal more effectively with their anxiety	
No, it seemed to make things worse	0 (0)
No, it didn’t really help	0 (0)
Yes, it helped somewhat	5 (63)
Yes, it helped a great deal	3 (38)
Q5. Whether they would use MindClimb again if they were to seek help again	
No, definitely not	0 (0)
No, I don’t think so	0 (0)
Yes, I think so	6 (75)
Yes, definitely	2 (25)
Therapists	
Q1. Satisfaction with the amount of contributions MindClimb made to each treatment	
Quite dissatisfied	1 (33)
Indifferent or mildly dissatisfied	0 (0)
Mostly satisfied	1 (33)
Very satisfied	1 (33)
Q2. Whether they would recommend MindClimb to a colleague who was looking to use an app in their clinical practice	
No, definitely not	0 (0)
No, I don’t think so	0 (0)
Yes, I think so	0 (0)
Yes, definitely	3 (100)
Q3. Whether they would use MindClimb in the future with other patients	
No, definitely not	0 (0)
No, I don’t think so	0 (0)
Yes, I think so	1 (33)

Question	Responses, n (%)
Yes, definitely	2 (67)
Q4. General satisfaction with using MindClimb as part of their clinical practice	
Quite dissatisfied	0 (0)
Indifferent or mildly dissatisfied	0 (0)
Mostly satisfied	1 (33)
Very satisfied	2 (67)

MindClimb experience interviews with adolescents identified that they all participated in the study to see if the app could help them manage their anxiety. All the adolescents felt confident in their ability to use the app, found it easy to use, and stated that experience was not needed to use it. One adolescent noted that some prior CBT sessions would be required so that treatment terminology and skills (eg, step ladders) were learned prior to MindClimb use. A selection of adolescent quotes on the MindClimb app, grouped by our a priori categories (experiences, barriers, facilitators), are provided in Textbox 1. All the adolescents felt that the app fit well with their usual approach to treatment activities and was useful to helping complete treatment activities outside of group sessions. Moreover, 6 of the adolescents reported that they did not experience any negative outcomes associated with using MindClimb. All 8 adolescents found the reminders for use (eg, "miss you" and "where have you been?") that were built into the app were helpful for use and necessary to ensure that they used the app regularly. Six adolescents found that they would forget to use the app if they had a busy day, if they "weren't in the right head space—like feeling mad," or if they had a lot of school homework.

Textbox 1. Adolescent perspectives on using the MindClimb app.

Experiences with MindClimb during treatment

- "I used it for a Step Ladder on how to be comfortable around new people. One of the steps was to sit beside someone on the bus. I felt accomplished when I did it." [Participant 2]
- "I used it when starting a new dance class. In the car I used the relaxation sounds/techniques." [Participant 5]
- "It is nice that it is on my phone and I can use it whenever I need it. It is easy and helpful." [Participant 3]
- "Helped a lot. Not needing to look at book [group materials]. Have it with me all the time on the phone." [Participant 4]
- "Pretty important. Step Ladders in group had reminders and would help me remember to do therapy homework which was helpful." [Participant 6]

Challenges and barriers to use

- "Sometimes worried if making stepladder wrong. Put two activities in same stepladder but figured it out." [Participant 6]
- "Low to mid was more helpful in terms of anxiety level using app. [When] very worried/nervous, [the app] did not help as anxiety too high." [Participant 8]
- "It can get boring with only Step Ladders. Being able to get different coloured characters and new accessories [rewards] would make me use it more." [Participant 2]

Facilitators for use

- "Reminders were good. Step Ladders helped. Homework [from the group] is stressful as is, [the app] helped remind to do it." [Participant 8]
- "Relaxed, planned out [prepared], organized." [Participant 7]
- "Reminded me of the things I learned in group. What they teach is there [in the app] all the time." [Participant 4]
- "… it was really easy to open up app anywhere and use it." [Participant 1]
- "Happy when got points and new [characters]." [Participant 6]

MindClimb experience interviews with therapists identified that they were interested in integrating technology into group sessions. One of the therapists described:

> *We think it would help in engaging in using skills. Also, the importance of having your skills and tools with you at all times, and youth always have their phones with them to refer to. In future, we should make it more part of the group and have youth keep their phones out to use during the group sessions.* [Therapist 3]

Two of the therapists noted that the views of their patients were important to app use:

> *If they didn't find it useful then I would be less likely to try again in the future.* [Therapist 2]

The 3 therapists described feeling confident in incorporating the app into treatment sessions, stating that the app was easy to use and intuitive. However, reminding patients to use it in sessions, including detailing app use in the treatment manuals and having therapists model use during group sessions, were described as approaches for ensuring routine use and maximizing app contributions, and 1 therapist noted that not having access to their patients' step ladders was challenging, as they liked to have copies on file ("ability to review/record information" [Therapist 1]). Selected quotes from the therapists' experiences with MindClimb use are presented in Textbox 2, as are quotes on perceived benefits and limitations to use.

Textbox 2. Therapist perspectives on using the MindClimb app with adolescents during group cognitive behavioral therapy (CBT).

Experience using MindClimb

- "So in group setting [I] was a little hesitant to ask and have them use it. Did not want to make them feel uncomfortable. Worried that would get distracted with navigating app instead of content. Would be great to have the app on computer or iPad on smart board—project on screen and follow on the large screen. This would be easier to help visually." [Therapist 1]
- "Sometimes forgot to remind people to use it. Need reminders for clinicians and possibly integrate in manual so that you remember to incorporate in each part of CBT." [Therapist 3]
- "No ethical or legal issues beyond what youth already record in their book or notes, need to keep it confidential and not able to [be] accessed by others. In most cases book/paper is less confidential as lying around." [Therapist 3]

Benefits to using MindClimb during treatment

- "Much more helpful for youth to use their skills in the moment- huge plus. Also most teens are digital oriented with their phone, so more likely to use than paper, and more fun. More [discreet] as well… Most teens will forget workbook but very few forget their phones." [Therapist 1]
- "Main benefits are that the material and skills are always accessible and anxiety can happen anywhere. Also great for reminding youth to practice. [Helps] the clinician in having skills and tools more readily available for practice (like not forgetting their homework sheet)." [Therapist 3]

Limitations to using MindClimb during treatment

- "Only disadvantage is not being able to share stepladders… Realistic thinking was [also] a challenge as the app did not fit as well with this skill." [Therapist 1]
- "I think having them put in the reminders in-session would have been better; to walk them through how to use it with their own examples. Again, easier done in an individual than a group setting." [Therapist 2]
- "[Need] preplanning with group to incorporate app into group structure. More info structure on how to do it in the group as to how they would use it. Also need to make sure they are using the app for group and not something else on their phone." [Therapist 1]
- "More screen time." [Therapist 3]

Discussion

Principal Results

In this study, a multiphasic user-centered design approach allowed our team to develop, design, and test the MindClimb app with input from designers, developers, adolescents, and therapists using multiple approaches. Guiding principles for app development were the promotion of autonomy and ownership of treatment goals outside of formal treatment sessions, the focus on real-time skills practice, the personalization of treatment, and the simplicity of use under stress. The design of the app incorporated core components in a successful digital mental health tool that was secure and co-designed by stakeholders (adolescent users and clinicians) and that integrated evidence-based skills and usability [28]. The app was designed with high rigor in 3 stages of development and had high usability ratings by users (adolescents) and app developers, with only minor errors that were corrected before the final version. In our real-world usability evaluation, we found that adolescents and therapists were satisfied with the MindClimb app and found it easy to use and helpful in practicing CBT skills during treatment. Adolescents found the relaxation strategies and the step ladders with reminders to practice to be the most helpful components of the app. Therapists identified the benefits of having this app to help youth practice skills between CBT sessions and encouraged further integration of app usage into CBT sessions to encourage regular app usage. Both therapists and adolescents found the tools on the app easy to use and the most relevant to CBT practice. Step ladders and relaxation strategies were the two most used components of the app, and adolescents reported using the app between CBT sessions.

Limitations

MindClimb was designed to be simple and easy to use between CBT sessions to plan and complete real-time exposure activities using skills (cognitive, relaxation, exposure practice, and reward) learned in treatment. As it was not intended to be a broad-scope, stand-alone treatment app, it includes some but not all of the recommended evidence-based treatment components for anxiety disorders: psychoeducation, self-monitoring, cognitive skills, problem solving, exposure activities, and contingency management. In a recent review of the content in smartphone apps marketed for child and adolescent anxiety, Bry et al [29] concluded that there are few comprehensive, evidence-based self-management apps available for use. We propose that fit for purpose also needs to be considered alongside app content to ensure that the proposed intent of an app is supported by its content and that evaluations

of an app are also consistent with its intent. In this regard, we feel that MindClimb has been designed for a specific purpose for use by adolescents (users) in treatment and CBT therapists, and its functions reflect this purpose.

The results of MindClimb implementation are derived from a small sample of adolescents and therapists and should be interpreted with this limitation in mind. We were unable to recruit our intended sample size for phase 3 of the project, which would have allowed us to study app use over a more diverse user population (including, potentially, male users). Recruitment was challenging, as there were not as many CBT groups offered for anxiety during the study period due to a decrease CBT clinician availability. Future implementation studies of MindClimb (and other mHealth apps) need to examine the penetrance of MindClimb in clinical practice and include more diverse populations of users (reflected in gender identity, type of anxiety disorder and treatment setting, etc).

Comparison With Prior Work

A recent review by Wozney and colleagues [30] indicated that acceptability is the most commonly investigated implementation outcome among studies of e-mental health care technology for adolescent depression and anxiety. The development and evaluation of MindClimb contributes additional methods and data on app learnability and initial usability. The field of e-mental health care technology will continue to benefit from studying a broader range of implementation outcomes. For apps such as MindClimb that are intended for use alongside therapy, important implementation outcomes include fidelity (the degree to which an app was implemented during CBT as intended), penetration (the extent of app use within a treatment setting), and sustainability (the extent to which an app's use was maintained in a treatment setting) [31]. Feedback from therapists in this study suggested that workflow and integration issues are critical for embedding new technology in clinical care. Like our study, in other mHealth studies, technology has been regarded as useful, easy, and relevant to CBT practice, but support during use, other commitments, perceived relevance, and difficulty of use can be factors in whether an adolescent uses and continues to use a technology [32,33]. Implementation studies examining fidelity, penetration, and sustainability can introduce new understanding of the impact of apps on the course of treatment and on therapist roles, responsibilities, and workflow and inform the development of tailored implementation strategies to support app use during therapy. Such evaluations can occur within studies testing the impact of the technology on adolescent health outcomes. This hybrid approach to studying intervention effectiveness and implementation [34] offers an efficient and comprehensive approach to studying mHealth technologies in real-world clinical settings.

Conclusions

MindClimb was developed and tested by adolescent users of CBT and therapists in cooperation with app developers. This approach resulted in an mHealth app that was relevant, accepted, and used by adolescents during CBT for anxiety. Evaluation of the use of this app in a clinical practice setting demonstrated that adolescents and therapists generally felt it was helpful and easy to use for CBT practice outside of therapy sessions. As exposure practice is the core component to anxiety CBT treatment and often reported as the most challenging to practice, the app shows promise in helping encourage adolescents to use these skills outside of therapy sessions. Having CBT skills and tools at all times on their phones provides adolescents the opportunity to practice skills more often between sessions. MindClimb increased access to evidence-based CBT skills outside of the formal therapy sessions in a format accepted and used by teenagers. Clinicians found the app helpful in encouraging practice but sometimes forgot to remind their patients; they recommended integrating app usage into the CBT treatment manual. Implementation studies with larger youth samples that examine the integration of technology in clinical care and the impact of the app plus CBT on clinical care processes and patient outcomes are now necessary.

Acknowledgments

The authors would like to acknowledge the contributions of the NYAC youth engagement facilitators who assisted with youth consultation in phase 1; Marcus O'Neill, who helped to coordinate the study; and Jason Murray from Optio Publishing Inc, who helped to develop MindClimb. The study team would also like to thank the Treatment of Anxiety group at the IWK Health Centre in Halifax, Nova Scotia, and the clinicians and adolescents who participated in this study and the creation of MindClimb. This study was funded by the Canadian Institutes of Health Research (CIHR KAL 139705).

Conflicts of Interest

AN, AB, JC, AH, JH, and LW disclose filing a copyright of commercialization of the MindClimb app through TEC Edmonton at the University of Alberta. All other authors have no conflicts to declare.

Multimedia Appendix 1
Questions asked during initial consultations to develop the MindClimb app.
[DOCX File , 76 KB - mhealth_v8i12e18439_app1.docx]

Multimedia Appendix 2
Questions asked during the Think Aloud activity.
[DOCX File , 106 KB - mhealth_v8i12e18439_app2.docx]

Multimedia Appendix 3
MindClimb usage survey.
[DOCX File , 73 KB - mhealth_v8i12e18439_app3.docx]

Multimedia Appendix 4
Satisfaction questions.
[DOCX File , 86 KB - mhealth_v8i12e18439_app4.docx]

Multimedia Appendix 5
User experience interview questions for therapists.
[DOCX File , 126 KB - mhealth_v8i12e18439_app5.docx]

Multimedia Appendix 6
User experience interview questions for adolescents.
[DOCX File , 102 KB - mhealth_v8i12e18439_app6.docx]

References

1. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
2. Glenn D, Golinelli D, Rose RD, Roy-Byrne P, Stein MB, Sullivan G, et al. Who gets the most out of cognitive behavioral therapy for anxiety disorders? The role of treatment dose and patient engagement. J Consult Clin Psychol 2013 Aug;81(4):639-649 [FREE Full text] [doi: 10.1037/a0033403] [Medline: 23750465]
3. Kazantzis N, Whittington C, Dattilio F. Meta-analysis of homework effects in cognitive and behavioral therapy: A replication and extension. Clin Psychol Sci Pract 2010;17:144-156. [doi: 10.1111/j.1468-2850.2010.01204.x]
4. Scheel MJ, Hanson WE, Razzhavaikina TI. The Process of Recommending Homework in Psychotherapy: A Review of Therapist Delivery Methods, Client Acceptability, and Factors That Affect Compliance. Psychother Theor Res Pract Training 2004;41(1):38-55. [doi: 10.1037/0033-3204.41.1.38]
5. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
6. Grist R, Porter J, Stallard P. Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
7. Pramana G, Parmanto B, Kendall PC, Silk JS. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed J E Health 2014 May;20(5):419-427 [FREE Full text] [doi: 10.1089/tmj.2013.0214] [Medline: 24579913]
8. Silk JS, Pramana G, Sequeira SL, Lindhiem O, Kendall PC, Rosen D, et al. Using a Smartphone App and Clinician Portal to Enhance Brief Cognitive Behavioral Therapy for Childhood Anxiety Disorders. Behav Ther 2020 Jan;51(1):69-84. [doi: 10.1016/j.beth.2019.05.002] [Medline: 32005341]
9. Clough BA, Casey LM. Technological adjuncts to enhance current psychotherapy practices: a review. Clin Psychol Rev 2011 Apr;31(3):279-292. [doi: 10.1016/j.cpr.2010.12.008] [Medline: 21382535]
10. Preziosa A, Grassi A, Gaggioli A, Riva G. Therapeutic applications of the mobile phone. Br J Guidance Counselling 2009 Aug;37(3):313-325. [doi: 10.1080/03069880902957031]
11. Gentili C, Zetterqvist V, Rickardsson J, Holmström L, Simons L, Wicksell R. ACTsmart - development and feasibility of digital Acceptance and Commitment Therapy for adults with chronic pain. NPJ Digit Med 2020;3:20 [FREE Full text] [doi: 10.1038/s41746-020-0228-4] [Medline: 32128450]
12. van Eijk D, van Kuijk J, Hoolhorst F, Kim C, Harkema C, Dorrestijn S. Design for Usability; practice-oriented research for user-centered product design. Work 2012;41(Suppl 1):1008-1015. [doi: 10.3233/WOR-2012-1010-1008] [Medline: 22316853]
13. Wilson K, Bell C, Wilson L, Witteman H. Agile research to complement agile development: a proposal for an mHealth research lifecycle. NPJ Digit Med 2018;1:46. [doi: 10.1038/s41746-018-0053-1] [Medline: 31304326]
14. Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability Inspection Methods. New York, NY: John Wiley & Sons; 1994:25-62.
15. Hale WW, Raaijmakers Q, Muris P, Meeus W. Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in the general adolescent population. J Am Acad Child Adolesc Psychiatry 2005 Mar;44(3):283-290. [doi: 10.1097/00004583-200503000-00013] [Medline: 15725973]
16. Rees H, Noyes JM. Mobile telephones, computers, and the internet: sex differences in adolescents' use and attitudes. Cyberpsychol Behav 2007 Jun;10(3):482-484. [doi: 10.1089/cpb.2006.9927] [Medline: 17594276]

17. Nielsen J, Mack RL, editors. Usability Inspection Methods. New York, NY: John Wiley & Sons; 1994.
18. InVisionApp Inc. URL: https://www.invisionapp.com [accessed 2020-11-05]
19. Heffernan OS, Herzog TM, Schiralli JE, Hawke LD, Chaim G, Henderson JL. Implementation of a youth-adult partnership model in youth mental health systems research: Challenges and successes. Health Expect 2017 Dec;20(6):1183-1188 [FREE Full text] [doi: 10.1111/hex.12554] [Medline: 28295940]
20. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
21. Tan W, Liu D, Bishu R. Web evaluation: Heuristic evaluation vs. user testing. Int J Ind Ergonomics 2009 Jul;39(4):621-627. [doi: 10.1016/j.ergon.2008.02.012]
22. TestFlight. Apple Store Connect. URL: https://developer.apple.com/testflight [accessed 2020-11-05]
23. 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]
24. Bangor A, Kortum PT, Miller JT. Determining what individual SUS scores mean: Adding an adjective rating scale. J Usability Stud 2009;4(3):114-123. [doi: 10.5555/2835587.2835589]
25. Nguyen TD, Attkisson CC, Stegner BL. Assessment of patient satisfaction: development and refinement of a service evaluation questionnaire. Eval Program Plann 1983;6(3-4):299-313. [doi: 10.1016/0149-7189(83)90010-1] [Medline: 10267258]
26. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
27. 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]
28. Wilhelm S, Weingarden H, Ladis I, Braddick V, Shin J, Jacobson NC. Cognitive-Behavioral Therapy in the Digital Age: Presidential Address. Behav Ther 2020 Jan;51(1):1-14 [FREE Full text] [doi: 10.1016/j.beth.2019.08.001] [Medline: 32005328]
29. Bry LJ, Chou T, Miguel E, Comer JS. Consumer Smartphone Apps Marketed for Child and Adolescent Anxiety: A Systematic Review and Content Analysis. Behav Ther 2018 Dec;49(2):249-261. [doi: 10.1016/j.beth.2017.07.008] [Medline: 29530263]
30. Wozney L, McGrath PJ, Gehring ND, Bennett K, Huguet A, Hartling L, et al. eMental Healthcare Technologies for Anxiety and Depression in Childhood and Adolescence: Systematic Review of Studies Reporting Implementation Outcomes. JMIR Ment Health 2018 Jun 26;5(2):e48 [FREE Full text] [doi: 10.2196/mental.9655] [Medline: 29945858]
31. 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;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
32. Spence SH, Donovan CL, March S, Gamble A, Anderson R, Prosser S, et al. Online CBT in the Treatment of Child and Adolescent Anxiety Disorders: Issues in the Development of BRAVE–ONLINE and Two Case Illustrations. Behav Cogn Psychother 2008 Jul 01;36(4):411-430. [doi: 10.1017/s135246580800444x]
33. Gerrits RS, van der Zanden RA, Visscher RF, Conijn BP. Master your mood online A preventive chat group intervention for adolescents. Aust e-Journal Adv Ment Health 2014 Dec 17;6(3):152-162. [doi: 10.5172/jamh.6.3.152]
34. 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]

Abbreviations

CBT: cognitive behavioral therapy
EMI: ecological momentary intervention
mHealth: mobile health
NYAC: National Youth Advisory Committee
PIN: personal identification number
SCARED: Screen for Child Anxiety Related Emotional Disorders
SRS: severity ranking scale
SUS: System Usability Scale
TDF: Theoretical Domains Framework

Edited by G Eysenbach, L Buis; submitted 26.02.20; peer-reviewed by J Rickardsson, T Ungar; comments to author 18.03.20; revised version received 13.10.20; accepted 28.10.20; published 08.12.20.

Please cite as:
Newton A, Bagnell A, Rosychuk R, Duguay J, Wozney L, Huguet A, Henderson J, Curran J
A Mobile Phone–Based App for Use During Cognitive Behavioral Therapy for Adolescents With Anxiety (MindClimb): User-Centered Design and Usability Study
JMIR Mhealth Uhealth 2020;8(12):e18439
URL: https://mhealth.jmir.org/2020/12/e18439
doi:10.2196/18439
PMID:33289671

Original Paper

User Experience of Mobile Personal Health Records for the Emergency Department: Mixed Methods Study

Su Min Kim[1], BSc; Taerim Kim[2], MD, PhD; Won Chul Cha[1,2,3], MD; Jae-Ho Lee[4,5], MD, PhD; In Ho Kwon[6,7], MD, PhD; Yuri Choi[6,7], MD; June-Sung Kim[4], MD

[1]Department of Digital Health, Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul, Republic of Korea
[2]Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[3]Health Information and Strategy Center, Samsung Medical Center, Seoul, Republic of Korea
[4]Department of Emergency Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[5]Department of Information Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[6]Department of Emergency Medicine, Dong-a University Hospital, Busan, Republic of Korea
[7]Department of Emergency Medicine, Dong-a University College of Medicine, Busan, Republic of Korea

Corresponding Author:
Taerim Kim, MD, PhD
Department of Emergency Medicine
Samsung Medical Center
Sungkyunkwan University School of Medicine
81 Irwon-ro Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 1021252342
Email: taerimi.kim@samsung.com

Abstract

Background: Personal health records (PHRs) can be useful in the emergency department, as they provide patient information in an accurate and timely manner and enable it to be used actively. This has an effect on patients' health outcomes and patient experience. Despite the importance of PHRs in emergencies, there are only a few studies related to PHRs in emergencies that evaluate patient experience.

Objective: This study aims to introduce the novel mobile PHR (mPHR) platform to emergency environments and assess user experience.

Methods: The study was conducted from October 2019 to November 2019. In total, 1000 patients or carers in the emergency departments of 3 hospitals were provided an application-based service called FirstER, which was developed to collect and utilize medical information for patients in the emergency department. This study was performed as a mixed methods study. After using FirstER, we investigated its usability and conducted a survey on the experience of obtaining medical information with a legacy system and with FirstER. Additionally, we interviewed 24 patients to gain insight into their experiences regarding medical information using FirstER. For the quantitative analysis, the survey results were analyzed using descriptive statistics (mean and standard deviation). For the qualitative analysis, we determined the keywords and their frequencies from each survey question and interview question.

Results: In total, 1000 participants, consisting of both patients and carers, were recruited in this study. Their mean age was 41.4 (SD 13.3) years. We ascertained participants' satisfaction with FirstER and their mPHR needs through a survey and an in-depth interview. With the current system, participants were not well aware of their health conditions and medical information, and they were passive in the use of their medical information and treatment. However, they wanted their medical information for several reasons, such as information sharing and managing their health conditions. FirstER provided participants with their needed information and an easy way to access it. The mean System Usability Scale (SUS) value was 67.1 (SD 13.8), which was considered very near to acceptable.

Conclusions: This study is the first to implement mPHRs in the emergency department of large tertiary hospitals in the Republic of Korea. FirstER was found to enhance user experience in emergencies, as it provided necessary medical information and proper

user experience. Moreover, the average SUS was 67.1, which means that participants found FirstER to be very near to acceptable. This is very encouraging in that FirstER was developed within a very short time, and it was a pilot study.

Trial Registration: Clinicaltrials.gov NCT04180618; https://clinicaltrials.gov/ct2/show/NCT04180618

(JMIR Mhealth Uhealth 2020;8(12):e24326) doi:10.2196/24326

KEYWORDS

personal health records; mobile health; patient engagement

Introduction

Importance of Information in the Emergency Department

Patient information is very important in the emergency department. Medical staff can treat patients appropriately when they have accurate patient information in a limited time, which affects the patients' health outcome [1-4]. From the patient's point of view, some studies show that a more positive experience was had in terms of improved health outcomes or emotional aspects during the treatment process when the patient received medical information about themselves [5-8]. However, it is relatively difficult to obtain or provide medical information for emergency patients as compared to other patients, such as outpatients; handling patient information is challenging because they visit the hospital unexpectedly and patient information is often managed by a separate hospital [4-6]. Therefore, it is important to develop ways for providing patient medical information to satisfy the special situations and needs of emergency patients, which, ultimately, contribute to the improvement of health outcomes [9].

Patient Health Records Enhance Care Effectiveness by Providing Patient Medical Information

Despite the importance of access to and provision of complete patient information, access to on-demand medical information is not well achieved [10]. In this situation, a feasible solution is the patient health record (PHR), which allows patients to generate and manage their overall information [11-16], as well as the mobile PHR (mPHR), which is linked with mobile phones and can access information from anywhere. These enable continuous care and follow-up by aggregating patient information that had previously been fragmented around medical institutions for the patient's focus [14,17]. In addition, by better understanding one's health through one's own medical information, patients can actively participate in treatment-related activities such as decision-making and medication compliance, which positively affect self-management effectiveness. Consequently, it improves care outcomes and patient status [11-16].

Importance of User Experience of PHR for the Emergency Department

We expected to improve the patient experience by providing information through mPHR in emergency situations. However, studies conducted thus far have focused on PHR research in nonemergency situations [13,15]. In the context of emergency situations, there were studies on instruction and education to improve the discharge process of patients, but this did not utilize PHR [5]. For PHR in emergency situations, some studies were conducted with respect to setting systems for PHR [18]. Additionally, a study reported that both patients and medical staff were willing to use PHR [18]. Although many studies have suggested that PHRs have a positive effect on patients, it is necessary to observe the user experience of mPHR in real emergency situations because the utility of a system depends on a specific environment [19]. We wanted to ensure that we provided easy-to-understand patient information that was genuine and updated [6,20,21] and that the patients could use the mPHR application well [6,20-22]. Research on these patients' experiences will lead to better patient participation and satisfaction, and ultimately, may achieve the goal of improving the quality of health care [20,23].

Objective

This study aimed to introduce the novel mPHR platform to emergency environments and to assess user experience.

Methods

Development of the Novel Mobile PHR: FirstER

FirstER, a mobile application-based platform, was developed to collect and utilize medical information, especially for emergency patients and medical staff. Our previous study described the content of FirstER [24]. We conducted interviews and surveys with various stakeholders in emergencies to identify PHR service requirements in emergency medical environments, services and functions based on PHRs, patients' willingness to provide information for PHR services, and items of medical value for PHR development. That study proved the validity of the need for PHRs in the emergency medical environment and was used as basic data before implementing practical services; therefore, it became the basis for organizing the information used in this study.

Further, medical staff reassessed the results to determine the most necessary details and organized the information. Later, the whole system was created, including a mobile application for patients and a web page for medical staff. Servers were set up in each hospital, and they were connected with a security cloud service so that the medical information from 3 hospitals could be gathered in a cloud assigned for each patient. With the mobile application and web page, patients and medical staff could access the medical information of the patients in the emergency department.

Structure of the FirstER

Overview

The following process was carried out: First, when a patient visited the emergency department, they downloaded the FirstER application on their mobile phone, after which the patient created an account and agreed to provide personal information. Second, the cloud received the subscriber's information (including patient ID) and sent it to the linked server in the hospital system. Third, the subscriber's emergency department data were extracted from the hospital information system server, and the subscriber's data were sent to the service-linked server. Fourth, the service-linked server sent the subscriber's emergency department data to the cloud. Fifth, in the cloud, the emergency department data were stored and sent to the application. Lastly, the patient could check their emergency department data on their mobile phone.

Once the patient agreed to reveal and share their medical information with medical staff through an emailed link to a medical information summary, the medical staff could connect to the web page showing the patient's medical information (Figure 1).

FirstER consists of 2 categories: health records, and information management and setting (Figure 2).

Figure 1. Overview of the FirstER system. ①~⑥ (line): the process in which patients obtained their medical information. ⓐ~ⓒ (dotted line): the process in which medical staff accessed the patient's medical information. When a user sent a request to the client, the webserver processed the command and sent an answer back to the user. The user requests that the webserver could not process were sent to the WebSphere Application Server (WAS), and the results were handed over to the users after they were received. It provided static content such as HTML, CSS, etc. The WAS server provided dynamic content, such as DB inquiry, processing logic, etc. ER: emergency room; DB: database.

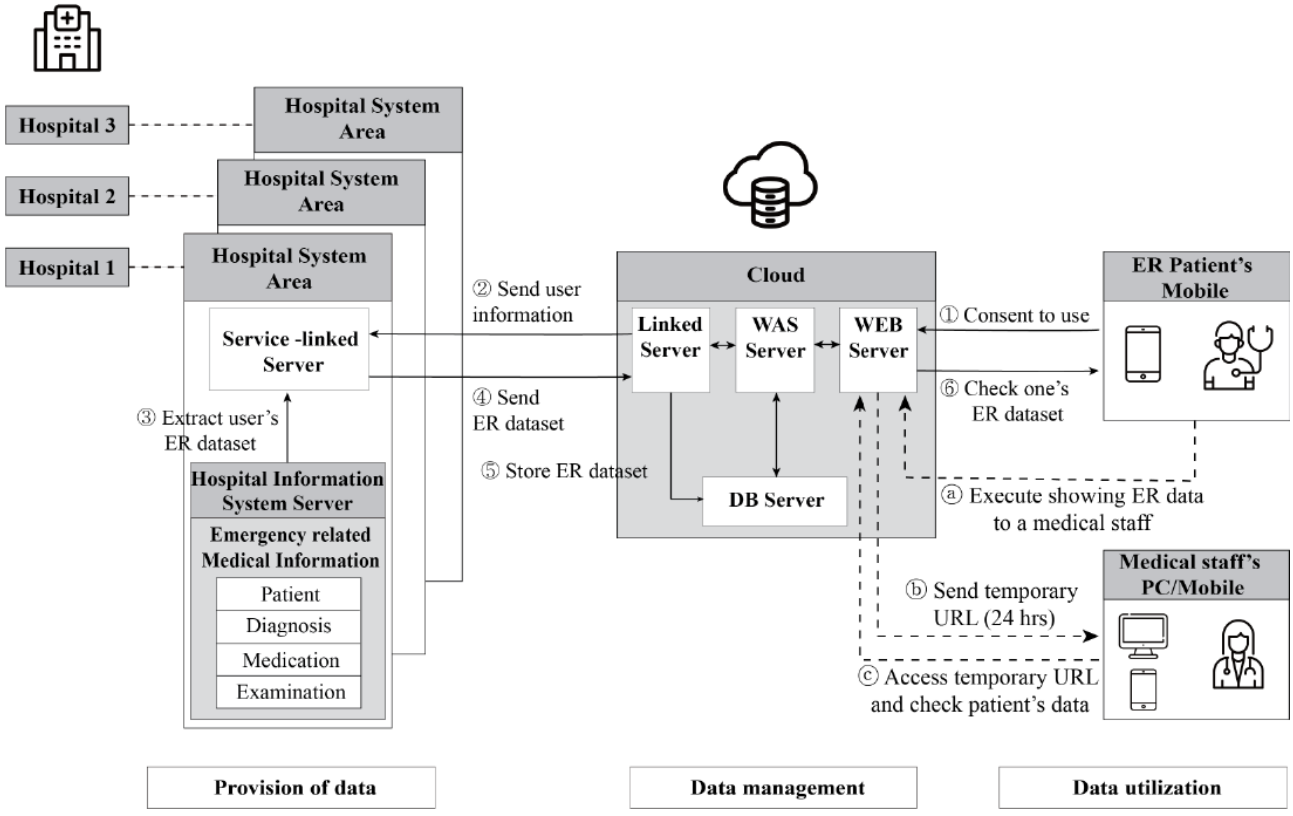
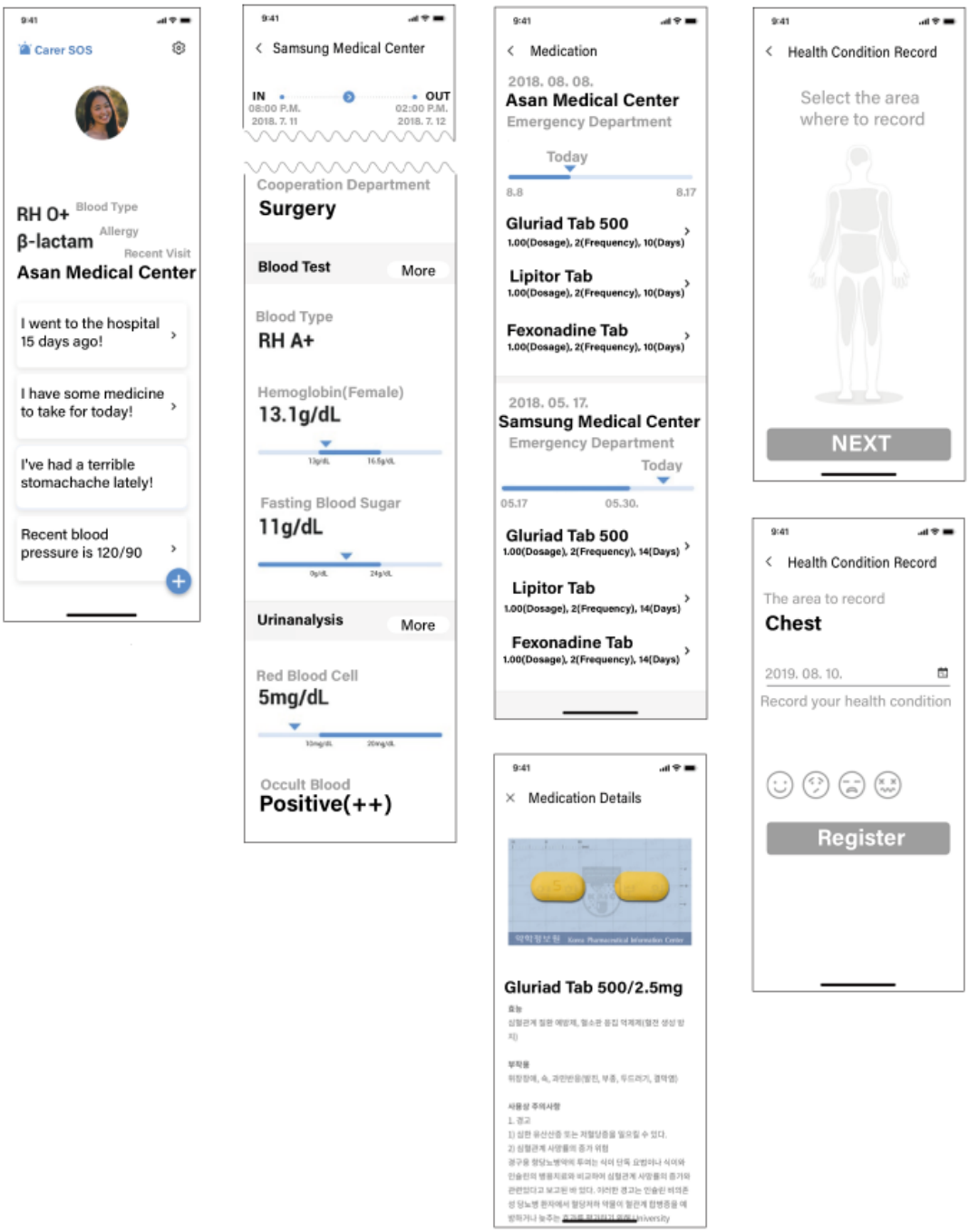


Figure 2. Screens of the FirstER application. (a) The main page shows the user's brief information for an emergency, such as name, gender, age, blood type, allergy, and recent hospital visits. Additionally, there is a button for sending a message to the user's carer in case of an emergency. Moreover, there are 4 sections for hospital visit records, medication, pain, and blood pressure. (b) Hospital visit records are presented; when the user clicks on a hospital visit record, the user can access specific information, such as lab test results. (c) Medication records allow the user to see records of prescribed medications. (d) FirstER has self-record components for the user to input their pain and blood pressure; on the Pain record screen, the user can choose where they feel pain and how it felt, and the user can input brief, descriptive text. The app screen language was translated to English.



Health Records

The health records comprised 2 sections: hospital records and self-reported information. The hospital records involved visit records and information on medication. When the patient who created an account in the service logged into the application, they could see the main page, which featured a summary of their records (name, gender, age, blood type, and recent hospital visits) at the top of the page. Additionally, there were 4 buttons showing hospital visits, medication, self-reports of pain, and

blood pressure. From the main page, the patient could download the emergency department information by entering the patient ID. Under "Hospital Visit Records," there were details of the reason for the visit, the patient's status at that time, and multiple lab results, such as blood, urine, and biochemical examination. Under "Medication Information," users could see the name and in-depth information about the prescribed drug.

In the "Self-report" section, the user could select body regions in an image of the human body to indicate where they feel pain and enter their symptoms with an icon expressing the severity of the pain. Additionally, the user could record their blood pressure. These self-reports could be used in medical visits or for the user's health self-management.

As the main feature of FirstER, the user could send the medical information summary to their carers and medical staff.

Information Management and Setting

The Information Management and Setting page allowed users to manage individual, health, and policy information and access help. Under "Individual Information Management," the patient could manage their information, including patient ID for each hospital and passwords. In the "Health Information Management" section, the patient could choose which data are visible (in case the patient believed that some information was unnecessary). Additionally, a user guide, contact information for asking questions, and policy information was featured. Moreover, the user could decide to withdraw from the service without any constraint in this section.

Study Design

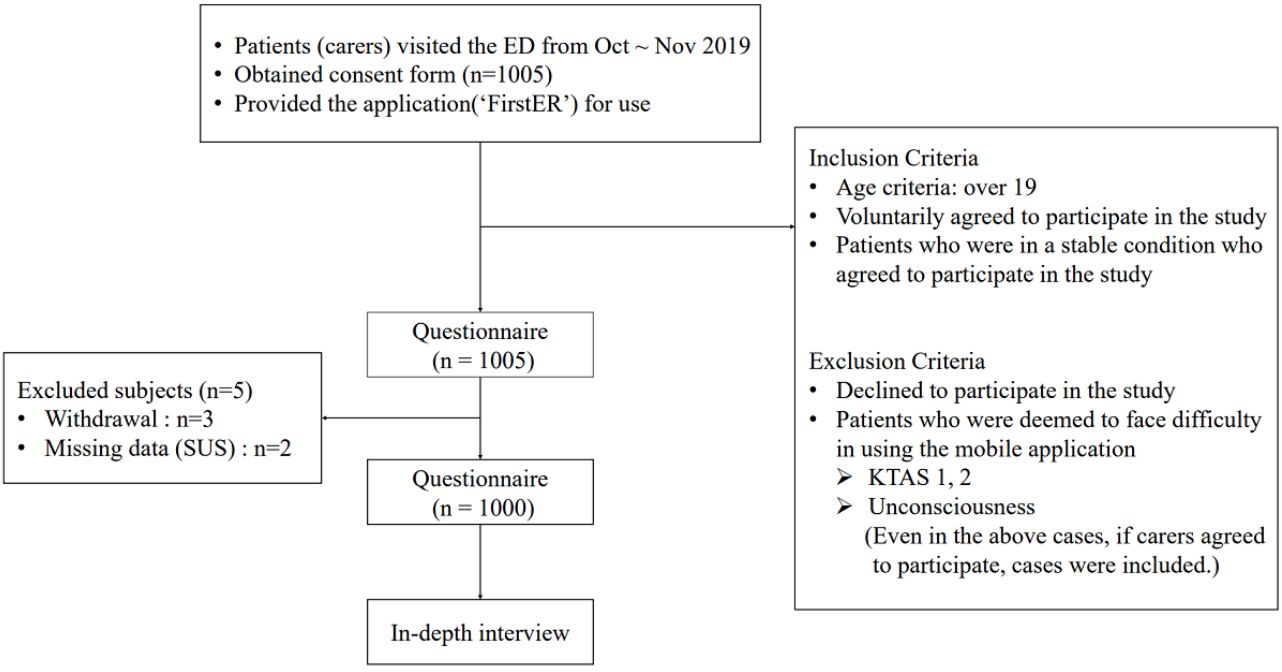
Participants

The participants recruited for this study had visited the Emergency Department of Samsung Medical Center (SMC), Asan Medical Center (AMC), or Dong-A Medical Center (DMC) from October 2019 to November 2019. These 3 medical institutes are large tertiary hospitals in Seoul and Busan in Korea; the number of beds in each hospital are 1989 in SMC, 2705 in AMC, and 1000 in DMC. A System Usability Scale (SUS) score of 68 was considered an average score, and a score of 70 was the basis for acceptability. Therefore, 69.5 points (ie, slightly above 68) was set as the target SUS value. The method for calculating the number of samples used a constant of (68) and a difference of (69.5-68=1.5). To analyze the mean values, a 1-sample *t* test was chosen. The effect size was 0.117, the significance level was 0.05, and the power was 0.95. Effect size was obtained using the following formula:

Under these conditions, a required sample size of 792 was calculated and 1000 people were recruited for multicenter research, taking a dropout rate of approximately 21% into account.

Participants were excluded if they were younger than 19 years, declined to participate, were deemed to face difficultly using mobile applications, or were unconscious while leaving the emergency department (eg, if the patient was disoriented or confused, not conscious, or whether they were in a state of shock or cardiac arrest). However, even in these cases, carers were able to be included as clinical trial subjects. Informed consent was obtained from all the participants. The study protocol (2019-07-066-010) was approved by the institutional review board (IRB) office at Samsung Medical Center in Seoul, Republic of Korea (Figure 3).

Figure 3. Flowchart for the study, featuring inclusion and exclusion criteria. ED: Emergency Department; KTAS: Korean Triage and Acuity Scale (a Korean emergency patient classification tool divided into levels 1-5 depending on severity); SUS: System Usability Scale.



Protocol

We conducted a usability study with a fully functional mobile phone prototype using a mixed methods approach. When the patients and carers visited the hospitals' emergency departments, they were informed of the study's purpose and provided with consent forms. Further, they could download the FirstER application, and they were informed about how to use it. After they installed the application, they were asked to complete a paper questionnaire, including information on demographic characteristics, a smartphone-use score, surveys, and a SUS score. After that, in-depth interviews were conducted with semistructured questions, including details of the survey's content. The interview was conducted via convenience sampling. All the interviews were recorded with the consent of each participant (Figure 3).

Outcome Measurement

The primary outcome was the application's usability as reported by the users (ie, patients and carers) and how they felt while using FirstER. For this, we used the survey method, an in-depth interview, and SUS scores.

Questionnaire Study

The questionnaire consisted of 4 categories: (1) demographic information, (2) smartphone-use status, (3) survey on the experience of obtaining medical information with a legacy system and FirstER, and (4) the SUS for assessing the usability of the application using a 5-point Likert scale (1-strongly disagree, 5-strongly agree) [25].

In the demographics category, participants were asked about their age, gender, marriage status, main carer, education level, and residential area. For smartphone-use status, we obtained 10 questions from a smartphone technology quotient and indices for smartphone usage developed in Korea, which consisted of 4 parts: recognition, access, usage, and capacity [26]. The questions were about the period and duration of smartphone use, how important the smartphone and its applications were in the participants' life, how well the participants could use applications, and so on. A 5-point Likert scale was used to respond to these questions (1-strongly disagree, 5-strongly agree), except for questions 7 and 8, for which a 4-point Likert scale was used. For each question, the responses were converted to a 10-point scale; each selection was multiplied by 2.5 for questions 7 and 8, and by 2 for the other questions, and each score was then summed up based on the highest score of 100.

The survey category consisted of surveys about previous or current medical information experience as well as the experience of using FirstER as a novel way of viewing medical information. First, we asked the reason for requiring medical information from hospitals, which medical information was needed, thoughts on obtaining medical information from the hospital with regard to convenience, and the reasons why the participants wanted to obtain their own medical information. Second, we asked how the participant felt about their medical information being saved on the internet, what the most useful and unnecessary functions of the application were, and for comments on more functions for improving FirstER.

In-depth Interview

To acquire more information about the subjects, interview questions about the survey's themes and the patient's medical information experiences in emergencies were asked. The interview questions were semiconstructed.

Data Analysis

First, we determined the constitution of demographic features and calculated the mean (SD) for the SUS. Further, we conducted linear regression for the SUS according to the smartphone-use score. Data analyses were conducted with R software (version 3.3.1; R Foundation).

For the survey and in-depth interview, we analyzed the frequency, tendency, and keywords of the answers to each question to determine and categorize the participants' thoughts on themes.

Results

Demographics and Smartphone-Use Score

In total, 1000 participants, which included patients and carers, were recruited in this study. The demographic information of these participants is summarized in Table 1. The study participants included people of different age groups, genders, and marital statuses, and differed in the relationship of the main carer, the area of residence, education level, and status as a patient or carer.

There were more female participants (618/1000, 61.8%) than male participants (382/1000, 38.2%). Furthermore, the average age was 41.4 (SD 13.3) years. Of the 1000 participants, there were 414 (41.4%) patients and 586 (58.6%) carers. The number of participants who self-reported as not married was 35.9% (359/1000), and 64.1% (641/1000) self-reported as married. The main carer of the patients was mostly a spouse or family member (962/1000, 96.2%). Of the 1000 participants, 722 (72.2%) had obtained an education higher than college graduation, and 702 (70.2%) lived in Seoul or the capital area. The mean smartphone-use score was 81.4 (SD 11.2; Table 1).

Table 1. Demographics and smartphone-use scores of study participants (n=1000).

Characteristics	Value, n (%)
Gender	
Male	382 (38.2)
Female	618 (61.8)
Age in years	
19-39	481 (48.1)
40-59	419 (41.9)
60+	100 (10.0)
Marital status	
Not married	359 (35.9)
Married	641 (64.1)
Patient/carer	
Patient	414 (41.4)
Carer	586 (58.6)
Respondent's main carer	
Spouse/family	962 (96.2)
Other	38 (3.8)
Education	
College or higher-level education	722 (72.2)
High school or lower-level education	278 (27.8)
Residential district	
Seoul/capital area	702 (70.2)
Other	298 (29.8)
Smartphone-use score	
≤60	45 (4.5)
61-80	356 (35.6)
≥81	599 (59.9)

Survey

Experience of Obtaining Medical Information With a Legacy System

With respect to requesting medical information from the hospitals they had visited, 34.3% (358/1043) of participants reported having no experience requesting their personal medical information, and 56% (584/1043) reported that the reason for requesting their personal medical information was for their employment company or insurance company. The participants needed various kinds of information about their previous conditions while visiting hospitals and wanted to possess their own medical information to use whenever they wanted (eg, to share it with related people; Table 2).

Table 2. Experience of obtaining medical information with a legacy system (n=1000).

Survey question and response options	Value, n (%)
1. Reasons for requiring medical information from hospitals[a]	
No experience	358 (34.3)
For submission to insurance companies	513(49.2)
For submission to company or school	71 (6.8)
For personal storage	17 (1.6)
Deliver to family or carers	50 (4.8)
Other	34 (3.3)
2. How convenient did you feel the process was for requesting medical information before?	
Inconvenient	395 (39.5)
Mediocre	395 (39.5)
Convenient	142 (14.2)
No response	68 (6.8)
3. Have the absence of any of the following kinds of medical information caused you inconvenience when visiting a hospital?[a]	
Diagnosis and test results	526 (33.7)
Medication information in progress (name, capacity, etc)	455 (29.2)
Types of tests carried out	416 (26.7)
The name of the hospital visited, or the department of medicine	71 (4.6)
Other	91 (5.8)
4. What are your reasons for wanting to get personal medical information?[a]	
I want to know all the information related to my health.	685 (28.8)
I want to have it at all times in case of an emergency.	561 (23.6)
I think it's my obvious right.	346 (14.5)
I want to show it to another hospital or another doctor.	295 (12.4)
I want to show it to my family and my carer.	250 (10.5)
I want to know my health condition in detail through an internet search or community, etc.	222 (9.3)
I don't want to collect more detailed medical information than I have now.	16 (0.7)
Other	6 (0.2)

[a]The respondent could select multiple responses.

Experience of Using FirstER

Regarding FirstER, the most helpful information in the application, as per the participants, was ranked in the following order: lab test results (327/1091, 30.0%), personal health self-record (280/1091, 25.7%), function showing data to medical staff (211/1091, 19.3%), and drug information (124/1091, 11.4%). Further, most participants (329/1018, 32.3%) responded that all the information showed on the application was necessary. Personal medical information is sensitive, but many participants (790/1000, 79.0%) agreed to store their medical information on the internet with their consent, while others (207/1000, 20.7%) felt uncomfortable. Table 3 shows participant comments on the application.

Table 3. Participant experience of using FirstER (n=1000).

Question and response options	Value, n (%)
1. The most helpful information in the application[a]	
Lab test results	327 (30.0)
Personal health self-records	280 (25.6)
Function showing data to medical staff	211 (19.3)
Medication information	124 (11.4)
Real-time emergency department records	99 (9.1)
Hospital visits record	29 (2.7)
Other	21(1.9)
2. The most unnecessary information in the application[a]	
None	329 (32.3)
Hospital visits record	111 (10.9)
Personal health self-records	109 (10.7)
Real-time emergency department records	79 (7.8)
Function showing data to medical staff	51 (5.0)
Lab test results	29 (2.8)
Medication information	23 (2.3)
Other	287 (28.2)
3. Opinion about the acceptability/unacceptability of personal medical information being stored on the internet when consent is given	
It does not matter	450 (45.0)
Not good or bad	340 (34.0)
Uncomfortable	207 (20.7)
No response	3 (0.3)
4. Information that would be useful if additionally provided (a subjective answer)	
Expanding nonemergency services	22
Imaging test results	15
Improvement in understanding of medical information	15

[a]The respondent could select multiple responses.

In-depth Interview

Through interviews, we wanted to compare the experience of obtaining medical information with a legacy system with that of using FirstER to see whether FirstER alleviated discomfort.

Experience of Obtaining Medical Information With a Legacy System

Interviews showed that participants were keen to know more about their medical information and to participate in their medical care. However, they were not well informed of their medical information, and they were passive in the use of medical information and their own medical care. The following is a description of the difficulties that patients were experiencing concerning their medical information: (1) difficulties getting accurate medical information in a timely manner; (2) difficulties obtaining medical information; (3) diverse reasons for desiring personal medical information.

Difficulties Getting Accurate Medical Information in a Timely Manner

The interviewees could not provide exact information when they were asked questions by medical staff because they could not remember, or it was difficult to arrange, medical information in an urgent situation.

> *There were many things I couldn't think of, and it was difficult to prepare documents about medical information while rushing to the emergency department.* [study participant (patient)]

Difficulty Obtaining Medical Information

In the present system, patients usually obtained their medical information from the hospital. This was often experienced as stressful, as it was time-consuming and many medical information documents or CDs must be prepared.

It took me a while to register a video, and I always had to go to the hospital 30 minutes and an hour early. [study participant (patient)]

Desire to Obtain Medical Information

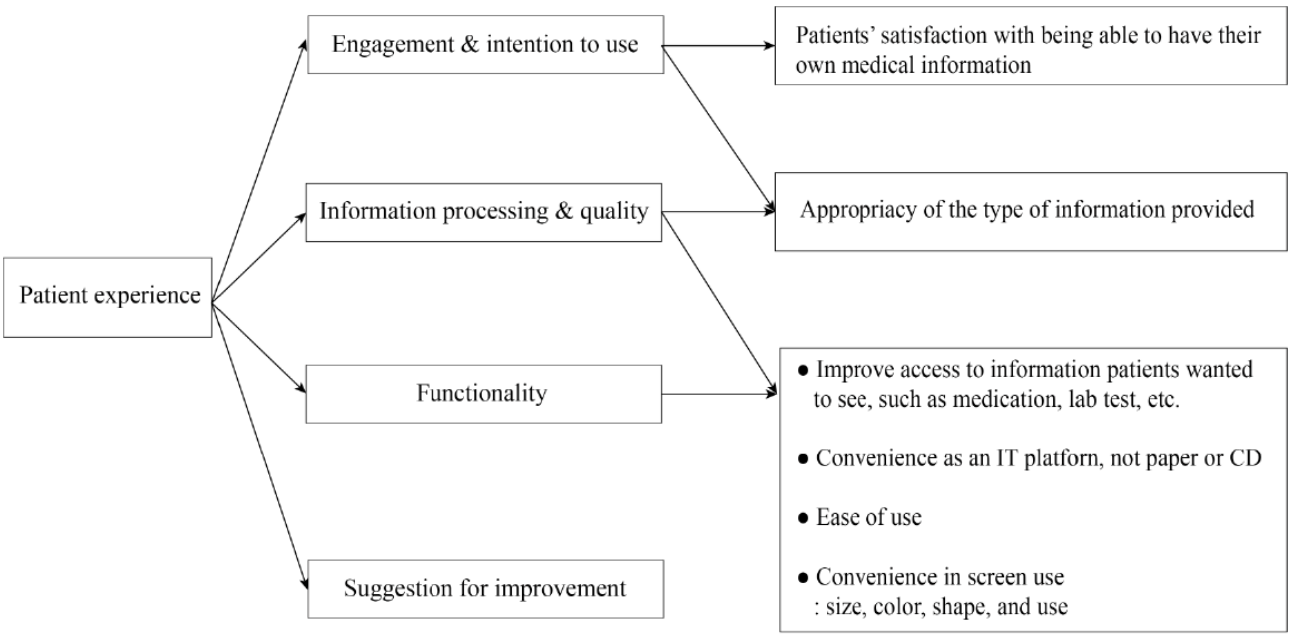
The reasons for wanting to obtain one's own medical information were diverse, and the results obtained were the same as those in the survey: participants wanted to use medical information directly in an emergency, and they needed exact information; also, participants wanted to concurrently look for a second opinion, considering its importance in cases of long-term disease.

I always want to obtain information on my physical condition. [study participant (patient)]

Experience of Using FirstER

We conducted interviews about FirstER to identify whether patients were satisfied with their experiences with it. We constructed interview questions with 4 categories: engagement, information processing and quality, functionality, and suggestions for improvement [19,23,27,28]. Each category related and interacted with others. Most participants were satisfied with the 4 categories of experience (Figure 4).

Figure 4. The construction of questions about the user experience of FirstER.



Engagement and Intention of Use

Questions about engagement examined whether FirstER had satisfied the users' needs for medical information so that they could continue using the application.

It's important that I have my medical information because I've been sick for a long time. [study participant (patient)]

I think I'll use it a lot because I can get a lot of information and it is convenient. [study participant (patient)]

Information Processing and Quality

We asked whether the patients had been provided with the information they wanted and whether the type of information or its contents were sufficient. Most users answered in the affirmative to these questions. Concretely, the participants reported that lab test results and medication information were the most useful. It meant that the participants used FirstER for viewing the medical information they needed. Some participants (7/20, 35.0%) reported that informational variety was lacking in the application, and some (4/20, 20.0%) reported that it was appropriate. This, once again, confirmed that user demand for information is high.

I'm satisfied with the basic medical treatment and prescription medicine. [study participant (carer)]

I don't think we have much information yet. The more, the better. [study participant (patient)]

Functionality

We asked if FirstER was convenient to use; provided increased access to necessary information; could help with communication with medical staff; and had an appropriate layout, screen color scheme, size features, and navigation. It seemed to be helpful in emergencies, and participants answered that it was more convenient to use an IT platform than to pack various documents. In addition, most of the respondents said that the application was easy to use and was properly configured, but there were opinions that it would be nice to have a screen-magnifying function for the elderly and more explanations about medical terms or figures.

I think it's best that I can show my medical information to the emergency department when I'm in a hurry. [study participant (patient)]

Just by several clicks, I can see my blood test result. It was more convenient than using the internet or seeing a document. [study participant (patient)]

> *I think it's good for the elderly too, because the font is big. Also, it's simple to navigate, so I think it'll be easy for the elderly as well as young people.* [study participant (patient)]

Suggestions for Improvement

Through the users' opinions for improvement, we wanted to gain insight into the factors that encourage users to continue using the application and to understand the kind of experience patients currently have. Responses to the service were positive. Various opinions were provided regarding improvements, such as a service extension for cooperative hospitals or outpatients and wards, simplifying the authentication and login process, providing additional information like imaging test results, and providing explanations of terminology and information.

> *I hope many hospitals participate so that I can see lots of other medical information.* [study participant (carer)]

> *General people don't know medical terminologies. It would be better with a brief explanation.* [study participant (carer)]

Multimedia Appendix 1 contains more information about the in-depth interviews.

SUS Analysis

We evaluated participant usability of FirstER with the System Usability Scale (SUS). The mean SUS score value was 67.1 (SD 13.8), which means participants evaluated FirstER as very near to acceptable [25]. When determining the relationship between smartphone-use scores and SUS values, the correlation coefficient was 0.8887, and it had a significant linear regression (P=.0015).

Summary of User Experience of the Legacy System and FirstER

A summary of the user experience with the legacy system and with FirtER is presented in Table 4.

Table 4. Summary of the user experience with the legacy system and with FirstER.

Key theme	Legacy system	FirstER
Engagement and intention to use	• Patients were not well aware of their health conditions and medical information, and they were passive in the use of their medical information and treatment. • Patients wanted their medical information for proof and for several other reasons, such as sharing with others, preparing for emergencies, etc.	• Patients were satisfied with owning their medical information. • FirstER improved patient communication with medical staff about their condition.
Information processing and quality	• Patients wanted their previous medical information, such as information on lab tests, medication, hospital visit records (hospital and department names), etc.	• FirstER provided the information patients had wanted, and the patients found the content to be overall sufficient. • Some participants reported that it would have been better if there had been explanations for better understanding of the medical terms.
Functionality	• Difficulty obtaining accurate medical information promptly • The hassle of getting medical information; time-consuming, many documents, etc.	• FirstER enabled better and easier access to medical information using patients 'mobile phones.

Discussion

Principal Results and Strengths of the Study

To the best of our knowledge, this was the first study to introduce mPHR to end-users such as patients and carers in real emergencies that were not covered by conventional PHR studies. It aimed to ensure that mPHRs improve patient access to information and improve patient experience in the emergency department. It is also meaningful to implement the essence of PHR in that it allows patients to collect their own medical records and seek continuous medical care by solving issues of high security, interoperability, and accuracy in patient record delivery with 3 large tertiary hospitals in Korea.

Studies have shown that patients were not well aware of their medical information and were passive in utilizing medical information and medical care. However, patients and carers wanted to own and better utilize medical information and were willing to be active in their medical activities. This study was the first PHR demonstration attempted in emergencies. The purpose was to observe the user experience of FirstER in emergencies; therefore, it is difficult to conclude whether FirstER outperformed the legacy system. However, FirstER fulfilled these patient needs by increasing access to information they previously needed and made it possible to sufficiently use this information, which showed that the introduction of mPHR is reasonable in real emergencies.

Suitability of Information in FirstER: Improving Patient Experience by Providing Information That Patients Need

Emergency departments function in almost all hospitals in one space, so they require faster and more accurate information and access to as much information as possible [3]. By providing patients with their necessary information, patient experience can be enhanced [5,23]. However, there is no standard for what

should be put into the emergency PHR. FirstER contained name, address, birth date, contact information, communication, allergy, and medical history related details [29]. Additionally, it also contained lab test results, medication-related information, etc, that cause information gaps in communication in emergency situations [4].

Medical staff selected the type of information reflected in FirstER, based on prior research on which information points are necessary for patients and medical staff in emergencies [24]. Imaging information was not included in FirstER because of system problems and general patient inability to read them independently. The survey results showed that 32.3% (329/1018) of participants reported no unnecessary information in FirstER, indicating that FirstER contained enough core data needed in emergency situations.

Reflections on the Low SUS Outcome

After receiving positive responses to FirstER in the survey and interview, the SUS was also expected to be above-average; however, it was not. The reason may be that SUS, which only deals with the ease of use of the application itself, does not contain the experience of medical information retrieval covered in surveys and interviews. Moreover, an unstable environment such as the emergency department may have made it more difficult to use FirstER rather than a typical application.

This demonstration was a pilot study to see the usability of mPHR in emergencies rather than to perfect the application. Usually, to make a better application, it takes a longer time period, and the application must go through several iterative processes. If we can further develop the application based on the results from this study, we can expect a better SUS.

Challenges in Application

First, interoperability is critical. In this study, we organized hospitals to participate and share data. However, for extension and everyday use of the service, detailed and extensive discussion is needed, as complexity increases exponentially with the increase in the number of participants. Even within an institution, multiple departments such as outpatient and operating rooms often lack sufficient standardization on patient data, which would result in challenges in a real-world application.

Second, data disclosure and privacy issues are critical. We designed this study using tight security protocols. Based on privacy legislation, the medical cloud zone, an service infrastructure (IaaS) provided by the Samsung Data System (Seoul, Korea), met the facilities and equipment standards necessary for the management and preservation of electronic medical records (EMRs), which is a requirement for the remote storage of EMRs. As for data transfer to each patient, we obtained strict written consent approved by the IRB and other consent forms in the application informing patients of which information we would collect from them. For registration, verification through the patient's email address was required, and a unique patient's hospital ID was used as a key to access the patient's information from the system [30]. In addition, we provided insurance in case of accidents like patient information leakage.

Regarding data disclosure and privacy for real implementation, we have to maintain confidentiality based on regulations. Intensive user verification such as biometric technologies can be used to this end. In addition, there should be a fundamental consensus regarding the range of information sharable with these systems. Providers and users need to actively discuss the types, amounts, and methods of data sharing. In addition, we are required to consider the range of subjects that patients may want to share their data with.

Limitation

First, this study was conducted in only 3 hospital emergency departments; therefore, there could be a bias in the selection of subjects, and the results may not reflect patients and carers in all emergencies. In addition, because we had interviewed some participants via a convenience sampling method, it was difficult to reflect the opinions of all patients in the emergency department.

Second, this study was conducted within a short period of time, and within that period, visits to the emergency departments were often one-time; therefore, we could not check the participants using FirstER for a longer period of time.

Third, in relation to the above, we have not seen a better medical outcome among the patients, such as medication compliance, self-management, etc. It would have been better to show the practical utility of FirstER if we had identified better medical outcomes than those who did not use it.

Conclusion

FirstER showed that mPHRs can potentially contribute to enhancing patient experience by providing patients and carers with conveniently accessible medical information in real emergencies. To the best of our knowledge, this is the first study to introduce mPHR to end-users such as patients and carers in real emergencies that were not covered by conventional PHR studies. It is also meaningful to implement the essence of PHR in that it allows patients to manage their own medical records and seek continuous medical care while solving the issues of high security, interoperability, and accuracy.

Acknowledgments

This work was supported by the Technology Innovation Program (20004503, The Development of a Personal-Health-Record Management Platform for Chronic Patients in Emergency Rooms based on Actual Demand) funded by the Ministry of Trade, Industry, and Energy (MOTIE, Korea).

Conflicts of Interest

None declared.

Multimedia Appendix 1
In-depth interview.
[DOCX File , 22 KB - mhealth_v8i12e24326_app1.docx]

References

1. Everson J, Kocher KE, Adler-Milstein J. Health information exchange associated with improved emergency department care through faster accessing of patient information from outside organizations. J Am Med Inform Assoc 2017 Apr 01:e103-e110. [doi: 10.1093/jamia/ocw116]
2. Finnell JT, Overhage JM, McDonald CJ. In support of emergency department health information technology. 2005 Presented at: AMIA Annu Symp Proc Medline; 2005; Washington, DC p. 246-250.
3. Kaelber DC, Bates DW. Health information exchange and patient safety. J Biomed Inform 2007 Dec;40(6 Suppl):S40-S45 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.011] [Medline: 17950041]
4. Stiell A, Forster AJ, Stiell IG, van Walraven C. Prevalence of information gaps in the emergency department and the effect on patient outcomes. CMAJ 2003 Nov 11;169(10):1023-1028 [FREE Full text] [Medline: 14609971]
5. Blackburn J, Ousey K, Goodwin E. Information and communication in the emergency department. International Emergency Nursing 2019 Jan;42:30-35. [doi: 10.1016/j.ienj.2018.07.002]
6. Frojd C, Swenne C, Rubertsson C, Gunningberg L, Wadensten B. Patient information and participation still in need of improvement: evaluation of patients' perceptions of quality of care. J Nurs Manag 2011 Mar; Medline 2011 Mar;19(2):e. [doi: 10.1111/j.1365-2834.2010.01197.x]
7. Agency FHR. Improving the Emergency Department Discharge Process; Environmental Scal Report. Rockville, MD: Agency for Healthcare Research and Quality; 2014:A.
8. Uscher-Pines L, Pines J, Kellermann A, Gillen E, Mehrotra A. Emergency department visits for nonurgent conditions: systematic literature review. Am J Manag Care 2013 Jan;19(1):47-59 [FREE Full text] [Medline: 23379744]
9. Adams R. Improving health outcomes with better patient understanding and education. RMHP 2010 Oct:61. [doi: 10.2147/rmhp.s7500]
10. Greene JC, Haun JN, French DD, Chambers SL, Roswell RH. Reduced Hospitalizations, Emergency Room Visits, and Costs Associated with a Web-Based Health Literacy, Aligned-Incentive Intervention: Mixed Methods Study. J Med Internet Res 2019 Oct 17;21(10):e14772. [doi: 10.2196/14772]
11. HealthIT. gov. What are the differences between electronic medical records, electronic health records, and personal health records?. URL: http://www.healthit.gov/providers-professionals/faqs/what-are-differences-between-electronic-medical-records-electronic [accessed 2019-12-27]
12. Ball M, Smith C, Bakalar R. Personal health recordsmpowering consumers. J Healthc Inf Manag 2007 Winter; Medline 2007;21(1):e.
13. Bouri N, Ravi S. Going mobile: how mobile personal health records can improve health care during emergencies. JMIR Mhealth Uhealth 2014 Mar;2(1):e8 [FREE Full text] [doi: 10.2196/mhealth.3017] [Medline: 25098942]
14. Endsley S, Kibbe DC, Linares A, Colorafi K. An introduction to personal health records. Fam Pract Manag 2006 May;13(5):57-62 [FREE Full text] [Medline: 16736906]
15. Kharrazi H, Chisholm R, VanNasdale D, Thompson B. Mobile personal health records: an evaluation of features and functionality. Int J Med Inform 2012 Sep;81(9):579-593. [doi: 10.1016/j.ijmedinf.2012.04.007] [Medline: 22809779]
16. 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;13(2):121-126 [FREE Full text] [doi: 10.1197/jamia.M2025] [Medline: 16357345]
17. Menon A, Greenwald S, Ma T, Kooshesh S, Duriseti R. Patient and Physician Willingness to Use Personal Health Records in the Emergency Department. WestJEM 2012 May 1;13(2):172-175. [doi: 10.5811/westjem.2011.11.6844]
18. Thummavet P, Vasupongayya S. A novel personal health record system for handling emergency situations. 2013 Presented at: International Computer Science and Engineering Conference (ICSEC) Sep; 2013; Bangkok p. 266-271. [doi: 10.1109/icsec.2013.6694791]
19. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016 May;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
20. Patient Engagement. World Health Organization. 2016. URL: https://apps.who.int/iris/bitstream/handle/10665/252269/9789241511629-eng.pdf [accessed 2020-02-12]
21. Lee G, Park JY, Shin S, Hwang JS, Ryu HJ, Lee JH, et al. Which Users Should Be the Focus of Mobile Personal Health Records? Analysis of User Characteristics Influencing Usage of a Tethered Mobile Personal Health Record. Telemed J E Health 2015 Oct 8. [doi: 10.1089/tmj.2015.0137] [Medline: 26447775]

22. Davis GT, Menon S, Parrish DE, Sittig DF, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. J Am Med Inform Assoc 2014 Jul;21(4):737-741. [doi: 10.1136/amiajnl-2013-002239] [Medline: 24154835]
23. Sonis JD, Aaronson EL, Castagna A, White B. A Conceptual Model for Emergency Department Patient Experience. Journal of Patient Experience 2018 Aug 21;6(3):173-178. [doi: 10.1177/2374373518795415]
24. Soh J. A Study on the Needs for Personal Health Record for Emergency Medicine Doctoral. Seoul: Sungkyunkwan University; 2019.
25. Bangor A, Kortum P, Miller J. Determining What Individual SUS Scores Mean; Adding an Adjective Rating Scale. J Usability Studies 2009 May 2009 May;4(3):114-123.
26. Choi Y, Yun H, Lee CC. Developing Indices for Smartphone Usage: Smartphone Technology Quotient. Entrue J Inf Technol 2013;12:87-104.
27. Walker D, Menser T, Yen P, McAlearney A. Optimizing the User Experience: Identifying Opportunities to Improve Use of an Inpatient Portal. Appl Clin Inform 2018 Feb 14;09(01):105-113. [doi: 10.1055/s-0037-1621732]
28. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
29. DeVore A. The Electronic Health Record for the Physician's Office with SimChart for the Medical Office. St. Louis, Missourilsevier Health Sciences 2015:E.
30. Choi Y, Kim J, Kwon IH, Kim T, Kim SM, Cha W, et al. Development of a Mobile Personal Health Record Application Designed for Emergency Care in Korea; Integrated Information from Multicenter Electronic Medical Records. Applied Sciences 2020 Sep 25;10(19):6711. [doi: 10.3390/app10196711]

Abbreviations

AMC: Asan Medical Center
DMC: Dong-A Medical Center
EMR: electronic medical record
IRB: institutional review board
mPHR: mobile personal health record
PHR: personal health record
SMC: Samsung Medical Center

Edited by L Buis; submitted 16.09.20; peer-reviewed by HS Kim, YC Chen; comments to author 08.10.20; revised version received 22.10.20; accepted 03.11.20; published 15.12.20.

Please cite as:
Kim SM, Kim T, Cha WC, Lee JH, Kwon IH, Choi Y, Kim JS
User Experience of Mobile Personal Health Records for the Emergency Department: Mixed Methods Study
JMIR Mhealth Uhealth 2020;8(12):e24326
URL: http://mhealth.jmir.org/2020/12/e24326/
doi:10.2196/24326
PMID:33320102

Original Paper

ColistinDose, a Mobile App for Determining Intravenous Dosage Regimens of Colistimethate in Critically Ill Adult Patients: Clinician-Centered Design and Development Study

Xueliang Hua[1*], BSc; Chen Li[2,3,4*], PhD; Jason M Pogue[5], PharmD; Varun S Sharma[4], MSc; Ilias Karaiskos[6], MD, PhD; Keith S Kaye[7], MD, MPH; Brian T Tsuji[8,9], PharmD; Phillip J Bergen[3,10], PhD; Yan Zhu[3,10], PhD; Jiangning Song[2,3,11], PhD; Jian Li[3,10], PhD

[1]Independent Researcher, Santa Clara, CA, United States
[2]Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia
[3]Infection and Immunity Program, Biomedicine Discovery Institute, Monash University, Melbourne, Australia
[4]Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, Switzerland
[5]Department of Clinical Pharmacy, University of Michigan College of Pharmacy, Ann Arbor, MI, United States
[6]1st Internal Medicine and Infectious Diseases Department, Hygeia Hospital, Marousi, Greece
[7]Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, United States
[8]Laboratory for Antimicrobial Dynamics, NYS Center of Excellence in Bioinformatics & Life Sciences, Buffalo, NY, United States
[9]School of Pharmacy and Pharmaceutical Sciences, University at Buffalo, Buffalo, NY, United States
[10]Department of Microbiology, Monash University, Melbourne, Australia
[11]Monash Centre for Data Science, Faculty of Information Technology, Monash University, Melbourne, Australia
[*]these authors contributed equally

Corresponding Author:
Jian Li, PhD
Department of Microbiology
Monash University
19 Innovation Walk, Monash University
Melbourne, VIC 3800
Australia
Phone: 61 3 9903 9702
Email: jian.li@monash.edu

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2020/12/e26593/

Abstract

Background: Determining a suitable dose of intravenous colistimethate is challenging because of complicated pharmacokinetics, confusing terminology, and the potential for renal toxicity. Only recently have reliable pharmacokinetic/pharmacodynamic data and dosing recommendations for intravenous colistimethate become available.

Objective: The aim of this work was to develop a clinician-friendly, easy-to-use mobile app incorporating up-to-date dosing recommendations for intravenous colistimethate in critically ill adult patients.

Methods: Swift programming language and common libraries were used for the development of an app, ColistinDose, on the iPhone operating system (iOS; Apple Inc). The compatibility among different iOS versions and mobile devices was validated. Dosing calculations were based on equations developed in our recent population pharmacokinetic study. Recommended doses generated by the app were validated by comparison against doses calculated manually using the appropriate equations.

Results: ColistinDose provides 3 major functionalities, namely (1) calculation of a loading dose, (2) calculation of a daily dose based on the renal function of the patient (including differing types of renal replacement therapies), and (3) retrieval of historical calculation results. It is freely available at the Apple App Store for iOS (version 9 and above). Calculated doses accurately reflected

doses recommended in patients with varying degrees of renal function based on the published equations. ColistinDose performs calculations on a local mobile device (iPhone or iPad) without the need for an internet connection.

Conclusions: With its user-friendly interface, ColistinDose provides an accurate and easy-to-use tool for clinicians to calculate dosage regimens of intravenous colistimethate in critically ill patients with varying degrees of renal function. It has significant potential to avoid the prescribing errors and patient safety issues that currently confound the clinical use of colistimethate, thereby optimizing patient treatment.

(JMIR Mhealth Uhealth 2020;8(12):e20525) doi:10.2196/20525

KEYWORDS

ColistinDose; colistimethate; colistin methanesulfonate; colistin; polymyxins; mobile app; renal function; renal replacement therapy; intermittent hemodialysis; sustained low-efficiency dialysis; continuous renal replacement therapy

Introduction

As the drug discovery pipeline for new antibiotics has dwindled [1,2], multidrug-resistant gram-negative pathogens have become a serious global health threat [3]. Colistin (ie, polymyxin E) has increasingly been used as a "last-line" therapy to treat infections caused by gram-negative "superbugs" unresponsive to other agents [4,5]. In clinical settings, colistin is most commonly available as an inactive prodrug, colistimethate (also known as colistin methanesulfonate [CMS]), for intravenous and inhalational administration [6]. Unfortunately, nephrotoxicity following intravenous administration of polymyxins (colistin or polymyxin B) can occur in up to 60% of patients and is the major dose-limiting factor [7-9]. Owing to earlier difficulties in determining concentrations of colistimethate and formed colistin, only relatively recently have studies reliably investigated the pharmacokinetics (PK) of colistimethate and formed colistin in patients [10-18]. These studies have revealed the extremely complicated PK of colistimethate and formed colistin, which in turn has made determining appropriate dosing regimens for intravenous colistimethate very challenging. The complex PK of both the prodrug and formed colistin in patients extend to those with different renal functions and those on different renal replacement therapies (RRT) given that the apparent clearance of colistin is dependent on renal function [10]. Adding to this complexity is that the dose units of colistimethate are expressed differently in different parts of the world, namely as either colistin base activity (CBA) or number of international units (IU) (approximately 33.3 mg of CBA=1 million IU=approximately 80 mg of CMS); these different expressions are known to have caused prescribing errors and patient safety issues, and substantially confound the clinical use of colistimethate [19]. Selecting an optimal dose of colistimethate in critically ill patients is thus a difficult process with serious consequences for both underdosing (treatment failure and development of resistance) and overdosing (toxicity).

Mobile devices have become commonplace in health care settings. Professional mobile apps have created paradigm shifts in modern medicine in a number of areas including information storage and access, patient management and monitoring, clinical decision making, and clinical practice transformation [20]. Given the role of colistimethate as one of the most important last-line therapies for multidrug-resistant gram-negative bacteria and the difficulties associated with accurately determining an appropriate dose across many patient groups (including those on RRT), a user-friendly smartphone app utilizing the latest clinical pharmacological findings to determine optimal, personalized dosing regimens of intravenous colistimethate in patients would provide valuable assistance to clinicians.

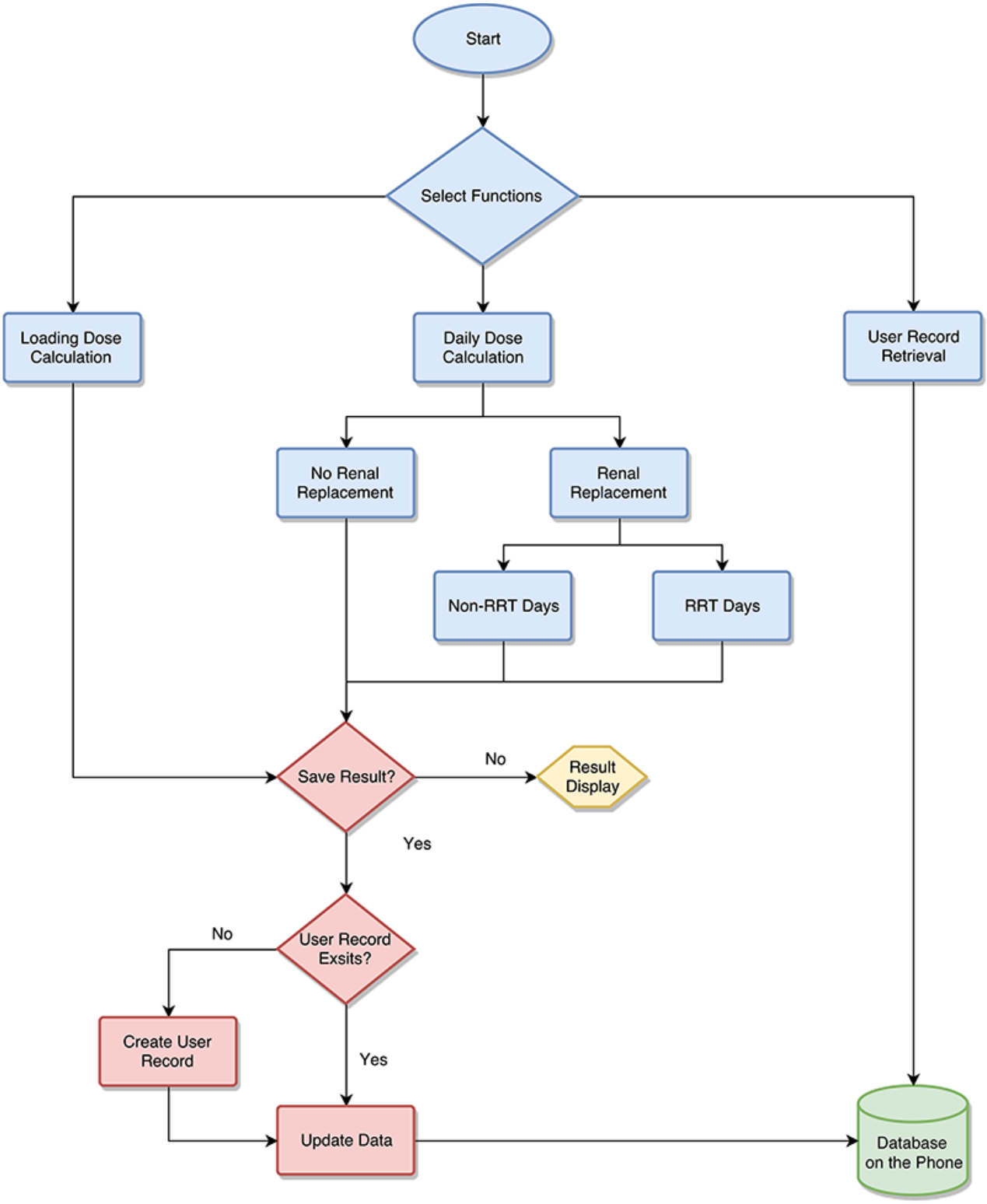
We recently published the most comprehensive population PK model to date with dosing recommendations based on data from a total of 214 critically ill patients that included 29 patients on different forms of RRT [11]. To the best of our knowledge, there are no apps and only two online tools available for the calculation of colistimethate dosage regimens: a colistin dosing calculator [21] and a colistin calculator [22]. Neither tool utilizes the latest population PK model [11] for calculating the dosage regimens of intravenous colistimethate. This study aimed to develop an app, ColistinDose, for the iPhone operating system (iOS; Apple Inc) to allow clinicians to accurately and conveniently determine the dosing regimens of intravenous colistimethate on a local iPhone or iPad. Our app provides clinicians with an easy-to-use tool for determining optimal dosing regimens of intravenous colistimethate in adult patients with varying degrees of renal function directly at the bedside.

Methods

Development of the ColistinDose Smartphone App

Xcode and Interface Builder (Apple Inc) were employed to build the app's human interaction interfaces. Several open-source libraries (including Charts, Former, Persei, and Realm; please refer to the Multimedia Appendix 1 and the Acknowledgments section in the app for detailed information) were also utilized to implement relevant user interface elements, animation, and, in particular, local data storage. The overall design of the workflow of ColistinDose is shown in Figure 1. We employed the Swift programming language and common libraries for iOS to ensure that ColistinDose was functional on different screen sizes and versions of the operating system on iPhones and iPads.

Figure 1. The workflow of ColistinDose. RRT: renal replacement therapy.



PK/Pharmacodynamics (PD) Models of Intravenous Colistimethate in Patients for ColistinDose

In 2017, we published dosing suggestions of intravenous colistimethate based on the largest population PK study conducted to date [11], including 214 critically ill patients (composed of 105 patients from an interim analysis [10] and 109 patients from an additional study site in Greece). In particular, the 2017 study included a larger number of patients with low creatinine clearances (CrCls), including 16 patients on hemodialysis, 4 on sustained low-efficiency dialysis (SLED), 7 on continuous veno-venous hemodialysis (CVVHD), and 2 on continuous veno-venous hemofiltration (CVVH); these additional patients receiving RRT improved the ability to predict doses for these patients. The structural model from the interim analysis provided the starting point for further analysis, with all relevant covariates retested. After the new population PK analysis, the final model remained similar to the earlier model.

For the calculation of the recommended loading dose (all patients), our final algorithm used ideal body weight (IBW) in all cases rather than the lower of actual body weight or IBW as per the interim analysis. The algorithm-derived maintenance (daily) dosing recommendations for patients with a CrCl >30 mL/min remained very similar to those proposed in the interim analysis. However, for patients with a CrCl <30 mL/min (including patients prescribed intermittent hemodialysis [IHD] who were on a nondialysis day), the new recommendations were approximately 100% higher than those suggested in the interim analysis and also substantially higher than the doses currently approved by the Food and Drug Administration [23]. By utilizing the most recent equations, ColistinDose provides the most accurate dosing recommendations, especially for patients with low CrCls.

Calculations of Dosage Regimens Using ColistinDose

Calculations were based on the equations for the loading dose (all patients) and daily dose (for patients not receiving RRT and for patients on various types of RRT) developed in our recent population PK study [11]. Patients' IBW (kg) was calculated using the following equation: IBW = 50 kg (45.5 kg for females) + 2.3 × number of inches over 5 feet (ie, 60 inches); CrCl (mL/min) was calculated using the Cockcroft-Gault equation [24]. Three major functionalities were implemented in ColistinDose, namely calculations of the loading dose and maintenance daily dose, plus retrieval of historical records saved on the device. Recommended doses generated by ColistinDose were validated by comparison against doses calculated manually using the appropriate equations from the population PK study [11].

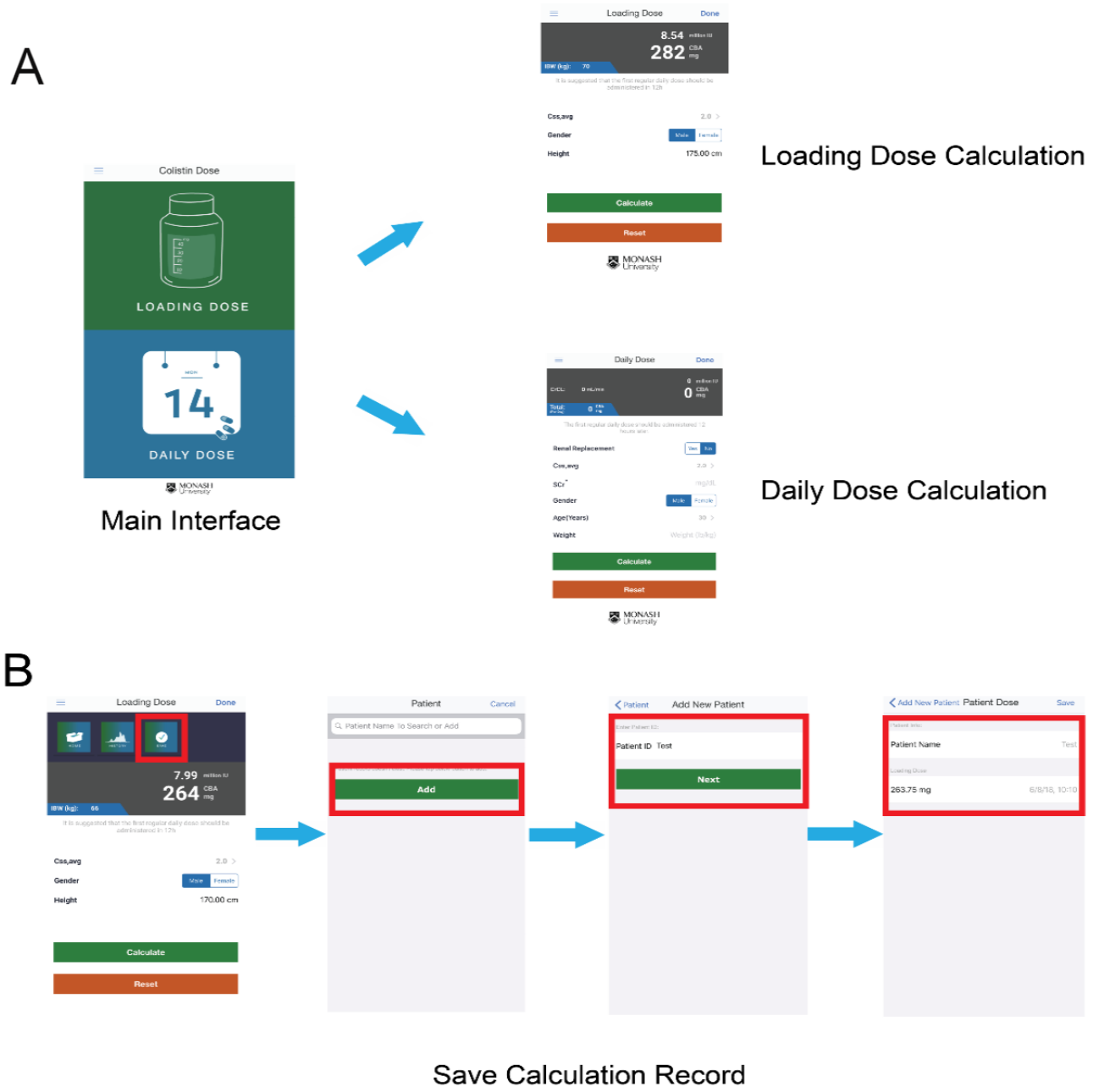
Results

The ColistinDose App

ColistinDose is simple to use and freely available for iOS (version 9 and above) at the Apple App Store. Upon opening the app to the main interface, the clinician is prompted to choose calculation of either a loading dose or a daily dose (Figure 2A). Prior to the publication of the final results of our population PK study [11], the importance of expediently achieving relatively high levels of formed colistin concentrations via a loading dose was already well known [10,11,17,18]. Following initiation of colistimethate treatment, the concentration of formed colistin increases slowly and may take up to 48 h before an acceptable average steady-state plasma concentration ($C_{ss,avg}$) of formed colistin is achieved. Achieving therapeutic concentrations rapidly in critically ill patients is important, as suboptimal PK/PD and delayed effective therapy are associated with increased mortality rates [25]. Low concentrations may also lead to development of resistance [26]. In ColistinDose, calculation of the loading dose is based on three key parameters: (1) the targeted $C_{ss,avg}$ of formed colistin, (2) gender (male/female), and (3) body height (inches or centimeters) (Figure 2A). Based on the recent population PK model, which found a colistin $C_{ss,avg}$ of 2 mg/L to be a suitable initial target concentration for intravenous treatment of bloodstream infections and some minor infections when the colistin minimal inhibitory concentration (MIC) was ≤2 mg/L [11], the selectable $C_{ss,avg}$ in ColistinDose has a default value of 2 mg/L and ranges from 0.5 mg/L to 4 mg/L, with a step of 0.5 mg/L. After pressing the "Calculate" button, the suggested loading dose is calculated using the following equation: CBA (mg) = $C_{ss,avg}$ (mg/L) × 2.0 × IBW (kg), where the IBW is calculated using the Devine formula described in "Methods" [27]. It is important to note that should the recommended loading dose exceed the maximum recommended daily dose of 300 mg CBA (ie, 9 million IU) [23], a warning message (headed "Immediate Attention") will pop up to alert the user that the loading dose has been capped at 300 mg CBA due to safety considerations. Importantly, to help clinicians from different parts of the world, the calculated dose is expressed in both CBA and million IU to avoid prescribing errors that have been known to occur due to confusion over the conversion [19].

Because colistimethate (the administered prodrug) is predominantly renally cleared [28], whereas the formed colistin is largely cleared by nonrenal pathways [29], the equations for the calculation of the daily (maintenance) dose are complicated and require information on the renal function of the patient and whether they are on RRT [11]. Once calculation of the daily dose is selected from the main interface of ColistinDose, the CrCl (mL/min) is determined via input of the patient's serum creatinine (mg/dL), gender (male/female), age (years), and weight (lbs or kg). A screenshot of the daily dose calculation is shown in Figure 2A. The targeted $C_{ss,avg}$ (mg/L) of formed colistin is also required [11]. Whether the patient is on RRT must also be selected. For patients not on RRT, the recommended daily dose is given in both CBA and million IU for 12-hourly administration (eg, 75 mg CBA per 12 h), with a suggestion that the first regular daily dose be administered 12 hours after the loading dose. This latter recommendation was taken from our population PK study [11]. For patients on RRT, three common dialysis options are provided: (1) IHD, (2) SLED, and (3) continuous RRT (CRRT). Based on our population PK analysis in patients receiving RRT, on a dialysis day of IHD, 20% or 50% of the baseline daily dose should be added after a 2-hour or 5-hour dialysis session, respectively [11]. For patients undergoing SLED or CRRT, 10% of the baseline daily dose should be supplemented for every hour of dialysis [11]. When the particular type of dialysis is chosen and a daily dose calculated, the user is reminded of this information with a screen prompt. For example, if CRRT is selected, the following message will appear below the dosing recommendations: "During CRRT, add 10% per 1 hour of CRRT to the baseline daily dose," as suggested by our population PK study [11]. As per the loading dose, a warning message (headed "Immediate Attention") pops up to alert the user whenever the calculated daily dose exceeds the maximum recommended 300 mg CBA (ie, 9 million IU) [11]. It should be noted that the actual calculated dose of colistimethate by ColistinDose can be higher, albeit with a cap of 300 mg.

Figure 2. Screenshots illustrating the use of ColistinDose. (A) The main interface and calculations of the loading dose and daily dose. (B) Steps for saving the calculation results.



Clinicians can choose to store the historical calculation records in the mobile device locally for each patient by using the "Save" option. Figure 2B demonstrates the saving of the dose calculations for a patient. With this function, clinicians can easily revisit previously calculated colistimethate doses as well as the types of renal dialysis, making dose adjustments where appropriate.

Comparisons of Loading and Maintenance Daily Doses With Published Equations

To evaluate the accuracy of ColistinDose against our published dosing algorithms, we performed mock calculations of loading and maintenance daily doses using the height, age, and body weight of 8 "patients," including 2 patients on each type of RRT (Multimedia Appendix 2). The calculated loading and daily doses for the 8 patients were almost identical to those calculated manually using the equations from the latest population PK model [11], the minor differences being due to number rounding. The rounded results are then converted to million IU. Overall, the results in Multimedia Appendix 2 demonstrate that the doses generated by ColistinDose accurately reflect doses recommended in patients with varying degrees of renal function and on different forms of RRT based on the most up-to-date published algorithms. In addition, coauthors JMP, IK, and KSK have evaluated and used ColistinDose in their clinical practices. The conversion function between two different dose definitions (ie, CBA versus IU) in ColistinDose provides a useful tool to prevent potential prescription errors.

Discussion

Comparison of ColistinDose With Existing Colistin Dosing Aids

Compared with the two existing dosing aids (ie, colistin dosing calculator [21] and colistin calculator [22]), the advantages of ColistinDose are twofold. First, ColistinDose utilizes equations derived from the most recent and comprehensive population PK model published in 2017 [11], whereas the colistin dosing calculator and colistin calculator utilize dosing algorithms based on an interim population PK analysis published in 2011 [10]. Prior to the publication of the latest population PK study [11], we had published an interim population PK analysis of intravenous colistimethate that examined 105 critically ill patients across 3 study sites (2 in the United States and 1 in Thailand), including 12 patients on hemodialysis and 4 patients on CRRT (3 on CVVHD and 1 on CVVH) [10]. A nonlinear mixed-effects modeling tool, S-ADAPT, was used to analyze plasma concentration-versus-time data for both colistimethate and colistin and found that their dispositions were best described by linear 2- and 1-compartment models, respectively. These models then formed the basis for the development of colistimethate dosing suggestions for various categories of patients based on the PK data and its integration with PD data for *Acinetobacter baumannii* and *Pseudomonas aeruginosa* in murine thigh and lung infection models [12]. Prior to the interim population PK analysis [10], information on the PK of colistimethate and formed colistin was scarce and had been reported in only 32 patients, all with CrCls >40 mL/min [18,30]. As previously discussed, while dosing recommendations for patients with a CrCl >30 mL/min were very similar between the interim and final population PK analyses, there were substantial differences for patients with lower CrCls, including patients on dialysis; in the final analysis, recommendations for these latter patients were approximately 100% higher than those suggested in the interim analysis. Use of the interim dosing equations thus risks exposing patients to subtherapeutic concentrations of the active component, colistin, decreasing the likelihood of efficacy and increasing the possibility of the emergence of colistin resistance. As colistimethate is a last-line therapy often used in critically ill patients when no other therapeutic options are available, attaining sufficient plasma concentrations is imperative given the potential consequences of treatment failure.

Second, a major advantage of ColistinDose is that once the app is downloaded and installed, all calculations are conducted on the local mobile device (iPhone or iPad) without the need for an internet connection. In contrast, the colistin dosing calculator [21] and colistin calculator [22] are only available online.

Patients' Privacy Protection

To protect the patient's privacy and comply with ethics and data safety requirements, all data are only stored within and accessible from the app on the clinician's mobile device; deleting or uninstalling the app will simultaneously delete all associated information permanently, as demonstrated in the disclaimer and terms of use issued by Monash University (refer to Multimedia Appendix 3). Clinicians are only required to input the patient's ID—which could be either their patient ID number, a nickname, or anything easy to remember—to save the calculation results. Were the app to crash, only the crash data, not the patient's information or previously saved calculation results, would be sent to the app developer via the iOS under the users' agreement.

Limitations and Outlook

The primary limitation of ColistinDose is that the user interface is in English. This may cause inconvenience for non–English-native clinicians to use the app. In light of this, we are considering adding multiple languages in the next version of ColistinDose. Because colistin has a narrow therapeutic window and different renal functions can cause substantial interpatient variability of the PK [31], therapeutic drug monitoring is highly recommended to ensure favorable clinical outcomes in patients. Incorporation of new therapeutic drug monitoring results will improve the precision of our population PK model, as well as the future version of ColistinDose.

Conclusions

Colistimethate is increasingly used as a treatment of last resort in medically complex patients for infections caused by gram-negative "superbugs". However, its complicated pharmacology and confusing dose definition can cause patient safety issues. Here, we have developed an easy-to-use mobile app, ColistinDose, that can be used at the bedside to facilitate the calculation of intravenous colistimethate dosage regimens for adult patients with varying degrees of renal function. To date, ColistinDose has been downloaded from 53 countries and regions around the globe. The potential for ColistinDose to improve patient care is significant, and its easy-to-use interface and functionalities can significantly assist clinicians worldwide to reduce prescribing errors, maximize efficacy, and minimize emergence of resistance and the likelihood of acute kidney injury in patients.

Acknowledgments

We are grateful to Gabrielle Hirsch from The Office of the General Counsel, Monash University, for the legal assistance, disclaimer, and Terms of Use for ColistinDose. We also deeply appreciate the efforts of a number of staff from Monash University, including Moh-Lee Ng and Kevin Sweeney for risk assessment; Sharnie Hewlett and Vicki Goodwin for app design and logo branding; Mathew Radford, Thusitha Kumarasinghe, and the Monash eSolution Team for IT assessment, iTunes Connect account, and Monash Email management; and Joy Hewitt, John Morrison, Chacko Joseph, and the Monash Insurance Service Team for the app administration, release approval, and insurance services.

JL is supported by research grants from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (R01 AI132154 and AI146160). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health. CL was supported by the Bridging Postdoctoral Fellowship of Faculty of Medicine, Nursing and Health Sciences, Monash University (BPF17-0021), and is supported by a CJ Martin Early Career Research Fellowship of the National Health and Medical Research Council (NHMRC) of Australia (1143366). JL is an NHMRC Principal Research Fellow (1157909).

Authors' Contributions

JL conceived and designed the project; XH and CL designed, constructed, and tested the ColistinDose app; JMP, VSS, IK, KSK, BTT, PJB, YZ, and JS tested the ColistinDose app and provided critical feedback. CL, PJB, and JL drafted and revised the manuscript that has been read and approved by all the authors.

Conflicts of Interest

KSK is a consultant of Xellia Pharmaceuticals. The other authors have no conflicts to declare.

Multimedia Appendix 1
Detailed information of the open-source libraries and fonts utilized in ColistinDose.
[DOCX File , 14 KB - mhealth_v8i12e20525_app1.docx]

Multimedia Appendix 2
Mock calculations of the maintenance daily doses in different patients.
[DOCX File , 19 KB - mhealth_v8i12e20525_app2.docx]

Multimedia Appendix 3
Disclaimer and terms of use.
[DOCX File , 15 KB - mhealth_v8i12e20525_app3.docx]

References

1. Jackson N, Czaplewski L, Piddock L. Discovery and development of new antibacterial drugs: learning from experience? J Antimicrob Chemother 2018;73(6):1452-1459. [doi: 10.1093/jac/dky019]
2. Luepke KH, Suda KJ, Boucher H, Russo RL, Bonney MW, Hunt TD, et al. Past, Present, and Future of Antibacterial Economics: Increasing Bacterial Resistance, Limited Antibiotic Pipeline, and Societal Implications. Pharmacotherapy 2017;37(1):71-84. [doi: 10.1002/phar.1868]
3. Boucher HW, Talbot GH, Benjamin DK, Bradley J, Guidos RJ, Jones RN, et al. 10 x '20 Progress--development of new drugs active against gram-negative bacilli: an update from the Infectious Diseases Society of America. Clinical Infectious Diseases 2013 Apr 17;56(12):1685-1694. [doi: 10.1093/cid/cit152]
4. Landman D, Georgescu C, Martin DA, Quale J. Polymyxins revisited. Clin Microbiol Rev 2008 Jul;21(3):449-465. [doi: 10.1128/cmr.00006-08]
5. Velkov T, Roberts KD, Thompson PE, Li J. Polymyxins: a new hope in combating Gram-negative superbugs? Future Med Chem 2016 Jun;8(10):1017-1025. [doi: 10.4155/fmc-2016-0091]
6. Bergen PJ, Li J, Rayner CR, Nation RL. Colistin methanesulfonate is an inactive prodrug of colistin against Pseudomonas aeruginosa. Antimicrob Agents Chemother 2006 Jun;50(6):1953-1958. [doi: 10.1128/aac.00035-06]
7. Azad MAK, Nation RL, Velkov T, Li J. Mechanisms of polymyxin-induced nephrotoxicity. Adv Exp Med Biol 2019;1145:305-319. [doi: 10.1007/978-3-030-16373-0_18]
8. Kubin CJ, Ellman TM, Phadke V, Haynes LJ, Calfee DP, Yin MT. Incidence and predictors of acute kidney injury associated with intravenous polymyxin B therapy. J Infect 2012 Jul;65(1):80-87. [doi: 10.1016/j.jinf.2012.01.015]
9. Rigatto MH, Behle TF, Falci DR, Freitas T, Lopes NT, Nunes M, et al. Risk factors for acute kidney injury (AKI) in patients treated with polymyxin B and influence of AKI on mortality: a multicentre prospective cohort study. J Antimicrob Chemother 2015 Jan 20;70(5):1552-1557. [doi: 10.1093/jac/dku561]
10. Garonzik SM, Li J, Thamlikitkul V, Paterson DL, Shoham S, Jacob J, et al. Population pharmacokinetics of colistin methanesulfonate and formed colistin in critically ill patients from a multicenter study provide dosing suggestions for various categories of patients. Antimicrob Agents Chemother 2011 May 09;55(7):3284-3294. [doi: 10.1128/aac.01733-10]
11. Nation RL, Garonzik SM, Thamlikitkul V, Giamarellos-Bourboulis EJ, Forrest A, Paterson DL, et al. Dosing guidance for intravenous colistin in critically-ill patients. Clin Infect Dis 2017;64(5):565-571. [doi: 10.1093/cid/ciw839]
12. Cheah S, Wang J, Nguyen VTT, Turnidge JD, Li J, Nation RL. New pharmacokinetic/pharmacodynamic studies of systemically administered colistin against Pseudomonas aeruginosa and Acinetobacter baumannii in mouse thigh and lung

infection models: smaller response in lung infection. J Antimicrob Chemother 2015 Aug 27;70(12):3291-3297. [doi: 10.1093/jac/dkv267]
13. Grégoire N, Mimoz O, Mégarbane B, Comets E, Chatelier D, Lasocki S, et al. New colistin population pharmacokinetic data in critically ill patients suggesting an alternative loading dose rationale. Antimicrob Agents Chemother 2014 Sep 29;58(12):7324-7330. [doi: 10.1128/aac.03508-14]
14. Jacobs M, Grégoire N, Mégarbane B, Gobin P, Balayn D, Marchand S, et al. Population pharmacokinetics of colistin methanesulfonate and colistin in critically ill patients with acute renal failure requiring intermittent hemodialysis. Antimicrob Agents Chemother 2016 Jan 04;60(3):1788-1793. [doi: 10.1128/aac.01868-15]
15. Karaiskos I, Friberg LE, Pontikis K, Ioannidis K, Tsagkari V, Galani L, et al. Colistin population pharmacokinetics after application of a loading dose of 9 MU colistin methanesulfonate in critically ill patients. Antimicrob Agents Chemother 2015 Sep 14;59(12):7240-7248. [doi: 10.1128/aac.00554-15]
16. Bergen PJ, Bulitta JB, Forrest A, Tsuji BT, Li J, Nation RL. Pharmacokinetic/pharmacodynamic investigation of colistin against Pseudomonas aeruginosa using an in vitro model. Antimicrob Agents Chemother 2010 Sep;54(9):3783-3789. [doi: 10.1128/aac.00903-09]
17. Mohamed AF, Karaiskos I, Plachouras D, Karvanen M, Pontikis K, Jansson B, et al. Application of a loading dose of colistin methanesulfonate in critically ill patients: population pharmacokinetics, protein binding, and prediction of bacterial kill. Antimicrob Agents Chemother 2012 May 21;56(8):4241-4249. [doi: 10.1128/aac.06426-11]
18. Plachouras D, Karvanen M, Friberg LE, Papadomichelakis E, Antoniadou A, Tsangaris I, et al. Population pharmacokinetic analysis of colistin methanesulfonate and colistin after intravenous administration in critically ill patients with infections caused by gram-negative bacteria. Antimicrob Agents Chemother 2009 Aug;53(8):3430-3436. [doi: 10.1128/aac.01361-08]
19. Li J, Coulthard K, Nation R. Labelling conventions and product package insert of parenteral polymyxins: factors causing potential medication errors and impeding optimal clinical use. Adv Exp Med Biol 2019;1445:133-141. [doi: 10.1007/978-3-030-16373-0_10]
20. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364. [Medline: 24883008]
21. McAuley D. Colistin dosing calculator (Beta). GlobalRPh. URL: https://globalrph.com/medcalcs/colistin-dosing-calculator-beta/ [accessed 2020-09-23]
22. Justo JA. Colistin calculator. ClinCalc. URL: https://clincalc.com/Colistin/ [accessed 2020-09-23]
23. U.S. Food and Drug Administration. Coly-Mycin M Parenteral (Colistmethate for Injection, USP). U.S. Food and Drug Administration.: U.S. Food and Drug Administration URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/050108s033lbl.pdf [accessed 2020-08-23]
24. Cockcroft DW, Gault H. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16(1):31-41. [doi: 10.1159/000180580]
25. Raghavan M, Marik PE. Management of sepsis during the early "golden hours". J Emerg Med 2006 Aug;31(2):185-199. [doi: 10.1016/j.jemermed.2006.05.008]
26. Lim LM, Ly N, Anderson D, Yang JC, Macander L, Jarkowski A, et al. Resurgence of colistin: a review of resistance, toxicity, pharmacodynamics, and dosing. Pharmacotherapy 2010 Dec;30(12):1279-1291. [doi: 10.1592/phco.30.12.1279]
27. Devine BJ. Gentamicin therapy. Drug Intell Clin Pharm 1974;8:650-655.
28. Li J, Milne RW, Nation RL, Turnidge JD, Smeaton TC, Coulthard K. Pharmacokinetics of colistin methanesulphonate and colistin in rats following an intravenous dose of colistin methanesulphonate. J Antimicrob Chemother 2004 Mar 24;53(5):837-840. [doi: 10.1093/jac/dkh167]
29. Li J, Milne RW, Nation RL, Turnidge JD, Smeaton TC, Coulthard K. Use of high-performance liquid chromatography to study the pharmacokinetics of colistin sulfate in rats following intravenous administration. Antimicrob Agents Chemother 2003 May;47(5):1766-1770. [doi: 10.1128/aac.47.5.1766-1770.2003]
30. Markou N, Markantonis SL, Dimitrakis E, Panidis D, Boutzouka E, Karatzas S, et al. Colistin serum concentrations after intravenous administration in critically ill patients with serious multidrug-resistant, gram-negative bacilli infections: A prospective, open-label, uncontrolled study. Clin Ther 2008 Jan;30(1):143-151. [doi: 10.1016/j.clinthera.2008.01.015]
31. Tsuji BT, Pogue JM, Zavascki AP, Paul M, Daikos GL, Forrest A, et al. International consensus guidelines for the optimal use of the polymyxins: endorsed by the American College of Clinical Pharmacy (ACCP), European Society of Clinical Microbiology and Infectious Diseases (ESCMID), Infectious Diseases Society of America (IDSA), International Society for Anti‐infective Pharmacology (ISAP), Society of Critical Care Medicine (SCCM), and Society of Infectious Diseases Pharmacists (SIDP). Pharmacotherapy 2019 Feb 02;39(1):10-39. [doi: 10.1002/phar.2209]

Abbreviations

CBA: colistin base activity
CMS: colistin methanesulfonate
CrCl: creatinine clearance
CRRT: continuous renal replacement therapy

Css,avg: average steady-state plasma colistin concentration
CVVH: continuous veno-venous hemofiltration
CVVHD: continuous veno-venous hemodialysis
IBW: ideal body weight
IHD: intermittent hemodialysis
IU: international units
MIC: minimal inhibitory concentration
PD: pharmacodynamics
PK: pharmacokinetics
RRT: renal replacement therapy
SLED: sustained low-efficiency dialysis

Edited by G Eysenbach, L Buis; submitted 21.05.20; peer-reviewed by M Scheetz, KL Mauco; comments to author 02.08.20; revised version received 23.09.20; accepted 28.10.20; published 16.12.20.

Please cite as:
Hua X, Li C, Pogue JM, Sharma VS, Karaiskos I, Kaye KS, Tsuji BT, Bergen PJ, Zhu Y, Song J, Li J
ColistinDose, a Mobile App for Determining Intravenous Dosage Regimens of Colistimethate in Critically Ill Adult Patients: Clinician-Centered Design and Development Study
JMIR Mhealth Uhealth 2020;8(12):e20525
URL: http://mhealth.jmir.org/2020/12/e20525/
doi:10.2196/20525
PMID:33325835

Original Paper

Opportunities and Challenges of a Self-Management App to Support People With Spinal Cord Injury in the Prevention of Pressure Injuries: Qualitative Study

Julia Amann[1,2], PhD; Maddalena Fiordelli[1,3], PhD; Anke Scheel-Sailer[4,5], MD; Mirjam Brach[1,5], MPH, MBA; Sara Rubinelli[1,5], PhD

[1]Swiss Paraplegic Research, Nottwil, Switzerland

[2]Health Ethics and Policy Lab, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

[3]Institute of Communication and Health, Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland

[4]Swiss Paraplegic Center, Nottwil, Switzerland

[5]Department of Health Sciences and Medicine, University of Lucerne, Lucerne, Switzerland

Corresponding Author:
Julia Amann, PhD
Swiss Paraplegic Research
Guido A Zäch Strasse 4
Nottwil, 6207
Switzerland
Phone: 41 44632 ext 4619
Email: julia.amann@hest.ethz.ch

Abstract

Background: Mobile health applications can offer tailored self-management support to individuals living with chronic health conditions. However, there are several challenges to the adoption of these technologies in practice. Co-design is a promising approach to overcoming some of these challenges by enabling the development of solutions that meet the actual needs and preferences of the relevant stakeholder groups.

Objective: Taking spinal cord injury as a case in point, the overall objectives of this study were to identify the perceived benefits of a co-designed self-management app that could promote its uptake and to explore the factors that may impede adoption.

Methods: We adopted a qualitative research approach guided by the Technology Acceptance Model. Data were collected through semistructured interviews with individuals with spinal cord injury (n=15) and two focus groups with health care professionals specialized in spinal cord injury (n=7, n=5). Prior to the interviews and focus groups, study participants were given time to explore the app prototype. All interviews were transcribed verbatim and analyzed using inductive thematic analysis.

Results: Findings of our analysis indicate that study participants perceived the app prototype as potentially useful for supporting individuals with spinal cord injury in preventing pressure injuries. In particular, we identified three concrete use cases highlighting the benefits of the app for different audiences: (1) a companion for newly injured individuals, (2) an emergency kit and motivational support, and 3) a guide for informal caregivers and family members. We also uncovered several challenges that might impede the adoption of the self-management app in practice, including (1) challenges in motivating individuals to use the app, (2) concerns about the misuse and abuse of the app, and (3) organizational and maintenance challenges.

Conclusions: This study adds to a growing body of research that investigates individuals' adoption and nonadoption behavior regarding mobile health solutions. Building on earlier work, we make recommendations on how to address the barriers to the adoption of mobile health solutions identified by this study. In particular, there is a need to foster trust in mobile health among prospective users, including both patients and health care professionals. Moreover, increasing personal relevance of mobile health solutions through personalization may be a promising approach to promote uptake. Last but not least, organizational support also plays an instrumental role in mobile health adoption. We conclude that even though co-design is promoted as a promising approach to develop self-management tools, co-design does not guarantee adoption. More research is needed to identify the most promising strategies to promote the adoption of evidence-based mobile health solutions in practice.

(JMIR Mhealth Uhealth 2020;8(12):e22452) doi:10.2196/22452

KEYWORDS

mHealth; eHealth; self-management; spinal cord injury; pressure injury; prevention; technology acceptance

Introduction

Mobile health applications promise to provide tailored self-management support to individuals living with chronic health conditions. However, there are several challenges to the adoption of these technologies in practice. First, while there is an abundance of health apps freely available, the majority are not evidence-based [1,2]. This is especially true for apps aimed at individuals with disabilities [3-5]. There are also no clear guidelines or quality standards when it comes to mobile health (mHealth) applications. This may, in turn, discourage health care professionals from recommending them to patients [6]. Second, research has shown that individuals' use of health applications often declines over time when the initial phase of curiosity has worn off [7,8]. A potential explanation for this may be that there is a misfit of the technology leading to a lack of integration into the person's life. Third, technologies are evolving at an ever-increasing pace, yet, in health care, moving from development to adoption is a long and complex process [9,10]. This means that once a pilot evaluation is concluded, the application's design and functionalities may already be outdated and may thus fail to reproduce the outcomes of a pilot evaluation in a real-life setting.

Involving relevant stakeholder groups in the early stages of the app development process constitutes a promising approach to address these challenges [11-14]. First, co-design can ensure an evidence-guided approach to translating different stakeholders' needs and preferences into digital solutions in a meaningful way [12,15-19]. Moreover, if decisions regarding content, functionality, and design are guided by real-life experiences of prospective users, this helps to ensure the appropriateness and relevance of self-management applications. Involving prospective users also fosters trust and the credibility of the technology, thus increasing the likelihood of system acceptance and, as a result, accelerating adoption into practice [20,21].

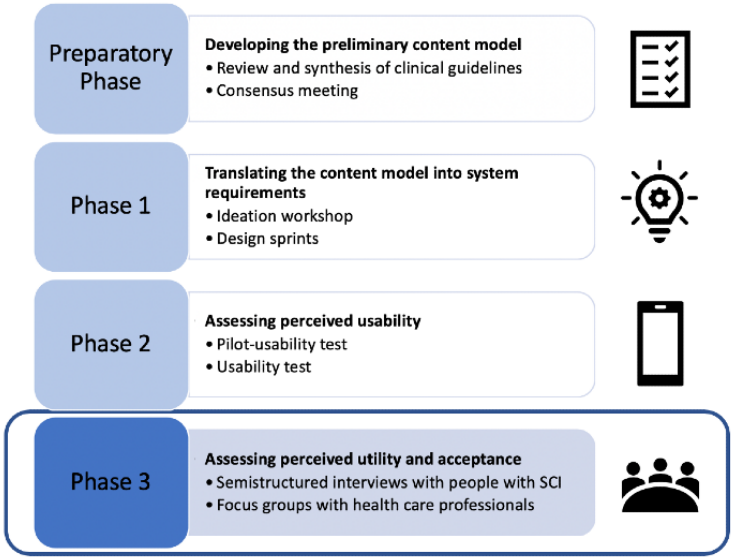
This study forms part of a larger co-design research project carried out in Switzerland. The project aimed to develop an evidence-based self-management app to support community-dwelling individuals with spinal cord injury (SCI) in the prevention of pressure injuries. SCI is a complex chronic health condition that makes individuals prone to incurring secondary health conditions. Pressure injuries are among the most common and serious secondary health conditions in people with SCI [22-24]. In addition to a number of nonmodifiable risk factors (eg, natural skin aging) [25], self-management plays a key role in the prevention of pressure injuries [26,27]. The overall objectives of this study were to identify the perceived benefits of the self-management app that could promote its uptake and to explore the factors that may impede its adoption.

Methods

Study Design

This study forms part of a larger mixed-methods study that combined qualitative research methods (ie, semistructured interviews) into user-centered design approaches (ie, ideation workshop, design sprints, usability tests). Figure 1 illustrates the methodological approach of the overall project. Detailed descriptions of the app development process have been published elsewhere [5,28]. This paper focuses on assessing the perceived utility and acceptance of the app prototype that resulted from the previous phases of this project. The study was approved by the cantonal ethics commission (EKNZ 2017-01787).

Figure 1. Overview of the study methodology. SCI: spinal cord injury.



Data Collection and Analysis

Following a convenience sampling approach, we conducted 15 semistructured interviews with individuals with SCI and two focus groups with health care professionals specialized in SCI (n=7; n=5). Both individuals with SCI and health care professionals were given time to explore the app prototype (see Multimedia Appendix 1 for screenshots of the app prototype) prior to the interviews/focus groups. People with SCI first took part in a usability test [28] that allowed them to explore the app prototype in more depth. Healthcare professionals received a brief introduction to the main functionalities and elements of the app by the first author (JA). They could then freely explore the app prototype in pairs for 15 minutes before the focus group discussion. The interview and focus group guides were informed by the constructs of the Technology Acceptance Model [29]. In particular, we aimed to capture participants' views regarding usability and perceived usefulness of the app prototype, as well as their attitudes and intentions toward using the app. In this paper, we focused our analysis on identifying concrete benefits of the app that could promote its uptake and exploring the factors that may impede its adoption into practice. Results of a preliminary usability test of the co-designed self-management app are reported elsewhere [28].

All interviews and both focus groups were conducted in German by the first author (JA), who is a native German speaker and communication scholar trained in qualitative research methods. All interviews and both focus groups were audio-recorded. Additionally, a research assistant was present during all interviews and focus groups to take notes. Interviews lasted 33 minutes on average, ranging between 17 and 69 minutes. Focus group 1 (FG1; n=7) lasted 100 minutes, and focus group 2 (FG2; n=5) lasted 97 minutes. In total, 687 minutes of audio material were transcribed verbatim for analysis.

Data were analyzed using inductive thematic analysis [30]. For this purpose, initial line-by-line coding was carried out by the first author (JA). In this initial coding phase, we aimed to identify instances where participants described situations where they either would personally use or could imagine someone else engaging with the self-management app. In the next step, we focused on uncovering the factors that may impede adoption by focusing on those instances where participants voiced concerns or challenges. Starting from the initial coding and in line with our research objectives, we identified preliminary themes and subthemes, which were continuously refined through discussion in regular meetings involving all coauthors.

The App Prototype

The interactive web-based app prototype evaluated in this study is the result of a multistage co-design process [5,28]. It includes five main features: (1) a smart camera, (2) a pressure injury diary, (3) an expert consultation, (4) reminders, and (5) a knowledge repository.

As such, the app prototype comprises features that can be used independently by people with SCI (eg, documenting pressure injuries, setting reminders), as well as a communication component that allows for interaction with Parahelp, a home care service provider specialized in SCI care. (See Multimedia Appendix 1 for screenshots of the app prototype.)

Participants

Within the scope of the overall project, we aimed to involve a balanced sample of participants in terms of age, gender, lesion level for people with SCI, and health care professionals' expertise [28]. Prospective participants were identified through the local rehabilitation center (Swiss Paraplegic Center), patient association (Swiss Paraplegic Association), and home care provider specialized in SCI (Parahelp). Additionally, we launched a call for participants on the Paraplegie Community [31], an online community for people with SCI initiated and funded by the Swiss Paraplegic Foundation. All participants received detailed information on the study and signed a consent form. Study participants' characteristics for the interviews and focus groups are presented in Table 1 and Table 2, respectively.

Table 1. Interview participant characteristics.

Characteristic	People with SCI (n=15)
Age (years), mean (range)	40.8 (28-58)
Gender, n	
Male	11
Female	4
Lesion level, n	
Paraplegic	7
Quadriplegic	8

Table 2. Focus group participant characteristics.

Characteristic	Total (n=12)	Focus group 1 (n=7)	Focus group 2 (n=5)
Work experience (years), mean (range)	22.9 (10-36)	25.4 (12-36)	19.4 (10-30)
Gender, n			
Male	1	1	0
Female	11	6	5
Professional role, n			
Wound expert	7	5	2
Administrator	1	1	0
Psychologist	1	0	1
Occupational therapist	1	0	1
Nutritionist	1	1	0
Therapy instructor	1	0	1

Results

Overview

Findings of our analysis indicate that both individuals with SCI and health care professionals saw potential value in the self-management app prototype. Study participants acknowledged its usefulness for different life situations and target groups, including newly injured individuals and experienced wheelchair users, as well as informal caregivers. In addition, we also identified some of the key factors that could impede the app's adoption into practice. In the Results section, we illustrate our findings using quotes from the interviews and focus groups, each followed by a participant identifier. With the exception of the theme "Organizational and maintenance challenges" (which was predominantly present in the focus groups with health care professionals), themes occurred fairly equally in both groups. Thus, when we refer to "study participants," this encompasses both individuals with SCI and health care professionals.

Benefits Promoting Uptake

A Companion for Newly Injured Individuals

Wheelchair users and health care professionals concurred that individuals who recently suffered a spinal cord injury, for example, those who are completing or have recently completed first rehabilitation, could particularly benefit from the app as a complementary measure to traditional patient education.

> *The app could be complementary [to traditional patient education], yes. So the nurses could introduce the app to patients, maybe those that are not so much into paper-based information sheets and prefer something more interactive.* [B, FG2]

One of the participants with SCI underlined how the app could support people who were recently injured in processing information at their own pace and in their own time. In recalling her own experiences from first rehabilitation, the participant described how the app could contribute to fostering patient autonomy and self-determination.

> *But then the problem is simply that you were forced to attend this information day [on pressure injuries]. And with an app like this, I have the feeling that you have a look at the information when you want to. And you take as much information as you can process at that particular moment. And the advantage is simply that I can just shut it down. And I can then, perhaps in the evening, have another look at it when I'm in bed. I can really adjust it to my pace. And then, I feel, you're more receptive. Because everybody has their own pace. And I think it's bad when a certain pace is forced on you. Even on important issues. But if it's forced on me and I block it out, then it's no good for me.* [P15]

Another important aspect that participants noted was that the app could help to lower individuals' inhibition thresholds when in need of medical assistance or advice. In other words, they described the app as a low-key means of getting in touch with evidence-based knowledge and a health care professional at times or in situations where one might otherwise feel uncomfortable. Given that newly injured individuals in particular may not immediately recognize the seriousness of an early-stage pressure injury or may be embarrassed to seek help, participants noted that the app could support these individuals in early detection before the situation worsens.

> *Especially on Sunday or late in the evening, then you don't dare to call, then you can just quickly ask, I think that's a very good thing. For example, if you get a bladder infection with a fever and you don't know exactly how/what/where - if you don't have that much experience. Certainly also for the relatives or caregivers [this would be useful]...at the onset, I think that's great.* [P2]

> *Or you can also ask the question anonymously if you are embarrassed, if the pressure injury is in an embarrassing place, for example. [Then it's probably easier to ask such a question], I could imagine. Maybe if you're embarrassed about having to take a picture [of a pressure injury] on your butt.* [P6]

Both individuals with SCI and health care professionals also pointed out that the app could help newly injured individuals to build and adopt new habits to better self-manage. Building these new habits after partial loss of functioning is key to effective self-management. Wheelchair users attributed particular value to the reminder function as a potentially powerful habit-building assistance that can provide structure and introduce new routines into someone's life.

> *I would target first rehab patients more. The ones who don't have the knowledge are the ones who really need to be told. So that it becomes routine.* [P13]

> *I should be able to work with that, but maybe someone who's just starting rehab. ...They'd have to use the app to get into the rhythm. That helps 100% in my opinion.* [P1]

> *So, to some extent, if you get the reminder, maybe it can help you that you get into a rhythm like, like I have developed one now over the years. And you say, "Okay, over lunchtime, in the evening I look at this and that," and yes, maybe that can help you. That in the end you'll have the rhythm internalized. You just get the reminder. So just as a support. Because I think that's important. That you still have a rhythm of your own, but you can adjust this here [in the app].* [P11]

An Emergency Kit and Motivational Support

Individuals with SCI highlighted that the app could be particularly useful under extraordinary conditions (ie, outside of their daily routine). According to them, these were situations where even experienced wheelchair users might be hesitant or unable to use their usual means of contact (ie, telephone call). Frequently mentioned examples of such extraordinary circumstances were during holidays abroad, where one might not trust the foreign health care system, or when fallen ill. Wheelchair users also described nighttime and weekends as situations where the app could potentially come in handy.

> *You can directly get in touch with people, with specialists, I would appreciate that. For example, I once had a bruise when I lived in the south of France, then I actually did it exactly like that: I took pictures and thought about which doctors I knew and sent the photo.* [P10]

> *If it looks like this, you should try this and that for two days, and if it doesn't get better, you have to do something. Now if someone didn't have a doctor while they are on holiday, let's say, if you're somewhere where you don't have that much confidence in medical assistance, I think it [the app] is great.* [P6]

These views were shared by the health care professionals. They also noted that the app could support wheelchair users in recognizing skin problems at an early stage. According to them, the app could help to reduce the risk of individuals waiting too long to get in touch by providing them with easy access to health care providers' expertise. In this context, health care professionals emphasized the benefits of being able to receive images (ie, using the smart camera feature), allowing them to get an idea of how serious the situation is.

> *A: Maybe this app would lower the inhibition threshold to ask questions. So if someone has the feeling "Ah, I have a scratch or something" and before going running to the doctor or saying "[It's not that bad,] I'm not going for just that," you could ask experts, send a photo. I think maybe this could help you to identify [a skin problem] early or act or react faster.*

> *B: That's for at home, when uncertainties arise, when you don't know exactly what you're supposed to do. So you can just quickly send a photo to Parahelp.*

> *D: Yes or maybe make an appointment and talk...*

> *B: Right, instead of having to drive 30 km [to see a health care professional in person].* [FG2]

Both wheelchair users and health care professionals described refreshing one's memory and extending knowledge on pressure injury prevention as key benefits that could promote self-management, particularly among more experienced wheelchair users. In addition, being able to learn from one's own experience by reflecting on past experiences or behaviors (ie, through self-documentation in the pressure injury diary) was mentioned as a potential motivational boost to become more self-aware and attentive.

> *You can experiment a little, browse around a little. Maybe refresh some forgotten knowledge. Watch a short video. How do others do it? The curiosity "What do the others do?" is always there, so it's actually a win-win situation.* [FG1, D]

> *I say now, long-term reminders I like because, for example, I didn't know that there was such thing as a cushion-check. I would do it if I no longer felt comfortable or was no longer sitting comfortably, then I would do a cushion-check. This can sometimes take two to three days or maybe even a weekend. I do a lot intuitively, so it's good to introduce certain things like the cushion-check, I think that's a good thing.* [P1]

> *Here I have the opportunity to see this every day when a photo comes in. To see what I should change, have I worn a different shoe, or are was it a hot day and the foot is swollen. That would be more obvious here, in any case. I could go through this process more thoroughly [in the app], in the sense of self-reflection. Through reflection we learn, we can also develop and grow. That has certain advantages, I think.* [P9]

A Guide for Informal Caregivers and Family Members

According to the study participants, informal caregivers could also benefit from using the app as a tool to learn more about spinal cord injury and the risk of pressure injuries. They suggested that the app could allow caregivers to better understand the manifold challenges that individuals with SCI face and propose ways to better support them in their daily life. In this context, both wheelchair users and health care professionals mentioned practical aspects that could help to increase knowledge and skills among caregivers and, as a result, improve support for people with SCI. In their remarks,

participants touched upon the importance of shared responsibility of care and the related need to provide caregivers with evidence-based information and practical support.

> *Yes, or a caregiver. I'd say a 60+ elderly lady, who looks after her husband. So if a person is really sleeping all day, that you remind the caregiver or the person themselves: "Okay, I have to change sides or lie on my stomach.*" [P14]

> *And also, and certainly, the family members. The next of kin. That they know what this is all about. Because it's not enough for the person in the wheelchair to know, instead the family should also know where the spot is and know if it's somehow normal-red, which means it can be pressed away.* [P15]

> *Above all, however, we notice that the relatives are most interested in what could change or has changed in terms of their [wheelchair user's] care and status, at least that's the case for first-rehab patients. Those [informational materials] are mostly used by the relatives, also at home, when suddenly questions arise that have not yet been a topic in everyday life.* [D, FG2]

When comparing how information was currently provided to family members (ie, a bulky informational brochure) to the possibilities that the app would offer, health care professionals agreed that it would make information retrieval easier. They argued that an app could easily be used on the go, at any time, even just out of boredom. According to them, this may not only help family members to save time when looking for information but may also incentivize them to do so. Most importantly, the app would provide them with curated, evidence-based information.

> *B: Yes, that's just the way it is, they hear, the relatives also hear everything, but it's perhaps a lot of information on such a day and yes...That's the issue with the patient manual, it's just an A4 folder, a fairly bulky folder that takes up space. You always have to open it first...And just like that, here [with the app] open it and bang, bang, bang, and then you are where you are...That's the difference maybe, which makes it more motivating to deal with it.*

> *G: You're probably more likely to have a look at it.*

> *B: Yes, when you're bored and then just browse.*

> *G: You click around more.*

> *F: It always saves time when I am looking for information.* [FG2]

One study participant with SCI also referred to the potential of the app to address emotional aspects that are of great importance for the psychological well-being of people with SCI and the people around them alike. In his view, the app could help to foster understanding and empathy for individuals with SCI.

> *Yes, for all those in wheelchairs and this could also be interesting for the family, provide a clear overview, or for people who are not directly but indirectly affected through someone they know. That they get help and tips on how to interact with my son or friend, that's what I find super cool. ...because a short example, I thought I had a huge family but since my accident three years ago they have all disappeared. So really disappeared or they are distancing themselves, etc.* [P5]

Challenges Impeding Adoption

Motivating Individuals to Use the App

Established attitudes and behavioral patterns of people with SCI toward pressure injuries constitute potential obstacles. A key challenge we identified relates to onboarding or, in other words, motivating both experienced and newly injured individuals to use the app. Healthcare professionals voiced concerns that wheelchair users may, in fact, only start using the app once they experience skin problems. This raises questions as to how to best engage newly injured individuals in early preventive measures.

> *There are many who come and say: Now I've been in a wheelchair for 20 years, I've had no problems with my skin and now all of a sudden I have a decu [pressure injury]. Well, that won't be someone who has regularly looked into such an app before. You probably only do that once you experience problems.* [C, FG2]

> *I see it very much as an acute approach, "If something happens to me, I have a red spot, uh. I know here [in the app] I might find an answer."* [F, FG2]

In line with this, some of the wheelchair users noted that they didn't consider themselves a primary target audience. In this context, some noted that frequent reminders in the form of pop-up messages could easily become annoying. Drawing on their own experiences, wheelchair users explained that they had internalized important preventive behaviors and knowledge. While they could see the potential benefits of the app for others, both newly injured and experienced, they themselves indicated that they would not need the app.

> *The reminders can also become too much, but in a way, I think it's actually good, because sometimes it takes a bit [of a push], doesn't it? Many, especially wheelchair users, who may have other things to do...It helps them if they are reminded a little bit of what they have to do, of things that might not be a lot of fun right now. But maybe it will give them the extra drive when you see it [the reminder]. And then you do it and then you're done.* [P3]

> *I don't really need reminders as a support, I actually do everything quite independently. Sometimes you notice that it wasn't enough pressure-release and that you have to do something else again - so, no, I don't need something like that [an app to remind me].* [P2]

Similarly, the secure contact function was viewed as beneficial, yet many wheelchair users noted that under ordinary conditions, they would rather see a health professional in person or call them in case of an emergency. A key consideration in case of an emergency was to choose the quickest and simplest way to

get professional assistance, and according to some, this may not necessarily be through the app's contact function.

> *Many people are afraid [to call]. But with me, if I want to know something, I usually call quickly. And I know I can find it somewhere on the website myself, but I don't have time. Just tell me quickly and done.* [P4]

> *If I had a pressure injury, I'd take a picture of it. I had one once and I was told to take a picture of it when I called. My friend took a picture of it quickly. The question is, how fast will it be handled when I put it into an app or when I call, what is the better or faster way, that's what I'm wondering.* [P6]

In relation to behavioral and attitudinal barriers, health care professionals raised another important issue they encountered in their daily work: "regular clients" (ie, patients who regularly have problems and often have to be hospitalized because of these problems). According to them, this group of patients would use pressure injuries as a means of attracting attention, sometimes even consciously manipulating wounds to receive care. They considered it unlikely that these individuals could be motivated to use the app, let alone benefit from the app to prevent pressure injuries.

> *B: I know wound patients who tamper with their wounds so that someone will come and take care of them.*
>
> *A: Exactly, but unfortunately we can't reach those people [with the app].*
>
> *B: But these are not necessarily the people who want to improve their situation. But these are the people who, let's say, when they experience healing, then they also experience a loss of care and this loss of care they just can't handle...So these are not the people we reach with something like that [the app].* [FG2]

Concerns About Misuse and Abuse of the App

Study participants expressed concerns that the app could potentially be misused and that both patients and health care professionals could fall victim to such misuse. On the one hand, they worried about privacy and data protection issues that prospective app users would be facing. In particular, participants raised questions regarding the storage, sharing, backup, and deletion of personal data.

> *On the one hand, that would be an advantage for me and certainly also an advantage if you could read it again. On the other hand, I simply have to ask myself, is the data even stored safely? Does it really stay where it is or is it passed on to third parties?* [P12]

> *Privacy is very important to me that people keep information to themselves and don't pass it on to anyone. ...Inside Parahelp I think it would be ok, but if it suddenly ends up with the plastic surgeons, I would really have a problem with that. It must really be a trusting relationship.* [P13]

> *If I delete the app again, I have seen that there is a registration, so I assume that it [data] remains stored there. Do I want to have this data available or do I want it effectively deleted so that it is no longer on the server or delete the entire registration? That would be interesting to know. Yes, the risk is, that something is done with the data. ...* [P14]

However, these concerns were not shared by all participants. For instance, one study participant with SCI noted that the stand-alone use of a smartphone with mobile data enabled would already compromise personal privacy. He thus did not consider the collection of data through the self-management app problematic but rather a necessary requirement to provide adequate support.

> *Either you live totally cut off or you can forget it because as soon as you have even one Facebook account, or just one Instagram account, or a smartphone with data volume, your privacy is limited. That's why it [the app] is not dangerous at all, on the contrary...The data helps you, that's why it needs certain permissions, like camera access, like audio access and position navigation, like GPS data coordinate access, that's what it needs.* [P5]

In addition to privacy concerns impacting the user, some study participants were concerned about the potential misuse of the app by patients and the negative impact this could have on health care professionals. In particular, they worried about excessive messaging from people with SCI without actual skin problems. According to the study participants, the fact that it would easier to initiate contact may lower individuals' inhibition thresholds, thus increasing the risk that some people might contact health care professionals for irrelevant information or just to talk to someone.

> *Maybe lonely people who put a lot of stress on the experts, that would be a disadvantage for the experts. Yes, there are many wheelchair users who are lonely. You can see that around here [specialized center], there are many who want to live here and try to get something [a place to live here]. Here it's a place where you don't feel lonely anymore. I can imagine that if you are bored, some people will just have a look [at who is online]. "There's Julia or Freddy. I'll ask."* [P10]

> *The fact is, there are many who feel alone and they just want to chat and start [a conversation] for every little thing or for nothing. But the aim is certainly not to overwhelm the professionals.* [P5]

> *The risk is just that you have people out there who are just lonely and would use the app in search of interaction, you know?* [F, FG1]

Organizational and Maintenance Challenges

When discussing what was needed to bring the app into practice, health care professionals also referred to organizational problems, including the difficulty of integrating the app with currently used systems. They feared that the adoption of the app might increase their workload (eg, having to reenter existing

information). Healthcare professionals were also unsure of how to securely identify their clientele and worried about managing resources. Here, it is important to note that unlike other home care providers, Parahelp has a specific care mandate, meaning that only individuals meeting predetermined criteria are entitled to receive assistance from Parahelp (ie, persons with spinal cord injury). Determining whether these conditions are met was highlighted as a core challenge by health care professionals.

> *You have to be able to make a query to determine: does this person belong to our target audience or not. Because if you can't have a professional make the selection, then we'll suddenly have everyone in the chat. And then it [workload] explodes.* [F, FG1]

With regard to the challenges of integrating the app into practice, some health care professionals also pointed to weaknesses in the app's content. For example, the lack of information on psychological aspects was criticized. The maintenance and further development of the app were also heavily discussed. In particular, health care professionals were concerned about how information would be kept up to date and who would be responsible for ensuring information accuracy. They agreed that the app content should be regularly checked to ensure that it contains the latest evidence-based information. However, it was unclear who would be responsible for this task.

> *Exactly, so I think the possibility for developments and new innovations - because this is also evolving- this would need to be included [in the app]. It just has to be taken up somehow.* [B, FG2]
>
> *A: But then who feeds the whole thing [app] with topics and content? So if I have such a question now as someone affected, that I go to Romania or something like that? Or I have the question...*
>
> *B: It won't be that specific...*
>
> *A: Somebody has to feed [content into] that [the app].*
>
> *B: Right, exactly. Of course, it can't be that specific, you can't provide resources for every country, it's not possible, but if for example, the person goes to a hot country, or to a cold country.* [FG2]

Discussion

Principal Findings

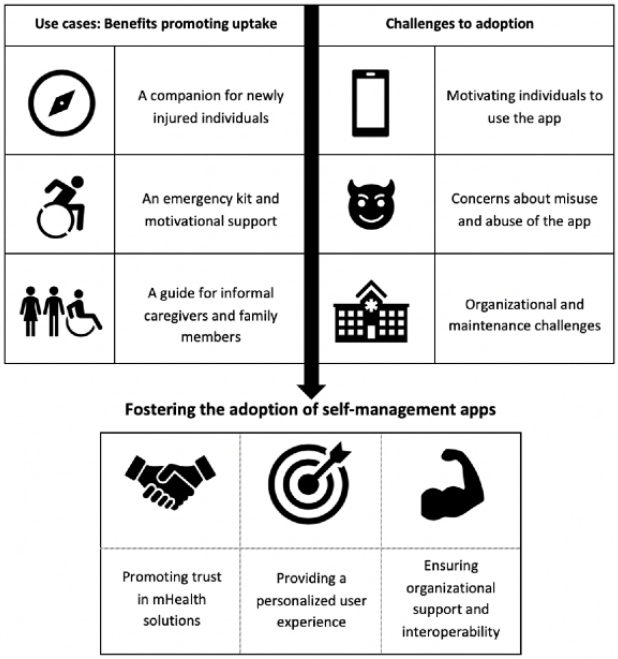
This study identified the perceived benefits of a co-designed self-management app for the prevention of pressure injuries that could promote its uptake and shed light on the factors that may impede its adoption in practice. It is one of the first studies on the perceived opportunities and challenges of a mobile health application to support individuals with spinal cord injury in the prevention of pressure injuries. By capturing both the perspectives of people with spinal cord injury and health care professionals, it complements earlier work, which has predominantly focused on assessing patient views [32] and compliance with app-based interventions [33].

Study participants perceived the app as potentially useful for supporting individuals with SCI in preventing pressure injuries. In particular, we identified three concrete use cases highlighting the benefits of the app: (1) a companion for newly injured individuals, (2) an emergency kit and motivational support, and (3) a guide for informal caregivers and family members. In addition, we also uncovered several challenges that might impede the adoption of the self-management app in practice, including (1) challenges in motivating individuals to use the app, (2) concerns about the misuse and abuse of the app, and (3) organizational and maintenance challenges.

Fostering the Adoption of Self-management Apps

It seems intuitive that a self-management app can be experienced as empowering and motivating by some while being deemed useless or even annoying by others. Similarly, some health care professionals may welcome an app as a promising alternative, while others may worry about increased workload. What can be done to capitalize on the app's perceived benefits and to reduce barriers to uptake? Several studies have addressed this question more generally by developing frameworks aimed at improving the uptake and impact of eHealth technologies [14,34]. Drawing on our findings and Greenhalgh et al's nonadoption, abandonment, scale-up, spread, and sustainability framework [34], we propose three promising strategies to address the barriers to adoption identified by this study (Figure 2).

Figure 2. Fostering the adoption of self-management apps.



Promoting Trust in mHealth Solutions

Findings of this study showed that both participants with SCI and health care professionals had concerns about the data protection and privacy of the app. Individuals with SCI in particular were worried that personal health information might be shared with third parties without their explicit consent. These findings indicate that trust in the app's commitment and adherence to data protection and privacy plays an important role when it comes to adoption. In this context, transparency plays a central role [35,36].

To create trust in an app's commitment and adherence to data protection, users need to be informed about the kind of data being collected, ownership of and control over the data, and their right to have their data deleted. All this information needs to be accessible to prospective users in a clear and concise format, informing them about their rights to access and control their data. However, how personal information is handled by mobile health apps often remains unclear [37-39]. Some apps, for example, track and share user data continuously by default, even when the app is not in active use, sometimes without explicitly informing users about these tracking practices in their terms of use [39]. A recent study [36] that investigated end users' opinions on what constitutes a trustworthy mobile health app found that users interpreted lengthy terms of service and privacy policies as an attempt by developers to obscure possible risks. Moreover, the study showed that users perceived apps that request "too much" personal information as untrustworthy.

Tools like the mHealth App Trustworthiness checklist [36] can help to guide developers in creating trustworthy apps to foster the adoption of mHealth solutions. Visualization techniques might be promising approaches to make privacy information more accessible, making it easier for users to understand how apps collect, use, and share data. Also, a privacy by design approach that allows users to opt in for data sharing and enables granular permissions could help to strengthen user privacy [39] and, as a result, foster trust. Last but not least, the credibility of the provider (ie, the organization behind the app) will also likely have an impact on prospective users' trust. To further strengthen prospective users' trust, authors have also proposed to include logos of reputable organizations endorsing the app (eg, universities or hospitals) [40].

Providing a Personalized User Experience

Our findings showed that established attitudes and behavioral patterns of individuals with SCI constitute potential barriers to adoption. "Why use an app to contact a health care professional when I can just WhatsApp them?" "Why would I need a reminder to check my skin after having done it for the past 20 years?" "Why do I need the app? I never had problems with my skin." These are the kind of questions prospective users might ask themselves. All of these questions hint at an important aspect: value. How can we create value for a diverse user group ranging from newly injured individuals to experienced wheelchair users, from those that constantly struggle with pressure injuries to those who have never experienced any problems with their skin? This is where personalization comes in.

Personalization technologies can help to better tailor the self-management app to an individual user's needs and preferences [41] by ensuring that recommendations regarding preventive measures are perceived as relevant and actionable rather than generic advice. In matching an individual's preferences and lifestyle, personalization can lower barriers to act on recommendations. Contextualization can ensure that interventions (eg, in the form of recommendations) are delivered at moments of need (eg, long time with low activity levels) or at an opportune moment when a particular recommendation is easy to follow (eg, when leaving work) [42,43]. While personalization is a familiar concept in health communication,

artificial intelligence and machine learning techniques offer unprecedented opportunities to take personalization to the next level [42,44,45]. These techniques can then be used to translate population-level data alongside individual user data to provide a personalized user experience, similar to recommendation systems like Netflix [41,42]. However, for personalization to be realized, users are required to disclose personal information [46]. Such information may include personal information that users actively enter themselves (eg, age, gender) but it may be complemented by information on app usage through tracking (eg, how often or when the app is used).

It is, however, important to acknowledge the limits of personalization. What personalization likely cannot achieve is motivating download and first-time use. This is where health care professionals have to step in to promote and endorse the app. However, in order to encourage health care professionals to endorse the app, it is first important to obtain their buy-in. To this end, it will be important to demonstrate that the app allows health care professionals to better support their patients in effectively reducing pressure injuries (ie, by supporting self-management and early detection). Moreover, it will be essential to show that the app will not impose an additional burden on health care professionals, leading to increased workloads.

Ensuring Organizational Support and Interoperability

It is important to consider that in the case of an app that enables communication between patients and health care providers, there are two prospective user groups: senders and receivers. Even though many of the app's features can be used by people with SCI independently, the communication component requires uptake from health care professionals as well. Findings of our study indicate that health care professionals were particularly concerned about challenges in integrating the proposed solution with currently used systems and worried about additional workload to reenter existing information. Additionally, there were also concerns about inappropriate use of the app that could lead to overburdening of staff.

Given the crucial role of health care professionals' acceptance of new technologies, it is of utmost priority to take these concerns seriously and to respond to them in an adequate manner. In addition to ensuring the technical requirements are met, including issues related to system operability, organizations also need to ensure that new systems can be smoothly integrated into existing workflows [47]. Organizations may also want to establish technical training and support to deal with cases of malfunction or inappropriate use [48]. This would allow health care professionals to feel that support is available and may help to reduce concerns over increased workloads.

Close to the completion of our own project, we realized that we had invested a lot of time and resources to understand what the user interface (ie, the front end of the application that individuals with SCI would interact with) should look like. However, our findings indicated that we had neglected to consider the administrative interface (ie, the back end of the application operated by health care professionals). In other words, we had not collected sufficient information on the internal workflows of the care provider Parahelp to determine whether and how the app could actually be integrated into their daily practice. In the current app prototype, we also did not consider reimbursement aspects, which constitute a relevant factor for adoption but, at the same time, trigger many other organizational questions.

Limitations

This study is subject to some limitations. First, we need to acknowledge that our research was focused on a narrow, specialized health care setting in Switzerland. However, while our findings may not necessarily be directly applicable to other health care settings or health conditions, they have relevant implications for mHealth research more broadly. In particular, the strategies we identified to promote the adoption of evidence-based mHealth solutions lend themselves to inform research on self-management apps in other areas. Second, two focus group participants (both affiliated with Parahelp) had been involved in the co-design process of the app prototype at an earlier stage of the project. This may have led them to hold more favorable attitudes toward the project and possibly steer the focus group discussion in this direction. To account for this potential bias, we made it clear that critical scrutiny would help us to identify potential weaknesses of the app prototype, which would in turn foster improvement. Lastly, we need to acknowledge that the characteristics of the wheelchair users with SCI who participated in this study may have had an impact on our findings. While we aimed for a balanced sample of wheelchair users in terms of age, gender, and lesion level, we cannot rule out the possibility that participating individuals may have had overall more favorable attitudes toward mHealth solutions than those who did not take part. While findings of this study may not reflect the views of the Swiss population with SCI, they provide the basis for further investigations at a population level.

Conclusion

Findings of this study indicate that even though co-design seems to be a promising approach to develop self-management tools that meet the needs of different stakeholder groups, it does not guarantee adoption. To fully harness the potential of a co-designed mHealth self-management solution and to overcome barriers to adoption, further efforts are needed. In particular, there is a need to foster trust in mHealth solutions among prospective users, including both patients and health care professionals. Moreover, increasing the personal relevance of mHealth solutions through personalization may be a promising approach to driving adoption. However, with data-driven approaches, such as machine learning, several ethical questions need to be considered, including but not limited to issues of autonomy, data protection, and privacy. Last but not least, organizational support and financial concepts also play a key role in mHealth adoption, particularly in the clinical context.

Acknowledgments

This project was funded by Suva, the Swiss Accident Insurance Fund.

The authors would like to thank Suva for providing funding for this project. We would also like to thank all study participants for their time, trust, and valuable contributions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the self-management app prototype.
[PDF File (Adobe PDF File), 398 KB - mhealth_v8i12e22452_app1.pdf]

References

1. Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2012 Aug 25;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
2. Yu JS, Kuhn E, Miller KE, Taylor K. Smartphone apps for insomnia: examining existing apps' usability and adherence to evidence-based principles for insomnia management. Transl Behav Med 2019 Jan 01;9(1):110-119. [doi: 10.1093/tbm/iby014] [Medline: 30590862]
3. Ramey L, Osborne C, Kasitinon D, Juengst S. Apps and Mobile Health Technology in Rehabilitation: The Good, the Bad, and the Unknown. Phys Med Rehabil Clin N Am 2019 May;30(2):485-497. [doi: 10.1016/j.pmr.2018.12.001] [Medline: 30954161]
4. 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. [doi: 10.3390/ijerph15030515] [Medline: 29538292]
5. Fiordelli M, Zanini C, Amann J, Scheel-Sailer A, Brach M, Stucki G, et al. Selecting Evidence-Based Content for Inclusion in Self-Management Apps for Pressure Injuries in Individuals With Spinal Cord Injury: Participatory Design Study. JMIR Mhealth Uhealth 2020 May 20;8(5):e15818 [FREE Full text] [doi: 10.2196/15818] [Medline: 32432559]
6. Byambasuren O, Beller E, Glasziou P. Current Knowledge and Adoption of Mobile Health Apps Among Australian General Practitioners: Survey Study. JMIR Mhealth Uhealth 2019 Jun 3;7(6):e13199 [FREE Full text] [doi: 10.2196/13199] [Medline: 31199343]
7. Tong HL, Coiera E, Tong W, Wang Y, Quiroz JC, Martin P, et al. Efficacy of a Mobile Social Networking Intervention in Promoting Physical Activity: Quasi-Experimental Study. JMIR Mhealth Uhealth 2019 Mar 28;7(3):e12181 [FREE Full text] [doi: 10.2196/12181] [Medline: 30920379]
8. Herrmann LK, Kim J. The fitness of apps: a theory-based examination of mobile fitness app usage over 5 months. Mhealth 2017 Jan 30;3:2-2. [doi: 10.21037/mhealth.2017.01.03] [Medline: 28293619]
9. Whittaker R. Issues in mHealth: Findings From Key Informant Interviews. J Med Internet Res 2012 Oct 02;14(5):e129 [FREE Full text] [doi: 10.2196/jmir.1989] [Medline: 23032424]
10. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3-283rv3. [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
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. Mortenson WB, Singh G, MacGillivray M, Sadeghi M, Mills P, Adams J, et al. Development of a Self-Management App for People with Spinal Cord Injury. J Med Syst 2019 Apr 22;43(6). [doi: 10.1007/s10916-019-1273-x] [Medline: 31011881]
13. MacGillivray MK, Sadeghi M, Mills PB, Adams J, Sawatzky BJ, Mortenson WB. Implementing a self-management mobile app for spinal cord injury during inpatient rehabilitation and following community discharge: A feasibility study. J Spinal Cord Med 2019 May 15;43(5):676-684. [doi: 10.1080/10790268.2019.1614343] [Medline: 31091160]
14. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A Holistic Framework to Improve the Uptake and Impact of eHealth Technologies. J Med Internet Res 2011 Dec 13;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
15. Revenäs Å, Opava CH, Martin C, Demmelmaier I, Keller C, Åsenlöf P. Development of a Web-Based and Mobile App to Support Physical Activity in Individuals With Rheumatoid Arthritis: Results From the Second Step of a Co-Design Process. JMIR Res Protoc 2015 Feb 09;4(1):e22 [FREE Full text] [doi: 10.2196/resprot.3795] [Medline: 25665589]
16. Crosby LE, Ware RE, Goldstein A, Walton A, Joffe NE, Vogel C, et al. Development and evaluation of iManage: A self-management app co-designed by adolescents with sickle cell disease. Pediatr Blood Cancer 2016 Aug 30;64(1):139-145. [doi: 10.1002/pbc.26177] [Medline: 27574031]

17. 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 [FREE Full text] [doi: 10.2196/mhealth.9856] [Medline: 30339132]
18. Jessen S, Mirkovic J, Ruland CM. Creating Gameful Design in mHealth: A Participatory Co-Design Approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
19. Peters D, Davis S, Calvo RA, Sawyer SM, Smith L, Foster JM. Young People's Preferences for an Asthma Self-Management App Highlight Psychological Needs: A Participatory Study. J Med Internet Res 2017 Apr 11;19(4):e113 [FREE Full text] [doi: 10.2196/jmir.6994] [Medline: 28400353]
20. Abelein U, Paech B. Understanding the Influence of User Participation and Involvement on System Success – a Systematic Mapping Study. Empir Softw Eng 2013 Dec 5;20(1):28-81. [doi: 10.1007/s10664-013-9278-4]
21. Tremblay M, Latulippe K, Giguere AM, Provencher V, Poulin V, Dubé V, et al. Requirements for an Electronic Health Tool to Support the Process of Help Seeking by Caregivers of Functionally Impaired Older Adults: Co-Design Approach. JMIR Aging 2019 Jun 07;2(1):e12327 [FREE Full text] [doi: 10.2196/12327] [Medline: 31518279]
22. Brienza D, Krishnan S, Karg P, Sowa G, Allegretti AL. Predictors of pressure ulcer incidence following traumatic spinal cord injury: a secondary analysis of a prospective longitudinal study. Spinal Cord 2017 Sep 12;56(1):28-34. [doi: 10.1038/sc.2017.96] [Medline: 28895575]
23. Gabbe BJ, Nunn A. Profile and costs of secondary conditions resulting in emergency department presentations and readmission to hospital following traumatic spinal cord injury. Injury 2016 Aug;47(8):1847-1855. [doi: 10.1016/j.injury.2016.06.012] [Medline: 27343134]
24. Carlson M, Vigen CLP, Rubayi S, Blanche EI, Blanchard J, Atkins M, et al. Lifestyle intervention for adults with spinal cord injury: Results of the USC-RLANRC Pressure Ulcer Prevention Study. J Spinal Cord Med 2019 Jan;42(1):2-19 [FREE Full text] [doi: 10.1080/10790268.2017.1313931] [Medline: 28414254]
25. Marin J, Nixon J, Gorecki C. A systematic review of risk factors for the development and recurrence of pressure ulcers in people with spinal cord injuries. Spinal Cord 2013 Apr 16;51(7):522-527. [doi: 10.1038/sc.2013.29] [Medline: 23588570]
26. Kruger EA, Pires M, Ngann Y, Sterling M, Rubayi S. Comprehensive management of pressure ulcers in spinal cord injury: Current concepts and future trends. J Spinal Cord Med 2013 Nov 19;36(6):572-585. [doi: 10.1179/2045772313y.0000000093] [Medline: 24090179]
27. Scheel-Sailer A, Wyss A, Boldt C, Post MW, Lay V. Prevalence, location, grade of pressure ulcers and association with specific patient characteristics in adult spinal cord injury patients during the hospital stay: a prospective cohort study. Spinal Cord 2013 Sep 3;51(11):828-833. [doi: 10.1038/sc.2013.91] [Medline: 23999106]
28. Amann J, Fiordelli M, Brach M, Bertschy S, Scheel-Sailer A, Rubinelli S. Co-designing a Self-Management App Prototype to Support People With Spinal Cord Injury in the Prevention of Pressure Injuries: Mixed Methods Study. JMIR Mhealth Uhealth 2020 Jul 9;8(7):e18018 [FREE Full text] [doi: 10.2196/18018] [Medline: 32673241]
29. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Swiss Paraplegic Group. Paraplegie Community. URL: http://www.community.paraplegie.ch [accessed 2020-05-11]
32. Shirai T, Bulandres P, Choi J, D'Ortenzio D, Moon N, Musselman K, et al. The use of a mobile educational tool on pressure injury education for individuals living with spinal cord injury/disease: a qualitative research study. Disability and Rehabilitation 2020 Jun 04:1-10. [doi: 10.1080/09638288.2020.1771780] [Medline: 32493089]
33. Hubli M, Zemp R, Albisser U, Camenzind F, Leonova O, Curt A, et al. Feedback improves compliance of pressure relief activities in wheelchair users with spinal cord injury. Spinal Cord 2020 Jul 21. [doi: 10.1038/s41393-020-0522-7] [Medline: 32694751]
34. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
35. Albrecht U. Transparency of Health-Apps for Trust and Decision Making. J Med Internet Res 2013 Dec 30;15(12):e277 [FREE Full text] [doi: 10.2196/jmir.2981] [Medline: 24449711]
36. van Haasteren A, Gille F, Fadda M, Vayena E. Development of the mHealth App Trustworthiness checklist. Digit Health 2019 Nov 21;5:205520761988646. [doi: 10.1177/2055207619886463] [Medline: 31803490]
37. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr 21;22(e1):e28-e33. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
38. O'Loughlin K, Neary M, Adkins EC, Schueller SM. Reviewing the data security and privacy policies of mobile apps for depression. Internet Interventions 2019 Mar;15:110-115. [doi: 10.1016/j.invent.2018.12.001] [Medline: 30792962]
39. Brandtzaeg PB, Pultier A, Moen GM. Losing Control to Data-Hungry Apps: A Mixed-Methods Approach to Mobile App Privacy. Social Science Computer Review 2018 May 31;37(4):466-488. [doi: 10.1177/0894439318777706]

40. Taki S, Russell CG, Lymer S, Laws R, Campbell K, Appleton J, et al. A Mixed Methods Study to Explore the Effects of Program Design Elements and Participant Characteristics on Parents' Engagement With an mHealth Program to Promote Healthy Infant Feeding: The Growing Healthy Program. Front. Endocrinol 2019 Jun 25;10. [doi: 10.3389/fendo.2019.00397] [Medline: 31293515]
41. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The Behavioral Intervention Technology Model: An Integrated Conceptual and Technological Framework for eHealth and mHealth Interventions. J Med Internet Res 2014 Jun 05;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
42. Rabbi M. Optimizing mHealth interventions with a bandit. In: Digital Phenotyping and Mobile Sensing. Cham: Springer; 2019:277-291.
43. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th international Conference on Persuasive Technology. 2009 Apr 26 Presented at: International Conference on Persuasive Technology; 2009; Claremont, California. [doi: 10.1145/1541948.1541999]
44. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
45. Zahia S, Garcia Zapirain MB, Sevillano X, González A, Kim PJ, Elmaghraby A. Pressure injury image analysis with machine learning techniques: A systematic review on previous and possible future methods. Artificial Intelligence in Medicine 2020 Jan;102:101742. [doi: 10.1016/j.artmed.2019.101742] [Medline: 31980110]
46. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. Electronic Commerce Research and Applications 2016 Mar;16:55-65. [doi: 10.1016/j.elerap.2015.11.001]
47. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a Modified Technology Acceptance Model to Evaluate Healthcare Professionals' Adoption of a New Telemonitoring System. Telemedicine and e-Health 2012 Jan;18(1):54-59. [doi: 10.1089/tmj.2011.0066] [Medline: 22082108]
48. Sezgin E, Özkan-Yildirim S, Yildirim S. Understanding the perception towards using mHealth applications in practice. Information Development 2016 Dec 19;34(2):182-200. [doi: 10.1177/0266666916684180]

Abbreviations

FG1: focus group 1
FG2: focus group 2
mHealth: mobile health
SCI: spinal cord injury

Edited by G Eysenbach, L Buis; submitted 14.07.20; peer-reviewed by J Yu, R Cruz Martínez; comments to author 01.08.20; revised version received 14.08.20; accepted 09.10.20; published 09.12.20.

Please cite as:
Amann J, Fiordelli M, Scheel-Sailer A, Brach M, Rubinelli S
Opportunities and Challenges of a Self-Management App to Support People With Spinal Cord Injury in the Prevention of Pressure Injuries: Qualitative Study
JMIR Mhealth Uhealth 2020;8(12):e22452
URL: http://mhealth.jmir.org/2020/12/e22452/
doi:10.2196/22452
PMID:33295876

Original Paper

The Utility of Predicting Hospitalizations Among Patients With Heart Failure Using mHealth: Observational Study

Susie Cartledge[1,2], BN (Hons), PhD; Ralph Maddison[2], BHSc, MSc (Hons), PhD; Sara Vogrin[3], MBBS, MBiostat; Roman Falls[3], BPharmSc; Odgerel Tumur[3], MBBS, PhD; Ingrid Hopper[4], MBBS, PhD; Christopher Neil[3,5], MBBS, PhD

[1]College of Nursing and Health Sciences, Flinders University, Adelaide, Australia

[2]Institution for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia

[3]Department of Medicine, Western Health, University of Melbourne, Melbourne, Australia

[4]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia

[5]Western Health Chronic Disease Alliance, Melbourne, Australia

Corresponding Author:
Susie Cartledge, BN (Hons), PhD
Institution for Physical Activity and Nutrition
School of Exercise and Nutrition Sciences
Deakin University
Locked Bag 20000
Geelong, 3220
Australia
Phone: 61 438263164
Email: susie.cartledge@deakin.edu.au

Abstract

Background: Heart failure decompensation is a major driver of hospitalizations and represents a significant burden to the health care system. Identifying those at greatest risk of admission can allow for targeted interventions to reduce this risk.

Objective: This paper aims to compare the predictive value of objective and subjective heart failure respiratory symptoms on imminent heart failure decompensation and subsequent hospitalization within a 30-day period.

Methods: A prospective observational pilot study was conducted. People living at home with heart failure were recruited from a single-center heart failure outpatient clinic. Objective (blood pressure, heart rate, weight, B-type natriuretic peptide) and subjective (4 heart failure respiratory symptoms scored for severity on a 5-point Likert scale) data were collected twice weekly for a 30-day period.

Results: A total of 29 participants (median age 79 years; 18/29, 62% men) completed the study. During the study period, 10 of the 29 participants (34%) were hospitalized as a result of heart failure. For objective data, only heart rate exhibited a between-group difference. However, it was nonsignificant for variability (P=.71). Subjective symptom scores provided better prediction. Specifically, the highest precision of heart failure hospitalization was observed when patients with heart failure experienced severe dyspnea, orthopnea, and bendopnea on any given day (area under the curve of 0.77; sensitivity of 83%; specificity of 73%).

Conclusions: The use of subjective respiratory symptom reporting on a 5-point Likert scale may facilitate a simple and low-cost method of predicting heart failure decompensation and imminent hospitalization. Serial collection of symptom data could be augmented using ecological momentary assessment of self-reported symptoms within a mobile health monitoring strategy for patients at high risk for heart failure decompensation.

(JMIR Mhealth Uhealth 2020;8(12):e18496) doi:10.2196/18496

KEYWORDS

cardiac failure; heart failure; readmission; hospitalization; risk prediction; mHealth

Introduction

Heart failure (HF) is a complex, chronic, and debilitating cardiac condition currently estimated to affect 38 million people internationally [1]. In Australia, 2014 prevalence estimates indicated that there were 480,000 adults living with HF, which represents 2.1% of the Australian population, with the prevalence forecast to significantly increase [2]. HF is caused by the inability of the heart to fill and eject sufficient blood to meet bodily demands, resulting in symptoms such as dyspnea, fatigue, and palpitations [3]. The exacerbation of HF symptoms, representing HF decompensation, is a major driver of hospitalization rates. HF hospitalizations represent a significant proportion of the total expenditure for HF in Australia annually [2]. Therefore, monitoring HF symptoms is essential in order to identify and prevent potential HF decompensation and subsequent hospitalization.

Collaboration between people living with HF and health care professionals (eg, heart failure nurses) is critical for monitoring HF symptoms and potential exacerbations [4]. While there are smartphone apps that focus on symptom monitoring, none currently provide risk prediction [4]. While HF mortality can be predicted with reasonable accuracy [5,6], risk prediction for HF hospitalization has demonstrated only modest performance in models reported to date. Variables used in HF risk predication models have included, in isolation or in combination [7], administrative data (such as Medicare claims data) [8], patient characteristics, clinical data, and geomapping [9]. However, the quantitation of self-reported subjective symptoms as an early indication of decline and therefore risk has on the whole been overlooked. Therefore, the aim of this study was to compare the predictive value of objectively and subjectively measured HF respiratory symptoms on imminent HF decompensation and subsequent hospitalization within a 30-day period.

Methods

Study Design, Setting, and Participants

We conducted a prospective observational pilot study with participants identified via cardiologist assessment as being at high risk of a HF hospitalization. Participants were recruited from a single-center HF outpatient clinic within a tertiary hospital in Melbourne, Australia. Eligible participants were older than 18 years, had a physician-documented HF diagnosis, had a previous hospital admission for HF exacerbation, were on maximum tolerated pharmacotherapy, and were able to read and understand English. Exclusion criteria included severe HF symptoms (New York Heart Association Class IV), advanced malignancy, cognitive impairment, and use of end-of-life care. This study was approved by the Western Health Human Research Ethics Committee (2016.071).

Measures

Participants were visited twice weekly by a research assistant (one a biomedical science graduate the other a medical doctor) for a 30-day period from June 2016 to May 2017. The research assistants collected all measurements, including the subjective respiratory scores, from patients. Study data and sources are described in Table 1.

Table 1. Variables collected during study period and associated data sources.

Variables	Data source
Objective measures	
Blood pressure	Validated study sphygmomanometer
Heart rate	Manual pulse
Weight	Validated study scales
B-type natriuretic peptide	Point-of-care testing
Subjective measures	
Dyspnea	5-point Likert scale[a]
Bendopnea	5-point Likert scale[a]
Orthopnea	5-point Likert scale[a]
Paroxysmal nocturnal dyspnea	5-point Likert scale[a]
Other variables	
Demographics	Medical record and participant survey
Medical history	Medical record
Hospitalization status	Medical record

[a]1 indicates no symptoms and 5 indicates severe symptoms.

Subjective symptoms of dyspnea, orthopnea, bendopnea, and paroxysmal nocturnal dyspnea (PND) were chosen because they are routinely used indicators of clinical status in HF, each of which feature in 2 key diagnosis criteria [10,11]. Likert scales to instantaneously quantify dyspnea in HF populations have been researched using 7- and 5-point scales [12,13]. Additionally, the 5-point Likert scale–quantified dyspnea has

previously demonstrated a relationship with subsequent emergency readmission [13].

Statistical Analysis

Given that this is a pilot study, no formal power calculations were undertaken. Baseline characteristics are presented as median (interquartile range) or frequency (percentage) and are compared between hospitalized and nonhospitalized patients using a rank sum test and Fisher exact test.

To assess whether objective measures (blood pressure, heart rate [HR], variation in HR, weight, and B-type natriuretic peptide [BNP]) were associated with hospitalization, their mean value, standard deviation, and slope of change were calculated over 7 days prior to hospitalization (for admitted patients) and over the whole observation period for others (with at least 7 days' clearance before and after any hospitalization). These were then compared using a rank sum test. The same technique was applied using a symptom severity score, and their variability was compared by calculating a median score and range of scores.

To determine the optimal severity cutoff value on the Likert scale for each respective symptom, a random day within the 7 days prior to hospitalization (for hospitalized patients) and any random day for others was chosen. The area under the receiver operating characteristics curve (AUC) and the Youden index (YI) were calculated. This process was repeated 1000 times with different combinations of random days. The cutoff value with the most frequent highest AUC and Youden index was chosen. This cutoff was then used for all analyses.

To determine which combination of symptoms best predicted hospitalization, we calculated AUC, YI, sensitivity, and specificity and compared them among all combinations of symptoms (eg, bendopnea and orthopnea, bendopnea and dyspnea, bendopnea and dyspnea and orthopnea—a total of 10 possible combinations) on 1000 combinations of randomly chosen days (as described above). In the next step, 2 random consecutive symptom measurements (usually 2 to 3 days apart) were chosen, and we determined the ability to predict hospitalization if the symptom was severe on either day or both days or if the symptom severity increased across the 2 days. This was performed separately for each symptom and for all 10 combinations of symptoms (as described above). Sensitivity, specificity, AUC, and YI were calculated. All analyses were performed using Stata 15.1 (StataCorp).

Results

Baseline Population Characteristics

A total of 30 participants met the study inclusion criteria and provided written informed consent; however, one participant withdrew from the study shortly after enrollment. During the study, 10 of the 29 participants (34%) were hospitalized as a result of decompensated HF, as adjudicated by the Boston criteria on file review (the comparator group). Another participant was admitted for infection without congestion.

As demonstrated in Table 2, the participants who were hospitalized had higher left ventricular ejection fractions but worse baseline Minnesota Living With HF scores and were on less angiotensin- converting enzyme inhibitors and angiotensin II receptor blocker therapies.

Table 2. Baseline demographics and comparison by heart failure hospitalization.

Characteristic	All participants (N=29)	Not hospitalized (19/29, 66%)	Hospitalized (10/29, 34%)	*P* value
Age (years), median (IQR)	79 (69-84)	74 (68-82)	81.5 (73-89)	.12
Sex, n (%)				**.69**
Male	18 (62)	11 (58)	7 (70)	
Female	11 (38)	8 (42)	3 (30)	
Ethnic group, n (%)				**.56**
Non-Indigenous Australian	7 (24)	6 (32)	1 (10)	
European	8 (28)	5 (26)	3 (30)	
Indian	1 (3)	1 (5)	0 (0)	
Pacific Islander	5 (17)	2 (11)	3 (30)	
Other	8 (28)	5 (26)	3 (30)	
Born in Australia, n (%)	11 (39)	9 (47)	2 (22)	.25
Education, n (%)				**.12**
Some high school	19 (66)	15 (79)	4 (40)	
Trade certificate	8 (28)	4 (21)	4 (40)	
Some university	2 (7)	0 (0)	2 (20)	
Work status, n (%)				**.69**
Part-time	1 (3)	1 (5)	0 (0)	
Full-time	2 (7)	2 (11)	0 (0)	
Retired	26 (90)	16 (84)	10 (100)	
Household income (Aus $)[a], n (%)				**.40**
<20,000	18 (62)	10 (53)	8 (80)	
20,000-30,000	10 (34)	8 (42)	2 (20)	
30,000-40,000	1 (3)	1 (5)	0 (0)	
Years since HF[b] diagnosis, n (%)				**.42**
≤5 years	18 (62)	13 (68)	5 (50)	
5-10 years	4 (14)	3 (16)	1 (10)	
10-20 years	5 (17)	2 (11)	3 (30)	
>20 years	2 (7)	1 (5)	1 (10)	
HFrEF[c], n (%)	16 (55)	12 (63)	4 (40)	.27
Baseline NYHA[d], n (%)				**.09**
NYHA II	10 (34)	9 (47)	1 (10)	
NYHA III	19 (66)	10 (53)	9 (90)	
LVEF[e], median (IQR)	35 (28-50)	30 (22-46)	50 (35-57)	.02
Baseline Minnesota Living With HF score[f], median (IQR)	58 (48-70)	53 (40-62)	72.5 (54-81)	.02
HF hospitalization in previous 6 months, n (%)	25 (86)	17 (89)	8 (80)	.59
Emergency department attendance in previous 12 months, n (%)	21 (75)	13 (68)	8 (89)	.37
Body mass index, median (IQR)	30.42 (25.65-38.72)	30.80 (25.65-39.82)	30.27 (25.13-32.25)	.65
Medical history, n (%)				
Hypertension	27 (93)	17 (89)	10 (100)	.53

Characteristic	All participants (N=29)	Not hospitalized (19/29, 66%)	Hospitalized (10/29, 34%)	*P* value
Ischemic heart disease	16 (55)	10 (53)	6 (60)	>.99
Hypercholesterolemia	22 (76)	16 (84)	6 (60)	.19
Atrial fibrillation	19 (66)	11 (58)	8 (80)	.41
Cerebrovascular accident	7 (24)	6 (32)	1 (10)	.37
Diabetes mellitus	17 (59)	10 (53)	7 (70)	.45
COPD[g] or asthma	11 (38)	6 (32)	5 (50)	.43
Use of therapies				
Loop diuretic, n (%)	27 (93)	17 (89)	10 (100)	.53
Daily loop diuretic dose, median (IQR)	80 (80-120)	80 (40-80)	100 (80-60)	.09
β blocker, n (%)	26 (90)	17 (89)	9 (90)	>.99
ACE-I[h] or ARB[i], n (%)	22 (76)	17 (89)	5 (50)	.03
Aldosterone antagonist, n (%)	12 (41)	8 (42)	4 (40)	>.99

[a]At the time of publication, a currency exchange rate of Aus $1=US $0.74 was applicable.

[b]HF: heart failure.

[c]HFrEF: heart failure with reduced ejection fraction.

[d]NYHA: New York Heart Association.

[e]LVEF: left ventricular ejection fraction.

[f]The Minnesota Living With Heart Failure scoring range is from 0 to 105, with higher scores indicating poorer health-related quality of life.

[g]COPD: chronic obstructive pulmonary disease.

[h]ACE-I: angiotensin-converting enzyme inhibitor.

[i]ARB: angiotensin II receptor blocker.

Objective Measures

Of the 4 objective measures collected, only HR demonstrated a significant difference between groups (median 80 vs 67 beats per minute [bpm]; *P*=.02) (Table 3). Variation in HR over the study period, however, was nonsignificant between groups (7 vs 6 bpm; *P*=.71). Participants admitted to the hospital demonstrated higher BN*P* values (median 1113 vs 546 pg/mL; *P*=.09), with a higher average daily increase prior to hospitalization (20 vs 0.05 pg/mL/d; *P*=.08).

Table 3. Objective and subjective measures.

Measures	Not hospitalized, median (IQR), (n=19)	Hospitalized, median (IQR), (n=10)	*P* value
Objective measures			
Weight			
Average (kg)	87 (73.8 to 95.6)	84.5 (64 to 96.1)	.52
Variability (standard deviation) (kg)	0.97 (0.63 to 1.30)	1.07 (0.61 to 2.86)	.57
Slope (kg/d)	0.00 (–0.04 to 0.02)	0.13 (–0.15 to 0.33)	.26
Systolic blood pressure			
Average (mmHg)	121.8 (108 to 128.7)	123.2 (113.8 to 131.7)	.65
Variability (standard deviation) (mmHg)	12.8 (8.3 to 15.4)	15.4 (11.1 to 17.6)	.29
Slope (mmHg/d)	0.25 (–0.37 to 0.64)	1.06 (–0.82 to 10.5)	.31
Heart rate			
Average (bpm[a])	67.4 (62.0 to 76.2)	80.4 (74.3 to 88.0)	.02
Variability (standard deviation) (bpm)	7.5 (4.3 to 10.0)	7.2 (3.6 to 14.4)	.71
Slope (bpm/d)	–0.04 (–0.17 to 0.36)	–0.61 (–1.59 to 0.87)	.38
B-type natriuretic peptide			
Average (pg/mL)	545.6 (237.4 to 891.7)	1112.5 (554.5 to 1565)	.09
Variability (standard deviation) (pg/mL)	99.9 (47.5 to 231.9)	164.4 (75.2 to 186.71)	.83
Slope (pg/mL/d)	0.05 (–4.85 to 5.81)	19.65 (4 to 41.45)	.08
Subjective measures[b]			
Dyspnea			
Median symptom scores	2 (1 to 3)	4 (3 to 4)	.009
Ranges in symptom score	2 (1 to 3)	0 (0 to 1)	.003
Orthopnea			
Median symptom scores	1 (1 to 2)	3.3 (2 to 4)	.007
Ranges in symptom score	1 (0 to 2)	0.5 (0 to 1)	.22
Bendopnea			
Median symptom scores	2 (1 to 3)	4 (3.5 to 5)	.007
Ranges in symptom score	2 (1 to 3)	0 (0 to 1)	.006
Paroxysmal nocturnal dyspnea			
Median symptom scores	1 (1 to 2)	3 (1.5 to 4)	.009
Ranges in symptom score	1 (1 to 3)	0 (0 to 1)	.004

[a]bpm: beats per minute.

[b]For subjective measures, the median is the median and IQR of the participants' median scores. Range is the median and IQR of participants' range of scores; for example, if the range was 1, that patient had a symptom of 3 and 4 only.

Subjective Measures

Symptom scores were examined, with hospitalized patients reporting higher median dyspnea, bendopnea, and PND with a lower 7-day range, implying consistently worse symptoms (all *P*<.01) (Table 3). Figure 1 shows the symptom measurements experienced at each severity level of the Likert scale. Orthopnea was worse in the HF hospitalization group (*P*=.01) but had similar variability (*P*=.22). A symptom score of at least 3 for dyspnea, 2 for orthopnea, and 4 for PND and bendopnea produced the highest AUC and YI for predicting HF hospitalization. Figure 2 shows the areas under the curve for respiratory symptoms.

Figure 1. Subjective symptom measurements experienced at each severity level of the Likert scale, showing the weighting of severity in the hospitalized versus nonhospitalized group.

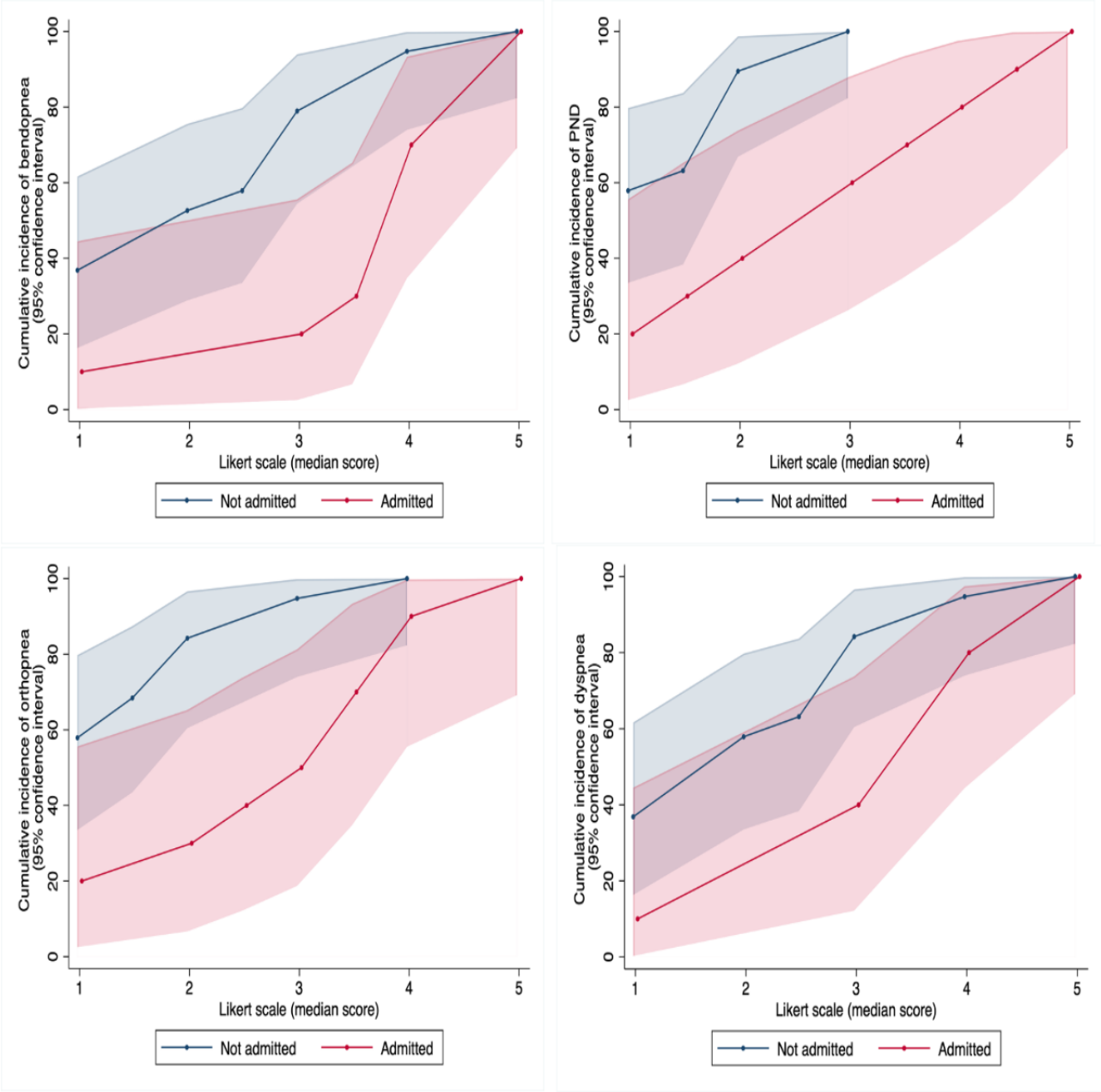
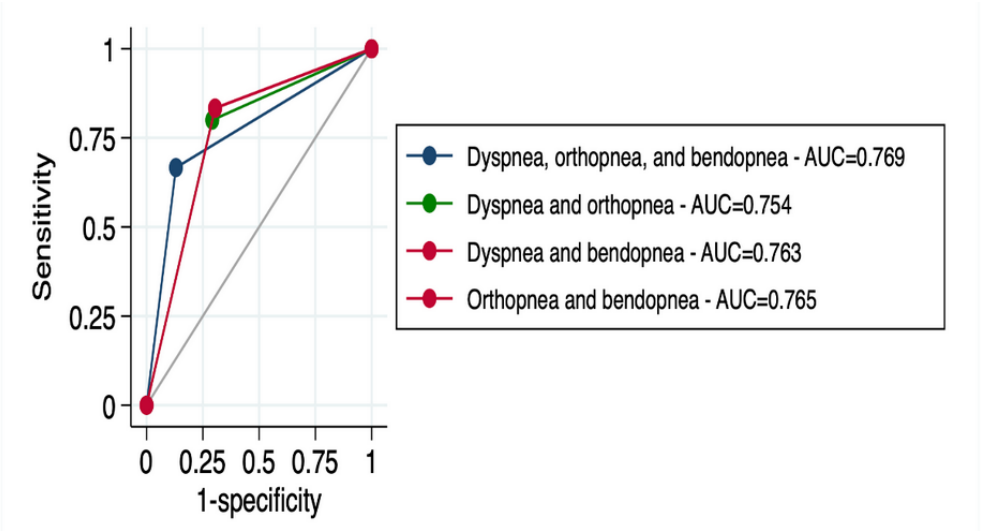


Figure 2. Areas under the curve for subjective respiratory symptoms. AUC: area under the curve.



The highest YI (0.54) and AUC (0.77) were observed when on any given day dyspnea, orthopnea, and bendopnea symptoms were severe. This combination predicted hospitalization with a sensitivity of 81% and a specificity of 73%. Similar results were observed when at least orthopnea and bendopnea were severe (YI=0.53; AUC=0.76; sensitivity of 81%; specificity of 73%). Higher sensitivities but lower specificities were observed when dyspnea and bendopnea (84% and 68%, respectively) or dyspnea and orthopnea (87% and 64%, respectively) were severe (see Multimedia Appendix 1 for graphical representation). All other combinations resulted in a YI <0.50 and an AUC <0.75.

Discussion

Principal Findings

In this analysis of the prediction of decompensation and HF hospitalization within 30 days, we demonstrated that compared with objective measures, a simplified system for quantifying respiratory symptom status may be an accurate and useful predictor. Specifically, the highest precision of HF hospitalization was observed when patients with HF experienced severe dyspnea, orthopnea, and bendopnea on any given day (AUC=0.77; sensitivity of 83%; specificity of 73%). Early detection of deterioration would allow the care team to provide agile HF care that may be able to prevent subsequent hospital admission.

The lack of a sound risk prediction tool for imminent HF hospitalization makes organization and prioritization of HF care challenging [5]. Current Australian HF guidelines call for systems of care with an “alert system to flag patients who are displaying signs of clinical deterioration and pathways for rapid medical review” [3]. Given that subjective data outperformed objective data in this pilot cohort, there are opportunities for patients to be able to regularly log respiratory symptoms. This eliminates the need for patients to regularly use medical equipment (eg, sphygmomanometer, scales) to collect prediction data.

Real-time regular data collection of a participant’s state, such as HF symptoms, can be conducted through ecological momentary assessment (EMA) [14]. This method allows a picture to be formed of a participant’s symptoms, reducing recall bias. EMA data are best collected by electronic means, such as mobile and wireless devices (mobile health [mHealth]), to ensure easy, timely, and compliant documentation [14,15].

The ease of collection of EMA self-reported respiratory symptom data via mHealth could lead to large data sets for analysis. Artificial intelligence and machine learning are increasingly being used to provide clinically meaningful predictive data analysis, especially with large data sets [16]. This technique can be applied to build automated clinical decision systems for problems such as hospitalization risk [16].

Limitations

The main limitation of this pilot study is its sample size and length of follow-up. The small sample size may affect the variability of some measures (eg, weight), and this should be addressed in future larger trials over a longer period. Future studies could establish baseline respiratory symptoms for patients and examine the timeline of changes in the severity of symptoms and of hospitalization. Additionally, there was selection bias, as patients were recruited from a single-center HF clinic, and this may not represent a typical HF population.

However, unlike previous studies using patients drawn from clinical trials [17], our patients were recruited from standard practice, which may increase the generalizability of the study. Another strength is the use of primary data collected during the study rather than secondary data, such as trial databases, registries, and administrative claims data, which have been used in other studies.

Conclusions

The use of patient-reported serial quantification of 4 key respiratory HF symptoms (dyspnea, bendopnea, orthopnea, and PND) may provide low-cost detection of imminent decompensation and therefore potential hospitalization. Future research should focus on testing and validating this model with a larger sample, augmenting the findings using an EMA of self-reported HF symptoms via mHealth, and using artificial intelligence data analysis techniques to increase risk prediction accuracy.

Acknowledgments

We would like to acknowledge the work of Tarveen Singh, Andrew Mulligan, and Abduwaysi Amie, who assisted with the data collection for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[DOCX File , 401 KB - mhealth_v8i12e18496_app1.docx]

References

1. Vos T, Flaxman A, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2163-2196 [FREE Full text] [doi: 10.1016/S0140-6736(12)61729-2] [Medline: 23245607]
2. Chan Y, Tuttle C, Ball J, Teng TK, Ahamed Y, Carrington MJ, et al. Current and projected burden of heart failure in the Australian adult population: a substantive but still ill-defined major health issue. BMC Health Serv Res 2016 Sep 21;16(1):501 [FREE Full text] [doi: 10.1186/s12913-016-1748-0] [Medline: 27654659]
3. NHFA CSANZ Heart Failure Guidelines Working Group, Atherton JJ, Sindone A, De Pasquale CG, Driscoll A, MacDonald PS, et al. National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand: Guidelines for the Prevention, Detection, and Management of Heart Failure in Australia 2018. Heart Lung Circ 2018 Oct;27(10):1123-1208 [FREE Full text] [doi: 10.1016/j.hlc.2018.06.1042] [Medline: 30077227]
4. 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]
5. Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk prediction in patients with heart failure: a systematic review and analysis. JACC Heart Fail 2014 Oct;2(5):440-446 [FREE Full text] [doi: 10.1016/j.jchf.2014.04.008] [Medline: 25194291]
6. Levy WC, Mozaffarian D, Linker DT, Sutradhar SC, Anker SD, Cropp AB, et al. The Seattle Heart Failure Model. Circulation 2006 Mar 21;113(11):1424-1433. [doi: 10.1161/circulationaha.105.584102] [Medline: 16534009]
7. Ross JS, Mulvey GK, Stauffer B, Patlolla V, Bernheim SM, Keenan PS, et al. Statistical models and patient predictors of readmission for heart failure: a systematic review. Arch Intern Med 2008 Jul 14;168(13):1371-1386. [doi: 10.1001/archinte.168.13.1371] [Medline: 18625917]
8. Keenan PS, Normand ST, Lin Z, Drye EE, Bhat KR, Ross JS, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. Circ Cardiovasc Qual Outcomes 2008 Sep;1(1):29-37. [doi: 10.1161/CIRCOUTCOMES.108.802686] [Medline: 20031785]
9. Huynh QL, Saito M, Blizzard CL, Eskandari M, Johnson B, Adabi G, MARATHON Investigators. Roles of nonclinical and clinical data in prediction of 30-day rehospitalization or death among heart failure patients. J Card Fail 2015 May;21(5):374-381. [doi: 10.1016/j.cardfail.2015.02.002] [Medline: 25724302]
10. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. N Engl J Med 1971 Dec 23;285(26):1441-1446. [doi: 10.1056/NEJM197112232852601] [Medline: 5122894]
11. Carlson KJ, Lee DC, Goroll AH, Leahy M, Johnson RA. An analysis of physicians' reasons for prescribing long-term digitalis therapy in outpatients. J Chronic Dis 1985;38(9):733-739. [doi: 10.1016/0021-9681(85)90115-8] [Medline: 4030999]
12. Solomonica A, Burger AJ, Aronson D. Hemodynamic Determinants of Dyspnea Improvement in Acute Decompensated Heart Failure. Circ Heart Fail 2013 Jan;6(1):53-60. [doi: 10.1161/circheartfailure.112.970335] [Medline: 23152491]
13. Weber C, Miglioranza M, Moraes M, Sant'anna RT, Rover M, Kalil R, et al. The five-point Likert scale for dyspnea can properly assess the degree of pulmonary congestion and predict adverse events in heart failure outpatients. Clinics (Sao Paulo) 2014;69(5):341-346 [FREE Full text] [doi: 10.6061/clinics/2014(05)08] [Medline: 24838900]
14. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
15. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012 May;17(Suppl 1):11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]

16. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. J Am Coll Cardiol 2017 May 30;69(21):2657-2664 [FREE Full text] [doi: 10.1016/j.jacc.2017.03.571] [Medline: 28545640]
17. Vader JM, LaRue SJ, Stevens SR, Mentz RJ, DeVore AD, Lala A, et al. Timing and Causes of Readmission After Acute Heart Failure Hospitalization-Insights From the Heart Failure Network Trials. J Card Fail 2016 Nov;22(11):875-883 [FREE Full text] [doi: 10.1016/j.cardfail.2016.04.014] [Medline: 27133201]

Abbreviations

AUC: area under the receiver operating characteristics curve
BNP: B-type natriuretic peptide
bpm: beats per minute
EMA: ecological momentary assessment
HF: heart failure
HR: heart rate
mHealth: mobile health
PND: paroxysmal nocturnal dyspnea
YI: Youden index

Edited by G Strudwick; submitted 11.05.20; peer-reviewed by C El Morr, N Lowres; comments to author 02.06.20; revised version received 28.07.20; accepted 08.09.20; published 22.12.20.

Please cite as:
Cartledge S, Maddison R, Vogrin S, Falls R, Tumur O, Hopper I, Neil C
The Utility of Predicting Hospitalizations Among Patients With Heart Failure Using mHealth: Observational Study
JMIR Mhealth Uhealth 2020;8(12):e18496
URL: http://mhealth.jmir.org/2020/12/e18496/
doi:10.2196/18496
PMID:33350962

Original Paper

Feasibility of a Digital Health Intervention to Improve Diet Quality Among Women With High Blood Pressure: Randomized Controlled Feasibility Trial

Dori M Steinberg[1,2], RD, MS, PhD; Melissa C Kay[2,3], RD, MPH, MS, PhD; Laura P Svetkey[4,5], MHS, MD; Sandy Askew[2], MPH; Jacob Christy[6], MSc; Jasmine Burroughs[2], RD, LDN, MSc; Hira Ahmed[2], MA; Gary G Bennett[2,7], PhD

[1]Duke University School of Nursing, Durham, NC, United States

[2]Duke Global Digital Health Science Center, Duke Global Health Institute, Durham, NC, United States

[3]Duke Center for Childhood Obesity Research, Department of Pediatrics, Duke University Medical Center, Durham, NC, United States

[4]Sarah W Stedman Nutrition and Metabolism Center, Duke University Medical Center, Durham, NC, United States

[5]Division of Nephrology, Department of Medicine, Duke University Medical Center, Durham, NC, United States

[6]Department of Family Medicine and Community Health, Duke University, Durham, NC, United States

[7]Department of Psychology and Neuroscience, Duke University, Durham, NC, United States

Corresponding Author:
Dori M Steinberg, RD, MS, PhD
Duke University School of Nursing
307 Trent Drive
DUMC 3322
Durham, NC, 27710
United States
Phone: 1 919 613 9439
Email: dori.steinberg@duke.edu

Abstract

Background: Over 100 million individuals have high blood pressure, and more than half of them are women. The Dietary Approaches to Stop Hypertension (DASH) dietary pattern is a proven lifestyle approach to lower blood pressure, yet population-level adherence is poor. Innovative strategies that promote DASH are needed.

Objective: This paper aims to improve adherence to the DASH diet among women with hypertension or prehypertension.

Methods: We conducted a 3-month randomized controlled feasibility trial comparing app-based diet tracking (active comparator) to app-based diet tracking plus feedback on DASH adherence via text message (intervention). The intervention platform extracted nutrient data from the app, compared it to DASH recommendations, and sent tailored feedback text messages. Outcomes included the number of days participants tracked their diet, changes in their DASH adherence score, and blood pressure.

Results: The women (N=59) had a mean age of 49.9 (SD 11.9) years and were primarily non-Hispanic White (41/59, 69%) and college educated (49/59, 83%). The mean baseline DASH score was 2.3 (SD 1.3). At 3 months, the intervention and active comparator participants had similar mean days tracked per week (4.2, SD 2.1 days vs 4.6, SD 2.7 days; *P*=.54) and mean changes in their DASH score (0.8, 95% CI 0.2-1.5 vs 0.8, 95% CI 0.4-1.2; *P*=.75). Intervention participants had lower systolic (mean difference: –2.8 mmHg, 95% CI –1.8 to 7.4; *P*=.23) and diastolic (mean difference: –3.6 mmHg, 95% CI –0.2 to 7.3; *P*=.07) blood pressure compared with active comparator participants. Most intervention participants (23/29, 79%) said they would recommend the DASH Cloud intervention to a friend or family member. However, only 34% (10/59) indicated that the feedback text messages helped them reach their diet goals.

Conclusions: A digital health intervention to improve DASH adherence is feasible and produces moderately high engagement among women with elevated blood pressure. The intervention did not enhance DASH adherence over diet tracking alone but resulted in greater reductions in blood pressure. Larger studies are needed to determine how digital health interventions can improve population-level adherence to DASH.

Trial Registration: ClinicalTrials.gov NCT03215472; https://clinicaltrials.gov/ct2/show/NCT03215472

(JMIR Mhealth Uhealth 2020;8(12):e17536) doi:10.2196/17536

KEYWORDS

hypertension; DASH dietary pattern; digital health; nutrition; women's health; mHealth

Introduction

More than 100 million Americans have high blood pressure [1], the primary risk factor for heart disease and stroke, which are two of the leading causes of death in the United States [2]. In 2017, the American Heart Association implemented new guidelines for the detection, prevention, and treatment of high blood pressure [3]. These guidelines lowered the thresholds for what is considered an optimal blood pressure, which resulted in about 30 million more Americans requiring treatment for high blood pressure, including 13 million more women [1,3,4]. Many of these women, however, present with atypical risk factors for high blood pressure and may not receive the treatment they need [5]. Participation of women in blood pressure trials is disproportionately lower, and there is a need for a lifestyle treatment that is tailored for women [6]. Lifestyle treatment is indicated for all individuals with high blood pressure regardless of whether medication is also indicated, but it is the first line of treatment for those in the elevated blood pressure category [3].

The Dietary Approaches to Stop Hypertension (DASH) dietary pattern is an evidence-based lifestyle treatment to lower blood pressure [7]. DASH emphasizes nutrient-rich foods, such as fruits and vegetables, whole grains, and low-fat dairy products, and limits red meat and processed foods [8]. DASH was specifically designed to increase nutrient consumption of calcium, magnesium, potassium, fiber, and protein and reduce the intake of saturated fat, total fat, and cholesterol [8]. The initial DASH trial, which compared a typical US diet to the DASH dietary pattern among adults with nonmedicated high blood pressure, resulted in a decrease of systolic and diastolic blood pressure by 5.5 and 3.0 mmHg, respectively [7]. The blood pressure–lowering effects of DASH have been replicated in multiple trials [9-11].

As such, DASH is recommended in national blood pressure [3,12] and dietary [13] guidelines. However, the adoption of DASH on a population level remains poor [14]. Fewer than 1% of US adults are fully adherent to DASH, and only 20% meet half of the recommendations of DASH [9,15]. This indicates a need for improved dissemination and support to increase population uptake and adherence to DASH.

Smartphones are a promising channel for disseminating and supporting DASH adherence. More than three-quarters of Americans have smartphones, with high rates of ownership across all socioeconomic strata [16]. Diet and fitness smartphone apps account for 43% [16,17] of mobile medical apps, with the most popular apps having 50 million users [18]. These publicly available apps are primarily diet trackers that ask users to enter all foods and beverages consumed and then link those data with nutritional databases. Although popular, these apps lack evidence-based components known to produce behavior change, such as consistent self-monitoring, goal setting, problem-solving skills, feedback, and social support [19,20]. To address this, we can create programs that combine these publicly available apps with evidence-based behavior change principles to improve dissemination of DASH.

In the current study, we created the DASH Cloud intervention. Using data collected through a commercial diet-tracking smartphone app, DASH Cloud provides tailored feedback via text messages about adherence to DASH, accompanied by motivational messages designed to support behavior change. We tested the feasibility of the intervention among women with high blood pressure using a randomized controlled study design. The 3-month trial compared the DASH Cloud intervention to diet tracking alone on adherence to DASH. Using the principles of feasibility outlined by Leon et al [21], feasibility was determined based on successful recruitment and retention and the successful implementation of the intervention. We hypothesized that women receiving the 3-month DASH Cloud intervention would have (1) greater rates of diet tracking and (2) greater adherence to DASH.

Methods

Study Design and Eligibility

We recruited women aged 21 to 70 years with a BMI of >18.5 kg/m^2. Women were eligible if they self-reported being diagnosed with hypertension, using medication for blood pressure, or having a recent systolic measurement of 120 to 159 mmHg or a diastolic blood pressure measurement of 80 to 99 mmHg. They were required to have a smartphone with the latest operating system or be willing to update to the current version, have a data plan, and be willing to receive daily text messages. Having an email address and fluency in English were also required. Participants were excluded if they had a cardiovascular event in the last 6 months or a current cancer diagnosis, had been institutionalized for a psychiatric disorder within the past year, or if they were pregnant, lactating, or enrolled in another dietary change study.

Eligible participants were redirected to an online consent form to review and sign, which included both a brief video that described the study and a written version. Participants then completed online surveys and attended baseline study visits at our study offices. During the first baseline visit, participants' height, weight, blood pressure, and depression status were measured.

At the first baseline visit, participants downloaded and were instructed on how to use the Nutritionix diet-tracking app (Syndigo LLC) and asked to track their diet for a week to assess their ability to track using the study app. Participants came back for a second baseline visit 1 week later if they tracked a minimum of 4 consecutive days. During the second baseline visit, participants were randomized using a permuted block design with random block sizes ranging from 4 to 8. An

allocation table was created by a biostatistician using Sealed Envelope and uploaded to a Research Electronic Data Capture (REDCap) project. Allocation was revealed to unblinded research staff upon random assignment by the REDCap software, allowing the research staff to orient the participant to either the DASH Cloud intervention or an active comparator arm. Participants had a follow-up visit after 3 months. Participants were compensated US $25 at the second baseline visit and the 3-month follow-up visit. All study protocols were in accordance with the ethical standards of Duke University and were approved by the Duke University Health Center institutional review board. The study was registered on ClinicalTrials.gov on July 12, 2017 (NCT03215472).

Recruitment

Participants were recruited using a multipronged approach, including (1) flyers distributed to gyms, community centers, grocery stores, and health and wellness clinics throughout Raleigh, Durham, and Chapel Hill, North Carolina; (2) study details posted on Research Match, a national clinical trials registry that matches participants to studies; (3) social media posts on Twitter, Facebook, and Nextdoor; and (4) contact with participants who were ineligible for other studies but showed interest in similar behavioral studies.

Intervention Description

The intervention included (1) daily diet tracking using the Nutritionix app; (2) daily or weekly text messages that included tailored feedback on adherence to DASH, motivational messages to boost adherence, and tips for specific dietary changes; (3) animated videos designed to increase skills around different topics related to DASH; and (4) the DASH booklet available from the National Heart, Lung, and Blood Institute (NHLBI) [22].

Daily Diet Tracking

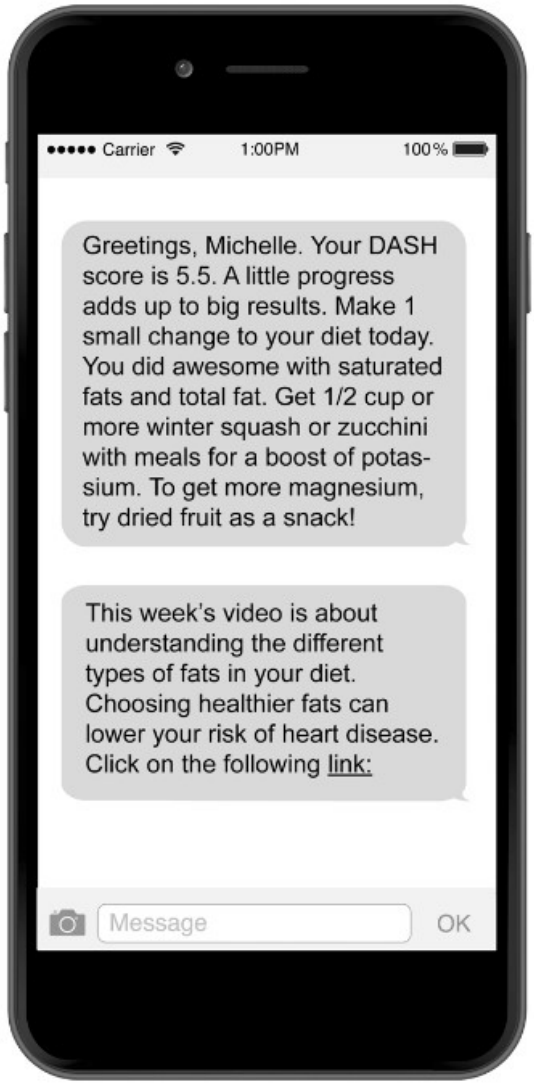
All participants were asked to use Nutritionix daily and enter all foods and beverages consumed. Like many commercial diet apps, Nutritionix primarily provides tracking capabilities. The app interface includes options to enter foods by meal and time of day and includes small pictures of each food. Nutritionix also has a comprehensive and easily accessible nutrient database supporting entries into the app. This database includes all data available from the United States Department of Agriculture (USDA) Food Composition Database, data from restaurant chains, and foods added by registered dietitians staffed by Nutritionix. Dietary data were automatically uploaded from the Nutritionix app daily via an application programing interface (API). The API connects the Nutritionix server to the DASH Cloud intervention delivery system, which uses servers stored in a platform operated out of the Duke Global Digital Health Science Center called Prompt. As is common with nutrient databases [23], not all foods available in the Nutritionix database had complete nutrient data. For example, magnesium and potassium were nutrients we used in our DASH algorithm that were often not available for many foods. As such, to ensure we had complete nutrient data, our registered dietitians on staff manually filled in any missing nutrient data that were received via the Nutritionix API.

Feedback Texts

Using an automated algorithm, the intervention platform sent daily feedback via text message at 12:00 PM EST for the first 2 weeks and weekly feedback via text message for the remainder of the study (Figure 1). We opted to send the texts daily during the first 2 weeks to provide feedback proximal to the specific food choice. After 2 weeks, we switched to giving feedback weekly to reduce the potential burden from texting daily. This decision was based on balancing the concept of proximal feedback with the perceived burden from frequent notifications. The feedback message began with a personalized greeting (eg, “Hello, [name]”) and the participant’s DASH diet adherence score from the previous day (eg, “Your DASH score yesterday was 5.5”). This score was calculated using the Mellen et al [15] index to evaluate DASH adherence based on nutrient targets. The score uses a 9-point scale based on the previous day’s intake for potassium, sodium, magnesium, calcium, saturated fat, total fat, total protein, cholesterol, and fiber. Our algorithm compared the total intake and the recommended targets for these nutrients in DASH. Similar to the Mellen et al [15] index, our software platform was coded to automatically apply a score of 0, 0.5, or 1 to each nutrient based on its difference from the recommended target. The texts included feedback on which nutrients met the target and tips for improving nutrient intake for those that were suboptimal (eg, “You did best with reducing saturated fat and boosting your fiber intake and seemed to struggle with getting enough potassium and magnesium”).

Feedback also included motivational messages regarding changes in the overall DASH score and prompted participants to reflect on which dietary choices most impacted their DASH score. Each week on Sunday afternoons, we also sent a separate text message about a different topic related to DASH (eg, “Dining out on DASH,” with links to animated educational videos). Overall, these text messages were intended to be motivational by offering behavioral tips to reinforce dietary change and providing social support for women at risk for hypertension and cardiovascular disease. We aimed to tailor the intervention content for women and, in particular, improve understanding of the risks specific to women regarding cardiovascular disease. Participants were not expected to respond, and an automatic reply was sent if participants did respond indicating that they could reach out for support to our staff via email.

Figure 1. Sample text message sent to participants in the DASH Cloud intervention. DASH: Dietary Approaches to Stop Hypertension.



Active Comparator Arm

Participants in the active comparator arm were asked to track their diet daily using the Nutritionix app and were shown a video that introduced the DASH dietary pattern. They were asked to try to follow DASH guidelines. However, unlike the intervention arm, they did not receive any of the text messages, which included tailored feedback and skills training videos. Active comparator participants received the same publicly available booklet from NHLBI with information about the DASH diet.

Measures

Baseline Sociodemographic and Psychosocial Variables

To characterize participants, we measured various sociodemographic and psychosocial variables. Sociodemographic measures were collected at baseline using standard survey questions used in previous studies. This included age, race and ethnicity, educational attainment, marital status, employment, insurance status, and the number of children within the household. Depression was assessed using the validated 8-item survey from the Patient Health Questionnaire [24]. The scale ranges from 0 to 24, and a score of >10 indicates depression.

Anthropometrics

Weight and height were collected by study staff and BMI was calculated for each participant. Height was measured twice and recorded to the nearest 0.1 cm using a wall-mounted height rod (Portrod; Health O Meter). Weight was measured twice and recorded to the nearest 0.1 kg using a scale (Scale-Tronix 5005; Welch Allyn).

Engagement With Diet Tracking

Engagement was operationalized as a valid day tracked, determined by looking at overall caloric intake. A day with missing data or a daily caloric level of <600 calories was considered to be a day that did not have complete tracking data and was marked invalid [25]. We assessed the average days tracked per week and calculated the proportion of participants who tracked at least 5 days per week.

DASH Diet Adherence

To assess change in DASH adherence, dietary intake was measured using the Automated Self-Administered 24-hour (ASA24) recall tool from the National Cancer Institute. The

ASA24 is an automated tool that uses the USDA's validated multiple-pass method to elicit intake throughout a given day [26]. Using an unannounced protocol, participants were asked via email to complete 1 weekend day and 1 weekday of dietary intake within a 2-week period. The ASA24 provides comprehensive nutrient data on all foods and beverages consumed during the previous 24-hour period. These data were used to calculate a nutrient-based DASH index score, developed by Mellen and colleagues [15], as mentioned above. Nutrient targets can be found in the "Change in DASH Adherence Across Arms" section and were standardized to total calorie intake. Individual nutrient scores were summed to calculate a total DASH adherence score. As such, the score range is 0 to 9, with higher scores indicative of greater adherence and a score of 9 indicating full adherence to the DASH dietary pattern.

Blood Pressure

Participants confirmed that they had not consumed coffee or used tobacco 30 minutes before the blood pressure measurement. We standardized our procedures to use the right arm unless a participant indicated that it was medically contraindicated, and we measured arm circumference to determine blood pressure cuff size. After resting for 5 minutes, blood pressure was measured 3 times with 30-second intervals in between each measurement using a blood pressure monitor (HEM-907XL; Omron Healthcare). Using the average blood pressure measurement, we categorized participants based on their severity of hypertension [3].

Intervention Satisfaction

At 3 months, using surveys adapted from our previous studies, we assessed perceived ease and usefulness of the Nutritionix app across both study arms and overall satisfaction with the DASH Cloud intervention within the intervention arm only. Using a 5-point Likert scale from strongly disagree to strongly agree, we asked participants about their perceptions of all elements of the intervention, including the ease and usefulness of the Nutritionix app, the DASH score in the feedback, the personalization of the texts, the timing of the texts, and whether an individual would recommend this program to others. The "Satisfaction With Intervention Activities" section describes the intervention satisfaction questions we asked the intervention arm only and the proportion that responded in agreement.

Analysis

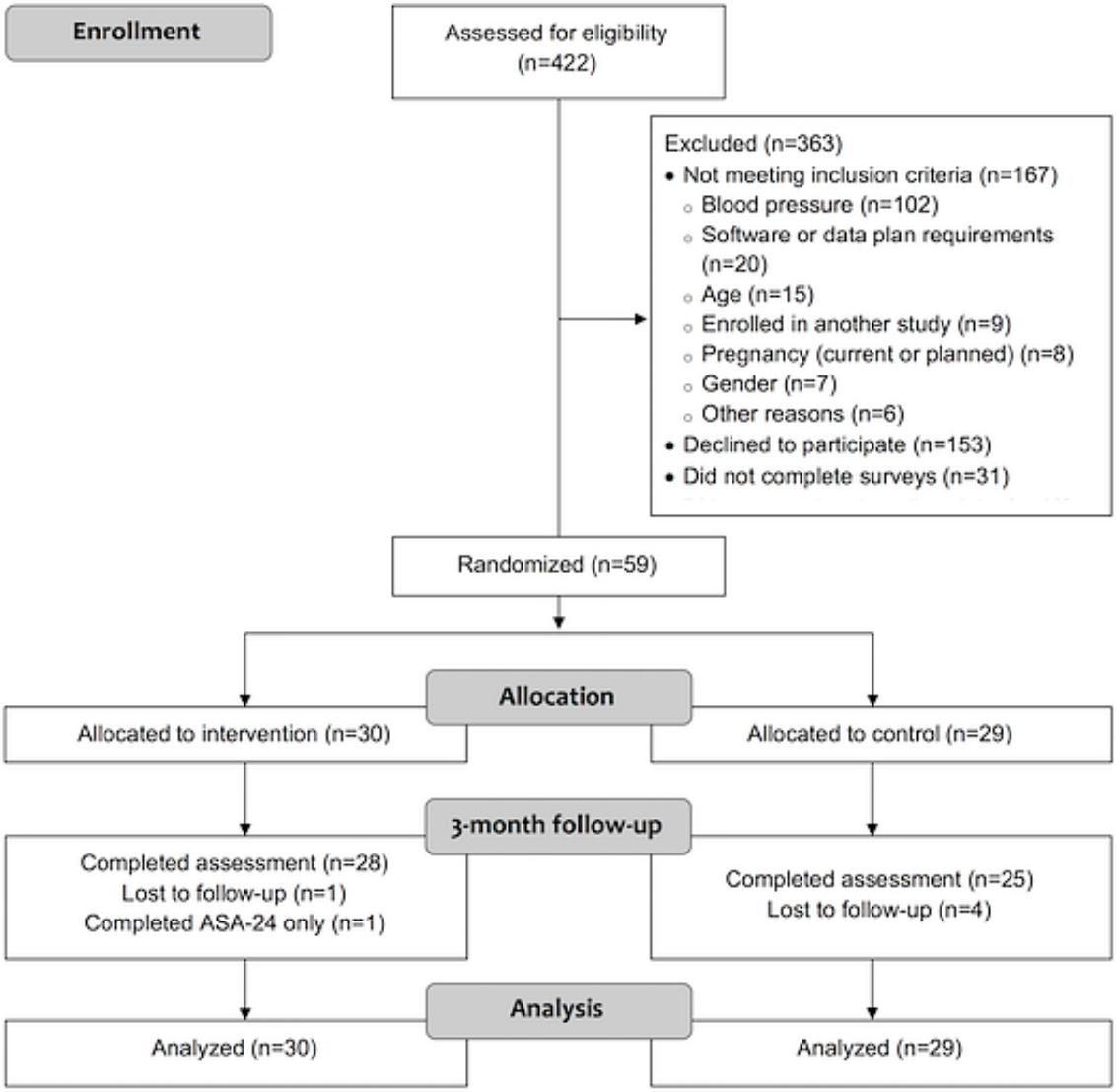
For descriptive analyses, normally distributed variables were summarized and reported as means and standard deviations. Engagement was analyzed using a 2-tailed *t* test to compare overall mean engagement over the course of the study. Engagement trends over time were analyzed using a random-intercept mixed model with unstructured covariance. The model included a continuous variable for time in weeks, a treatment group variable, and a time-by-group interaction term. To assess changes in the DASH adherence score, we conducted a primary analysis that treated invalid baseline and 3-month data as missing. Using standard protocols for obtaining valid dietary intake [23], we defined invalid assessments as those having a mean daily caloric intake of <600 or >3500 calories or those that did not include at least 2 recalls from 1 weekend day and 1 weekday or that included recalls that were collected more than 2 weeks apart. To examine the effect of the treatment group over time on DASH adherence, controlling for baseline levels of DASH adherence, we conducted linear regression models assessing 3-month DASH scores and DASH component values by study arm and including baseline DASH adherence as a covariate. Within-group changes in DASH scores and components were estimated using repeated-measures analysis of variance methods. Sensitivity analysis models were conducted using the same methods while including measurements previously excluded as invalid in the per-protocol models. Between-group differences in 3-month systolic and diastolic blood pressure were also assessed using linear regression models and adjusting for baseline values. Missing 3-month values in the DASH and blood pressure models were addressed using intent-to-treat principles with maximum likelihood estimations. Analyses were conducted using SAS 9.4 (SAS Institute) software and *P* values with an α <.05 were considered statistically significant.

Results

Recruitment and Retention Rates

The recruitment period lasted from July 21, 2017, to November 5, 2017. The CONSORT (Consolidated Standards of Reporting Trials) diagram (Figure 2) shows the study flow for both recruitment and retention; 422 individuals filled out the screening on the DASH Cloud website, and 363 of those individuals were ineligible, declined participation, or did not complete baseline activities. The primary reason for ineligibility was not meeting the blood pressure criteria. Of the remaining eligible participants, 59 were then randomized to receive the DASH Cloud intervention (n=30) or the active comparator arm (n=29). At the end of the study, 90% (28/30 intervention, 25/29 control) of participants attended the final 3-month visit. The ASA24 surveys were collected outside the in-person 3-month visit and had different retention rates. Of the 59 total participants, 53 (90%; 29/30 intervention, 24/29 control) completed the ASA24 per protocol at baseline, and 46 (78%; 25/30 intervention, 21/29 control) completed it at 3 months; 43 (73%; 24/30 intervention, 19/29 control) completed it per protocol at both time points. All participants completed at least 1 ASA24 at baseline, and 55 of the 59 (93%) participants completed at least 1 diet recall at 3 months.

Figure 2. Study flow and CONSORT diagram. CONSORT: Consolidated Standards of Reporting Trials. ASA24: Automated Self-Administered 24-hour recall tool.



Baseline Characteristics

As is shown in Table 1, participants (N=59) were all women with a mean age of 49.9 (SD 11.9) years and a mean BMI of 33.9 (SD 7.6) kg/m^2 (Table 1). Most participants were non-Hispanic White (41/59, 69%), insured (57/59, 97%), college educated (49/59, 83%), employed (42/59, 71%), and married (36/59, 61%). At baseline, the mean systolic blood pressure was 122.9 (SD 14.2) mmHg and the mean diastolic blood pressure was 80.2 (SD 8.8) mmHg. The majority of participants (41/59, 69%) were classified as having elevated blood pressure according to updated blood pressure guidelines from the American Heart Association and the American College of Cardiology, which is defined as any measurement of systolic blood pressure over 120 mmHg or a diastolic blood pressure measurement over 80 mmHg. These updated guidelines have lowered the threshold for elevated blood pressure, and lifestyle treatments, such as DASH, are a first line of treatment for those meeting these criteria. There were no differences in baseline characteristics by study arm.

Table 1. Sociodemographic and clinical baseline characteristics of participants enrolled in the DASH Cloud intervention trial (N=59).

Characteristic	Total
Age (years), mean (SD)	49.9 (11.9)
BMI (kg/m^2), mean (SD)	33.8 (7.6)
Systolic blood pressure (mmHg), mean (SD)	122.9 (14.2)
Diastolic blood pressure (mmHg), mean (SD)	80.2 (8.8)
Normal blood pressure category (defined as <120/80 mmHg), n (%)[a]	18 (31)
Self-reported use of blood pressure medications, n (%)	29 (49)
Married or living with partner, n (%)	36 (61)
Children in household, n (%)	23 (41)
Currently employed, n (%)	42 (71)
Currently insured, n (%)	57 (97)
Education, n (%)	
Vocational or trade school or some college after high school	5 (8)
Associate degree from college or university	5 (8)
College degree	15 (25)
Postgraduate degree from college or university	34 (58)
Race/ethnicity, n (%)	
Non-Hispanic White	41 (69)
Non-Hispanic Black	10 (17)
Hispanic Black	2 (3)
Hispanic White	2 (3)
Other (Asian, multiracial, or unreported)	4 (7)
Depression, n (%)[b]	
Minimal depression (score of 0-4)	37 (63)
Mild depression (score of 5-9)	18 (31)
Moderate depression (score of 10-14)	4 (6)

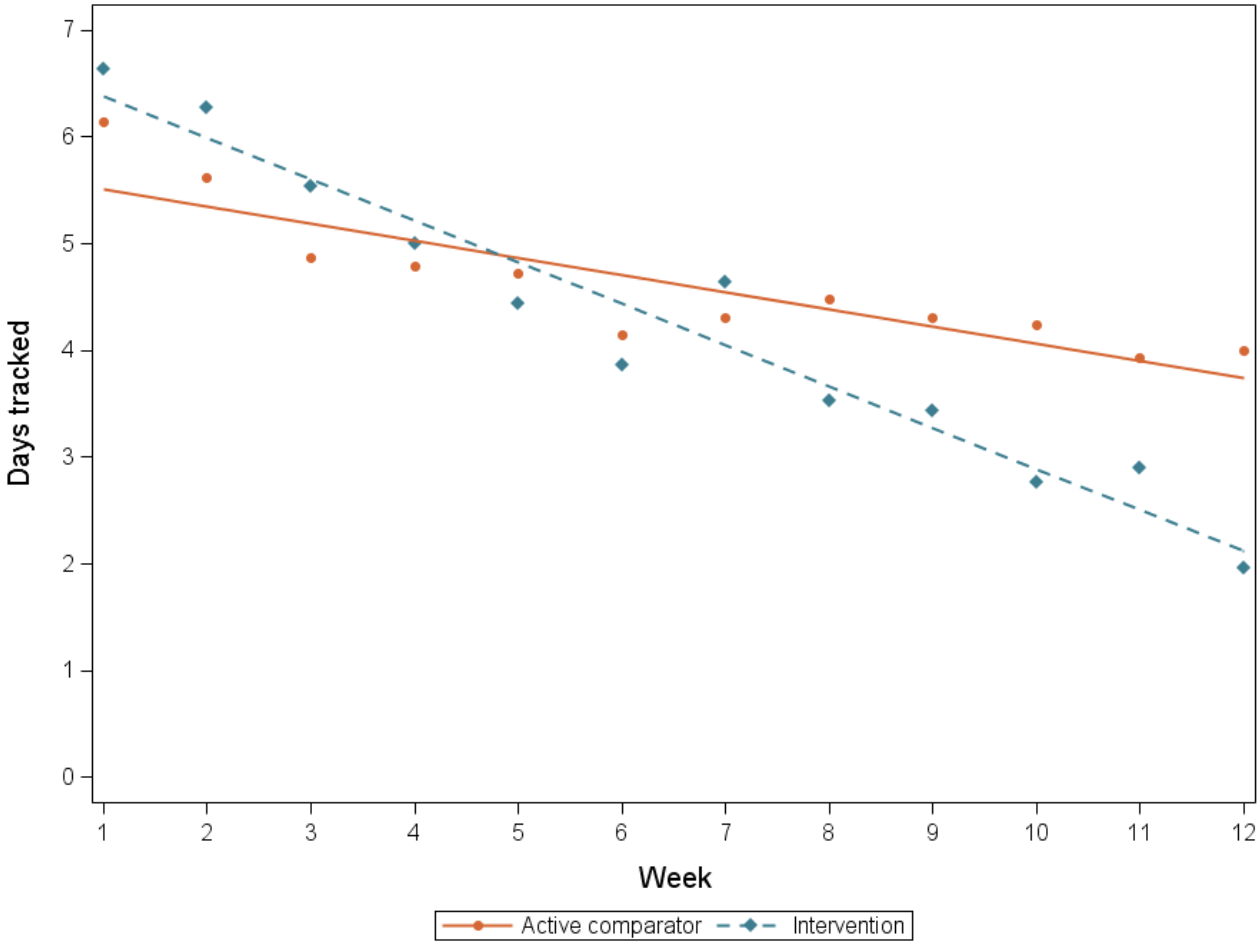
[a]Using the blood pressure treatment categories from the 2017 American Heart Association and the American College of Cardiology guidelines.

[b]Using the 8-item Patient Health Questionnaire scale, with scores ranging from 0 to 24.

Engagement With Diet Tracking

Figure 3 shows the actual and predicted days tracked per week across study arms. The mean days tracked per week overall was slightly higher, although not statistically significant, in the active comparator arm compared with the intervention arm (mean 4.6, SD 2.7 days vs 4.2, SD 2.1 days; *P*=.54). Likewise, a higher proportion of active comparator participants tracked 5 days or more per week, on average, compared with intervention participants (18/29, 63% vs 14/30, 47%; *P*=.24). However, the intervention participants experienced a steeper reduction in diet-tracking engagement over time, with rates decreasing by 0.23 (95% CI 0.16-0.29) days per week (*P*<.001), which is about a day per month faster than the active comparator group.

Figure 3. Predicted mean days of diet tracking over time by study arm.



Change in DASH Adherence Across Arms

As is shown in Table 2, the mean DASH score at baseline was 2.2 (SD 1.3), and there were no significant differences across study arms (2.2, SD 1.3 in intervention vs 2.4, SD 1.3 in active comparator; *P*=.85). After we adjusted for baseline DASH scores, participants in the intervention arm had DASH scores that were 0.01 (95% CI –0.7 to 0.7) points lower than the active comparator arm at 3 months (*P*=.97). Both groups experienced significant increases in DASH scores at 3 months (intervention mean change: 0.8, 95% CI 0.2-1.5; *P*=.02; active comparator mean change: 0.8, 95% CI 0.4-1.2; *P*<.001). No significant differences were observed between groups on any nutrient component within the DASH score after adjusting for baseline between-group differences. Both groups experienced significant or marginally significant increases in fiber intake over the 3 months. Both groups also reported significant decreases in saturated fat. Within the intervention group, there was a significant increase in magnesium intake and a significant decrease in total fat consumption. Within the active comparator group, there was a significant decrease in total fat consumption and an increase in calcium.

Sensitivity analysis that included ASA24 data that were not valid per protocol produced results that were generally consistent with the per-protocol models using valid data. Data from all 59 participants were included, with 93% (55/59) of participants contributing 3-month data to these models.

Table 2. Change in DASH adherence score and DASH score components within and across study arms among participants in the DASH Cloud intervention trial (N=59).[a]

Nutrient	Standards	Intervention				Active comparator				Between-group difference	
		Baseline, mean (SD) (n=29)	3 months, mean (SD) (n=25)	Estimated mean difference (95% CI) (n=30)	*P* value	Baseline, mean (SD) (n=24)	3 months, mean (SD) (n=21)	Estimated mean difference (95% CI) (n=29)	*P* value	Adjusted mean difference (95% CI) (N=59)	*P* value
DASH[b] score	0-9	2.2 (1.3)	3.1 (1.4)	0.8 (0.2 to 1.5)	.02	2.3 (1.3)	3.1 (1.3)	0.8 (0.4 to 1.2)	.001	–0.01 (–0.7 to 0.7)	.97
Total fat (% of total calories)	<27% of total daily kcal	37.9 (7.4)	37.6 (7.2)	–0.2 (–2.7 to 2.4)	.89	39.4 (7.1)	34.9 (6.8)	–3.9 (–6.7 to –1.0)	.01	3.1 (–0.2 to 6.3)	.06
Saturated fat (% total calories)	<6% of total daily kcal	12.6 (3.5)	11.4 (2.9)	–1.4 (–2.6 to –0.2)	.03	12.9 (2.8)	11.2 (3.0)	–1.5 (–2.9 to –0.1)	.03	0.1 (–1.4 to 1.5)	.90
Protein (% total calories)	>18% of total daily kcal	16.0 (3.9)	17.4 (5.2)	1.2 (–1.1 to 3.4)	.30	16.5 (4.1)	16.4 (3.1)	–0.01 (–1.7 to 1.7)	.99	1.0 (–1.3 to 3.4)	.39
Cholesterol (mg/1000 kcal)	<71.4 mg/1000 kcal per day	172.2 (84.8)	160.9 (101.1)	–13.2 (–61.9 to 35.4)	.58	167.1 (80.8)	133.2 (67.4)	–31.5 (–71.5 to 8.4)	.12	23.8 (–25.2 to 72.8)	.34
Fiber (g/1000 kcal)	>14.8 g/1000 kcal per day	9.1 (4.9)	11.0 (4.3)	2.1 (0.7 to 3.5)	.004	9.5 (3.7)	11.6 (4.8)	1.9 (–0.04 to 3.8)	.054	0.1 (–2.0 to 2.1)	.96
Magnesium (mg/1000 kcal)	>238 mg/1000 kcal per day	142.2 (32.0)	171.3 (45.7)	28.3 (7.5 to 49.0)	.01	157.3 (51.1)	173.6 (44.1)	14.0 (–6.9 to 34.9)	.18	5.7 (–19.1 to 30.5)	.65
Calcium (mg/1000 kcal)	>590 mg/1000 kcal per day	494.4 (170.9)	489.5 (222.9)	–12.5 (–96.9 to 71.9)	.76	433.6 (137.2)	544.3 (238.3)	112.3 (3.6 to 221.1)	.04	–96.5 (–221.1 to 28.1)	.13
Potassium (mg/1000 kcal)	>2238 mg/1000 kcal per day	1318.0 (420.3)	1489.8 (439.8)	168.7 (–23.4 to 360.9)	.08	1388 (323.0)	1559.2 (446.0)	163.8 (–33.4 to 360.9)	.10	–25.0 (–252.7 to 202.7)	.83
Sodium (mg)	<2400 mg per day	3223.6 (912.1)	3078.4 (975.1)	–187.3 (–602.8 to 228.1)	.36	2775.6 (860.4)	3024.1 (737.1)	241.4 (–205.5 to 688.4)	.28	–101.5 (–576.5 to 373.4)	.68

[a]Participants with an invalid ASA24 were treated as missing.

[b]DASH: Dietary Approaches to Stop Hypertension.

Change in Blood Pressure Across Arms

After adjusting for baseline, participants in the intervention arm had systolic blood pressures that were, on average, 2.8 (95% CI –1.8 to 7.4) mmHg lower than the active comparator arm at 3 months, though not significantly (*P*=.23). Participants' 3-month average diastolic blood pressure was lower in the intervention arm, on average, by 3.6 (95% CI –0.2 to 7.3) mmHg compared with the active comparator arm after adjusting for baseline, though not significantly (*P*=.07).

Association Between Change in DASH Adherence and Change in Blood Pressure

Overall, changes in DASH scores and changes in blood pressure were inversely correlated. A 1-unit improvement in the DASH score was associated with a 2.5 (95% CI 0.5 to 4.5) mmHg decrease in systolic blood pressure (*r*=–0.34; *P*=.02) and a 1.6 (95% CI 0.05 to 3.3) mmHg decrease in diastolic blood pressure (*r*=–0.27; *P*=.05). The observed correlation was consistent between groups (*P*=.63). On average, the intervention group decreased by 2.7 (95% CI 0.4 to 5.0) mmHg in systolic blood pressure (*r*=–0.44; *P*=.03) and 1.3 (95% CI –1.0 to 3.6) mmHg in diastolic blood pressure (*r*=–0.23; *P*=.26) for each unit improvement in the DASH score over 3 months. The association

in the active comparator group was slightly weaker, with a 1.7 (95% CI –2.1 to 5.4) mmHg decrease in systolic blood pressure (r=–0.20; P=.37) and a 1.8 (95% CI –0.8 to 4.4) mmHg decrease in diastolic blood pressure (r=–0.29; P=.16) per unit of DASH score improvement.

Satisfaction With Intervention Activities

Across study arms, 82% (24/29) of intervention participants indicated with agreement or strong agreement that the Nutritionix app was easy to use. Half of participants (14/29, 50%) said that they would use the app frequently, and only 15% (4/29) indicated that the app was cumbersome to use. Within the intervention arm, the majority of intervention participants (23/29, 79%) said they would recommend the DASH Cloud intervention to a friend or family member (Table 3). Many participants (23/29, 79%) felt the texts were sent at a convenient time, and many (22/29, 76%) preferred a consistent time to receive the texts. Most (23/29, 79%) participants felt the DASH dietary pattern tips were easy to understand. Notably, only about half (16/29, 55%) of participants felt the feedback and DASH score were helpful; 45% (13/29) felt that the DASH score was motivating, and 41% (12/29) indicated that the DASH score reflected their dietary pattern. In addition, 76% (22/29) felt that a DASH score of 10 (reflecting full adherence) was difficult to achieve. About one-third (10/29, 34%) reported that the text messages helped them reach their diet goals, and only 28% (8/29) felt the text messages were personalized.

Table 3. Perceived usefulness and ease of use of intervention components among participants receiving the DASH Cloud intervention (n=29).

Statements about DASH[a] Cloud intervention components	Agreement with statement, n (%)
The feedback received on the automated text messages was helpful.	16 (55)
The DASH text messages helped me reach my personal diet goals.	10 (34)
The DASH text messages felt personalized.	8 (28)
The DASH text messages were sent at a convenient time each day.	21 (72)
I found the DASH score of 1-10 helpful.	16 (55)
I found the DASH score motivating.	13 (45)
The DASH score accurately reflected my diet pattern.	12 (41)
A DASH score of 10 was difficult to achieve.	22 (76)
I found the diet tips easy to understand.	23 (79)
I found the diet tips helpful to improve my DASH score.	16 (55)
I found the videos helpful.	19 (66)
I enjoyed watching the videos.	16 (55)
I learned a lot from the videos.	14 (48)
I applied the skills I learned from the video in my routine	14 (50)
Would you recommend this program to a friend or family member looking to eat healthy?	23 (79)

[a]DASH: Dietary Approaches to Stop Hypertension.

Discussion

Principal Findings

Adoption of the DASH dietary pattern can help with blood pressure reduction for the 100 million Americans with suboptimal blood pressure [3]. This study aimed to develop and test the feasibility of using a digital health tool that leverages smartphone diet-tracking apps to improve population-level adoption of DASH. We found that it was feasible to use a commercially available diet-tracking app in our intervention platform and add a behavioral intervention aimed at improving DASH adherence. We were successful at recruiting and retaining women with high blood pressure and achieved moderate to high rates of engagement with diet tracking across both study arms. Similarly, both study arms saw a small increase in DASH adherence at 3 months. However, we found that adding a digital behavioral DASH intervention to a diet-tracking app did not increase DASH adherence compared to diet tracking alone.

These findings are notable for various reasons. First, they signal that participants can achieve and sustain moderate to high engagement with diet-tracking apps even if no feedback is provided on app entries and that these apps can be used to improve their behavioral efficacy. Studies that have examined diet-tracking engagement without added behavioral intervention components demonstrated lower rates of tracking than the average of 4 days per week that we found in this study. In a study by Laing and colleagues [27], rates of diet tracking using the popular app My Fitness Pal dropped significantly after the first month. Patel and colleagues [28] found a similar steep decline in diet tracking after the first month. It may be that our focus on DASH rather than fitness or weight control motivated all participants regardless of whether they received feedback to engage more frequently.

Second, these findings also signal that any feedback is not always better than no feedback for improving engagement and overall dietary quality, with the caveat that we were not powered to evaluate those differences statistically. Behavior change

theories and relevant empirical studies support that giving feedback, particularly tailored feedback, can improve engagement and subsequent behavior change [29,30]. However, the structure of the feedback, the level of personalization, and the relevancy of the feedback are important. Given the study design, we cannot know for sure why the intervention group tracked their diet less often as a result of the feedback; responses to the intervention satisfaction questionnaire give us some insight. Most notable is that less than half of participants found the DASH score used in the feedback motivating and reflective of what they ate. This finding does not preclude the use of "summary type" scoring feedback for improving dietary intake; rather, this suggests that we need to reconsider how to make it more helpful and interpretable to individuals. Additionally, as in any technology, we experienced bugs or issues with coding that did not line up with the logic outlined in the feedback algorithm. We aimed to test for these bugs before and during the execution of the intervention. However, it is possible that the presence of bugs could have impacted intervention engagement and perceptions about helpfulness of the feedback messages. We did not measure this in the study, so we are not able to quantitatively or qualitatively indicate how often this occurred or how much impact any bugs had on outcomes. Future studies should aim to consider measuring the impact of these bugs on outcomes.

Additionally, many participants stated that achieving full DASH adherence was difficult. Given this perception, the DASH score may not have provided positive reinforcement to support dietary behavior change. We are not aware of other studies that gave feedback on nutrients to improve diet quality, and we decided to use nutrient intake as part of the feedback because of its clear link to DASH adherence and the subsequent benefits on blood pressure. Participants may prefer feedback on adherence using food groups or some other measure of adherence. Overall, it would have been helpful to conduct a formal qualitative study after the intervention to better understand the discrepancies between the groups in engagement outcomes and determine how we could have improved satisfaction and perceptions of ease and helpfulness of the intervention. Future studies testing the feasibility of digital health interventions such as DASH Cloud should consider how best to understand the factors that will improve sustained engagement. We did not include a coach or any human support, and it may be that to change dietary behaviors, additional supports are needed beyond what can be provided via technology alone.

Third, these findings indicate that we may need to reconsider how to best provide personalized feedback to improve overall DASH adherence. Only 28% (8/29) of participants said they felt the texts were personalized, despite the use of an algorithm designed to personalize messages about intake of specific DASH nutrients. Our understanding of the most effective formula for personalizing feedback is mixed, but many studies support that personalizing interventions and feedback improves engagement [29,30]. We suggest that the personalization also reflect individual circumstances (eg, including feedback tips specific to one's family structure) and attitudes and beliefs (eg, self-efficacy for changing behavior) and how those characteristics then impact the individual's dietary behaviors. Future studies should work toward testing these different approaches when personalizing feedback. Despite the low ratings for the use of the DASH score and the personalization of the texts, 79% (23/29) of participants said they would recommend this program. This suggests that the overall concept of using diet-tracking apps to promote adherence to DASH is appealing.

The high rating of the program overall is important because of the substantial need to improve population-level adoption of DASH. Mellen and colleagues [15] used population-level data to assess DASH adherence shortly before and after the integration of DASH into national dietary guidelines. They found that less than 1% of the population is fully adherent to DASH and only 20% meet half of the DASH recommendations [15]. DASH adherence remained poor in subsequent cross-sectional population cohort studies [9]. As is supported by the current study, it may be that full DASH adherence is difficult to achieve. Indeed, trials testing intensive behavioral interventions to promote DASH have also struggled to help individuals achieve full adherence [31,32]. The PREMIER trial [33,34] was a comprehensive, multicomponent, behavioral intervention to improve adoption of DASH. The PREMIER intervention included standard behavior change components delivered via frequent face-to-face counseling with a registered dietitian and group meetings [33,34]. This high-intensity approach was effective in improving DASH adherence, but not full adoption of DASH [32]. This intervention also included other behavioral components, including sodium reduction, increased physical activity, and weight loss. Although comprehensive, focusing on multiple behaviors at a given time may have made it more difficult to fully adopt all DASH recommendations. Similarly, the ENCORE (Exercise and Nutritional Interventions for Cardiovascular Health) trial tested a comprehensive behavioral approach to adopting DASH, assessing the comparative efficacy of DASH to DASH plus behaviors for weight management on changes in blood pressure [35]. They found that both groups doubled their DASH adherence score, but the average score post intervention remained suboptimal [31]. Notably, the ENCORE trial found a linear relationship between DASH adherence and change in blood pressure. Full DASH adherence may be optimal, but partial adherence to DASH can be effective for lowering blood pressure [31].

The intensive intervention approaches tested in both PREMIER and ENCORE were effective in improving the adoption of DASH to clinically meaningful levels, but these approaches are not accessible to the broader population. To our knowledge, this was one of the first studies to test the feasibility of using mobile technologies and smartphone apps to disseminate and improve the adoption of DASH. Mann and colleagues [36] developed a similar intervention called DASH Mobile that consisted of easy tracking of DASH food portions; integrated Bluetooth blood pressure, weight, and pedometer monitoring; goal setting; simple data visualizations; and multimedia video clips to train patients in the basic concepts of the lifestyle change plan. This intervention has some notable distinctions from the current study. First, the intervention included more than just diet tracking and text message feedback and videos. It included

other behavioral goals and behaviors to track, as well as the use of a coach to support behavior change efforts. Second, the research team developed the app rather than using commercially available ones to start. As such, it went through prototype testing and the researchers had to consider many elements of user design that for many commercially available apps have already been considered. Mann and colleagues [36] discuss lessons learned in the development of their DASH Mobile platform that are important for others looking to develop digital health programs to increase adoption of DASH. Our findings add to that evidence base but suggest that a lower-intensity approach improves DASH adoption only slightly. To truly extend the reach of evidence-based behavioral interventions to disseminate DASH, we need to continue to test ways to replicate what was done in PREMIER and ENCORE while maintaining the potential for dissemination we aimed for in DASH Cloud.

Strengths and Limitations

Strengths of this study included the purposeful choice to use commercial apps in the intervention platform. This choice allows for more flexibility in future studies and increases the potential that this platform could be adapted for other technologies or outcomes. For example, if a new technology is developed that uses a different tool for tracking diet (eg, voice-assisted tracking), the intervention could be easily implemented with that new technology. This design choice increases the potential for dissemination. The use of a randomized controlled design allowed us to isolate the feasibility of the feedback and disentangle the effects of the tracking. Our results provide good insight about how feedback may not always improve engagement with diet tracking. Further, we achieved strong recruitment and retention rates, which supports the feasibility of this approach.

Despite these strengths, there are several limitations worth noting. This was a feasibility trial with a small sample size, so it is difficult to interpret the findings of this study because it was not powered to show an effect on changes in DASH adherence, engagement, or blood pressure. The primary focus was to assess the feasibility of developing the DASH Cloud platform by examining recruitment and retention rates and engagement with intervention activities, then exploring the effects of the intervention on DASH diet adherence and changes in blood pressure. As such, we did not conduct any power analyses. This is consistent with guidelines from Leon et al [21], which state that conducting power analyses is not needed when testing for feasibility. We also did not limit our study sample to only participants with hypertension or participants not taking hypertension medications. This was purposeful, given the focus on feasibility. However, this may have attenuated any differential effects on change in blood pressure. A common limitation of dietary studies is that the accuracy of the data collected is subject to potential recall and response biases [25,26]. Since participants were aware that their diet data would be collected, it may have led them to overestimate or underestimate dietary intake. Similarly, since we operationalized engagement using a valid day of tracking based on caloric intake, there may have been nondifferential misclassification of participants. Participants may have logged one meal or one food item to reach this threshold. Furthermore, the results of this study are not generalizable beyond the study population, which included mostly non-Hispanic White, educated women. The focus on recruiting only women was to better understand ways to reduce women's disproportionate cardiovascular risk using lifestyle approaches to blood pressure management.

Conclusion and Future Directions

We developed a digital program, DASH Cloud, that leverages commercially available diet-tracking apps that millions of Americans are using. In the current feasibility study, we found moderate to high engagement with diet tracking, an important predictor of successful behavior changes. However, this study also provided insights into the manner and type of feedback that should be given to participants using these diet-tracking apps. We need to better understand how to provide both personalized and relevant feedback to improve the uptake of DASH. To truly impact population health, we need to continually think not just about what best promotes the adoption of a healthy diet but also how to best disseminate and extend the reach of evidence-based treatments like the DASH dietary pattern.

Acknowledgments

The authors would like to thank the staff and researchers at the Duke Global Digital Health Science Center and the Sarah W Stedman Nutrition and Metabolism Center at Duke University. We are grateful to the Building Interdisciplinary Careers in Women's Health career development program and all the women who participated in the DASH Cloud study.

DMS was supported by the Building Interdisciplinary Research Careers in Women's Health Career Development Award (K12HD043446) from the National Institutes of Health. The funder had no role in the preparation, review, or approval of the manuscript and decision to submit the manuscript for publication.

Authors' Contributions

DMS had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. DMS conceived of the study, acquired study funding, led the study design and supervised its coordination, and drafted the manuscript for publication. JB, JC, and HA managed the study execution and contributed to drafting the manuscript. LPS and GGB consulted on data safety, study design, and execution and contributed to drafting the manuscript. JC, HA, and JB coordinated the intervention design and contributed to drafting the manuscript. SA participated in study design, conducted statistical analysis, and contributed to drafting the manuscript. MCK participated in statistical analysis and contributed to drafting

the manuscript. No writing assistance other than copy editing was provided in the preparation of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

DMS reports serving on the clinical advisory board for Omada Health. GGB reports being a shareholder in Coeus Health and serving on a scientific advisory board for Nutrisystem. The remaining authors declare no conflicts of interest.

References

1. Bundy J, Mills K, Chen J, Li C, Greenland P, He J. Estimating the Association of the 2017 and 2014 Hypertension Guidelines With Cardiovascular Events and Deaths in US Adults: An Analysis of National Data. JAMA Cardiol 2018 Jul 01;3(7):572-581 [FREE Full text] [doi: 10.1001/jamacardio.2018.1240] [Medline: 29800138]
2. Centers for Disease Control and Prevention. Vital signs: awareness and treatment of uncontrolled hypertension among adults--United States, 2003-2010. Morbidity and Mortality Weekly Report 2012;61(35):703-709 [FREE Full text]
3. Whelton P, Carey R, Aronow W, Casey Jr DE, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018;71(6):e13-e115. [doi: 10.1161/hyp.0000000000000076]
4. Muntner P, Carey RM, Gidding S, Jones DW, Taler SJ, Wright JT, et al. Potential US Population Impact of the 2017 ACC/AHA High Blood Pressure Guideline. Circulation 2018 Jan 09;137(2):109-118. [doi: 10.1161/circulationaha.117.032582]
5. Abramson BL, Srivaratharajah K, Davis L, Parapid B. Women and Hypertension: Beyond the 2017 Guideline for Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. American College of Cardiology. 2018 Jul 27. URL: https://www.acc.org/latest-in-cardiology/articles/2018/07/27/09/02/women-and-hypertension [accessed 2020-10-16]
6. Garcia M, Mulvagh SL, Bairey Merz CN, Buring JE, Manson JE. Cardiovascular Disease in Women. Circ Res 2016 Apr 15;118(8):1273-1293. [doi: 10.1161/circresaha.116.307547]
7. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/nejm199704173361601]
8. Sacks FM, Obarzanek E, Windhauser MM, Svetkey LP, Vollmer WM, McCullough M, DASH Investigators. Rationale and design of the Dietary Approaches to Stop Hypertension trial (DASH). Annals of Epidemiology 1995 Mar;5(2):108-118. [doi: 10.1016/1047-2797(94)00055-x]
9. Kwan MW, Wong MC, Wang HH, Liu KQ, Lee CL, Yan BP, et al. Compliance with the Dietary Approaches to Stop Hypertension (DASH) diet: a systematic review. PLoS One 2013;8(10):e78412 [FREE Full text] [doi: 10.1371/journal.pone.0078412] [Medline: 24205227]
10. Saneei P, Salehi-Abargouei A, Esmaillzadeh A, Azadbakht L. Influence of Dietary Approaches to Stop Hypertension (DASH) diet on blood pressure: a systematic review and meta-analysis on randomized controlled trials. Nutr Metab Cardiovasc Dis 2014 Dec;24(12):1253-1261. [doi: 10.1016/j.numecd.2014.06.008] [Medline: 25149893]
11. Svetkey LP, Simons-Morton DG, Proschan MA, Sacks FM, Conlin PR, Harsha D, et al. Effect of the dietary approaches to stop hypertension diet and reduced sodium intake on blood pressure control. J Clin Hypertens (Greenwich) 2004 Jul;6(7):373-381. [doi: 10.1111/j.1524-6175.2004.03523.x] [Medline: 15249792]
12. James P, Oparil S, Carter B, Cushman WC, Dennison-Himmelfarb C, Handler J, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). JAMA 2014 Feb 05;311(5):507-520. [doi: 10.1001/jama.2013.284427] [Medline: 24352797]
13. US Department of Health and Human Services, US Department of Agriculture. Dietary Guidelines for Americans: 2015-2020. 8th ed. Health.gov. URL: https://health.gov/dietaryguidelines/2015/guidelines [accessed 2020-10-16]
14. Steinberg D, Bennett GG, Svetkey L. The DASH Diet, 20 Years Later. JAMA 2017 Apr 18;317(15):1529-1530 [FREE Full text] [doi: 10.1001/jama.2017.1628] [Medline: 28278326]
15. Mellen PB, Gao SK, Vitolins MZ, Goff DC. Deteriorating dietary habits among adults with hypertension: DASH dietary accordance, NHANES 1988-1994 and 1999-2004. Arch Intern Med 2008 Feb 11;168(3):308-314. [doi: 10.1001/archinternmed.2007.119] [Medline: 18268173]
16. Mobile fact sheet. Pew Research Center. 2019 Jun 12. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-10-16]
17. Percentage of Mobile Medical Application Categories Used by U.S. Adults at Least Once as of 2017. Statista. URL: www.statista.com/statistics/378850/top-mobile-health-application-categories-used-by-us-consumers/ [accessed 2018-05-15]
18. Most popular health and fitness apps in the United States as of May 2018, by monthly active users (in millions). Statista. URL: https://www.statista.com/statistics/650748/health-fitness-app-usage-usa/ [accessed 2018-05-14]
19. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]

20. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
21. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
22. Your guide to lowering your blood pressure with DASH: DASH eating plan. National Heart, Lung, and Blood Institute. URL: https://www.nhlbi.nih.gov/files/docs/public/heart/new_dash.pdf [accessed 2020-10-26]
23. Kapsokefalou M, Roe M, Turrini A, Costa HS, Martinez-Victoria E, Marletta L, et al. Food Composition at Present: New Challenges. Nutrients 2019 Jul 25;11(8):1714 [FREE Full text] [doi: 10.3390/nu11081714] [Medline: 31349634]
24. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
25. Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires : the Eating at America's Table Study. Am J Epidemiol 2001 Dec 15;154(12):1089-1099. [doi: 10.1093/aje/154.12.1089] [Medline: 11744511]
26. Subar AF, Kirkpatrick SI, Mittl B, Zimmerman TP, Thompson FE, Bingley C, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet 2012 Aug;112(8):1134-1137 [FREE Full text] [doi: 10.1016/j.jand.2012.04.016] [Medline: 22704899]
27. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a Smartphone Application for Weight Loss Compared With Usual Care in Overweight Primary Care Patients. Ann Intern Med 2014 Nov 18;161(Suppl 10):S5. [doi: 10.7326/m13-3005]
28. Patel ML, Hopkins CM, Brooks TL, Bennett GG. Comparing Self-Monitoring Strategies for Weight Loss in a Smartphone App: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Feb 28;7(2):e12209 [FREE Full text] [doi: 10.2196/12209] [Medline: 30816851]
29. de Vries H, Kremers SPJ, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]
30. 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]
31. Epstein DE, Sherwood A, Smith PJ, Craighead L, Caccia C, Lin P, et al. Determinants and consequences of adherence to the dietary approaches to stop hypertension diet in African-American and white adults with high blood pressure: results from the ENCORE trial. J Acad Nutr Diet 2012 Nov;112(11):1763-1773 [FREE Full text] [doi: 10.1016/j.jand.2012.07.007] [Medline: 23000025]
32. Lin P, Appel LJ, Funk K, Craddick S, Chen C, Elmer P, et al. The PREMIER intervention helps participants follow the Dietary Approaches to Stop Hypertension dietary pattern and the current Dietary Reference Intakes recommendations. J Am Diet Assoc 2007 Sep;107(9):1541-1551. [doi: 10.1016/j.jada.2007.06.019] [Medline: 17761231]
33. Appel LJ, Champagne CM, Harsha DW, Cooper LS, Obarzanek E, Elmer PJ, Writing Group of the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA 2003;289(16):2083-2093. [doi: 10.1001/jama.289.16.2083] [Medline: 12709466]
34. Svetkey LP, Harsha DW, Vollmer WM, Stevens VJ, Obarzanek E, Elmer PJ, et al. Premier: a clinical trial of comprehensive lifestyle modification for blood pressure control: rationale, design and baseline characteristics. Ann Epidemiol 2003 Jul;13(6):462-471. [doi: 10.1016/s1047-2797(03)00006-1]
35. Hinderliter AL, Sherwood A, Craighead LW, Lin P, Watkins L, Babyak MA, et al. The long-term effects of lifestyle change on blood pressure: One-year follow-up of the ENCORE study. Am J Hypertens 2014 May;27(5):734-741 [FREE Full text] [doi: 10.1093/ajh/hpt183] [Medline: 24084586]
36. Mann DM, Quintiliani LM, Reddy S, Kitos NR, Weng M. Dietary Approaches to Stop Hypertension: Lessons Learned From a Case Study on the Development of an mHealth Behavior Change System. JMIR Mhealth Uhealth 2014 Oct 23;2(4):e41 [FREE Full text] [doi: 10.2196/mhealth.3307] [Medline: 25340979]

Abbreviations

API: application programming interface
ASA24: Automated Self-Administered 24-hour
CONSORT: Consolidated Standards of Reporting Trials
DASH: Dietary Approaches to Stop Hypertension
ENCORE: Exercise and Nutritional Interventions for Cardiovascular Health
NHLBI: National Heart, Lung, and Blood Institute
REDCap: Research Electronic Data Capture
USDA: United States Department of Agriculture

Edited by G Eysenbach; submitted 18.12.19; peer-reviewed by NC Tan, P Hsiao, A Pfammatter, B Green; comments to author 11.03.20; revised version received 03.08.20; accepted 12.09.20; published 07.12.20.

Please cite as:
Steinberg DM, Kay MC, Svetkey LP, Askew S, Christy J, Burroughs J, Ahmed H, Bennett GG
Feasibility of a Digital Health Intervention to Improve Diet Quality Among Women With High Blood Pressure: Randomized Controlled Feasibility Trial
JMIR Mhealth Uhealth 2020;8(12):e17536
URL: https://mhealth.jmir.org/2020/12/e17536
doi:10.2196/17536
PMID:33284116

Original Paper

Experiences, Attitudes, and Needs of Users of a Pregnancy and Parenting App (Baby Buddy) During the COVID-19 Pandemic: Mixed Methods Study

Alexandra Rhodes[1], BSc, MSc; Sara Kheireddine[1], BSc, MSc; Andrea D Smith[1], BSc, MPhil, PhD

Department of Behavioural Science and Health, Institute of Epidemiology and Health Care, University College London, London, United Kingdom

Corresponding Author:
Alexandra Rhodes, BSc, MSc
Department of Behavioural Science and Health
Institute of Epidemiology and Health Care
University College London
Gower Street
London, WC1E 6BT
United Kingdom
Phone: 44 20 7679 2000
Email: alexandra.rhodes.15@ucl.ac.uk

Abstract

Background: The COVID-19 pandemic has impacted the lives of expectant parents and parents of young babies, with disruptions in health care provision and loss of social support.

Objective: This study investigated the impact of the COVID-19 pandemic and its associated lockdown on this population through the lens of users of the UK National Health Service–approved pregnancy and parenting smartphone app, Baby Buddy. The study aims were threefold: to gain insights into the attitudes and experiences of expectant and recent parents (with babies under 24 weeks of age) during the COVID-19 pandemic; to investigate whether Baby Buddy is meeting users' needs during this time; and to identify ways to revise the content of Baby Buddy to better support its users now and in future.

Methods: A mixed methods study design combining a web-based survey with semistructured telephone interviews among Baby Buddy users in the United Kingdom was applied. Data were collected from April 15 to mid-June 2020, corresponding to weeks 4-13 of the lockdown in the United Kingdom.

Results: A total of 436 expectant (n=244, 56.0%) and recent (n=192, 44.0%) parents responded to the web-based survey, of which 79.1% (n=345) were aged 25-39 years and 17.2% (n=75) spoke English as their second language. Of the 436 respondents, 88.5% (386/436) reported increased levels of anxiety around pregnancy, birth, and being a new parent, and 58.0% (253/436) were concerned about their emotional and mental health. Of the 244 pregnant respondents, 43.4% (n=106) were concerned about their physical health. Telephone interviews with 13 pregnant women and 19 recent parents revealed similarly increased levels of anxiety due to reduced health care provision and loss of support from friends and family. Although a minority of respondents identified some positive outcomes of lockdown, such as family bonding, many telephone interviewees reported feeling isolated, disregarded, and overwhelmed. Recent parents were particularly anxious about the impact of the lockdown on their baby's development and socialization. Many interviewees were also concerned about their physical health as a consequence of both limited access to face-to-face medical appointments and their own poorer dietary and physical activity behaviors. Across both samples, 97.0% (423/436) of respondents reported that Baby Buddy was currently helping them, with many commenting that its role was even more important given the lack of face-to-face support from health care and parenting organizations. Greater speed in updating digital content to reflect changes due to the pandemic was suggested.

Conclusions: The COVID-19 pandemic has created heightened anxiety and stress among expectant parents and those with a young baby, and for many, lockdown has had an adverse impact on their physical and mental well-being. With reductions in health care and social support, expectant and new parents are increasingly relying on web-based resources. As a free, evidence-based app, Baby Buddy is well positioned to meet this need. The app could support its users even more by actively directing them to the wealth of existing content relevant to their concerns and by adding content to give users the knowledge and confidence to meet new challenges.

(JMIR Mhealth Uhealth 2020;8(12):e23157) doi:10.2196/23157

KEYWORDS

pregnancy; parenting; app; COVID-19; pregnancy support; postnatal support; perinatal; mental well-being; physical well-being; support; well-being; experience; attitude; needs

Introduction

The COVID-19 pandemic has impacted people's lives in ways that few could have predicted. With health care services channeling resources to patients with COVID-19 and face-to-face contacts being reduced to limit the risk of contagion, antenatal, birth and postnatal care have all been affected. Moreover, due to the COVID-19 pandemic and its associated lockdown, expectant parents and parents of young babies have lost their usual support networks.

In recent years, the use of pregnancy and parenting apps has become widespread [1,2]. A 2016 study in Australia found that three-quarters of a sample (n=410) of pregnant women and women who had had a baby in the last three years had used a pregnancy app, and over half had used a parenting app [3]. In terms of the people who use pregnancy apps, research suggests that women who are pregnant for the first time and have owned a smartphone for longer have higher uptake [4], while women with lower income and nonnative speakers of the app language have lower uptake [2]. Access to information on pregnancy health and fetal development has been reported to be the primary reason for using a pregnancy app, along with access to information on nutrition [2,5]. Although data on the use of pregnancy and parenting apps in the United Kingdom are not available, smartphone ownership is reported to be in excess of 90% among UK women of childbearing age [6].

Launched in the United Kingdom in 2014, Baby Buddy is a pregnancy and parenting app approved by the National Health Service (NHS) that is designed to support expectant and new parents through pregnancy and the first 24 weeks of parenthood and to reduce inequalities [7]. Baby Buddy was developed by taking a proportionate universalism approach: it was designed and developed for use by all UK parents but was created to be particularly relevant and engaging for parents from communities whose voices are seldom heard and whose children are at higher risk of poorer outcomes [8]. The app is free to download and is available in all app stores. In addition, Best Beginnings, the charity behind Baby Buddy, works with NHS trusts and local authorities to embed the app into maternity care pathways. Baby Buddy provides trusted, evidence-based information and self-care tools to help expectant and new parents build their knowledge and confidence and effect positive behavior changes. Its content and functionality were co-created with parents and professional health organizations. With low literacy level requirements and extensive video content (over 300 short video clips of parents sharing their own stories and professionals sharing advice), Baby Buddy ensures inclusivity and is easily accessible to people not in education, training, or employment and those whose first language is not English (see Multimedia Appendix 1 for more details about the app). Contrary to other pregnancy apps [2], Baby Buddy data show disproportionate usage by young women (under 25 years of age), women from Black, Asian, and minority ethnic communities, and people who speak English as a second language (ESL) [9].

As a virtual resource, Baby Buddy has been able to continue to provide support to its users during the pandemic. However, Best Beginnings was quick to recognize that the needs of Baby Buddy users and the role the app plays for them might have shifted during this period. Therefore, the decision was made to conduct a service evaluation study to identify how Baby Buddy is currently meeting its users' needs and how to increase the value of the app to its users during this time. With a new version of Baby Buddy due to launch in 2021, the findings from this research could also feed into the future development of the app. This study acted as the precursor to a UK population-wide web-based survey exploring the experiences of expectant parents and parents of babies under 2 years of age, Babies in Lockdown [10].

The aims of this study were threefold: first, to gain insights into the attitudes and experiences of expectant parents and parents of very young babies during the COVID-19 pandemic; second, to investigate whether Baby Buddy is meeting the needs of its users during this time; and third, to identify ways in which the content of Baby Buddy can be added to or updated to better support its users both now and in the future.

Methods

Ethical Approval

As a service evaluation study, this project was exempt from University College London (UCL) ethical approval. However, procedures for informing respondents about the research and how their data would be handled, maintaining their anonymity, and obtaining respondent consent complied with the General Data Protection Regulation (GDPR) and were consistent with standards demanded by UCL ethics.

Given the potential mental fragility of respondents caused by the COVID-19 pandemic, a strategy for providing follow-up support was planned, including links to support within the app (eg, the 24/7 Baby Buddy Crisis Messenger clinically supervised text service) and telephone and web-based support from other charitable organizations.

Study Design

The study used a mixed methods approach comprising an anonymous web-based survey and a series of semistructured telephone interviews among a sample of survey respondents who expressed willingness to take part in further research.

The web-based survey was designed to capture the impact of COVID-19 on the respondents' views and experiences of being pregnant or parenting a young baby, mood, levels of anxiety, key concerns, and additional worries; the perceived impact of the pandemic on their baby; contact with health care

professionals and perinatal support; and their use and views of Baby Buddy. The survey was hosted on SurveyMonkey, with collection settings to anonymize responses and disallow multiple responses from the same IP address. The survey comprised 12 questions (over 6 pages), all but 3 of which were mandatory and 8 of which included free text boxes for respondents to record their opinions in their own words. The questions were worded to reflect the status of the respondent (ie, pregnant, gave birth within the last 6 months, or partner of either) established in question 1. Minimal demographic information (age, ESL, and region) was collected to reduce respondent burden. The survey was piloted with expectant and new parents (n=4) and professionals working in the perinatal field (n=4). Data collection occurred between April 15 and May 31, 2020.

The topic guide for the telephone interviews was designed by AR and approved by Best Beginnings (see Multimedia Appendix 2). All interviews were conducted by AR, an experienced qualitative researcher with many years of experience conducting sensitive research, and SK, who received training and guidance on interviewing approaches and techniques from AR. Telephone interviews took place from April 29 to June 18, 2020.

Sample and Recruitment Strategy

A convenience sample of Baby Buddy app users was recruited via the app. A notification and link to the survey was incorporated into many pieces of daily information within the Today's Information feature in the app from April 15 to May 31, 2020 (see Multimedia Appendix 3 for the recruitment message). The Today's Information feature delivers a new piece of gestational stage–appropriate information each morning to the user's home screen. The link was added to every seventh piece of daily information to ensure that it would be delivered to all users once over the course of a week irrespective of their gestational stage or baby age. The link took potential respondents to a survey information page that outlined the anonymous nature of the survey and how the data would be reported; the page also provided a link to the Best Beginnings privacy statement as well as an email address for queries. Participants' consent was sought before they proceeded to the survey.

Telephone interview respondents were selected from the list of those who agreed to participate in further research (requested at the end of the web-based survey) and had provided contact details. Selection was initially random but was later based on age, postcode, and ESL to ensure a representative sample. An a priori decision was made to conduct a minimum of 20 interviews, after which AR and SK would decide whether the saturation point had been reached and, if not, among which respondent types further investigation was required.

Neither the web-based survey nor the telephone interview respondents were offered any incentives.

Data Analysis

Quantitative data analyses were undertaken in SPSS version 26 (IBM Corporation). Pearson chi-square tests were used to determine whether responses varied according to age or ESL. Qualitative data (telephone interviews and comments from the free text boxes) was analyzed thematically [11]. The 5-stage analysis process was conducted by hand. First, AR and SK listened to and transcribed their own interviews as well as 30% of each other's interviews. Second, AR developed an initial coding framework (see Multimedia Appendix 4) based on key topics from the interview guide and free text boxes as well as additional, unanticipated topics that emerged from the first five interviews [12]. Third, AR and SK coded their own and 30% of each other's interview transcripts and AR coded the free text comments from the survey. The coding framework continued to be refined throughout the coding process, with any discrepancies between AR and SK over codes or coding definitions being resolved through discussion. Fourth, emerging themes were identified, and AR and SK agreed on the final themes through an iterative process of discussion, refinement, and development. Fifth, AR and SK revisited the transcripts to confirm the legitimacy of the final themes. AS verified the coding and theme development process for accuracy and appropriateness.

Results

Participant Characteristics

Web-Based Survey

A total of 543 Baby Buddy app users opened the survey, and 436 (80.7%) completed it; however, 23 users (4.2%) declined to answer the demographic questions, of whom 14 (2.6) also did not answer the last three survey questions. The characteristics of the final sample are shown in Table 1. Of the sample, 55.9% respondents were pregnant (244/436) and 44.0% were recent parents (192/436). Only 2.1% (9/436) of the sample were partners of pregnant women or women who had recently given birth. The majority of the sample was aged between 25 and 39 years (345/436, 79.1%) and spoke English as their first language (361/436, 82.8%).

Table 1. Descriptive characteristics of the web-based survey respondents by age group (N=436), n (%).

Characteristic	Pregnant	Partner pregnant	Given birth in last 6 months	Partner given birth in last 6 months	Total
Age (years)					
<21	12 (2.8)	0 (0)	3 (0.7)	0 (0)	15 (3.4)
21-24	15 (3.4)	0 (0)	8 (1.8)	0 (0)	23 (5.3)
25-29	64 (14.7)	0 (0)	40 (9.2)	1 (0.2)	105 (24.1)
30-34	76 (17.4)	1 (0.2)	83 (19.0)	2 (0.5)	162 (37.1)
35-39	41 (9.4)	1 (0.2)	35 (8.0)	1 (0.2)	78 (17.9)
40-44	12 (2.8)	0 (0)	14 (3.2)	0 (0)	26 (6.0)
≥45	2 (0.5)	1 (0,2)	1 (0.2)	0 (0)	4 (0.9)
Did not answer	17 (3.9)	2 (0.5)	4 (0.9)	0 (0)	23 (5.3)
Total	239 (54.8)	5 (1.1)	188 (43.1)	4 (0.9)	436 (100)
ESL[a]	44 (10.1)	1 (0.2)	28 (6.4)	2 (0.5)	75 (17.2)

[a]ESL: English as a second language.

Telephone Interviews

A total of 32 telephone interviews were conducted (AR, n=19, 59%; SK, n=13, 41%). With the exception of one interview that lasted 12 minutes, the interviews lasted between 21 and 55 minutes, averaging 30 minutes (median interview length 28 minutes). Table 2 summarizes the telephone interview sample characteristics. The majority of respondents were women who were married or cohabiting. Of these 32 respondents, 14 (44%) were pregnant women (between 8 and 38 weeks), 2 of whom (14%) had another child or children. The remaining 18 respondents (56%) had a baby under 6 months of age (between 3 and 24 weeks), 2 of whom (11%) had another child and 1 of whom (6%) was a father. Of the 18 postnatal respondents, 2 (11%) had given birth during the lockdown. The majority of the sample was of White British ethnicity (21/32, 66%). Of the 32 respondents, 6 (19%) were ESL respondents.

Table 2. Descriptive characteristics of the telephone interviewees by age group (N=32), n (%).

Age (years)	Pregnant (n=13)			Postnatal (n=19)		
	White British	White other	Black, Asian, minority ethnicity	White British	White other	Black, Asian, minority ethnicity
21-24	1 (3)	1 (3)	1 (3)	0 (0)	0 (0)	0 (0)
25-29	0 (0)	0 (0)	2 (6)	2 (6)	1 (3)	1 (3)
30-34	2 (6)	1 (3)	0 (0)	6 (19)	2 (6)	1 (3)
35-39	2 (6)	0 (0)	1 (3)	3 (9)	0 (0)	0 (0)
40-44	2 (6)	0 (0)	0 (0)	2 (6)	0 (0)	0 (0)
≥45	0 (0)	0 (0)	0 (0)	1 (3)	0 (0)	0 (0)
Total	7 (22)	2 (6)	4 (13)	14 (44)	3 (9)	2 (6)

Overview of Quantitative and Qualitative Findings

Given the alignment of the quantitative and qualitative findings and to create a holistic picture, the results of the web-based survey and telephone interviews are presented together under the four main themes that emerged from the analysis of the qualitative data (see Multimedia Appendix 5 for theme definitions and examples). The first three themes are interconnected and consider the impact of COVID-19 on expectant parents and parents of very young babies. The first theme explores how the pandemic has resulted in increased levels of anxiety and stress, identifying three underlying causes: current disruption, future uncertainties, and fear of contracting COVID-19. The second and third themes expand on these underlying causes, focusing on the aspects of life during the pandemic that have most affected pregnant and postnatal women—reduced levels of support and life under lockdown. The final theme considers the use of Baby Buddy during the pandemic and respondents' ideas for additional content.

No statistically significant differences were detected according to age or ESL.

Increased Levels of Anxiety and Stress

Of the 436 web-based survey respondents, 88.5% (n=386) reported that the pandemic had increased their levels of anxiety around pregnancy, birth, and being a new parent, with 42.8% (n=187) claiming the pandemic had made them a lot more anxious (see Table 3). The negative impact of the pandemic on mental well-being was also reflected in the web-based survey responses to the question "Please can you give us three words

to describe your mood over the past 5 days," with *anxious* appearing alongside *happy* and *tired* as the three most commonly used words. In response to the question "What are your main concerns right now?" which included 11 options as well as "none of the above" and "other" (see Multimedia Appendix 6), over 60% of respondents identified their own emotional and mental health to be a main concern. For the 244 pregnant respondents, this was the highest scoring main concern at that moment (61.3%, 144/235), while for postnatal respondents, it was second (58.0%, 109/188) to concerns about their baby's health (63.3%, 119/188). The qualitative findings indicated a range of negative emotional states in addition to anxiety and stress, including loneliness, irritability, sadness, and depression.

> *I am a lot more tense and anxious some days more than others and then I am very irritable. I cannot go out much and don't really have motivation to talk to people.* [Free text]

> *I have had days when I break down and cry….my entire day is lying on the sofa or in bed doing absolutely nothing.* [Pregnant 21 weeks, age 20-24 years, Black, Asian, minority ethnicity]

> *I have started to struggle in the last two weeks…. I had had a fight with my boyfriend because I was so tired* [Baby 16 weeks, age 30-34 years, White other]

> *I have become lazy – I can't be bothered…lethargic. I don't feel hungry because I am not moving enough…I can't be bothered with cooking or cleaning now….the boys have lost their appetite – one boy is losing weight….they stay in their PJs all day, say there is no point in brushing their teeth as no one is going to smell their breath….the boys are worried that when I go to hospital I am going to die.* [Pregnant 34 weeks, age 25-29 years, Black, Asian, minority ethnicity]

Table 3. Levels of anxiety about pregnancy and birth (pregnant) and having a new baby (postnatal) during weeks 1-2 and weeks 3-4 (N=436), n (%).

Anxiety level[a]	Pregnant			Postnatal		
	Overall (n=244)	Weeks 1-2 (April 15-29, n=116)	Weeks 3-4 (May 6-20, n=75)	Overall (n=192)	Weeks 1-2 (April 15-22, n=74)	Weeks 3-4 (May 6-20, n=65)
A lot more anxious	108 (44.2)	56 (48.3)	31 (41.3)	79 (41.1)	21 (28.4)	32 (49.2)
A little more anxious	111 (45.5)	50 (43.1)	33 (44)	88 (45.8)	43 (58.1)	25 (38.5)
No different	23 (9.4)	10 (8.6)	9 (12)	18 (9.3)	8 (10.8)	6 (9.2)
A little less anxious	1 (0.5)	0 (0)	1 (1.3)	7 (3.6)	2 (2.7)	2 (3.1)
A lot less anxious	1 (0.5)	0 (0)	1 (1.3)	0 (0)	0 (0)	0 (0)

[a]Evaluated with a 5-point Likert scale. Questions asked: "How has coronavirus made you feel about pregnancy and birth?" (pregnant) and "How has coronavirus made you feel about being a new parent or supporting a new parent?" (postnatal).

Not surprisingly, people already suffering from mental health issues had been hit particularly hard by the pandemic.

> *I have struggled with postnatal depression, OCD [obsessive-compulsive disorder] and anxiety. The disruption to normal life has affected the progress I had been making.* [Free text]

A small minority of respondents painted a more positive picture of their state of mind, finding their home-based life relaxing and less stressful than "normal" life.

> *Not having to go out the house as much has made me more relaxed.* [Free text]

The qualitative findings indicated three main causes of these negative emotional states: fear of contracting COVID-19, current disruption, and future uncertainties. The contribution of these causes to anxiety and stress levels appeared to change over the course of the research; current disruption and fear of contracting the virus were responsible for raised anxiety and stress levels during the early weeks (April), but future uncertainties showed more impact later (May). Similar shifts were apparent in the web-based survey, which showed levels of anxiety decreasing slightly over time among pregnant respondents but increasing slightly over time among postnatal respondents (see Table 3). Possible explanations for this finding are discussed in the following sections.

Fear of Contracting COVID-19

During April, when the pandemic was peeking in the United Kingdom, anxiety over contracting COVID-19 was high. Pregnant women expressed anxiety over the impact their contracting the virus might have on their unborn child. Those close to delivery were particularly anxious about contracting COVID-19 in hospital when giving birth. Indeed, over two-thirds of the pregnant web-based survey respondents (152/235, 64.7%) said that staying safe when giving birth was worrying them more than normal at the moment, and over half (125/235, 53.2%) were worried about staying safe at antenatal appointments (see Multimedia Appendix 7).

> *I'm scared of the unknown- how much does it affect pregnant women and unborn babies, will my new-born be at huge risk?* [Free text]

While the most common concern for postnatal respondents was their baby's health, qualitative insights showed that respondents' concerns did not revolve around their baby contracting the virus but around the potential impact on their baby should they themselves contract COVID-19, such as whether they would still be able to breastfeed and who would look after their baby.

> *I don't want my baby to catch anything while he is so little. I don't want to catch coronavirus and leave my child without a parent.* [Free text]

Qualitative findings suggested that for many respondents, fear of contracting COVID-19 diminished during May, perhaps because death rates fell and respondents felt safer because lockdown and social distancing measures were well-established.

> *I am happy with lockdown as I feel my baby and I are protected.* [Free text]

The exception was respondents whose risk of contracting COVID-19 was above average, including respondents who were of Black, Asian, or minority ethnicity, those with underlying health conditions, and those whose partners were health care workers.

> *It's a worrying time to be health care professionals and being pregnant. I worry that my husband will get coronavirus and infect me and that could result in FGR [fetal growth restriction] or worse.* [Free text]

Current Disruption

During the first weeks of the research (weeks 4-6 of lockdown), the respondents' focus appeared to be on current disruptions caused by the pandemic and proximate uncertainties such as accessibility to food and medicines, job insecurity, and financial worries. Current disruption seemed to be greater for pregnant respondents for a number of reasons. For some pregnant women, there was considerable confusion as to whether they should be going to work if employers were not necessarily aware of or adhering to government advice.

> *I'm a community care worker. I've had problems with work as they won't pay me sick pay and they haven't fired me, they just aren't giving me any shifts so I've lost my income.* [Pregnant 8 weeks, age 21-24 years, White British]

Many pregnant respondents stated that the changes to their antenatal care schedule were unsettling, with some appointments being cancelled. Some women struggled to make contact with their midwives, especially in the early days of lockdown. For those whose due date was imminent, disruption of birth plans was a major source of anxiety, with the single greatest fear being having to go through labor and give birth without their partner. However, later survey responses and interviews suggested that some of these fears had subsided as reports circulated of good birth experiences, quiet labor wards, and partners being allowed to be present. Although 69.0% (78/113) of pregnant respondents said they were more worried than usual about staying safe when giving birth in the first two weeks of the survey, this proportion dropped to 57.9% (22/38) in the last two weeks of the survey (see Multimedia Appendix 7).

Future Uncertainties

Over the course of the research, a shift appeared to occur from focusing on current disruption to future uncertainties. Postnatal respondents were particularly anxious about the longer-term impact of the pandemic, such as limited access to support services to monitor their baby's development, the impact of lockdown on their baby's socialization, and how both they and their baby would cope with returning to work. Whereas 42% (78/187) of all respondents said they were more worried than usual about caring for a new baby under the "stay at home advice" in the first two weeks of the survey, this had risen to 55% (44/80) in the survey's last two weeks (see Multimedia Appendix 7).

> *Worried about the world – thinking ahead how's it going to be – a massive change that we are all going to have to go through – it's quite an anxious time really.* [Baby 3 months, age 30-34 years, Black, Asian, minority ethnicity]

Emotional and Physical Impact of Reduced Support

Free text comments from the web-based survey and telephone interviews showed that the impact of loss of support on the respondents' mental and physical well-being was substantial. Several sources of support were identified by respondents: family and friends, health care professionals, and government.

Support from Family and Friends

Although modern technology enables video calls with family and friends, the respondents, especially those with young babies, felt that they were suffering considerably from the loss of face-to-face interaction. Without antenatal and postnatal groups and visits to family and friends, respondents felt that they were missing out on opportunities to share experiences and learn from others. Many reported increased usage of digital pregnancy and parenting resources, and some had found online support groups and classes. In practical terms, the loss of informal childcare support (such as grandparents babysitting) meant that mothers no longer had time to themselves for mental revival or physical exercise. Several postnatal respondents remarked that their intended postpartum fitness regime had been put on hold as a consequence of having no help with childcare.

A frequently voiced concern for postnatal respondents was the stress of 24/7 parenting. They were not only struggling with the strain of having no time to themselves but were also worrying about having to be more self-reliant and responsible for monitoring their baby's health and development. This impact was particularly magnified for those whose partners were working long hours at home or out of the home, as well as for single mothers.

> *My partner is working from home, so I am trying to keep him (baby) quiet. I'd just like a break. He is literally with me 100% of the time.* [Baby 6 months, age 35-39 years, White British]

For many, the loss of face-to-face interaction with family and friends created feelings of isolation and loneliness. For those with pre-existing mental health issues, this could be particularly challenging.

> *My maternity leave has not been what I had hoped. I have not been able to access help with the baby from family or friends because of social distancing. I have not had the benefits of meeting up with other mums and have felt isolated and very lonely at times.* [Free text]

Support from Health Care Professionals

The survey findings showed that for pregnant respondents, there had been an equal number of face-to-face and telephone midwife

appointments (face-to-face: 134/235, 57.0%, vs telephone: 131/235, 55.7%). For postnatal respondents, health visitor appointments had been largely by telephone (telephone: 107/188, 56.9%, vs face-to-face: 24/188, 12.8%). Over 53% of pregnant respondents (125/235, 53.2%) and postnatal respondents (110/188, 58.5%) reported that they were worrying more than usual about being able to see their health care professional if they needed to (see Multimedia Appendix 7). Many respondents reported the negative impact of reduced support from health care professionals, with face-to-face antenatal and postnatal appointments being replaced with telephone appointments or simply being cancelled.

Pregnant Respondents

In the telephone interviews and free text sections of the survey, pregnant respondents often expressed their concern and disappointment at receiving what they regarded as suboptimal antenatal care. Although some respondents pointed out that telephone appointments reduced their risk of contracting COVID-19, many worried about the aspects of care they might be missing. Indeed, a common sentiment was that of being disregarded or sidelined.

> *My antenatal care is now much more rushed and stressful. There is no time to talk to the midwife and it very much feels like a quick process to check my urine and blood pressure only. This makes me feel less connected and more anxious.* [Free text]

> *I've had days when I break down and cry. We aren't getting as much support from the midwives. I am out of vitamins and the children's centres are closed. I've had 6 different midwives and when I call them I never get a reply, just a message, sometimes saying they are no longer working.* [Pregnant 21 weeks, age 21-24 years, Black, Asian, minority ethnicity]

> *I want to hear my baby's heartbeat – a phone call's not the same.* [Pregnant 7 weeks, age 35-39 years, White other]

While there was no evidence that pregnant women themselves chose to avoid antenatal appointments, many talked about their initial reluctance to visit a hospital, especially since their partner was unable to accompany them. However, their experiences had generally been positive, and separate areas or entrances for antenatal patients had reassured respondents.

In most instances, reduced health care professional support did not appear to have impacted the safety or health of the pregnant women. However, in a number of more concerning cases, the absence of face-to-face consultation had impacted the respondents both mentally and physically.

> *I had a bladder infection. They were like you can't come in to the doctor's, so everything was by phone and they didn't test it and then I had to make a choice, if I wanted to just take the lower tablets which would cover the bladder infection but not kidney, or the kidney ones but the kidney ones had a chance of miscarriage … stuff like that I find quite scary.* [Pregnant, 13 weeks, age 30-34 years, White British]

Postnatal Respondents

Qualitative findings suggested that the reduced support from health care professionals had a greater impact on postnatal respondents. A widespread lack of face-to-face appointments with health visitors and general practitioners (GPs) meant that many respondents were worried about their baby's health and development.

> *They were worried that he wasn't putting on weight…he was referred to the GP for acid reflux but we only had a telephone consultation…. the health visitor isn't coming any more. When I asked what I could do, she said check whether he is growing out of his baby grows [all-in-one sleep suits].* [Baby 8 weeks, age 40-44 years, White British]

> *My baby was on a feeding plan but I can't have her weighed anymore and the doctor has diagnosed reflux over the phone but the medication we've tried isn't working. I feel quite tearful and less supported than at first.* [Free text]

Although some respondents had been able to speak to their health visitor on the telephone, others reported having no health visitor service at all. Concerns about the baby's weight gain were frequently voiced, as were breastfeeding issues by those who had given birth just before or during lockdown.

> *I was struggling breastfeeding. I would have gone to breastfeeding group, but that's been cancelled…. I was in pain and I felt let down, but I felt really guilty asking for help.* [Baby 7 weeks, age 30-34 years, White British]

> *Feeding is a challenge at the moment….it's been painful. I was engorged and he was crying, and I was upset that he was upset, so my partner went to get some formula, but Tesco's isn't 24 hours anymore.* [Baby 7 weeks, age 30-34 years, White British]

Several respondents talked about their threshold for seeking medical or developmental advice for their baby being higher, both as a result of their concerns about burdening an already stressed health service and the safety of visiting a GP surgery or hospital. The resulting greater self-reliance in monitoring and diagnosing issues with their baby was worrying and stressful.

> *You can't get things now – you can't get a thermometer…. if you are a bit anxious because the baby is unwell that ramps up 20 times because you can't do anything.* [Baby 7 weeks, age 30-34 years, White British]

> *No-one has seen him in 2 months. What if there's something he's doing that's not right?* [Baby 5 months, age 25-29 years, White British]

A small but meaningful number of respondents had experienced postponed or cancelled hospital appointments for themselves or their baby, which had caused substantial anxiety.

> *He's got a cows' milk allergy and kept being sick…with what's going on, I think he's been forgotten – I feel as if it's been left. We were supposed to have*

> *an appointment with the pediatrician in March, they were going to weigh him, but I still haven't had my appointment….I think its affected my mental health a bit. I feel overwhelmed.* [Baby 6 months, age 35-39 years, White British]

> *I've been recommended to attend the 'birth afterthoughts' service to deal with my feelings over difficult birth but am wary of attending any hospital appts that are not absolutely necessary. Similarly baby is supposed to be referred to paediatric cardiac specialist- I've heard nothing...and hesitant to chase for same reasons.* [Free text]

Many postnatal respondents questioned the cancellation of 6-week postnatal checks. Although for most respondents, this had little impact beyond reinforcing the sense of being disregarded by the NHS, a small number of respondents felt that their recovery had been hindered by not being able to see their GP.

> *I had an episiotomy and I'm still in pain 7 weeks later. The doctor said no six-week checks … eventually the doctor gave me a call but I haven't had an examination. I feel as if I have been short changed.* [Baby 7 weeks, age 30-34 years, White British]

> *I didn't get the aftercare I was expecting and I'm in agony but not able to get real help from the NHS. Just feels like I had my baby and then abandoned.* [Free text]

One exception to postnatal respondents' general dissatisfaction with health care services was baby vaccinations. Respondents reported positive experiences of visiting GP surgeries, remarking on reassuringly high levels of hygiene and social distancing precautions.

Support from Government

A lack of clear guidelines from the government led some respondents to feel that they were a forgotten sector of the population. Although initial guidelines had suggested that pregnant women should shield, there was confusion as to whether this meant staying indoors or being allowed out to exercise and whether it was still applicable (by May) or had been superseded by new advice. Postnatal respondents were not aware of any advice directed at them as parents of young babies, and many were critical of its absence.

> *I just feel like new moms and parents have just been completely left out of the Government's mindset and their exit strategy really.* [Baby 3 months, age 30-34 years, White British]

> *I don't think I am being supported. The Government should be giving us more information, like where do you go to get your baby weighed…there were 8 weeks when she wasn't weighed – quite a worry really.* [Baby 12 weeks, age 30-34 years, White British]

As talk of easing lockdown restrictions began (mid-May), postnatal respondents were particularly frustrated by the lack of consideration given to their circumstances.

> *I would like someone to be representing our needs to politicians so that priority is given to measures like "bubbles" that would help new families SO much rather than people being allowed to have a cleaner come!* [Free text]

Life Under Lockdown

Unsurprisingly, lockdown had an immense impact on all aspects of the respondents' lives. Within the context of pregnancy and parenting, four subthemes emerged from the qualitative analysis: missing out, lifestyle disruption, couple relationships, and impact on baby.

Missing Out

The greatest disappointment for pregnant respondents appeared to be that their partners were missing out on antenatal appointments. Women worried that less involvement might adversely affect their partner's bonding with the baby. Missing out on antenatal groups was also disappointing for women, although some were attending existing groups on the web or had found new online support groups or forums to compensate.

> *We've looked at the videos (on Baby Buddy) of the baby moving in the womb. It's good for my husband as he can't go to my scans. He hasn't even heard the heartbeat …. I worry he isn't bonding with the bump.* [Pregnant 21 weeks, age 21-24 years, Black, Asian, minority ethnicity]

The sense of missing out was expressed much more strongly by postnatal respondents who felt that they were being deprived of all the fun activities they had looked forward to doing during maternity leave, such as play dates, baby classes, and seeing more of family and friends. Some respondents thought a UK-wide extension to maternity leave was needed, while those who were able to extend their maternity leave were considering doing so.

> *I'm due to go back in November, but I am thinking of taking the full 12 months as I haven't been able to do stuff with her.* [Baby 3 months, age 30-34 years, Black, Asian, minority ethnicity]

A small minority expressed more positive views of the benefits of a quieter life.

> *No pressure to go to groups, no need to go places at specific times, no stress of public feeding or changing nappies in public toilets. He has 100% of my attention and is very happy and settled.* [Free text]

Lifestyle Disruption

Both pregnant and postnatal respondents talked about disruption to routine adversely affecting their physical health. Their physical health was the third (106/235, 45.1%) and fourth (65/188, 34.6%) most highly rated concern at the moment (see Multimedia Appendix 6) for pregnant and postnatal respondents, respectively. Most people were getting far less exercise than normal, as work and activities outside the home had ceased and many were reluctant to go outside at all due to fear of contracting the virus. This was particularly true for those living

in urban areas, where it was difficult to conform to social distancing rules.

> *I don't get out of the house for days, weeks on end….there are too many people outside….it's getting difficult walking from room to room.* [Pregnant 34 weeks, age 25-29 years, Black, Asian, minority ethnicity]

> *We're not getting out much – we are super cautious because of the hole in her heart….so many people around it wouldn't have been possible to do any distancing.* [Baby 16 weeks, age 30-34 years, White British]

At the time of the research, the government guideline was to go out to exercise only once per day, and many parents were choosing to take their baby out over exercising themselves. For some respondents, the lack of physical activity had become self-perpetuating, with energy levels decreasing to a concerningly low level.

> *I hadn't been out (from mid-March) until the end of April. I went for a very short walk – I was really tired.* [Baby 4 months, age 30-34 years, White other]

Similarly, many respondents talked about changes to their diet. On the positive side, some respondents reported doing more home cooking during lockdown, although they were more likely to bake than cook nutritious meals. However, 35.3% (83/235) of pregnant respondents and 28.2% (53/188) of postpartum respondents were concerned about healthy eating. A reduction in consumption of fruit and vegetables was reported widely in the telephone interviews, as many respondents were attempting to visit shops as infrequently as possible and delegating grocery shopping to their partner. Reports of more frequent snacking and treat-eating due to boredom, misery, or simply a lack of other treat options were widespread.

> *Physically I am a wreck. My back hurts, I never do any exercise as my baby doesn't sleep during the day. My ligaments hurt, my breasts hurt. I want to lose some weight, but I am stress eating – I'm not hungry I just eat because I want something sweet.* [Baby 16 weeks, age 30-34 years, White other]

Couple Relationships

Here, experiences were divisive; many respondents with a partner working from home really appreciated the additional time spent together. Postnatal respondents pointed out the practical benefits of sharing childcare and household tasks as well as the emotional benefits of the greater opportunity to bond as a family.

> *My baby is having the best start to life! Full undivided access to her mummy and her daddy!! For 2 months!* [Free text]

However, for a small number of respondents, lockdown had created tensions within the couple's relationship. Those whose partners were working hard, either from home or outside, could find themselves resenting their partner's seemingly "normal" lives when they were feeling isolated, alone, and bored. Respondents also talked about arguments about levels of hygiene and adherence to lockdown guidelines; this seemed more prevalent in couples in which the partner was working outside the home.

> *There have been a couple of occasions where there have been cross words…. I am sat in this house all day alone ….with the hours he is out of the house I don't have another adult to speak to – it's made it a bit more isolating.* [Baby 7 weeks, age 30-34 years, White British]

Impact on Baby

Less than half of pregnant respondents (47%, 89/190) felt that the changes that the pandemic was causing in their lives were affecting their baby. Several telephone interviewees expressed concern that their stress might transfer to their baby in the womb and adversely affect its development. In contrast, over three-quarters of postnatal respondents (76%, 118/155) felt that the changes were affecting their baby, and this high level of concern was voiced strongly by telephone interviewees. They worried that their baby was missing out on interactions with other people, experiences, and activities, which would have a detrimental impact on the baby's development and socialization. Respondents frequently expressed concerns that their baby might forget their grandparents and other family members or become clingy and reluctant to interact with others.

> *My baby has been unable to meet family and I am anxious that she will be fearful of others when it comes to the time she needs to go to nursery.* [Free text]

Lockdown had also affected some babies' routines, with changes in sleeping patterns reported most frequently. Although some postnatal respondents viewed lockdown to be an ideal time to develop new routines, others found it challenging to cope with the changes and felt that their baby had become more unsettled and cried more frequently.

> *The only way I used to be able to get him to nap at lunch time was to go for a walk in the buggy … I was unable to go out at all at the start because he had a temperature so I couldn't get him to sleep so there were about 3 weeks that were kind of hellish.* [Baby 6 months, age 35-39 years, White British]

Use of the Baby Buddy App During the Pandemic

While the majority of survey and telephone interview respondents felt they were using the Baby Buddy app about the same amount during the pandemic (pregnant: 70.5%, 172/244; postnatal: 74.5%, 143/192), nearly a quarter of the survey respondents said they were using the app more (pregnant 25%, 61/244; postnatal 19.8%, 38/192). Comments from the free text boxes and telephone interviewees showed that many women were finding Baby Buddy to be particularly valuable in the absence of support from health care professionals and baby groups.

> *I have had very little contact with my midwife and I really like how much information I can get from the app given that I can't seem to get any support from anywhere else.* [Free text]

> *Without groups to go to I've found having these apps to be even more useful and helpful.* [Free text]
>
> *I have enjoyed using Baby Buddy throughout pregnancy and these early months and find it even more reassuring now that external input is so limited.* [Baby 7 weeks, age 30-34 years, White British]
>
> *I look forward to the little bit of advice every day. It is written in such a non-judgemental way, encouraging but not patronising. I feel like it's giving me some of the general knowledge I'd be picking up from other mums at baby groups.* [Free text]

Table 4 summarizes the ways in which Baby Buddy helped its users, with 90.0% (392/436) of respondents identifying at least one way in which the app was helping them during the pandemic. Access to reliable information was cited by 82.6% (360/436) of survey respondents, and telephone interviewees frequently remarked that the main strength of the app was the provision of trustworthy and easy-to-understand information. Most interview respondents had either been recommended Baby Buddy by their midwife or had seen a poster or leaflet for the app at an antenatal appointment. Therefore, they felt confident that the app met with NHS approval and that its content would be accurate and trustworthy.

> *Because it's linked to the NHS it's not like googling.* [Baby 3 months, age 30-34 years, Black, Asian, minority ethnicity]
>
> *Reliable, trustworthy and ad-free source of information. Great videos.* [Free text]

Table 4. Proportions of respondents' answers to the question "How is Baby Buddy helping you at the moment? (Tick all that apply to you)" (N=436), n (%).

Type of help	Pregnant (n=244)	Postnatal (n=192)	Total (N=436)
Giving me access to reliable information	193 (79.1)	167 (87.0)	360 (82.6)
Helping me bond with my baby	60 (24.6)	64 (33.3)	124 (28.4)
Helping me with my emotional and mental health	50 (20.5)	53 (27.6)	103 (23.6)
Helping me with my physical health	26 (10.7)	27 (14.1)	53 (12.2)
Helping me with my relationships	19 (7.8)	20 (10.4)	39 (8.9)
Baby Buddy isn't helping me	31 (12.7)	13 (6.8)	44 (10.1)
Helping me in other ways	24 (9.8)	23 (12.0)	46 (10.6)

Of the respondents who reported which Baby Buddy features they were finding most helpful at the moment, 97% (377/436) identified Today's Information (see Multimedia Appendix 8). Similarly, interview respondents consistently praised the Today's Information feature for being interesting, useful, and time-appropriate. Indeed, several respondents remarked on the seemingly uncanny timing of the messages.

> *It's reliable – like the people who are writing it are knowledgeable. Today's message is always more or less in line with what I am thinking.* [Baby 16 weeks, age 30-34 years, White other]
>
> *I'm not the best at researching stuff and reading a big book in advance so this gives me nice little bite-size chunks of things I need to know.* [Pregnant 16 weeks, age 40-44 years, White British]

Other features were helping a smaller number of Baby Buddy users at the moment. This could be explained by the finding that interview respondents frequently reported that they rarely explored the app beyond the Today's Information section. For those who did, the videos were singled out as being very helpful. Typically, respondents had viewed videos that had been highlighted in Today's Information. Fewer respondents had explored the videos by category, although those who had appreciated the number and variety of videos available.

> *The information is so useful – I share it with my husband. We've looked at the videos – the baby moving in the womb. It's really good for my husband because he didn't get to go to the scan and I worry he is not bonding with the bump. He hasn't heard the heartbeat.* [Pregnant 21 weeks, age 21-24 years, Black, Asian, minority ethnicity]
>
> *I like to see the videos – they've helped a lot to see that other women are struggling too.* [Baby 16 weeks, age 30-34 years, White other]

Other app features specifically praised by interview respondents included Ask Me, Your Appointments, You Can Do It, and You and Your Partner. Only one respondent was aware of Get Help. The idea of a 24/7 text helpline was widely appreciated; however, Get Help was not necessarily felt to be a good descriptor of this feature.

> *I love the partner's advice. A lot of apps don't bring the partner into it…it's brilliant. It's (pregnancy) a stressful time and two people don't always react the same.* [Pregnant 8 weeks, age 21-24 years, White British]

The main criticism from the 10.1% of users (44/436) who felt that Baby Buddy was not helping them during the pandemic was its slowness to adapt to the new situation. Recommending courses of action that were no longer available under current circumstances, such as ask your health visitor, join a play group, and have a baby shower, were both unhelpful and a sad reminder of activities that respondents were missing out on.

> *One of the things that came through was a video saying keep your distance from people with colds and 'flu and there was this pregnant woman sitting on a bench next to a man coughing. I laughed at it, but not*

> *to acknowledge when we are in a middle of a pandemic is a bit of a blunder I would say.* [Pregnant 16 weeks, age 40-44 years, White British]
>
> *Sometimes I won't check because it reminds me of all the things I'd like to be doing but can't because of social distancing.* [Free text]

Views on whether Baby Buddy should be providing COVID-19–specific information were mixed. While some respondents felt that this information should be prominent within the app, others expressed a keen desire that Baby Buddy remain a safe haven from the pandemic. Clarity around the guidelines and links to websites providing recommendations and advice were generally thought to be helpful.

Several respondents identified additional information and advice needs as a consequence of the pandemic. This included how to ensure your baby is gaining the right amount of weight and meeting developmental milestones (in the absence of health visitors); ideas for stimulating and entertaining babies (in the absence of classes); ways to calm a crying or fractious baby; ways to reduce one's own anxiety and stay calm; and making sure you are getting enough vitamin D with limited time outdoors.

> *Things to do with him when you are stuck inside – keeping him entertained, what's best for his development.* [Baby 5 months, age 25-29 years, White British]

An additional suggestion was the inclusion of self-made videos from Baby Buddy users with their tips for coping with pregnancy, birth, and parenting a very young baby during the pandemic.

Beyond content relating to the current circumstances, another notable demand for additional information came from respondents who had one or more other children. These respondents were keen to receive advice specific to their circumstances, such as how to cope with tiredness in pregnancy when you have an active toddler, how to prepare a child for the birth of their sibling, and how to avoid sibling rivalry.

> *The thing that I would find even more useful would be if they designed a section for 2nd or 3rd time mums…how to explain to your kids they've got another brother or sister coming, things like that.…that would be amazing.* [Pregnant 34 weeks, age 30-34 years, White other]

First-time mothers suggested adding new functionalities, usually those they had found in other parenting apps. These included features to allow mothers to record feeding information (eg, time and length of feed, which breast) or diaper records and chat rooms for parents to share tips and provide mutual support.

Discussion

Principal Findings

This service evaluation study aimed to understand the attitudes, experiences, and needs of Baby Buddy app users in the United Kingdom during the COVID-19 pandemic, with a view to identifying ways in which Baby Buddy could further support its users. The web-based survey in conjunction with the follow-up telephone interviews revealed heightened levels of anxiety and stress around being pregnant, giving birth, and parenting a new baby during the pandemic. Reduced support, especially from health care professionals, and the impact of the lockdown clearly affected the mental and physical well-being of pregnant women and parents of young babies. Baby Buddy users appreciated having access to a source of easily accessible, reliable, and trustworthy advice, and some users felt that the app helped to compensate for the loss of face-to-face advice and support. A small minority felt that greater speed in adapting the content to reflect the new situation would have been beneficial.

Reports of insufficient support for new parents in the United Kingdom with health and parenting issues have been reported elsewhere [13]. This research showed that with more restricted access to health care services and support, accessing reliable and trustworthy information has become a key concern for pregnant women and those with young babies. Previous research has highlighted concerns women have about the reliability of pregnancy and parenting advice on the web [14] as well as the importance of trustworthiness in health apps in general [15] and pregnancy apps in particular [16]. As an NHS-approved app, and one that is often recommended to users by their health care professionals, Baby Buddy is well positioned to meet this increased demand for digital pregnancy and parenting support.

The heightened levels of anxiety and stress during the COVID-19 pandemic found among Baby Buddy users are consistent with research among the general UK population [17] and perinatal populations from other countries [18]. A study of over 4000 pregnant women in China reported a significantly higher rate of depressive symptoms (measured on the Edinburgh Postnatal Depression Scale), suggesting an increased risk of mental illness, including self-harm behaviors [19]. Studies in Turkey [20] and Canada [21,22] have also found higher rates of anxiety and depression in pregnant women during the pandemic and called for clinical surveillance and psychosocial support to mitigate adverse impacts on mothers and babies. One study was able to compare levels of anxiety and stress in the same pregnant women before and during the pandemic, and it was found that the COVID-19 pandemic is responsible for significant increases in stress [23]. There is an established body of evidence that anxiety and stress in pregnancy are risk factors for adverse outcomes for mother and baby, including longer term cognitive, emotional, and behavioral development issues in childhood [24]. Studies have indicated negative affect to be associated with preterm birth and low birth weight [24], although some surprising data are emerging showing dramatic reductions in preterm births from January to April 2020 compared to previous years [25,26]. Maternal anxiety and stress has also been shown to predict postnatal depression, potentially leading to impaired parenting quality and effectiveness [27,28]. Although face-to-face care from maternity and mental health services is clearly the optimum solution, Baby Buddy is in a position to provide support to expectant and new parents through its existing content, including the perinatal mental health films, the feature for couples (You and Your Partner), the Get Help feature, signposting to many other charities, and the Baby Buddy

Crisis Messenger for 24/7 confidential and clinically assured support. The potential role of web-based resources in ameliorating COVID-19–related anxieties in perinatal women is echoed in a study among obstetricians in India [29].

This study also found physical health to be an important concern, especially among pregnant women. With lockdown limiting women's ability to exercise, restricting their access to healthy foods, and leading to more snacking on foods high in fat, sugar, and salt, there may be a long-term impact on mothers and babies as a result of increased risks associated with excessive gestational weight gain [30,31]. These findings are consistent with data from the general UK population [32,33]. Once again, Baby Buddy can support its users by directing them to existing film and written content on exercising at home and healthy meal and snacking ideas.

The vast majority (90%) of Baby Buddy users in this study reported finding the app helpful during this period, especially given the reduced support from health care professionals, family, and friends. However, while Baby Buddy provides a wealth of information to support its users, many users were not finding all this information. It is especially important during these times of reduced support and heightened anxiety to actively encourage Baby Buddy users to explore beyond the daily information feature, Today's Information, to find additional content and functionality. The addition of new video content featuring expectant and new parents during the COVID-19 pandemic may help to support users by normalizing feelings and experiences they are having during these unprecedented times. It may also help reduce feelings of "missing out" by exemplifying that many other young families are facing the same challenges.

With the pandemic limiting access to their usual support networks and health care professionals, Baby Buddy users have had to become more self-reliant in their roles of parenting a young baby. This has created additional information needs, particularly around baby stimulation, socialization, growth, and development. Although the app already contains written and film content covering these issues, adding new content to further meet these needs could improve the support Baby Buddy provides to its users and help increase new parents' confidence in monitoring their baby's well-being and development while reduced access to professional support continues.

This study was a precursor to a UK-wide survey of expectant parents and parents of children under 2 years of age, Babies in Lockdown, commissioned by Best Beginnings, Home-Start UK, and the Parent-Infant Foundation and undertaken by Critical Research [10]. The findings of the two studies are consistent and aligned. With a sample size of over 5000 and more demographic data, Babies in Lockdown was able to identify differences in response according to age, income, and ethnicity. This survey concluded that families already at risk (lower income, younger age, and from Black, Asian, minority ethnicity communities) have suffered the most as a result of the pandemic and the associated lockdown. Previous research has reported that the COVID-19 pandemic may be accentuating inequalities [34]. With one of Baby Buddy's strengths being its disproportionately high usage by Black, Asian, minority ethnicity, and ESL parents and Best Beginnings' remit of reducing inequalities in child health care, the Baby Buddy app has an important and pertinent role to play as the COVID-19 pandemic continues.

Strengths and Limitations

A strength of this study was that it was conducted during the height of the COVID-19 pandemic, with data collection beginning within three weeks of the start of lockdown in the United Kingdom. As a mixed methods study, the research was able to deliver statistically robust findings alongside detailed qualitative insights.

A limitation of this study is because the survey focused on Baby Buddy users, the findings cannot be generalized to all pregnant and new parents in the United Kingdom, although its findings were highly consistent with the follow-up survey described above [10]. A further limitation of this study is self-selection bias, as respondents chose whether to respond to the survey and take part in a telephone interview. Moreover, anxiety, stress, and depression levels were self-reported rather than being measured using validated scales. A final limitation of this study was the decision to collect only age and language demographics in the web-based survey. At the time the study was designed, there was no awareness of the disproportionate risk of poor COVID-19 outcomes among Black, Asian, and minority ethnicity communities [35], and the main consideration was to reduce barriers to completion. This omission was amended for the telephone interviews and in the follow-up study referred to previously [10].

Conclusion

The COVID-19 pandemic has created heightened anxiety and stress among expectant parents and those with a young baby. This study found that for many expectant and new parents, lockdown has had an adverse impact on their physical and mental well-being. With reductions in health and social supports, expectant and new parents are relying even more on web-based resources. As a free, evidence-based app, Baby Buddy is well positioned to meet this need. The app could support its users even more by actively directing them to the wealth of existing content relevant to their concerns and by adding content to give users the knowledge and confidence to meet the new challenges facing them. As the longer-term impact of the COVID-19 pandemic becomes apparent, it will be necessary to add to and update the content of the Baby Buddy app to meet new needs.

Acknowledgments

This work was supported by the Economic and Social Research Fund (ES/P000592/1) and the UK charity Best Beginnings (charity number 1120054).

Conflicts of Interest

AR is supported by the ESRC through a UCL, Bloomsbury, East London Doctoral Training Partnership (UBEL DTP). The studentship is cofunded by the UK charity Best Beginnings (charity number 1120054).

Multimedia Appendix 1
Baby Buddy Features.
[PPTX File , 5138 KB - mhealth_v8i12e23157_app1.pptx]

Multimedia Appendix 2
Topic Guides.
[DOCX File , 23 KB - mhealth_v8i12e23157_app2.docx]

Multimedia Appendix 3
Recruitment message.
[DOCX File , 12 KB - mhealth_v8i12e23157_app3.docx]

Multimedia Appendix 4
Coding Framework.
[DOCX File , 21 KB - mhealth_v8i12e23157_app4.docx]

Multimedia Appendix 5
Theme definitions.
[DOCX File , 17 KB - mhealth_v8i12e23157_app5.docx]

Multimedia Appendix 6
Respondents' main concerns right now.
[DOCX File , 16 KB - mhealth_v8i12e23157_app6.docx]

Multimedia Appendix 7
Healthcare issues worrying respondent more than usual.
[DOCX File , 15 KB - mhealth_v8i12e23157_app7.docx]

Multimedia Appendix 8
Baby Buddy features.
[DOCX File , 15 KB - mhealth_v8i12e23157_app8.docx]

References

1. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
2. Hughson J, Daly J, Woodward-Kron R, Hajek J, Story D. The Rise of Pregnancy Apps and the Implications for Culturally and Linguistically Diverse Women: Narrative Review. JMIR mHealth uHealth 2018 Nov 16;6(11):e189 [FREE Full text] [doi: 10.2196/mhealth.9119] [Medline: 30446483]
3. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]
4. Lee Y, Moon M. Utilization and Content Evaluation of Mobile Applications for Pregnancy, Birth, and Child Care. Healthc Inform Res 2016 Apr;22(2):73-80 [FREE Full text] [doi: 10.4258/hir.2016.22.2.73] [Medline: 27200216]
5. Wang N, Deng Z, Wen L, Ding Y, He G. Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study. JMIR Mhealth Uhealth 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: 10.2196/12631] [Medline: 31215516]
6. Smartphone ownership penetration in the United Kingdom (UK) in 2012-2020, by age. Statista. 2020 May 25. URL: https://www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age/ [accessed 2019-07-12]
7. Baby Buddy App. Best Beginnings. URL: https://www.bestbeginnings.org.uk/baby-buddy [accessed 2019-11-23]
8. Social determinants of health. World Health Organization. URL: http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review [accessed 2020-12-04]
9. Impact and evaluation of Baby Buddy app. Best Beginnings. URL: https://www.bestbeginnings.org.uk/evidence-impact-and-evaluation [accessed 2020-07-06]

10. Babies in Lockdown. URL: https://babiesinlockdown.info/ [accessed 2020-07-31]
11. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
12. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
13. Vazquez-Vazquez A, Dib S, Rougeaux E, Wells J, Fewtrell M. The impact of the Covid-19 lockdown on the experiences and feeding practices of new mothers in the UK: Preliminary data from the COVID-19 New Mum Study. Appetite 2021 Jan 01;156:104985 [FREE Full text] [doi: 10.1016/j.appet.2020.104985] [Medline: 33038477]
14. Javanmardi M, Noroozi M, Mostafavi F, Ashrafi-Rizi H. Internet Usage among Pregnant Women for Seeking Health Information: A Review Article. Iran J Nurs Midwifery Res 2018;23(2):79-86 [FREE Full text] [doi: 10.4103/ijnmr.IJNMR_82_17] [Medline: 29628953]
15. van Haasteren A, Gille F, Fadda M, Vayena E. Development of the mHealth App Trustworthiness checklist. Digit Health 2019 Nov 21;5:2055207619886463 [FREE Full text] [doi: 10.1177/2055207619886463] [Medline: 31803490]
16. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. Matern Child Nutr 2019 Jul 06;15(3):e12768 [FREE Full text] [doi: 10.1111/mcn.12768] [Medline: 30569549]
17. Fancourt D, Steptoe A, Dphil D, Bu F. Trajectories of depression and anxiety during enforced isolation due to COVID-19: longitudinal analyses of 59,318 adults in the UK with and without diagnosed mental illness. medRxiv. Preprint posted online on November 03, 2020. [doi: 10.1101/2020.06.03.20120923]
18. Genova VL, Lucas E, Camargo2 J, Schneider1 G, Pereira G, Da SR, et al. Repercussions of the COVID-19 Pandemic on the Mental Health of Pregnant and Puerperal Women: A Systematic Review. medRxiv. Preprint posted online on April 21, 2020. [doi: 10.1101/2020.08.17.20176560]
19. Wu Y, Zhang C, Liu H, Duan C, Li C, Fan J, et al. Perinatal depressive and anxiety symptoms of pregnant women during the coronavirus disease 2019 outbreak in China. Am J Obstet Gynecol 2020 Aug;223(2):240.e1-240.e9 [FREE Full text] [doi: 10.1016/j.ajog.2020.05.009] [Medline: 32437665]
20. Durankuş F, Aksu E. Effects of the COVID-19 pandemic on anxiety and depressive symptoms in pregnant women: a preliminary study. J Matern Fetal Neonatal Med 2020 May 18:1-7. [doi: 10.1080/14767058.2020.1763946] [Medline: 32419558]
21. Berthelot N, Lemieux R, Garon-Bissonnette J, Drouin-Maziade C, Martel É, Maziade M. Uptrend in distress and psychiatric symptomatology in pregnant women during the coronavirus disease 2019 pandemic. Acta Obstet Gynecol Scand 2020 Jul 03;99(7):848-855. [doi: 10.1111/aogs.13925] [Medline: 32449178]
22. Davenport M, Meyer S, Meah V, Strynadka M, Khurana R. Moms Are Not OK: COVID-19 and Maternal Mental Health. Front Glob Womens Health 2020 Jun 19;1. [doi: 10.3389/fgwh.2020.00001]
23. Ayaz R, Hocaoğlu M, Günay T, Yardımcı OD, Turgut A, Karateke A. Anxiety and depression symptoms in the same pregnant women before and during the COVID-19 pandemic. J Perinat Med 2020 Nov 26;48(9):965-970. [doi: 10.1515/jpm-2020-0380] [Medline: 32887191]
24. Dunkel Schetter C, Tanner L. Anxiety, depression and stress in pregnancy: implications for mothers, children, research, and practice. Curr Opin Psychiatry 2012 Mar;25(2):141-148 [FREE Full text] [doi: 10.1097/YCO.0b013e3283503680] [Medline: 22262028]
25. Hedermann G, Hedley P, Bækvad-Hansen M, Hjalgrim H, Rostgaard K, Poorisrisak P, et al. Danish premature birth rates during the COVID-19 lockdown. Arch Dis Child Fetal Neonatal Ed 2020 Aug 11:a [FREE Full text] [doi: 10.1136/archdischild-2020-319990] [Medline: 32788391]
26. Philip R, Purtill H, Reidy E, Daly M, Imcha M, McGrath D, et al. Unprecedented reduction in births of very low birthweight (VLBW) and extremely low birthweight (ELBW) infants during the COVID-19 lockdown in Ireland: a 'natural experiment' allowing analysis of data from the prior two decades. BMJ Glob Health 2020 Sep;5(9):a [FREE Full text] [doi: 10.1136/bmjgh-2020-003075] [Medline: 32999054]
27. Milgrom J, Gemmill AW, Bilszta JL, Hayes B, Barnett B, Brooks J, et al. Antenatal risk factors for postnatal depression: a large prospective study. J Affect Disord 2008 May;108(1-2):147-157. [doi: 10.1016/j.jad.2007.10.014] [Medline: 18067974]
28. Bos SC, Macedo A, Marques M, Pereira AT, Maia BR, Soares MJ, et al. Is positive affect in pregnancy protective of postpartum depression? Braz J Psychiatry 2013 Mar;35(1):5-12 [FREE Full text] [doi: 10.1016/j.rbp.2011.11.002] [Medline: 23567594]
29. Nanjundaswamy M, Shiva L, Desai G, Ganjekar S, Kishore T, Ram U, et al. COVID-19-related anxiety and concerns expressed by pregnant and postpartum women-a survey among obstetricians. Arch Womens Ment Health 2020 Aug 25 [FREE Full text] [doi: 10.1007/s00737-020-01060-w] [Medline: 32839898]
30. Ren M, Li H, Cai W, Niu X, Ji W, Zhang Z, et al. Excessive gestational weight gain in accordance with the IOM criteria and the risk of hypertensive disorders of pregnancy: a meta-analysis. BMC Pregnancy Childbirth 2018 Jul 04;18(1):281 [FREE Full text] [doi: 10.1186/s12884-018-1922-y] [Medline: 29973178]

31. Josey MJ, McCullough LE, Hoyo C, Williams-DeVane C. Overall gestational weight gain mediates the relationship between maternal and child obesity. BMC Public Health 2019 Aug 07;19(1):1062 [FREE Full text] [doi: 10.1186/s12889-019-7349-1] [Medline: 31391077]
32. Khan MA, Moverley Smith JE. "Covibesity," a new pandemic. Obes Med 2020 Sep;19:100282 [FREE Full text] [doi: 10.1016/j.obmed.2020.100282] [Medline: 32835125]
33. The silent pandemic: How lockdown is affecting future health. COVID Symptom Study. 2020 Jul 28. URL: https://covid.joinzoe.com/post/lockdown-weight-gain [accessed 2020-08-01]
34. Wright L, Steptoe A, Fancourt D. Are we all in this together? Longitudinal assessment of cumulative adversities by socioeconomic position in the first 3 weeks of lockdown in the UK. J Epidemiol Community Health 2020 Sep 05;74(9):683-688 [FREE Full text] [doi: 10.1136/jech-2020-214475] [Medline: 32503892]
35. Pan D, Sze S, Minhas J, Bangash M, Pareek N, Divall P, et al. The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. EClinicalMedicine 2020 Jun;23:100404 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100404] [Medline: 32632416]

Abbreviations

ESL: English as a second language
FGR: fetal growth restriction
GDPR: General Data Protection Regulation
GP: general practitioner
NHS: National Health Service
OCD: obsessive-compulsive disorder
UCL: University College London

Edited by L Buis; submitted 04.08.20; peer-reviewed by C Jacob, C Smith; comments to author 02.10.20; revised version received 08.10.20; accepted 19.10.20; published 09.12.20.

Please cite as:
Rhodes A, Kheireddine S, Smith AD
Experiences, Attitudes, and Needs of Users of a Pregnancy and Parenting App (Baby Buddy) During the COVID-19 Pandemic: Mixed Methods Study
JMIR Mhealth Uhealth 2020;8(12):e23157
URL: http://mhealth.jmir.org/2020/12/e23157/
doi:10.2196/23157
PMID:33264100

Original Paper

Improving Heart Disease Risk Through Quality-Focused Diet Logging: Pre-Post Study of a Diet Quality Tracking App

Bum Chul Kwon[1], PhD; Courtland VanDam[2], PhD; Stephanie E Chiuve[3,4], SCD; Hyung Wook Choi[5], MA; Paul Entler[6], DO; Pang-Ning Tan[7], PhD; Jina Huh-Yoo[5], PhD

[1]IBM Research, Cambridge, MA, United States

[2]MIT Licoln Laboratory, Lexington, MA, United States

[3]AbbVie Inc, Chicago, IL, United States

[4]Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

[5]College of Computing and Informatics, Drexel University, Philadelphia, PA, United States

[6]Sparrow Health System, Lansing, MI, United States

[7]Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, United States

Corresponding Author:
Jina Huh-Yoo, PhD
College of Computing and Informatics
Drexel University
3675 Market St
Philadelphia, PA, 19104
United States
Phone: 1 2158952474
Email: jinahuhyoo@drexel.edu

Abstract

Background: Diet-tracking mobile apps have gained increased interest from both academic and clinical fields. However, quantity-focused diet tracking (eg, calorie counting) can be time-consuming and tedious, leading to unsustained adoption. Diet quality—focusing on high-quality dietary patterns rather than quantifying diet into calories—has shown effectiveness in improving heart disease risk. The Healthy Heart Score (HHS) predicts 20-year cardiovascular risks based on the consumption of foods from quality-focused food categories, rather than detailed serving sizes. No studies have examined how mobile health (mHealth) apps focusing on diet quality can bring promising results in health outcomes and ease of adoption.

Objective: This study aims to design a mobile app to support the HHS-informed quality-focused dietary approach by enabling users to log simplified diet quality and view its real-time impact on future heart disease risks. Users were asked to log food categories that are the main predictors of the HHS. We measured the app's feasibility and efficacy in improving individuals' clinical and behavioral factors that affect future heart disease risks and app use.

Methods: We recruited 38 participants who were overweight or obese with high heart disease risk and who used the app for 5 weeks and measured weight, blood sugar, blood pressure, HHS, and diet score (DS)—the measurement for diet quality—at baseline and week 5 of the intervention.

Results: Most participants (30/38, 79%) used the app every week and showed significant improvements in DS (baseline: mean 1.31, SD 1.14; week 5: mean 2.36, SD 2.48; 2-tailed *t* test t_{29}=−2.85; *P*=.008) and HHS (baseline: mean 22.94, SD 18.86; week 4: mean 22.15, SD 18.58; t_{29}=2.41; *P*=.02) at week 5, although only 10 participants (10/38, 26%) checked their HHS risk scores more than once. Other outcomes, including weight, blood sugar, and blood pressure, did not show significant changes.

Conclusions: Our study showed that our logging tool significantly improved dietary choices. Participants were not interested in seeing the HHS and perceived logging diet categories irrelevant to improving the HHS as important. We discuss the complexities of addressing health risks and quantity- versus quality-based health monitoring and incorporating secondary behavior change goals that matter to users when designing mHealth apps.

(JMIR Mhealth Uhealth 2020;8(12):e21733) doi:10.2196/21733

KEYWORDS

mHealth; diet monitoring; diet tracking; food tracking; CVD; heart disease risk; health risk communication; human-computer interaction; user study; mobile phone

Introduction

Background

An increasing number of mobile apps have explored ways to monitor and improve health behavior efficiently and effectively [1-3]. Among these mobile health (mHealth) apps, diet monitoring is one of the most popular domains, as diabetes and obesity are shown to lead the top 2 fields producing revenue in the mHealth market [4]. A systematic review of mobile apps showed that mHealth apps on obesity and nutrition increased adherence to diet monitoring and effectively improved primary clinical outcomes, such as weight loss and maintenance of reduced blood glucose levels [1,5].

However, for effective dietary monitoring, ideally, users would need to quantify their food consumption down to the level of the number of grams of each nutrient [6]. Focusing on such a quantification of diet can bring several challenges. Food journaling can be *too much effort, time-consuming, or tedious* [7,8]. Food journaling with detailed entries can be challenging as users often might not remember or know what and how much they have eaten [9]. Users also feel that the dietary information in the database is unreliable, calories burnt seemed random and *did not line up* [10], and entering unhealthy food consumption in detail makes people feel guilty in general. As these barriers lead to limited engagement with diet-tracking apps, researchers have attempted lightweight approaches of diet tracking, and such attempts have been shown to be successful by providing users with a photo-based food-tracking app and encouraging them to track only one food type per day [11].

A 2018 study published in the *Journal of American Medical Association* showed the effectiveness of focusing on diet quality over quantity—with a focus on restricting low-quality foods, such as processed foods, added sugar, or refined grains—rather than calorie counting [12]. However, mobile apps on dietary monitoring focused on the quantification of diet (eg, calorie counting) and other health behaviors (eg, steps). This quantification approach does not necessarily address the needs of broader groups of individuals. Numeracy and literacy in general can be barriers. People show increased confusion around the serving size [13]; however, for these apps to work appropriately, it would require accurate calculations of these very nuanced behavior choices. For instance, one might have eaten a mixed salad, but the system needs to know how many grams of spinach versus carrots and which salad dressing were consumed to calculate accurate calories and nutritional content. Sophisticated, detailed, quantified tracking practices are not popular for all user groups [14]. Tracking detailed health information is a user burden, affecting sustained tracking behavior [8].

The effectiveness of mHealth includes seeing the effect of behavior change. The knowledge of risk level helps individuals understand how urgently they need to change their behavior. Individuals at higher risk are more motivated to change if they know that they are at high risk [15]. A mobile app allowing users to observe how their risks are affected by their day-to-day choices relating to health and wellness (eg, such as their choice of food that day) can greatly help in the prevention of chronic diseases. The awareness of heart disease risk is one of the most critical methods and strategies to change behavior. Numerous mobile apps have been designed to bring awareness directly or indirectly about heart disease [16,17]. However, these apps rarely show how lifestyle behavior changes related to risk factors—smoking, diet quality, or alcohol consumption—affect their outcomes to prevent heart disease [1,16-22]. Although understanding future risks increases motivation of individuals to change behavior, whether individuals will actually change their behavior is a more complicated, sophisticated problem to solve than just *getting the message across* [23].

Objectives

Our goal was to design and test a mobile app that would help users focus on improving diet quality with the help of real-time feedback on future heart disease risk as a result of their diet quality patterns. In this way, we could increase individuals' awareness of cardiovascular risks based on daily dietary choices. Thus, users can focus on the behavior that is present and immediate, rather than an uncertain future [24,25]. Users can log simplified categories that have a high-quality diet—for example, vegetables, fruits, and whole grains—to help them focus on the quality of food rather than the detailed nutritional value, calories, and quantity of food. Our research questions (RQs) are as follows:

- RQ 1: How feasible was logging diet quality?
- RQ 2: How feasible was communicating risk to motivate behavior change?
- RQ 3: How effective was the app in changing health outcomes?

Our study demonstrated that (1) monitoring simple diet quality can have a significant effect on dietary behavior change and (2) regardless of participants' interest in heart disease risk, the app reduced the risk.

Methods

Study Design

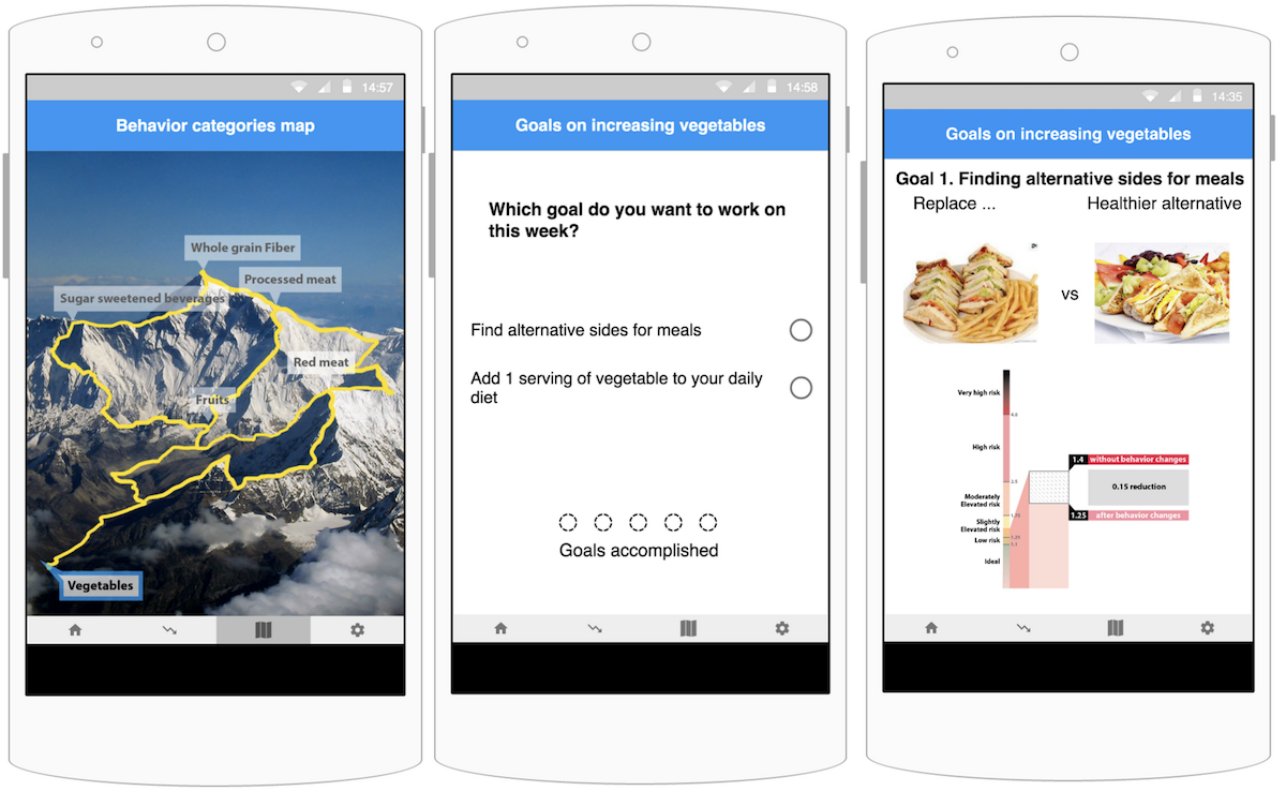
We designed the app based on behavior change techniques (BCTs) [26]. We used focus groups to iteratively improve the paper prototypes and developed an Android-based app as a result. We then conducted a 5-week pre-post study with a follow-up at 2 weeks after the study to evaluate the app's efficacy of clinical and behavioral outcome changes as well as app usage patterns.

Focus Groups for App Development

We conducted 3 focus groups in a sequence (n=13, with 3-5 people for each group) to iteratively improve the initial digital paper prototype (Figure 1). The participants were at risk for

heart disease and were recruited from a weight management clinic in the US Midwest. During the focus groups, the participants were presented with images from the initial prototype to test the usability and learnability of each screen (Figure 1). We revised the design iteratively based on the feedback. We then developed a mobile app on an Android platform.

Figure 1. Screens from the prototype presented to the focus group. Users can select which goal to work on using the mountain climbing metaphor (left). As users accomplish the goals, they can unlock the next category of goals. Selecting a category on the behavior category map will direct the user to the goal selection screen (center). The screen on the right shows choice for sides and how future cardiovascular risks might differ, if the user were to repeat the behavior for a week.



Final App Design

The BCT suggests 4 core components in designing an intervention: environmental contexts, goals, feedback and monitoring, and reinforcement. The app contains 5 screens: *main menu*, *profile*, *goals*, *meal calendar* (food logging screen), and *cardiovascular risk* (screen showing heart disease risk score). We designed the profile page to incorporate environmental context, the goals menu for users to personalize goals, meal calendar to log diet quality for feedback and monitoring, and cardiovascular risk screen to reinforce and reward positive diet change. A first-time user is directed to the profile screen to provide demographic information related to calculating their risk.

Diet Quality and Healthy Heart Score

The definition of high-quality diet in this study was based on the Healthy Heart Score (HHS), a risk score system for heart disease risk, developed at Harvard University (Figure 2) [15]. Among several heart disease risk models (eg, Framingham) [27], HHS is uniquely useful for middle-aged adults who do not have elevated clinical factors, such as high blood pressure or cholesterol, but may still be at high risk for developing cardiovascular disease (CVD). The HHS model builds on lifestyle factors, such as smoking status; level of physical activity; alcohol intake; and a diet score (DS) based on the consumption of fruits and vegetables, nuts, cereal fiber, sugar-sweetened beverages, and red and processed meats. HHS measures diet quality using the DS factor (Figure 2). A high DS indicates that the individual is eating more *healthy foods*, including fruits, vegetables, nuts, and white meat. Consumption of *unhealthy foods*, including red meat, processed meat, and sugary drinks, will lead to lower DS.

Figure 2. Description of the Healthy Heart Score and calculation of diet score for women and men.

Women

$$\text{20-year CVD risk (\%)} = 1 - 0.9660^{\exp(W-6.57301)} \times 100\%$$

$$\text{where X} = 0.10820 \times \text{age} + 0.15285\ (\text{if past smoker}) + 0.90138\ (\text{if current smoker}) + 0.04676 \times \text{BMI} - 0.01923 \times \frac{\text{grams}}{\text{d}}\ \text{of alcohol} + 0.0004 \times \left(\frac{\text{grams}}{\text{d}}\ \text{of alcohol}\right)^2 - 0.02951 \times \frac{\text{hours}}{\text{week}}\ \text{of exercise} - 0.05113 \times \text{diet score}^*$$

$$^*\text{Diet Score (Women)} = \left(0.03326 \times \frac{\text{grams}}{\text{d}}\ \text{of cereal fiber} + 0.18283\ \left(\text{if fruits + vegetables} \geq 3\frac{\text{servings}}{\text{d}}\right) + 0.14522\ \left(\text{if nuts 0.1 to 1}\frac{\text{servings}}{\text{d}}\right) + 0.24444\ \left(\text{if nuts} > 1\frac{\text{servings}}{\text{d}}\right) - 0.14631 \times \frac{\text{servings}}{\text{d}}\ \text{of sugar-sweetened beverages} - 0.15624 \times \frac{\text{servings}}{\text{d}}\ \text{of red and processed meats}\right) \times 10$$

Men

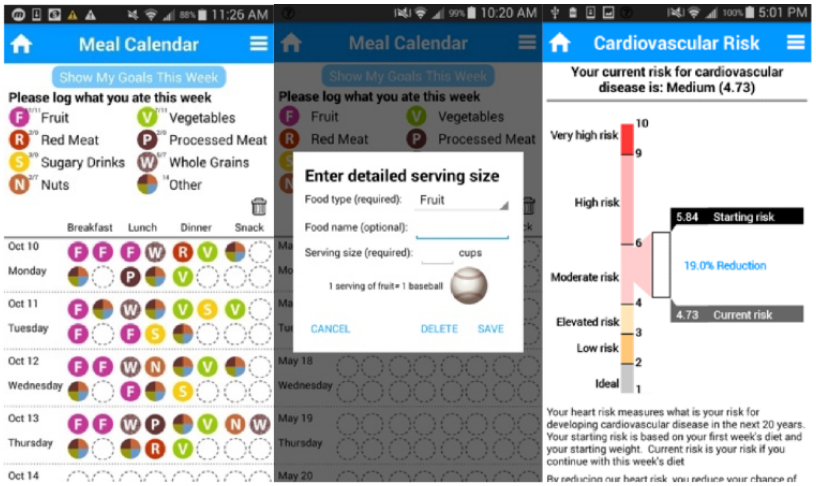
$$\text{20-year CVD risk (\%)} = 1 - 96368^{\exp(M-7.2437)} \times 100\%$$

$$\text{where X} = 0.13580 \times \text{age} - 0.0005 \times \text{age}^2 + 0.06979\ (\text{if past smoker}) + 0.42305\ (\text{if current smoker}) + 0.07424 \times \text{BMI} - 0.00898 \times \frac{\text{grams}}{\text{d}}\ \text{of alcohol} + 0.0001 \times \left(\frac{\text{grams}}{\text{d}}\ \text{of alcohol}\right)^2 - 0.01755 \times \frac{\text{hours}}{\text{week}}\ \text{of exercise} - 0.06691 \times \text{diet score}^*$$

$$^*\text{Diet Score (Men)} = \left(0.01816 \times \frac{\text{grams}}{\text{d}}\ \text{of cereal fiber} + 0.08819\ \left(\text{if fruits + vegetables} \geq 3\frac{\text{servings}}{\text{d}}\right) + 0.00535\ \left(\text{if nuts 0.1 to 1}\frac{\text{servings}}{\text{d}}\right) + 0.14285\ \left(\text{if nuts} > 1\frac{\text{servings}}{\text{d}}\right) - 0.14734 \times \frac{\text{servings}}{\text{d}}\ \text{of sugar-sweetened beverages} - 0.07112 \times \frac{\text{servings}}{\text{d}}\ \text{of red and processed meats}\right) \times 10$$

In the diet-monitoring screen (Figure 3, left), users can enter up to 4 food categories for each meal they eat each day: breakfast, lunch, dinner, and snack. Following HHS, users can log the overall quality of diet through the 7 food category items noted by HHS: 4 *healthy categories*, that is, fruits, vegetables, whole grains, and nuts, and 3 *unhealthy categories*, that is, red meat, processed meat, and sugary drinks. The app also allowed the selection of other categories to log foods not included in the provided categories. The goals screen showed the default number of servings suggested for each food category. Users can either drag a food category icon, for example, fruit, to one of the meal slots, which counts as one serving of that category to that meal, or tap the calendar and work on the pop-up window to increase or decrease the number of servings and add the name of the food they consumed. The definition of a serving was not defined—any consumption counted as a serving, following the antiquantification approach. In the goal screen (Figure 3, center), the default suggestions on the intake amount of unhealthy food categories were set to 0 servings. The combined total for fruits and vegetables should be at least 3 servings per day or, equivalently, 21 servings per week.

Figure 3. The left panel shows the meal calendar screen, where users can enter simple quality-oriented diet categories. The central panel shows options to add more details about the food, if the user desires. The right panel shows the screen that updates Healthy Heart Score as the user enters diet information.



Future Cardiovascular Risk

The cardiovascular risk (Figure 3, right) screen shows the current HHS, the user's calculated cardiovascular risk score in real time. The screen compares the risk at the time the user started using the app with the risk at the current week. In this risk screen, we rescaled the HHS to a range from 1 to 10 from its original unit, 0% to 100%, following the suggestion provided by the focus groups and in consultation with the expert who developed the HHS. The focus groups complained that the percentage was confusing—for example, it was unclear whether 50% meant 50% higher risk than others or half of the risk compared with others (or compared with the current status). In the rescaled scoring range, the ideal risk score for a healthy

individual is between 1 and 2, and if one has a risk score of 9 or above, the person is 4 times or more likely to develop heart disease than an individual with a healthy lifestyle.

Goals

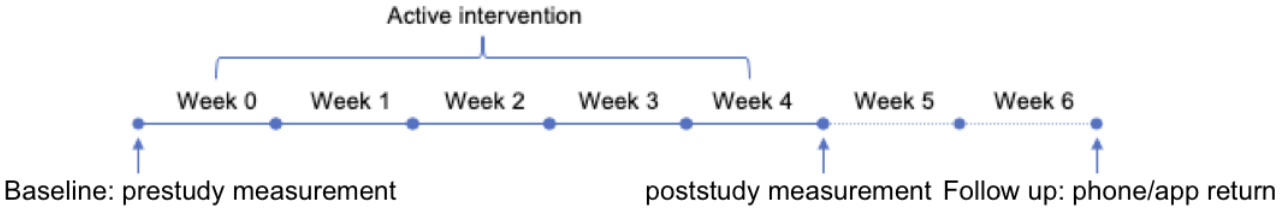
At the beginning of each week, the app prompts users to set their goals and directs them to the goals screen. Users can press the goal icon of the food categories they want to actively work on. Users can deactivate a goal by tapping it again, and the goal will appear grayed out. If one of the goals for unhealthy categories is active, users will be notified on the meal calendar if they exceed the maximum number of servings stated by the goal. The goals screen includes a checkbox that shows whether the user has met the goal.

Pre-Post Study: Recruitment and Procedures

The participants were recruited from a weight management clinic at a major hospital in the US Midwest. A research coordinator waited in the waiting room of the clinic with a recruitment flyer and screened interested potential participants for the following criteria: (1) aged above 18 years, (2) self-reported as overweight or obese, and (3) self-reported as at risk for diabetes. A total of 38 participants were enrolled to start the intervention between June and September 2016 for a 5-week period (denoted as week 0-4), with a follow-up meeting at 2 weeks after the end of week 4. The participants were asked to use the app for at least 6 days a week for the 5-week period of the study. The participants had the option to continue using the app until the follow-up meeting. Initially, the participants were asked to log their diet to establish a baseline. Starting at the beginning of week 1, the app started prompting the participants to set goals for each week based on HHS recommendations, either by keeping the default suggestion (ideal diet) or changing it to personalized goals.

At baseline and at the end of week 4, participants visited the clinic for a clinician to measure their weight and fasting blood sugar levels. At the end of week 4, participants were reminded that they were no longer required to use the app. In addition, an exit interview was held at the follow-up visit to discuss the participants' experiences with the app. Figure 4 shows the study procedure.

Figure 4. Timeline of prestudy and poststudy measurements and follow-up and the notation of the weeks.



All participants received cash compensations of up to US $50. Participants received partial or full compensation depending on how much they completed the following: 3 web-based surveys, measuring health outcomes twice, and using the app for at least 6 days a week during week 0 to 4. The app was provided to the participants in 2 ways. If the participants had an Android phone, the app was installed on their phones. Otherwise, the participants were provided with a Samsung Galaxy S3 phone, with the app installed, for the duration of the study. These participants were required to return their phones at the follow-up.

The study was approved with a full review by the institutional review board (IRB) of the institution that collected the data. All other researchers of this study were given access to deidentified data with the permission of the IRB.

RQs and Analyses

We wanted to answer 3 RQs regarding the feasibility of logging diet quality, motivating behavior change through feedback on future heart disease risks, and the app's efficacy of behavior change.

RQ 1: How Feasible Was Logging Diet Quality?

We recorded and analyzed the time, frequency, and screen activity of the participants' tapping events on the app. We analyzed how often participants went to each screen and which food categories were logged over time. We also analyzed user logs of food names to understand diet logging behavior.

RQ2: How Feasible Was Communicating Risk to Motivate Behavior Change?

We analyzed the participants' usage of the risk screen. We then tested the correlation between the participants' usage of the screen and HHS.

RQ3: How Effective Was the App in Changing Health Outcomes?

We conducted a paired sample *t* test to compare the outcome changes in diet quality, HHS, and in-clinic measurements (weight and fasting blood glucose levels) between pretest (at the beginning of week 0) and posttest (at the end of week 4) measurements. We used the statsmodels package in Python to run the analysis.

Results

Overview

In this section, we first report the demographic information of the participants and the recruitment outcome. We then report results on diet quality, risk score checking, health outcomes and DS, and the association between app use and DS.

Participants and App Use

A total of 84% (32/38) of the recruited participants completed the follow-up. Table 1 summarizes the demographic information of the participants. Overall, 58% (22/38) of participants used the study phones provided and the rest used their own phones.

In total, 8% (3/38) of the participants who were Android phone users stopped using the app and stopped responding to the researchers. Another participant (1/38, 3%) withdrew as the app was too cumbersome. Furthermore, 5% (2/38) other participants withdrew as they decided that the study no longer applied to them. The remaining 32 participants (female: 26/32, 81%; age (years): mean 57.48, SD 11.85) who completed the study had a wide range of age, weight, smoking status, and experience of using a smartphone. A total of 3% (1/32) of the participants were smokers, 38% (12/32) were former smokers, and 59% (19/32) were nonsmokers. A total of 53% (17/32) of participants were diagnosed with diabetes. Overall, 13% (4/32) of participants were overweight (BMI between 25 and 29.9 kg/m^2) and the remaining 88% (28/32) of participants were obese (BMI≥30) [28].

Table 1. Demographic information of participants.

Measure	Value
Participants who completed follow-up[a], n (%)	32 (85)
Participants who used the provided phones[a], n (%)	22 (58)
Age of participants who completed follow-up[b], years, mean (SD)	57.48 (11.85)
Female participants[b], n (%)	26 (81)
Participants who smoked, n (%)	
Current	1 (3)
Former	12(38)
Nonsmoker	19 (59)
Participants with diabetes[b], n (%)	17 (53)
Weight of participants, n (%)	
Overweight	4 (13)
Obese	28 (88)

[a]The total number of participants is 38.

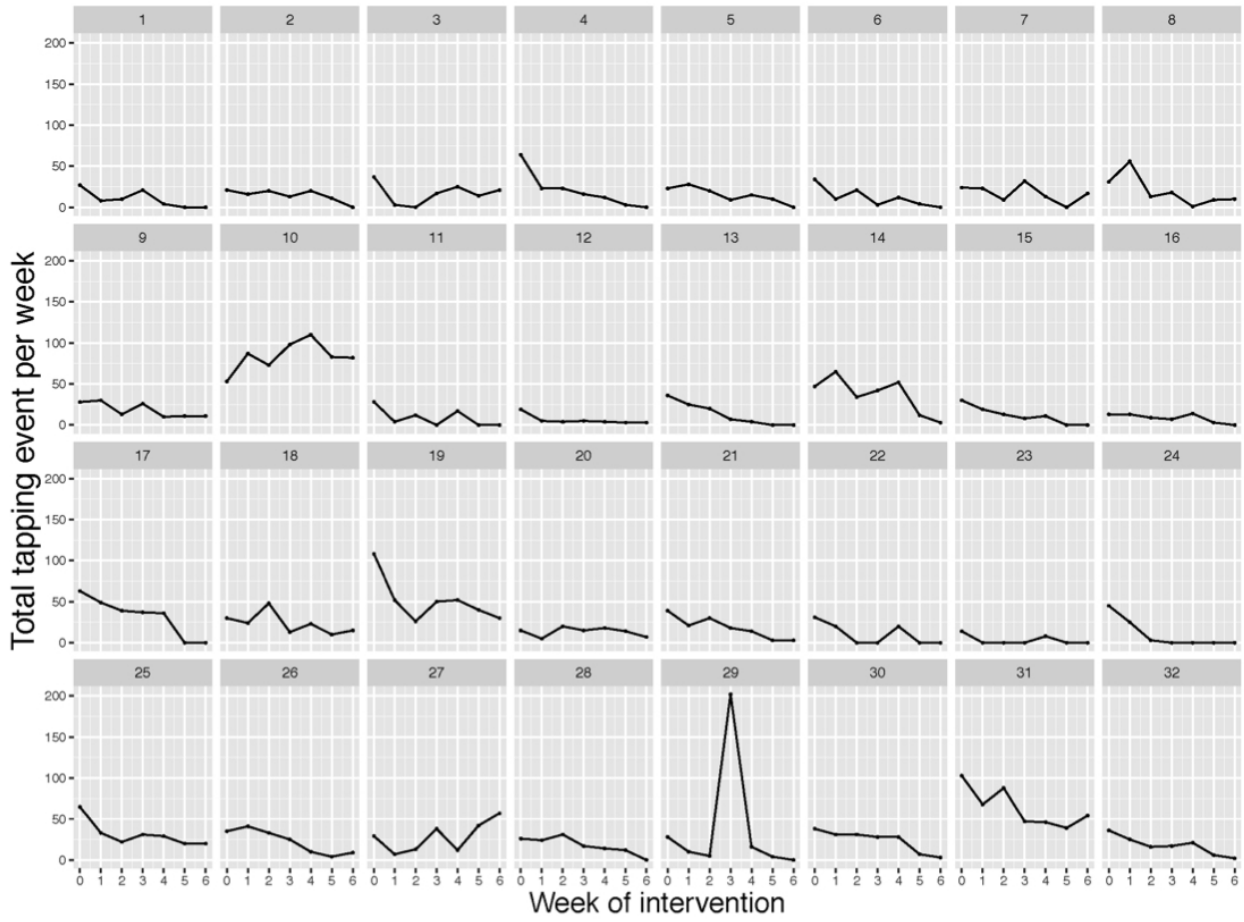
[b]The total number of participants who enrolled is 32.

RQ 1: How Feasible Was Logging Diet Quality?

App Use

During the active intervention when the participants were required to use the app (baseline to week 4), participants *tapped on the app*, that is, clicked to view contents or to make food entries, 27 times on average (SD 25.6) per week. After week 4 and until follow-up, participants tapped on the app 11 times on average (SD 18.3) per week. Figure 5 shows each participant's overall app use over the weeks.

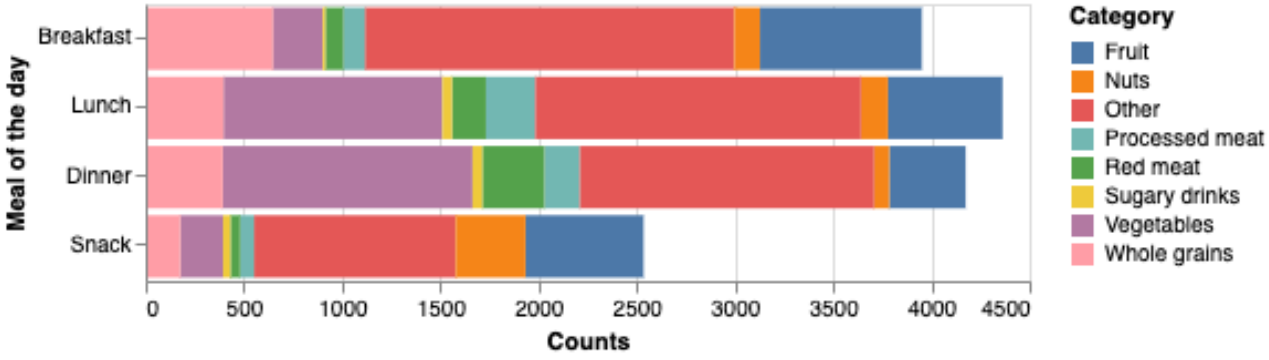
Figure 5. Each small graph shows each participant's total number of tapping events over the entire 7-week period, including the 2-week follow-up (between the baseline and follow-up). The x-axis shows the week of intervention (0 indicating the frequency accumulated between the baseline and at the end of week 0). The y-axis shows the total number of tappings for each week. A total of 27 participants visited the screen approximately every week for the intervention duration (weeks 0-4).



Diet Logging

Overall, 88% (28/32) of participants logged food consumption data every week between the baseline and the end of week 4, at least once a week. About half of the participants logged food consumption data approximately every day. As seen in Figure 6, during the active weeks, the 32 participants altogether logged *other* the most (6066/15,025, 40.37%), followed by *vegetables* (2857/15,025, 19.01%), *fruit* (2398/15,025, 15.96%), *whole grains* (1614/15,025, 10.74%), *nuts* (698/15,025, 4.65%), *red meat* (627/15,025, 4.17%), *processed meat* (614/15,025, 4.09%), and *sugary drinks* (151/15,025, 1.00%). Figure 6 shows when these food categories were logged to meal slots during the course of the day—breakfast, lunch, dinner, or snack. Fruits and whole grains were logged proportionally larger during breakfast meals than during other meals, and vegetables were logged proportionally larger during lunch and dinner meal times. *Other* categories were logged equally over all meal slots.

Figure 6. All 32 participants' logging per food category and for which meal of the day the logging occurred during the active intervention period (between baseline and the end of week 4).



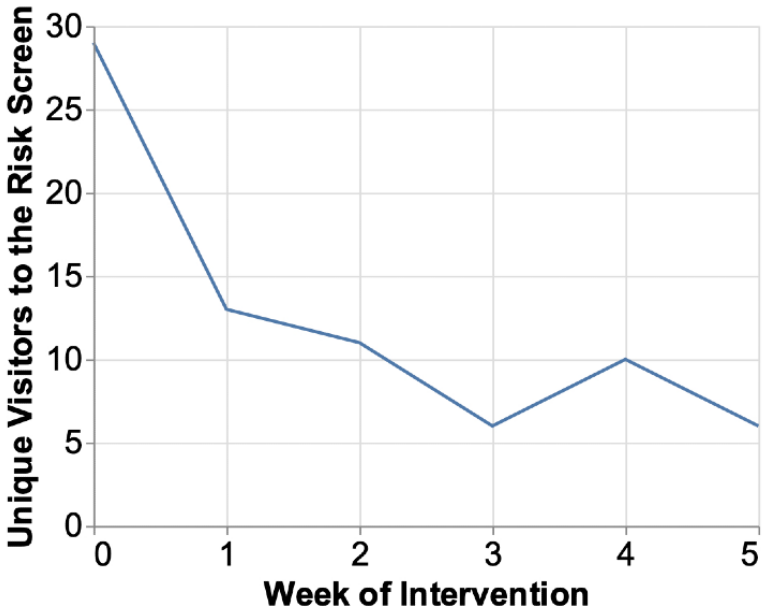
Participants entered a qualitative description of the food in the *food name* field for 38.14% (5730/15,025) of all diet logging instances. Overall, 48.87% (2800/5730) of these instances were entered when logging to the other category, 18.00% (504/2800) for vegetables, 15.00% (420/2800) for fruit, 8.00% (224/2800) for whole grains, 4.00% (112/2800) for nuts, 3.00% (84/2800) for red meat, 3.00% (84/2800) for processed meat, and 0.68% (19/2800) for sugary drinks. For all categories except other, participants gave a sample description of the food category they entered. For instance, the fruit category included descriptions such as *strawberries* or *grilled fruit salad* and the vegetable category included *arugula* or *grilled squash and zucchini with lemon and olive oil*.

When participants entered *other*, 98.89% (2769/2800) included detailed descriptions of the food. The qualitative analysis of these descriptions, together with the exit interviews, revealed several reasons for why the food was logged as *other*. The given food categories did not capture all the food categories the participants attempted to log, such as their current dietary goals (eg, to reduce dairy). The participants were instructed to only log what is related to heart disease risks, but they still captured other categories not affecting healthy heart risk, including dairy, dessert, or other protein foods (eg, 338/2769, 12.20% were proteins such as eggs, tofu, and beans and 584/2769, 21.09% were dairy such as milk, cheese, and Greek yogurt). Participants also logged foods as belonging to *other* category when the food was a mix of various food categories, which might have been difficult to capture in 1 or 2 food categories (eg, California roll, sandwich). A total of 0.69% (19/2769) showed red meat food, such as pork, and vegetables being logged as *other*, showing how the users might have been confused on what food categories these foods belonged to. Even though pork was red meat, the fact that it was logged as *other* matched with the exit interview content that the participants considered pork as white meat.

Risk Screen

As Figure 7 shows, at baseline, most participants checked their risk scores (n=29). However, in the following weeks after the baseline, most participants did not return to the risk screen to view the changes in their HHS. A total of 41% (13/32) people checked the risk screen in week 1, 34% (11/32) in week 2, 19% (6/32) in week 3, 31% (10/32) in week 4, and 19% (6/32) in week 5 until follow-up.

Figure 7. This figure shows the participants' use of the risk screen (loading frequency) over the weeks. A total of 29 of 32 participants checked their risks during week 1, and then, only a few checked again at week 4 (n=10). Most participants did not return to the risk screen to recheck it after the baseline.



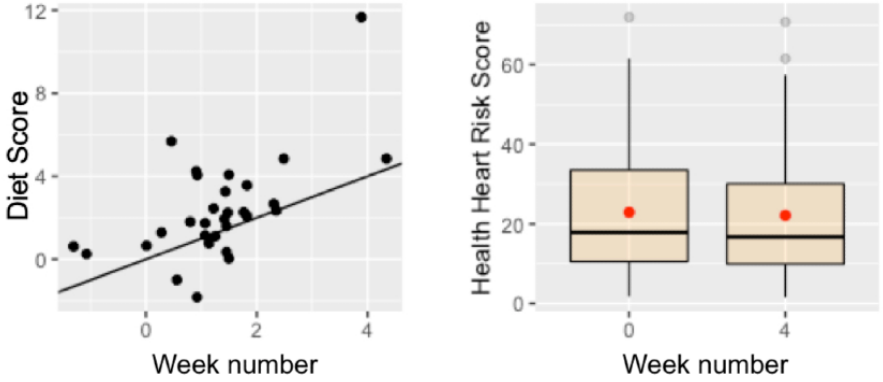
RQ 3: How Effective Was the App in Changing Health Outcomes?

Health Outcomes

DS

All but 6% (2/32) of participants logged their diet during the active intervention. Among the 94% (30/32) of participants who logged their diet during the active intervention, the DS showed a significant difference between baseline (mean 1.31, SD 1.14) and posttest during week 4 (mean 2.36, SD 2.48; t_{29}=−2.85; *P*=.008; Figure 8).

Figure 8. Changes in diet score (left) and Healthy Heart Score (right) between prestudy and poststudy measurements of the participants.



HHS

HHS also showed significant difference between baseline (mean 22.94, SD 18.86) and posttest measurements at the end of week 4 (mean 22.15, SD 18.58; t_{29}=2.41; *P*=.02).

There was no statistical association between food logging frequency and the 3 measures—DS, risk, and BMI.

In-Clinic Measurements

Weight (lbs) did not show a significant difference between pretest (mean 241.7, SD 61.17) and posttest (mean 242.6, SD 61.9) measurements, t_{29}=−1.043, *P*=.31. Blood sugar level (mg/dL) also did not show a significant difference between the pretest (mean 130.2, SD 76.62) and posttest (mean 123.3, SD 48.8) measurements, t_{28}=−0.95, *P*=.35.

Discussion

Principal Findings

The study showed feasibility for logging diet quality (RQ 1) but not for communicating risk (RQ 2). However, the app was effective in changing health outcomes (RQ 3), showing that logging simplified diet quality significantly improved dietary scores and future cardiovascular risk scores. The following are the key takeaways:

- The study showed no association between the frequency of logging and improved dietary scores, showing the importance of separating frequency of use in measuring health outcomes.
- The participants were not interested in monitoring the risk scores, but they still significantly decreased their risk scores by focusing on the target behavior. This finding has implications for health risk communication in mHealth app design.
- The study showed that users mostly logged irrelevant dietary behaviors to the target behavior. This finding shows the need to balance reducing monitoring items for efficiency versus what matters to users to support user experience.

Opportunities and Challenges of Quality-Focused Diet Monitoring

Previous literature shows that logging diet is highly associated with improved diet [29]. At the same time, studies have shown that not all users can benefit from sophisticated diet logging apps. Users often find diet logging a tedious, cumbersome activity, which leads to abandonment [7]. In addition, people do not always accurately estimate food proportions and nutritional contents [13]. Automated techniques, including calorie calculations and artificial intelligence–based food detection, can reduce such user burden in detailed diet logging [30-32]. However, these methods are still limited and error prone, which leads to increased user frustration and abandonment.

To address this gap, we implemented the HHS [15] into a mobile app, which simplified the diet-monitoring process to focus on improving diet quality over quantity. This incorporates a lenient approach toward logging food proportion and nutritional details in calculating the risk. By allowing users to focus on logging simplified diet quality that does not require logging detailed nutritional and caloric breakdown of each meal and focusing on whether a gross food group was consumed (fruits and vegetables), we showed that users steadily used the app even after the required period, for which they were not incentivized to use it. One participant asked if they could continue using the app even after the study had completed.

To our knowledge, our study is the first to investigate the feasibility of an mHealth solution for logging diet quality. Existing diet-monitoring apps have focused on logging the granular details of diet at a nutritional level [1]. Our findings on participants' use of our app are comparable with the user groups of other studies that showed a successful adoption of an mHealth tool. Individuals who complied with the study protocol regularly entered their food consumption details between 2 and 4 times a day [6,7,33].

At the same time, the study showed no association between frequency of use and increase in DS. This finding shows the need to separate quantitative measures of use from health outcomes. This implication aligns with discussions about whether sustained use of an mHealth app is positive or not—discontinuing to use an app might mean that the user no longer needs the app as they have achieved the health goal or have become more independent [34]. Furthermore, there can be other factors that influence the participants' healthier choices. According to Achananuparp et al [35], food journalers do not necessarily eat healthier food than others. Healthy eating behaviors are often affected by other sociodemographic factors, such as gender and regions of residence.

One challenge we faced in logging diet quality was that, even at the gross level of food categories, some participants were confused regarding categorizing food to the right categories (eg, confused pork as white meat, avocado as not being vegetable). Studies report that people struggle with journaling food regularly as the task is often time-consuming and laborious [8]. Therefore, it is desirable to implement a positive incentive for food journaling and to remove barriers for users. One idea is to help users go over their health management progress with their doctors, as Kim et al [33] tested in their study. We can implement photo-based food journaling with some automatic support to identify the category and to measure the portion of food [36]. Such photos can be distributed via social media, which can play a role in providing social support and encouragement for users as well.

Implications for Health Risk Communication in mHealth Design

The initial goal of this app was to increase individuals' awareness of cardiovascular risks based on daily dietary choices. We expected that users would check their risk scores as they changed their dietary patterns to understand how their risks were impacted by their dietary choices, thus making behavioral changes. However, although logging the diet quality was positively accepted by the participants, the participants rarely visited the risk screen during the study period. The participants mainly visited the risk screen at the very beginning to check their initial risk score, and a few came back for a second check after a week, and most did not return. The follow-up interview revealed that the participants noted that their scores did not seem to visibly change, so they did not think to check more often. At the same time, the HHS results showed that the participants still significantly improved their HHS at week 4. The HHS is associated with many clinical disease risks, such as diabetes mellitus, hypertension, and hypercholesterolemia [37]. In other words, studies report that patients with higher HHS show higher incidences of such diseases [37,38]. When we converted the HHS to 20-year CVD risk, the participants' initial risk resulted in 100%. This is because we recruited participants at high risk for heart disease due to their overweight or obesity condition. An HHS higher than 13 is calculated as 100% for the 20-year CVD risk. Accordingly, the change in the HHS from 22.94 to 22.15 will not decrease the CVD risk to 100%. Even a slight reduction of the HHS can decrease the CVD risks at this point. Despite the limited samples and duration, this study shows that the participants on average reduced their HHS by 0.78%, which is encouraging. An unanticipated design challenge that emerged in this study was regarding visualizing risks for those who were already well above the risk and for whom the risk would have remained at 100% despite any behavior changes they made. One idea is to augment a forecasting trajectory to the risk score, focusing on the changes toward reducing risks. The prediction can be designed so that it adjusts more sensitively to users' recent efforts for more motivation.

Communicating future risks is known to alert and motivate people to change behavior [39-41]. At the same time, if risk is too distant in the future, people may feel the risk is irrelevant, making it difficult to make behavior changes [42,43]. This study showed that the participants were initially motivated by their risk score; however, the behavior change was not related to their checking of the risk score over time. Although users did not check their risk scores, they decreased their overall risk scores. This finding has implications for the role of health risk communication in consumers using mHealth apps, in which continuous monitoring is the strength. The risk scores can serve as initial motivation to set up goals; however, users could focus on monitoring and improving the target risk behavior (in this case, diet quality), and the improvement with the risk can be a positive side effect.

Implications of Other in Monitoring Apps

The findings on the largest logging activity of *other* food categories provided implications for balancing between simplification and accommodation of users' *other* needs. The HHS discourages or encourages certain food categories to be consumed. This instruction—to focus on improving consumption of certain food categories—was reassured to the participants during the instruction. In the app design, we also specifically allowed users to only log the relevant food categories to improve the HHS. However, most diet logs were under *other*, where it included irrelevant food categories, such as dairy. According to the follow-up interview, this came from having a concurrent diet goal of participants on their own. When designing a monitoring app to improve health behavior, one needs to consider the gap between the chosen clinical approach and individuals' concurrent goals and considerations. Although simplifying the design to only monitor necessary information can improve efficiency and reduce user burden, this design approach might lose incorporating users' concurrent needs and focus. One should not consider what matters to users as *other* because it is considered irrelevant to what we designed as a target goal.

Limitations

This study did not include a control group, and the duration was only 5 weeks, which is not sufficient to show true behavior change. The data did not include information on whether the participants discontinued checking risk scores because of the lack of usability (eg, legibility of the visualization) or their disinterest in risks.

Conclusions

Our study showed the feasibility and efficacy of simplified diet quality monitoring in an mHealth app. Future work should further test the app's efficacy with a larger, focused population that is disinterested in using existing quantity-focused monitoring apps. Despite some known limitations on research design and duration, the findings provided significant contributions to understand the implications on the opportunities and challenges in designing a simplified, diet quality–focused monitoring app and how health risk communication can be effectively integrated into an mHealth design. The study also sheds light on finding the balance between affording users to focus on simplified target behavior, reducing user burden versus further incorporating what matters to users in designing a health monitoring app.

Conflicts of Interest

SEC is currently an employee of AbbVie and may hold stock or stock options. The work was completed when she was at Harvard. The remaining authors declare no conflict of interest.

References

1. DiFilippo KN, Huang W, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
2. Chiarini G, Ray P, Akter S, Masella C, Ganz A. Mhealth technologies for chronic diseases and elders: a systematic review. IEEE J Select Areas Commun 2013 Sep;31(9):6-18. [doi: 10.1109/JSAC.2013.SUP.0513001]
3. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 6;3:250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
4. Silva B, Rodrigues J, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
5. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mhealth interventions for obesity and diabetes treatment and self-management. Adv Nutr 2017 May;8(3):449-462 [FREE Full text] [doi: 10.3945/an.116.014100] [Medline: 28507010]
6. Hunt L, Dashti H, Chan Q, Gibson R, Vetter C. Quantifying diet intake and its association with cardiometabolic risk in the UK airwave health monitoring study: a data-driven approach. Nutrients 2020 Apr 22;12(4):- [FREE Full text] [doi: 10.3390/nu12041170] [Medline: 32331378]
7. Cordeiro F, Bales E, Cherry E, Fogarty J. Rethinking the Mobile Food Journal: Exploring Opportunities for Lightweight Photo-Based Capture. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18, 2015; Seoul, Republic of Korea. [doi: 10.1145/2702123.2702154]
8. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, Abowd GD, et al. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2015 Apr Presented at: ACM'15; April 18, 2015; Seoul, Republic of Korea URL: http://europepmc.org/abstract/MED/26894233 [doi: 10.1145/2702123.2702155]
9. Bentley F, Tollmar K. The Power of Mobile Notifications to Increase Wellbeing Logging Behavior. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: ACM'13; April 27, 2013; Paris, France. [doi: 10.1145/2470654.2466140]
10. Epstein D, Caraway M, Johnston C, Ping A, Fogarty J, Munson S. Beyond Abandonment to Next Steps: Understanding and Designing for Life After Personal Informatics Tool Use. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2016 Presented at: ACM'16; May 7, 2016; San Jose, California, United States. [doi: 10.1145/2858036.2858045]
11. Epstein D, Cordeiro F, Fogarty J, Hsieh G, Munson S. Crumbs: Lightweight Daily Food Challenges to Promote Engagement and Mindfulness. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2016 Presented at: ACM'16; May 7, 2016; San Jose, California, United States. [doi: 10.1145/2858036.2858044]
12. Gardner C, Trepanowski J, del Gobbo LC, Hauser M, Rigdon J, Ioannidis J, et al. Effect of low-fat vs low-carbohydrate diet on 12-month weight loss in overweight adults and the association with genotype pattern or insulin secretion: the dietfits randomized clinical trial. J Am Med Assoc 2018 Feb 20;319(7):667-679 [FREE Full text] [doi: 10.1001/jama.2018.0245] [Medline: 29466592]
13. HOGBIN MB, HESS MA. Public confusion over food portions and servings. J Am Diet Assoc 1999 Oct;99(10):1209-1211. [doi: 10.1016/s0002-8223(99)00297-7]
14. Majmudar MD, Colucci LA, Landman AB. The quantified patient of the future: opportunities and challenges. Healthc (Amst) 2015 Sep;3(3):153-156. [doi: 10.1016/j.hjdsi.2015.02.001] [Medline: 26384227]
15. Chiuve S, Cook N, Shay C, Rexrode K, Albert C, Manson J, et al. Lifestyle-based prediction model for the prevention of CVD: the Healthy Heart Score. J Am Heart Assoc 2014 Nov 14;3(6):e000954 [FREE Full text] [doi: 10.1161/JAHA.114.000954] [Medline: 25398889]
16. Chow CK, Ariyarathna N, Islam SM, Thiagalingam A, Redfern J. Mhealth in cardiovascular health care. Heart Lung Circ 2016 Aug;25(8):802-807. [doi: 10.1016/j.hlc.2016.04.009] [Medline: 27262389]
17. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, Herreros-González J. Mobile apps in cardiology: review. JMIR Mhealth Uhealth 2013 Jul 24;1(2):e15 [FREE Full text] [doi: 10.2196/mhealth.2737] [Medline: 25098320]
18. Azar K, Lesser L, Laing B, Stephens J, Aurora M, Burke L, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
19. 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]

20. Eyles H, McLean R, Neal B, Doughty R, Jiang Y, Ni Mhurchu C. Using mobile technology to support lower-salt food choices for people with cardiovascular disease: protocol for the SaltSwitch randomized controlled trial. BMC Public Health 2014 Sep 12;14:950 [FREE Full text] [doi: 10.1186/1471-2458-14-950] [Medline: 25217039]
21. Piette JD, List J, Rana GK, Townsend W, Striplin D, Heisler M. Mobile health devices as tools for worldwide cardiovascular risk reduction and disease management. Circulation 2015 Nov 24;132(21):2012-2027. [doi: 10.1161/circulationaha.114.008723]
22. Urrea B, Misra S, Plante TB, Kelli HM, Misra S, Blaha MJ, et al. Mobile health initiatives to improve outcomes in primary prevention of cardiovascular disease. Curr Treat Options Cardiovasc Med 2015 Dec;17(12):59. [doi: 10.1007/s11936-015-0417-7] [Medline: 26474892]
23. 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]
24. Pozolotina T, Olsen SO. Consideration of immediate and future consequences, perceived change in the future self, and health behavior. Health Mark Q 2019;36(1):35-53. [doi: 10.1080/07359683.2019.1567003] [Medline: 30848998]
25. Wood W, Neal D. Behav Sci Policy 2016;2(1):71-83. [doi: 10.1353/bsp.2016.0008]
26. 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]
27. Siontis G, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. Br Med J 2012 May 24;344:e3318. [doi: 10.1136/bmj.e3318] [Medline: 22628003]
28. Overweight and Obesity. National Institute of Health. 2017. URL: https://www.nhlbi.nih.gov/health-topics/overweight-and-obesity [accessed 2020-12-04]
29. Prestwich A, Kellar I, Parker R, MacRae S, Learmonth M, Sykes B, et al. How can self-efficacy be increased? Meta-analysis of dietary interventions. Health Psychol Rev 2014;8(3):270-285. [doi: 10.1080/17437199.2013.813729] [Medline: 25053214]
30. Beijbom O, Joshi N, Morris D, Saponas S, Khullar S. Menu-match: Restaurant-specific Food Logging From Images. In: IEEE Winter Conference on Applications of Computer Vision. 2015 Presented at: IEE'15; January 6, 2015; Waikoloa, Hawaii, United States. [doi: 10.1109/WACV.2015.117]
31. Bruno V, Resende S, Juan C. A survey on automated food monitoring and dietary management systems. J Health Med Inform 2017;8(3):- [FREE Full text] [doi: 10.4172/2157-7420.1000272] [Medline: 30101038]
32. Terzimehi? N, Leipold N, Schaefer H, Madenach M, Böhm M, Groh G, et al. Can an Automated Personalized Nutrition Assistance System Successfully Change Nutrition Behavior?-Study Design. In: International Conference on Information System. 2016 Presented at: CIS'16; June 1-6, 2016; Dublin, Ireland.
33. Kim Y, Ji S, Lee H, Kim J, Yoo S, Lee J. 'My Doctor is Keeping an Eye on Me!' Exploring the Clinical Applicability of a Mobile Food Logger. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2016 Presented at: ACM'16; May 7, 2016; San Jose, California, United States. [doi: 10.1145/2858036.2858145]
34. Epstein D, Caraway M, Johnston C, Ping A, Fogarty J, Munson S. Beyond Abandonment to Next Steps: Understanding and Designing for Life After Personal Informatics Tool Use. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2016 Presented at: ACM'16; May 7, 2016; San Jose, California, United States. [doi: 10.1145/2858036.2858045]
35. Achananuparp P, Lim E, Abhishek V. Does Journaling Encourage Healthier Choices? Analyzing Healthy Eating Behaviors of Food Journalers. In: International Conference on Digital Health. 2018 Presented at: CDH'18; April 23, 2018; Lyon, France. [doi: 10.1145/3194658.3194663]
36. Chung C, Agapie E, Schroeder J, Mishra S, Fogarty J, Munson S. When Personal Tracking Becomes Socialxamining the Use of Instagram for Healthy Eating. In: ACM SIGCHI Conference on Human Factors in Computing Systems. 2017 Presented at: ACM'17; May 6, 2017; Denver, Colorado. [doi: 10.1145/3025453.3025747]
37. Sotos-Prieto M, Mattei J, Hu FB, Chomistek AK, Rimm EB, Willett WC, et al. Association between a healthy heart score and the development of clinical cardiovascular risk factors among women. Circ Cardiovasc Qual Outcomes 2016 Feb 23;9(2 suppl 1):S77-S85. [doi: 10.1161/circoutcomes.115.002372]
38. Sotos-Prieto M, Mattei J, Cook NR, Hu FB, Willett WC, Chiuve SE, et al. Association between a 20-year cardiovascular disease risk score based on modifiable lifestyles and total and cause-specific mortality among us men and women. J Am Heart Assoc 2018 Nov 6;7(21):e010052 [FREE Full text] [doi: 10.1161/JAHA.118.010052] [Medline: 30373451]
39. Jensen JD, King AJ, Carcioppolo N, Davis L. Why are tailored messages more effective? A multiple mediation analysis of a breast cancer screening intervention. J Commun 2012 Oct;62(5):851-868 [FREE Full text] [doi: 10.1111/j.1460-2466.2012.01668.x] [Medline: 26405350]
40. Unruh HK, Bowen DJ, Meischke H, Bush N, Wooldridge JA. Women's approaches to the use of new technology for cancer risk information. Women Health 2004;40(1):59-78. [doi: 10.1300/J013v40n01_04] [Medline: 15778132]
41. Kushi LH, Byers T, Doyle C, Bandera EV, McCullough M, McTiernan A, American Cancer Society 2006 Nutrition and Physical Activity Guidelines Advisory Committee. American cancer society guidelines on nutrition and physical activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2006;56(5):254-81; quiz 313 [FREE Full text] [doi: 10.3322/canjclin.56.5.254] [Medline: 17005596]

42. Margolis PA, Stevens R, Bordley WC, Stuart J, Harlan C, Keyes-Elstein L, et al. From concept to application: the impact of a community-wide intervention to improve the delivery of preventive services to children. Pediatrics 2001 Sep;108(3):E42. [doi: 10.1542/peds.108.3.e42] [Medline: 11533360]
43. Edwards A, Thomas R, Williams R, Ellner AL, Brown P, Elwyn G. Presenting risk information to people with diabetes: evaluating effects and preferences for different formats by a web-based randomised controlled trial. Patient Educ Couns 2006 Nov;63(3):336-349. [doi: 10.1016/j.pec.2005.12.016] [Medline: 16860964]

Abbreviations

BCT: behavior change technique
CVD: cardiovascular disease
DS: diet score
HHS: Healthy Heart Score
IRB: institutional review board
mHealth: mobile health
RQ: research question

Edited by L Buis; submitted 23.06.20; peer-reviewed by D Epstein; comments to author 24.07.20; revised version received 15.09.20; accepted 20.11.20; published 23.12.20.

Please cite as:
Kwon BC, VanDam C, Chiuve SE, Choi HW, Entler P, Tan PN, Huh-Yoo J
Improving Heart Disease Risk Through Quality-Focused Diet Logging: Pre-Post Study of a Diet Quality Tracking App
JMIR Mhealth Uhealth 2020;8(12):e21733
URL: http://mhealth.jmir.org/2020/12/e21733/
doi:10.2196/21733
PMID:33355537

Original Paper

Availability of Spanish-Language Medical Apps in Google Play and the App Store: Retrospective Descriptive Analysis Using Google Tools

Inmaculada Grau-Corral[1,2], PhD; Pau Gascon[1,2], MPP, PhD; Francisco J Grajales III[1,3], MD; Belchin Kostov[4], PhD; Antoni Sisó Almirall[4,5,6], PhD

[1]Fundación iSYS, Barcelona, Spain

[2]Hospital Clínic, Barcelona, Spain

[3]Centre for Social Innovation & Impact Investing, Sauder School of Business, University of British Columbia, Vancouver, BC, Canada

[4]Primary Healthcare Transversal Research Group, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[5]Consorci d'Atenció Primària de Salut Barcelona Esquerre (CAPSBE), Barcelona, Spain

[6]University of Barcelona, Barcelona, Spain

Corresponding Author:
Inmaculada Grau-Corral, PhD
Fundación iSYS
C/Mallorca 140, 2-4
Barcelona, 08036
Spain
Phone: 34 692241233
Email: Imma.grau.corral@gmail.com

Abstract

Background: The number of medical and health apps in the App Store and Google Play repositories has been increasing in the recent years, and most of these apps are in English. However, little is known about the domain of Spanish health apps and their evolution.

Objective: The aim of this study was to perform a retrospective descriptive analysis of medical apps for patients in the Spanish language by using Google search tools over a 5-year period and to compare the results by using a reproducible methodology to obtain a better knowledge of the medical apps available in the Spanish Language.

Methods: Over a 5-year period, medical apps were catalogued using a Google-based methodology. Keywords of the first 14 categories of the International Classification of Diseases, Tenth Revision, were selected, and in December of each year, searches of the URLs of Google Play and the App Store were conducted using Google Advanced Search. The first 10 results were taken, and apps meeting the inclusion criteria were selected and rated with the iSYScore method.

Results: Out of a sample of 1358 apps, 136 met the inclusion criteria. The 3 main categories of the medical apps were in the fields of endocrinology (diabetes), respiratory (chronic obstructive pulmonary disease, asthma, and allergies), and neurology (multiple sclerosis, Parkinson disease, and Alzheimer disease). Few apps were maintained over the 5 years. Only 10 of the 136 apps were maintained for 3 years or more. There was a large number of original apps in other languages that were translated into Spanish (56/136, 41.2%). In the last year of the study, the main reason (73/280, 26.1%) for discarding an app was the date of the last update.

Conclusions: The market of Spanish apps is poor; only few apps have appeared repeatedly over 5 years. Differences were found with the international market in terms of apps related to mental health, heart and circulatory system, and cancer, and coincidences were found in the relevance of apps for diabetes control.

(JMIR Mhealth Uhealth 2020;8(12):e17139) doi:10.2196/17139

KEYWORDS

apps; mobile health; mHealth assessment; evaluation studies; health apps; ratings; mobile apps, Spanish apps; patient apps; Google Advanced Search; mobile phone

Introduction

Background

At its early stage of development, access to smartphone apps is limited to specific provider-customer relationships and the number of smartphone users is less. However, the smartphone market has accelerated over time, reaching 86% of the population in Europe [1]. Indeed, since the launch of mobile app platforms in 2008, including Google Play and App Store (for Android and iPhone operating system [iOS], respectively), the number of apps available has increased [2,3]. In the Apps store market, it is estimated that about 325,000 apps are dedicated to health [4].

The level of complexity associated with mobile health (mHealth) app products is heterogeneous, for example, a simple app that promotes a charity is not the same as a more complex one with Food and Drug Administration or European Union accreditation, which allows the monitoring of a patient with a heart disease. Likewise, there is also heterogeneity in the type of disease that is covered. For example, there are a large number of apps for monitoring patients with diabetes and a much lesser number of apps for patients who have undergone a transplant [4]. Moreover, there is also a cultural gap; most studies are conducted with apps that are developed in English. The inclusion criterion of most of the studies analyzing apps is that they were in English. There are only few studies referring to apps only in Spanish such as clinical trials, literature reviews, or meta-analyses. A simple search in PubMed for this type of studies with the terms “App” and “Spanish” yielded 23 results, while the search for “English” and “App” returned 997 results at the end of 2019.

Interventions aimed at Spanish cultural adaptation of an original development in English are more frequent. While there are official sites in Spain with rigorous selection criteria involving a small number of apps [5-7], multilingual platforms with a much larger number of apps such as the ORCHA Health App Library [8], Healthy living apps of Vic Health Foundation [9], and My Health Apps [10] dominate the market.

Regarding the use of mHealth in health care, there are interesting studies on the opportunities offered by its implementation for a safe and effective deployment in Latin America [11,12]. They describe the need to implement a regulation that guarantees the safety of developments [13,14], the protection of personal data [15], and all ethical aspects such as the European Union's “medical device” regulation [16]. However, the incidence and implementation of mHealth in the health system is not the focus of this study.

The objective of this study was to analyze the proposals that Spanish-speaking citizens can find in 2 large app stores, that is, Google Play and App Store, in Spanish. Users currently access the apps through 2 approaches: by recommendation or by filter. The “recommendation” can come from the health professional, a government, or public health authorities. The “filter” approach is conditioned by the results of the search engines of companies such as Google or Apple. Studies on policies and regulations set their point of view on access to apps by “recommendation.” In this study, we focused on the access to patient apps by “filter;” specifically, we used the “Google Advanced Search” tool.

Prior Work

In 2014, a research group of the iSYS Foundation developed the iSYScore [17] with the objective of obtaining indicators to help patients choose health apps. This score is similar to a triage scale, which allows agile selection and recommendations, but is not an accreditation. It does not delve into elements such as security that accreditations give nor in the efficiency of the app as the regulations of the Food and Drug Administration or European Union do. Once the iSYScore scale was ready, it was applied to a sample of apps obtained on the best results shown by Google Advanced Search in various groups of diseases. These annual collections have been recognized by the European mHealth Hub [18]. Since 2014, the iSYS Foundation has been collecting about 280 apps every year with Google Advanced Search, which verifies that they have met all the inclusion criteria and have applied the iSYScore in order to recommend a sample of the best results. However, the total number of apps analyzed annually is higher than that obtained by the methodology used by our group because recommendations of patient associations and registrations are also collected. Therefore, in order to achieve homogeneous results, only apps obtained with Google tools were included in this study.

Objectives

The aims of this study were (1) to analyze the findings obtained with a sample of medical apps per group of disease for 5 years that were selected by Google Advanced Search, (2) to describe the evolution of Spanish health apps over a 5-year period, and (3) to compare the health app studies available in Spanish with those in English or other languages.

Methods

Methodological Framework

The methodology in this study has 2 sections. The first section explains how the apps included in the sample were extracted and selected. The second section explains how the analysis on the selected apps was conducted.

Extraction and Selection of Apps

Every year, since 2014, researchers from the iSYS Foundation collected a sample of apps in Spanish for analysis. The analysis of this work corresponds to the samples obtained in 2014, 2015, 2016, 2017, and 2018.

To collect apps, researchers used the following methodology.

1. Selection of keywords by a set of diseases defined in the International Classification of Diseases, Tenth Revision (ICD-10) [19]: In the first year, the keywords selected were the most frequently used in the ICD-10 text using a word counter. In the following year, the researchers added words from websites belonging to patient associations. In the following years, the same set of keywords was used, with some small variation at the discretion of the researchers.

2. After selecting the keywords, the researchers used Google Advanced Search to obtain apps. For each set of diseases (for example, "Infectious and parasitic illnesses"), researchers put the keywords separated by spaces (OR), chose the Spanish language option, and entered the URL of Google Play for the first search and the one of Apple Store for the second one. The first 10 results of each platform were selected.
3. By having 14 groups of ICD-10 diseases with their keywords and targeting the search of 2 platforms, the expected result was 14×10×2=280 apps to analyze each year.

For the selection of apps, the following inclusion criteria were established: (1) Spanish language; (2) adequacy: e-books and podcasts were not accepted, as well as apps where cancer was a horoscope sign and apps with the absence of medical device accreditation if it was an app for medical diagnosis, treatment, or monitoring; (3) targeted audience: not for professionals; (4) availability: no password or geographic filter or other problems of availability; and (5) exceed a cut-off score according to the iSYScore (11 points out of a maximum of 48). In 2016, 2 new criteria were added: (1) to have a minimum number of downloads (N≥500) and (2) to have had an update the year before the sample was captured.

The exclusion criteria were duplicated apps and apps that promised miracle cures. The total sample of apps to be analyzed that met the inclusion criteria would be presented by ICD-10 disease groups in time series (one per year) to observe trends.

Analysis of the Selected Apps

To analyze evolution, the researchers considered the following parameters:

1. The number of apps per group of diseases in the sample during the study period and how many apps appeared recurrently over the years and which ones did not.
2. The main reason for discarding per year, given the observation of changes that have occurred during the capture of apps, was obtaining many results in English, despite the language filter.
3. Durability of the analyzed apps: In 2019, the researchers investigated the apps recommended in the previous years to observe their evolution.
4. Finally, the researchers investigated the possible factors influencing the durability of the apps depending on whether they were in native Spanish or translated or whether it depended on the promoter of the app.

Details to make these 4 sections:

1. The researchers would complete a table with the apps selected each year and group them by group of diseases and year. The table would allow them to obtain the totals and to keep track of those apps that appeared for more than 1 year.
2. To collect the main reasons for discard by year, the researchers had to add in tables the reason for discard by platform (Google Play and App Store) and by year, and detect if the main reason for the change varied.
3. To analyze the durability of the apps, the researchers searched the most established categories. For the analysis of the stability of the developments by disease group, a proxy variable was established and defined as a ratio: the relation between the number of repeated appearances by the number of individual apps of the category during the study period. The researchers considered that for a more "stable" disease group, apps would appear for more years in the annual sample. For example, if a disease group only has an app, and this app appears 5 years in a row, the ratio would be 5 (maximum value), and if it only appears 1 year, the ratio would be 1 (minimum value). Additionally, in 2019, apps from previous years would be revalidated to determine if they continue to meet the inclusion requirements or have disappeared. This will allow one to observe those that are no longer available. This will also allow us to observe which groups of diseases have a higher rate of "disappearance." To select the most consolidated categories, it was agreed to select those apps with a stability ratio higher than the average and a percentage of disappearance lower than the average. To observe the most unstable categories, the criterion used was those apps that had a ratio lower than the average and a percentage of disappearance higher than the average.
4. To study the possible factors influencing the durability of the apps, the researchers explored the developers' localization whether they were native apps in Spanish or translated from other languages. In addition, the promoter of the initiative was noted, in case this factor had an influence on durability.

Regarding the promoter of the apps, 6 categories were established: (1) health professionals that included individuals, health providers, and universities; (2) companies (neither pharmaceutical industry nor start-up); (3) nonprofit organizations, including foundations, associations, and scientific societies; (4) pharmaceutical industry; (5) patients and patient associations; and (6) projects and start-ups.

Statistical Analysis

The analysis of the potency and statistical significance used in studies that test a hypothesis does not apply to retrospective descriptive studies such as this one. Therefore, the results were summarized using descriptive statistical techniques such as percentages and means.

Results

Extraction and Selection of Apps

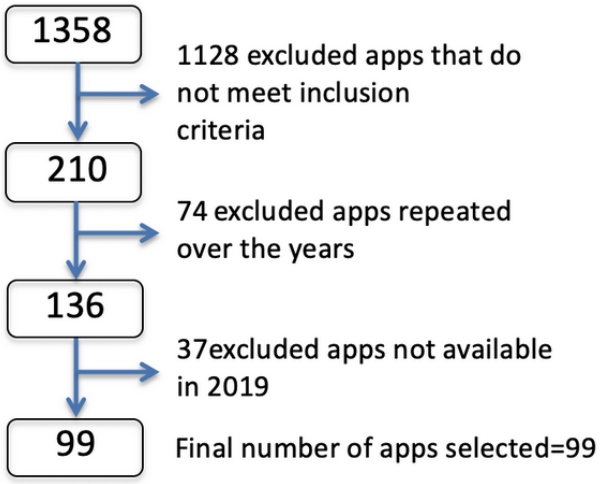
A sample of apps was collected over a 5-year period; 1358 apps (annual average 271.6, mean 263.5) were found using the described methodology. Every year, there was a significant decrease in the inclusion of the number of apps as these apps did not meet the inclusion criteria: of the 1358 obtained by our methodology, only 210 met the inclusion criteria (Table 1).

Table 1. Details of the number of apps collected over the years in this study.

Description	Year 2014	Year 2015	Year 2016	Year 2017	Year 2018
Apps collected (n=1377)	247	280	280	271	280
Apps that met inclusion criteria (n=210)	26	55	42	48	39

The searches by disease groups did not reach, in some cases, the 20 results expected in the methodology. Reviewing the 210 outcomes that met the inclusion criteria over 5 years, 74 repetitions of apps were found over the years and extracted, leaving a total of 99 different apps (Figure 1).

Figure 1. Flowchart of the selection of apps.



Analysis of the Selected Apps

During the study period, in which each group of diseases could have up to 20 different apps per year, a maximum of 8 and a minimum of 0 were found (Table 2).

Table 2. Evolution of the health apps included each year, classified by 14 ICD-10[a] disease groups (n=210).

Apps as per ICD-10	Number of apps in each year				
	Year 2014 (n=26)	Year 2015 (n=55)	Year 2016 (n=42)	Year 2017 (n=48)	Year 2018 (n=39)
Infectious and parasitic diseases (n=16)	1	6	3	2	4
Cancer (n=21)	4	8	5	2	2
Blood diseases (n=8)	1	2	2	1	2
Endocrine, nutritional and metabolic diseases (n=32)	5	6	5	11	5
Mental and behavioral disorders (n=8)	1	2	2	2	1
Diseases of the nervous system (n=28)	4	8	5	5	6
Eye diseases (n=11)	1	4	2	2	2
Ear diseases (n=12)	0	4	2	3	3
Circulatory system diseases (n=12)	2	3	2	3	2
Diseases of the respiratory system (n=18)	—[b]	3	3	6	6
Diseases of the digestive system (n=12)	1	3	3	3	2
Skin diseases (n=11)	3	2	3	3	0
Diseases of the musculoskeletal system (n=12)	3	2	2	3	2
Genitourinary system diseases (n=9)	0	2	3	2	2

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

[b]Not available.

Most apps were associated with the endocrine and nervous systems and cancer-related diseases. Apps dedicated to diabetes were included under endocrine diseases, and these apps were predominant over apps dedicated to obesity. With regard to diseases of the nervous system, tracking apps for multiple sclerosis and Parkinson disease, symptoms of meningitis, or information for relatives of people with Alzheimer disease were selected. With respect to cancer, the majority of the apps obtained by Google Advanced Search were related to skin cancer and the follow-up of skin lesions. There were also results involving apps providing information about breast cancer. In 2014, no data were collected in the "Diseases of the respiratory system." Despite this fact, this domain was in the fourth place in terms of total number of apps. Pollen alerts, asthma, and chronic obstructive pulmonary disease follow-up were the themes of the apps in the diseases of the respiratory system domain.

Reasons for Excluding the Apps

The main reason for the exclusion of apps was that they were not in Spanish (323/1358, 23.8%), although the language filter was activated in the search (Table 3).

Table 3. Reasons for app exclusion.[a]

Reasons, operating systems	Number of apps in each year				
	Year 2014	Year 2015	Year 2016	Year 2017	Year 2018
It was not in Spanish (n=323)					
Android (n=182)	36	64	15	40	27
iOS[b] (n=141)	5	51	43	24	18
Total	41	115	58	64	45
It was for health professionals (n=231)					
Android (n=114)	36	20	47	7	4
iOS (n=117)	54	37	16	7	3
Total	90	57	63	14	7
It was not suitable (n=227)					
Android (n=106)	54	11	22	10	9
iOS (n=121)	30	19	28	15	29
Total	84	30	50	25	38
Not available (n=156)					
Android (n=80)	54	7	6	2	11
iOS (n=76)	56	3	10	2	5
Total	110	10	16	4	16
Repetitions (n=114)					
Android (n=37)	0	1	2	5	29
iOS (n=14)	0	1	4	4	5
Both (n=63)	2	6	22	12	21
Total	2	8	28	21	55
Last update date (minimum the year prior to the annual collection) (n=152)					
Android (n=88)	—[c]	—	21	36	31
iOS (n=64)	—	—	2	20	42
Total	—	—	23	56	73
Downloads <500 (n=84)					
Android (n=76)	—	—	19	33	24
iOS (n=8)	—	—	0	6	2
Total	—	—	19	39	26
Did not rate 11 points on the iSYScore (n=127)					
Android (n=57)	13	23	12	5	4
iOS (n=70)	22	24	6	11	7
Total	35	47	18	16	11

[a]There may be more than one reason for exclusion of apps.

[b]iOS: iPhone operating system.

[c]Not available.

A significant number of apps (231/1358, 17.0%) were discarded because they were aimed at health professionals mostly in the genitourinary and musculoskeletal disease fields. Adding the number of downloads (minimum 500) and the last update date (minimum the year prior to the annual collection) in 2016 led to an increase in discard rates in successive years. In parallel, the number of discards decreased due to not reaching the minimum score with the iSYScore.

Stability of the Results Over Time

On revalidation of the apps from previous years in 2019, it was found that 37 out of 136 apps (27.2%) were no longer available

(Table 4). Only 10 apps appeared 3 or more times over the 5-year period.

Table 4. Stability of the results through the years.

Number of apps	Year 2014	Year 2015	Year 2016	Year 2017	Year 2018	Total	Repeated apps	Apps without repetition
Annual results (n)	26	55	42	48	39	210	74	136
Apps disappeared in 2019, n (%)	12 (46)	18 (33)	10 (24)	4 (8)	0 (0)	44 (20.9)	7 (10)	37 (27.2)
Apps remaining in 2019 (n)	14	37	32	44	39	166	N/A[a]	99

[a]N/A: Not applicable.

The apps most repeated over the 5-year study period were those related to the genitourinary diseases, whereas the least repeated were those associated with eye diseases. The most consolidated categories were those with a higher than average ratio and a lower than average percentage of disappearance and included the following 3 categories: endocrine, nutritional and metabolic diseases, genitourinary diseases, and infectious and parasitic diseases (mainly HIV infection). Three categories were the most unstable, having a lower than average ratio and a higher than average disappearance percentage: diseases of the digestive system, cancer, and mental and behavioral disorders (Table 5).

Table 5. Stability of the results through the years by group of disease (n=210).

ICD-10[a] group	Repeated apps (n=74)	Unique apps without repetition (n=136)	Stability ratio	Disappeared apps, n (%)	Apps in 2019 (n=99)
Infectious and parasitic diseases (n=16)	6	10	1.6	2 (20)	8
Cancer (n=21)	8	13	1.6	7 (54)	6
Blood diseases (n=8)	3	5	1.6	2 (40)	3
Endocrine, nutritional and metabolic diseases (n=32)	16	16	2.0	2 (13)	14
Mental and behavioral disorders (n=8)	1	7	1.1	3 (43)	4
Diseases of the nervous system (n=28)	13	15	1.9	5 (33)	10
Eye diseases (n=11)	2	9	1.2	0 (0)	9
Ear diseases (n=12)	1	11	1.1	2 (18)	9
Circulatory system diseases (n=12)	3	9	1.3	4 (44)	5
Diseases of the respiratory system (n=18)	6	12	1.5	2 (17)	10
Diseases of the digestive system (n=12)	4	8	1.5	5 (63)	3
Skin diseases (n=11)	3	8	1.4	1 (13)	7
Diseases of the musculoskeletal system (n=12)	3	9	1.3	1 (11)	8
Genitourinary system diseases (n=9)	5	4	2.3	1 (25)	3

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

Possible Factors Influencing the Durability of the Apps

Most of the apps found were originally developed in Spanish and were from Spain and Latin America (80/136, 58.8%). However, those that had been translated from another language showed a tendency to be more durable (Table 6).

Table 6. Native Spanish apps and apps that were translated from other languages to Spanish (n=136).

Language type of apps, availability	Values, n (%)
Spanish apps	
Available	55 (69)
Not available	25 (31)
Subtotal	80 (58.8)
Apps translated from other languages	
Available	44 (79)
Not available	12 (21)
Subtotal	56 (41.2)

Apps According to Promoter

The most stable app developments (with less disappearance of products) were those in the Company category (Table 7), while the most spurious ones were those related to projects and start-ups, as well as those led by the pharmaceutical industry.

Table 7. Apps selected and classified by promoter (n=136).

Promoter type and availability	Values, n (%)
Health care professionals	
Available	7 (70)
Not available	3 (30)
Subtotal	10 (7.3)
Company	
Available	36 (82)
Not available	8 (18)
Subtotal	44 (32.3)
Pharmaceutical industry	
Available	7 (64)
Not available	4 (36)
Subtotal	11 (8.1)
Nonprofit public foundation	
Available	25 (69)
Not available	11 (31)
Subtotal	36 (26.5)
Patients	
Available	7 (70)
Not available	3 (30)
Subtotal	10 (7.3)
Start-up+project	
Available	15 (60)
Not available	10 (40)
Subtotal	25 (18.4)

Discussion

Principal Findings

The search for the best health apps in the Spanish language using the Google Advanced Search algorithm resulted in only 136 apps meeting the inclusion criteria. Of these, a significant proportion (56/136, 41.2%) corresponded to English translations and a significant percentage disappeared over the 5-year study period. Only 10 apps appeared for 3 or more years (the most stable). Apps dedicated to diabetes were the most common [15] and had frequent medical device accreditation. A recent meta-analysis on the use of mHealth to support patients with diabetes showed positive results [20]. Despite the number of apps found in neurology (Alzheimer disease, epilepsy, multiple sclerosis, etc), there were only few studies on the effectiveness of those apps [21].

In the cancer section, the keyword "cancer" had to be discarded since 2014, as it offered better-positioned results related to horoscopes. Results related to possible nonevidence-based cures were also ruled out. According to the iSYScore results, the best positioned app was the Spanish translation of the ASCO App Cancer.net [22]. There is a lack of patient follow-up apps besides those related to skin or breast cancer, which have been shown to be useful in some studies [23-26]. One possible explanation may be that these apps are not available on commercial platforms (Google Play or App Store). Reviewing reasons for the apps that were discarded, researchers found that the main reason was that apps selected were not in Spanish (323/1358, 23.8%), despite Spanish language being selected in Google Advanced Search.

Indeed, in 2014, the "Spanish language" was added as a keyword in order to obtain more adequate results. The tendency to present results in other languages by these platforms has decreased over time (Table 3). Cancer, skin, and digestive system diseases were the most affected by this problem. Regarding discarded apps because of inadequacy (Table 3), apps related to mental and behavioral diseases showed the most interesting results. These had nonvalidated treatments more frequently (eg, cure of schizophrenia with phone vibrations) or were not available to the public because they were included as part of a controlled study for which an individual had to be included in the study in order to have access to the app. Other results not related to health (singers or horoscope signs) were also found in other categories.

Comparison With Studies in Other Languages

We compared our results with those of other studies such as the report by the American IQVIA, "The Growing Value of Digital Health" [4] and the German Research2Guidance "mHealth Economics 2017-Current Status and Future Trends in Mobile Health" [27]. IQVIA performed an analysis on the number of apps available by category. The category of "Health condition management" and specifically under the section "Disease Specific" showed a number in which the apps dedicated to mental disorders, diabetes, and the heart and circulatory system dominated the top positions in terms of the number of apps. The same report suggested that the evidence of the effectiveness of the apps described in 571 published studies in 5 patient populations on reductions in the utilization of acute care: prevention of diabetes, diabetes, asthma, cardiac rehabilitation, and pulmonary rehabilitation.

With regard to Research2Guidance, we found that comparison of our results was not adequate since their work was based on expert surveys while our results were based on searches using Google tools. That study reported that the most attractive app development fields are related to physicians (30%), diabetes (27%), the heart, blood, and circulatory system (24%), medications (24%), healthy lifestyles (22%), hospital efficiency (19%), and mental health (17%). Both reports coincide in the categories of diabetes, heart problems, and mental health as those that aroused most interest, although in different order of priority.

In this study, a large number of apps were found related to diabetes, cancer, and diseases of the nervous system. The apps found to be the most stable over time were those belonging to the categories of diabetes, infectious diseases, and kidney diseases, thereby disagreeing with the previously mentioned studies with the exception of the main category, that is, diabetes.

Analyzing the Evolution

The number of apps discarded because the Google search showed results in English and other languages significantly decreased in the last year of the study. However, repetitions in the same category and platform increased the number of discards in the same year. Small changes to the URL of the repeated apps might explain this fact. Over the years, there was a trend toward an increase in the number of apps ruled out due to lack of updates. However, the number of apps ruled out by a low iSYScore decreased over time, indicating a higher sample quality that exceeds the inclusion criteria. Regarding the different categories, the robustness of the endocrine (mostly diabetes) and nervous system categories over time was of note, while apps dedicated to cancer decreased.

Limitations of This Study

An obvious limitation of this study was the dependence on the results of the Google algorithm for the selection of the most representative results. This results in a volatile and dependent return. The decrease in the results on cancer and the digestive system and mental health disorders suggests that developments occur outside major markets such as in research fields [28] or payment software.

Conflicts of Interest

None declared.

References

1. Number of Mobile Subscribers Worldwide Hits 5 Billion. GSMA Intelligence. London, United Kingdom URL: https://www.gsma.com/newsroom/press-release/number-mobile-subscribers-worldwide-hits-5-billion/ [accessed 2019-05-27]
2. Statista. App stores: number of apps in leading app stores 2019. Internet: Mobile Internet & Apps. Hamburg Germany: Statista URL: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2019-07-17]
3. 325,000 mobile health apps available in 2017 – Android now the leading mHealth platform. research2guidance. Berlin, Germany: Research2Guidance; 2017 Nov. URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2019-05-31]
4. The Growing Value of Digital Health. Evidence and Impact on Human Health and the Healthcare System. Institute Report. Durhan, USA: The IQVIA Institute; 2017 Nov 07. URL: https://www.iqvia.com/en/institute/reports/the-growing-value-of-digital-health [accessed 2019-09-12]
5. López Seguí F, Pratdepàdua Bufill C, Rius Soler A, de San Pedro M, López Truño B, Aguiló Laine A, et al. Prescription and Integration of Accredited Mobile Apps in Catalan Health and Social Care: Protocol for the AppSalut Site Design. JMIR Res Protoc 2018 Dec 21;7(12):e11414 [FREE Full text] [doi: 10.2196/11414] [Medline: 30578234]
6. Distintivo AppSaludable - Catálogo de aplicaciones. Sevilla, Spain: Agencia de Calidad Sanitaria de Andalucía URL: http://www.calidadappsalud.com/distintivo/catalogo [accessed 2019-08-25]
7. AppSalut. Barcelona, Spain: Fundació TIC Salut Social URL: https://ticsalutsocial.cat/serveis/mhealth/serveis-acreditacio-i-segell-tics/ [accessed 2019-08-25]
8. Find the Best Health Apps. ORCHA. Daresbury, United Kindom: Orcha URL: https://appfinder.orcha.co.uk/ [accessed 2019-08-25]
9. Healthy living apps. VicHealth. Melbourne. Australia: VicHealth foundation URL: https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps [accessed 2019-08-25]
10. MyHealth Apps tried and tested by people like you. London, United Kingdom: Patient View URL: http://myhealthapps.net/ [accessed 2019-08-25]
11. Ruiz EF, Proaño A, Ponce OJ, Curioso WH. Tecnologías móviles para la salud pública en el Perú: lecciones aprendidas. Rev Peru Med Exp Salud Publica 2015 Jun 19;32(2):364. [doi: 10.17843/rpmesp.2015.322.1634]
12. Curioso WH. Building Capacity and Training for Digital Health: Challenges and Opportunities in Latin America. J Med Internet Res 2019 Dec 18;21(12):e16513 [FREE Full text] [doi: 10.2196/16513] [Medline: 31850849]
13. Alonso Arévalo J, Mirón-Canelo J. Aplicaciones móviles en salud: potencial, normativa de seguridad y regulación. Revista Cubana de Información en Ciencias de la Salud. 2017 Oct 11. URL: http://www.acimed.sld.cu/index.php/acimed/article/view/1136/690 [accessed 2018-10-11]
14. Guillen Pinto EP, Ramirez López L, Cifuentes Sanabria YP. Modelo de evaluación de requerimientos de privacidad, seguridad y calidad de servicio para aplicaciones médicas móviles. Rev. Univ. salud 2017 Aug 30;19(2):280. [doi: 10.22267/rus.171902.90]
15. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. Eur-lex.: The European Parliament and the Council of the European Union; 2016 Apr 27. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC [accessed 2019-07-08]
16. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. EUR-lex.: The European Parliament and the Council of the European Union; 2017 Apr 05. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745 [accessed 2019-06-26]
17. Grau I, Kostov B, Gallego JA, Grajales Iii F, Fernández-Luque L, Sisó-Almirall A. [Assessment method for mobile health applications in Spanish: The iSYScore index]. Semergen 2016;42(8):575-583. [doi: 10.1016/j.semerg.2015.12.001] [Medline: 26879598]
18. Health Apps repositories in Europe. European mhealth hub. Seville, Spain: WHO/ITU/Andalusian Ministry of Health URL: https://mhealth-hub.org/health-apps-repositories-in-europe [accessed 2020-01-11]
19. eCIE10ES - Edición electrónica de la CIE-10-ES Diagnósticos. eCIEmaps. Madrid, Spain: Ministerio de Sanidad, Consumo y Bienestar. Gobierno de España URL: https://eciemaps.mscbs.gob.es/ecieMaps/browser/index_10_mc.html [accessed 2019-09-15]
20. Bonoto BC, de Araújo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of Mobile Apps to Support the Care of Patients With Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
21. Dorsey ER, Glidden AM, Holloway MR, Birbeck GL, Schwamm LH. Teleneurology and mobile technologies: the future of neurological care. Nat Rev Neurol 2018 May;14(5):285-297. [doi: 10.1038/nrneurol.2018.31] [Medline: 29623949]
22. Doctor-approved patient information from ASCO. Cancer.net.: American Society of Clinical Oncology (ASCO) URL: https://www.cancer.net/ [accessed 2019-09-15]

23. Fu MR, Axelrod D, Guth AA, Rampertaap K, El-Shammaa N, Hiotis K, et al. mHealth self-care interventions: managing symptoms following breast cancer treatment. Mhealth 2016 Jul;2. [doi: 10.21037/mhealth.2016.07.03] [Medline: 27493951]
24. Osborn J, Ajakaiye A, Cooksley T, Subbe CP. Do mHealth applications improve clinical outcomes of patients with cancer? A critical appraisal of the peer-reviewed literature. Support Care Cancer 2020 Mar;28(3):1469-1479 [FREE Full text] [doi: 10.1007/s00520-019-04945-4] [Medline: 31273501]
25. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-Tested Mobile Apps for Breast Cancer Care: Systematic Review. JMIR Mhealth Uhealth 2019 Feb 11;7(2):e10930. [doi: 10.2196/10930] [Medline: 30741644]
26. Rat C, Hild S, Rault Sérandour J, Gaultier A, Quereux G, Dreno B, et al. Use of Smartphones for Early Detection of Melanoma: Systematic Review. J Med Internet Res 2018 Apr 13;20(4):e135 [FREE Full text] [doi: 10.2196/jmir.9392] [Medline: 29653918]
27. mHealth Economics 2017 – Current Status and Future Trends in Mobile Health. research2guidance Report. Berlin, Germany: research2guidance URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-09-12]
28. 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 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]

Abbreviations

ICD-10: International Classification of Diseases, Tenth Revision
iOS: iPhone operating system
mHealth: mobile health

Edited by C Lovis; submitted 20.11.19; peer-reviewed by W Curioso, L Quinlan; comments to author 15.04.20; revised version received 11.09.20; accepted 27.09.20; published 03.12.20.

Please cite as:
Grau-Corral I, Gascon P, Grajales III FJ, Kostov B, Sisó Almirall A
Availability of Spanish-Language Medical Apps in Google Play and the App Store: Retrospective Descriptive Analysis Using Google Tools
JMIR Mhealth Uhealth 2020;8(12):e17139
URL: https://mhealth.jmir.org/2020/12/e17139
doi:10.2196/17139
PMID:33270031

Review

eHealth Delivery of Educational Content Using Selected Visual Methods to Improve Health Literacy on Lifestyle-Related Diseases: Literature Review

Azusa Aida[1,2], MD; Thomas Svensson[1,3,4], MD, PhD; Akiko Kishi Svensson[1,2,3], MD, PhD; Ung-Il Chung[1,4,5], MD, PhD; Toshimasa Yamauchi[2], MD

[1]Precision Health, Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[2]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[3]Department of Clinical Sciences, Lund University, Skåne University Hospital, Malmö, Sweden

[4]School of Health Innovation, Kanagawa University of Human Services, Kawasaki-shi, Japan

[5]Clinical Biotechnology, Center for Disease Biology and Integrative Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Thomas Svensson, MD, PhD
Precision Health
Department of Bioengineering
Graduate School of Engineering, The University of Tokyo
7-3-1 Hongo
Bunkyo-ku
Tokyo, 113-8656
Japan
Phone: 81 358414737
Email: thomas.svensson@med.lu.se

Abstract

Background: Lifestyle-related diseases, such as stroke, heart disease, and diabetes, are examples of noncommunicable diseases. Noncommunicable diseases are now the leading cause of death in the world, and their major causes are lifestyle related. The number of eHealth interventions is increasing, which is expected to improve individuals' health literacy on lifestyle-related diseases.

Objective: This literature review aims to identify existing literature published in the past decade on eHealth interventions aimed at improving health literacy on lifestyle-related diseases among the general population using selected visual methods, such as educational videos, films, and movies.

Methods: A systematic literature search of the PubMed database was conducted in April 2019 for papers written in English and published from April 2, 2009, through April 2, 2019. A total of 538 papers were identified and screened in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. Finally, 23 papers were included in this review.

Results: The 23 papers were characterized according to study characteristics (author and year of publication, study design and region where the study was conducted, study objective, service platform, target disease and participant age, research period, outcomes, and research method); the playback time of the educational videos, films, and movies; and the evaluation of the study's impacts on health literacy. A total of 7 studies compared results using statistical methods. Of these, 5 studies reported significant positive effects of the intervention on health literacy and health-related measures (eg, physical activity, body weight). Although most of the studies included educational content aimed at improving health literacy, only 7 studies measured health literacy. In addition, only 5 studies assessed literacy using health literacy measurement tools.

Conclusions: This review found that the provision of educational content was satisfactory in most eHealth studies using selected visual methods, such as videos, films, and movies. These findings suggest that eHealth interventions influence people's health behaviors and that the need for this intervention is expected to increase. Despite the need to develop eHealth interventions, standardized measurement tools to evaluate health literacy are lacking. Further research is required to clarify acceptable health literacy measurements.

(JMIR Mhealth Uhealth 2020;8(12):e18316) doi:10.2196/18316

KEYWORDS

application; educational; eHealth; health literacy; lifestyle-related disease; mHealth; review

Introduction

Lifestyle-Related Diseases

Lifestyle-related diseases, such as stroke, heart disease, cancer, diabetes, and chronic respiratory disease, are leading examples of noncommunicable diseases and are now the leading causes of death in the world [1]. Of the top 10 causes of death worldwide in 2016, 6 are considered noncommunicable diseases and accounted for 71% of all deaths [2]. Lifestyle behaviors (eg, smoking, harmful consumption of alcohol, overeating, lack of exercise) or conditions (eg, overweight or obesity, hypertension, abnormal lipid metabolism, and hyperglycemia) are common risk factors for lifestyle-related diseases [3,4]. Lifestyle improvement (eg, high-quality diet, increased exercise) plays an important role in the prevention of lifestyle-related diseases [4,5], and improvement for some conditions can be achieved using eHealth-based interventions [3]. Furthermore, the association between lifestyle behavior and health literacy has been widely recognized [4,6].

Health Literacy

Health literacy is defined as "people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course" [7]. Low health literacy is associated with poor health outcomes and delayed diagnosis and treatment [8]. For example, people with limited health literacy may experience a distrust of providers, pessimism about treatment, and poor satisfaction with the quality of care, probably due to communication difficulties (understanding verbal directions, signs, and placards as well as complexity of instructions) [8]. Occasionally, these patients find it difficult to navigate their way in health care facilities and are therefore unable to receive primary prevention [9]. Conversely, an improvement in health literacy is associated with better health outcomes, such as changes in risk for chronic disease, a reduction in reported disease severity, and decreases in the number of emergency department visits and hospitalizations [10]. Improving health literacy could even expand and extend people's lives in the social, cultural, and work dimensions [7,11]. High rates of low health literacy in populations have prompted governments and national agencies in affected countries to develop national strategies and targets aimed at improving the health literacy of the general population [10].

Information and Communication Technologies, eHealth, and Mobile Health

The popularity of mobile technologies remains high and the number of users of mobile technologies is increasing [12]. The growing usability of information and communication technologies (ICTs), including mobile apps and web-based applications, can increase the accessibility of health support systems [13-15]. eHealth includes medical information services, including public health services, that are distributed via the internet and related technologies [16]. Mobile health (mHealth) is an expanding area within eHealth and involves the use of mobile computing in the fields of medicine and public health [12]. The use of mHealth services, including smartphone-based services, may benefit health care providers by exerting positive effects on patient education, diagnosis, and management as components of the health delivery processes [12,15]. Smartphones provide a range of functions, including telephone calls, text messages (SMS), photos, video, and web access [12,17].

Educational content to improve individuals' health literacy on lifestyle-related diseases can be offered in many ways, such as group learning, questionnaires, internet-based information searches, and downloadable apps [18-22]. Additionally, smartphone apps provide information through visual methods, such as graphics, videos, and pictures, which facilitate user understanding [14,21-23]. Many children and adults play video games, which include messages and entertaining formats and may lead users to change their health behaviors [18,24]. Accordingly, we wanted to investigate whether videos, games, and pictures are effective in increasing health awareness.

Objectives

The aims of this review are to identify literature published in the past decade on eHealth interventions that aimed to improve health literacy on lifestyle-related diseases among the general population using selected visual methods, such as videos, films, and movies. Specifically, our review seeks to categorize study characteristics and the evaluation of the studies' impacts on health literacy. Four main themes are discussed in this review: target age, measurement of health literacy, dietary health behavior, and the future directions of eHealth interventions.

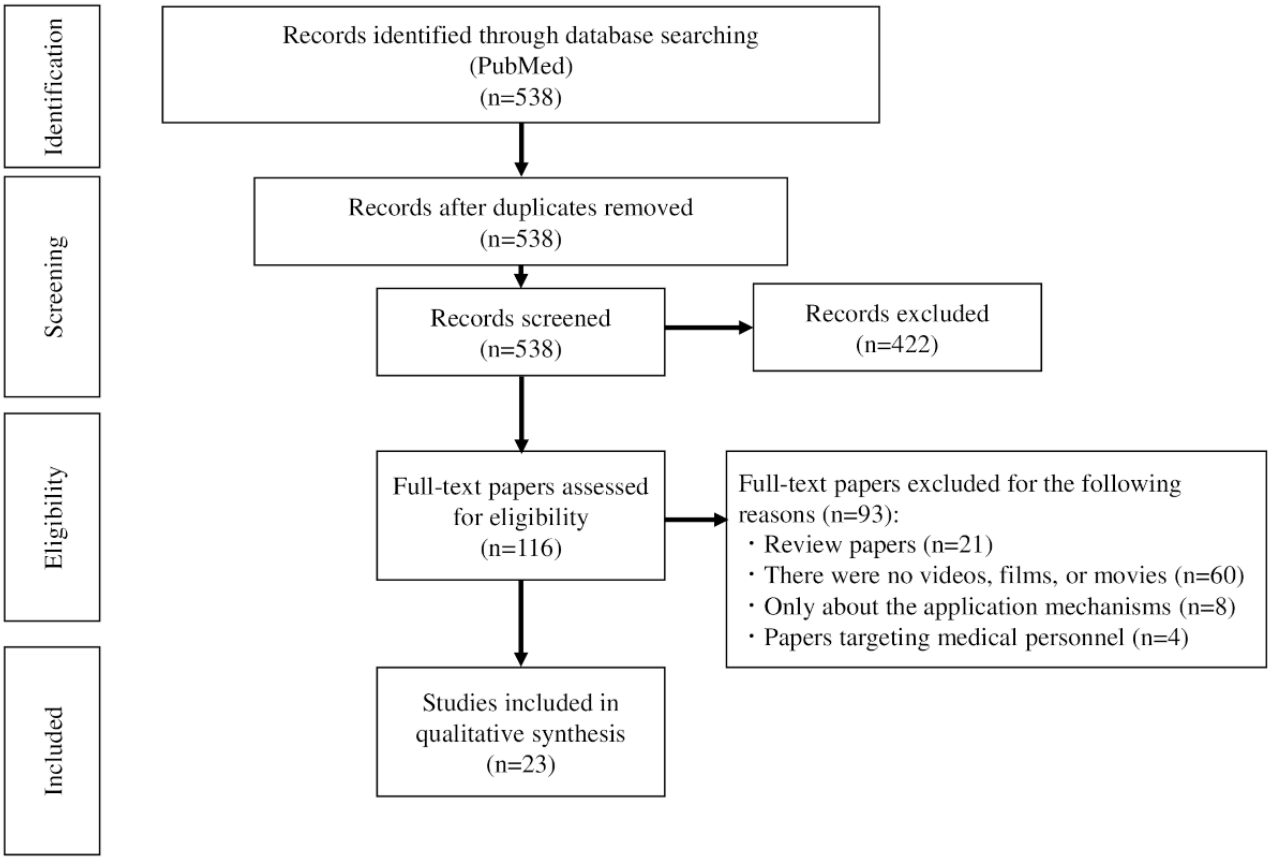
Methods

This literature review was performed using a systematic search and was conducted to emphasize the integration of studies across broader topics [3], with reference to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [25,26]. The search strategy was developed on April 2, 2019, by AA and TS, who also conducted the search.

Details of the search strategy, study selection, and extraction of information can be found in Multimedia Appendix 1. In brief, PubMed was searched in the title and abstract fields using search terms that covered 5 main domains: "digital health," "mHealth," "education," "health literacy," and "visual methods." The search identified 538 papers, with no duplicates (Figure 1). All papers were screened using a 2-stage process. In the first stage, 373 papers that did not meet the inclusion criteria were excluded. In the second stage, an additional 49 papers were excluded. The full texts of the 116 remaining papers were further surveyed,

resulting in the exclusion of 93 papers, thereby leaving a total of 23 papers to review.

Figure 1. PRISMA flow diagram of the paper selection process. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



A data-charting form was created to categorize the included studies, the study characteristics (author and year of publication, study design and region where the study was conducted, study objective, target disease and participant age, research period, outcomes, and research method); the playback time of the educational videos, films, and movies; and the evaluation of the study's impacts on health literacy (measurement to assess health literacy).

Results

Characteristics of the Included Studies

The included studies are characterized in Multimedia Appendix 2. Of the 23 papers we explored, 13 [18,27-38] were conducted in 2 stages, namely app development followed by intervention research. Two studies conducted randomized controlled trials (RCTs) [27,39], of which 1 [27] described the study protocol only. In addition, 8 studies were pilot studies [28-32,40-42], 6 studies were feasibility studies [33-38], 1 was an intervention study [43], and 1 was a cohort study using qualitative and quantitative methods [44]. As shown in Multimedia Appendix 3, primary outcomes were measured and the results were reported in 1 RCT [39], 1 intervention study [43], 5 pilot studies [30,31,40-42], and 2 feasibility studies [34,36]. A total of 7 studies [30,34,39-43] investigated the results statistically, of which 5 studies [31,34,39,40,43] reported significant positive effects on health literacy [34,39,43], and health-related measures (eg, physical activity, body weight) [30,41].

In addition, 3 noninterventional studies [18,45,46] were included because their main topics were video games, which was consistent with our search term of apps that included "games."

Study Objectives and Target Diseases

The objectives of the included studies were categorized into 3 groups: (1) health promotion, (2) disease prevention, and (3) disease management related to the target diseases.

One study [40] was conducted for health promotion and was an innovative mHealth cardiovascular health promotion program, 9 studies [18,27,30,33-35,45-47] aimed to prevent disease by reducing the risks for the diseases, 4 studies [18,34,45,46] promoted smoking cessation for smokers, 2 studies promoted weight loss to address obesity [30,33], 2 studies promoted diet and exercise to lower blood glucose levels in order to manage prediabetes [27,35], and 1 study [47] promoted physical activity to prevent lifestyle-related diseases.

A total of 15 studies [27-29,31,32,35-39,41-44,48] aimed to manage a disease, including diabetes (n=6) [27,31,35,38,42,43] (1 study targeted type 1 diabetes [42] and the remainder targeted type 2 diabetes [27,31,35,43,47]), heart failure (n=3) [41,44,48], cardiovascular disease (n=2) [28,32], stroke (n=2) [36,37], gout (n=1) [29], chronic obstructive pulmonary disease (n=1) [38], and osteoarthritis of the knee (n=1) [39].

Outcomes

Outcomes were investigated in 18 studies [18,27,29-36,38-44,48] and categorized into 2 types: (1) changes in measured values and (2) changes in dietary health behaviors.

Studies used a number of measured values: body weight [27,29,30,33,41,44,48], waist circumference [27,30], blood pressure [27,44], oxygen usage [38,44], hemoglobin A_{1c} [27,38,42], serum uric acid [29], and knowledge about health conditions (eg, diabetes knowledge, self-efficacy score, cardiovascular health knowledge, smoking knowledge) [31,34,39,40,42,43]. Studies also investigated behavioral changes, smoking attitude [18,34], nutrition [33,35,38,40], physical activity [33,35,38,41], stress [33], medication [32,38], specific health behavior [27], and usability and acceptability of the intervention [30,36].

To assess changes in measured values, individuals' health literacy was tested in 7 studies [31,34,39,40,42,43,48], and measurement tools were used in 5 studies only [31,34,40,42,43]. A 27-item questionnaire based on the Norwegian National Health Informatics' diabetes quiz to test theoretical knowledge was conducted for individuals aged 13 to 19 years with type 1 diabetes. The eHealth Literacy Scale (eHEALS) [40] and smoking knowledge score [34] were used for participants 18 years or older. The Rapid Estimate in Adult Literacy in Medicine (REALM) [31], the Diabetes Self-Efficacy Scale [31], and 2 types of study knowledge tests [39,48] were used for participants older than 40 years.

Platform of Development of eHealth Service

We categorized the included studies into 3 types of platforms: (1) applications (web-based applications or mobile apps), (2) websites, and (3) others.

Various apps were used in 22 studies [18,27-37,39-48]. Apps allow users to use interactive content [30,32,39,43], telephone interviews [44], face-to-face video conferencing [33,40,44], and social network service messages [42]. User satisfaction was evaluated for the intervention itself and for opinions on future development in 16 studies [18,30-32,34-37,39,41,42,44-48]. The 14 apps [18,28-32,34,37,39,41,42,44,45,48] receiving the highest satisfaction and appreciation ratings included those providing educational content (about diseases [30-32,34,39,41,48], through a diary program [29,42], and in game content [18,45,46]) using pictures, graphics [27,29,30,34,37,42,44], videos [18,27-31,33-37,39,41-45,47], icons, drawings, animations [47], and an avatar [48]. Only one app providing educational video [30] was considered unacceptable because the process of downloading and viewing the video was too difficult or took too much time.

By connecting with a wearable device [28], sensing devices (eg, blood pressure monitor, weight scales, pulse oximeter) [44], and Bluetooth [42], health information (eg, physical activity, blood glucose values) could be provided remotely.

Websites were based on a textual design to provide educational messages but, unlike apps, were unable to include a gaming component to provide some type of feedback-based reward [45]. Another platform [38] developed internet-enabled home programs that provided educational videos, individual consultations, and a health diary remotely using the patient's television at home and a connection to a computer.

Discussion

In this review, we aimed to identify and characterize the features of existing literature describing eHealth interventions that aimed to improve health literacy on lifestyle-related diseases among the general population using selected visual methods, such as videos, films, and movies. Through this research, we identified 4 themes, which we discuss here with their strengths and limitations.

Target Age

We found a difference in target age compared with previous studies focused on health literacy. Kim and Xie [14] found that health literacy was limited in individuals older than 65 years. In this review, 19 studies [27-36,38-42,44,45,47,48] included participants younger than 65 years. Although it appears that people with low health literacy are older [49,50], young adults also lack health literacy, including eHealth literacy [51,52].

In our review, we identified various ideas to facilitate the use of educational content to improve participants' health literacy and found that the kind of educational content provided was required to change as the age range of the target population widened. The use of icons and avatars facilitated usage for both younger and older individuals with low health literacy [41,48]. The younger the target age, the greater the acceptability of games [46].

Measurement of Health Literacy

Measurement tools for health literacy have yet to be established. Indeed, in this review, among the 22 included studies that developed apps for education [18,27-37,39-48], only 5 studies [31,34,40,42,43] used instruments to measure health literacy.

According to a review [14] that identified the relevant literature and examined the instruments used to measure individual health literacy levels, most studies used the eHEALS [53] and the Short Test of Functional Health Literacy (S-TOFHLA) [54]. The S-TOFHLA is a shortened version of the Test of Functional Health Literacy [55], which correlates with the REALM [56]. In another systematic review of health literacy using web-based health information environments [26], the Newest Vital Sign was used most often [57], followed by the REALM [58]. There is a lack of standard measurement tools to evaluate health literacy [17,22,59-62]. Additional research is required to identify measurement methods suitable for evaluating levels of health literacy.

Dietary Health Behavior Change

Improvements to lifestyle play an important role in lifestyle-related diseases [4,5]. Although 12 included studies [18,27,30,32-36,38,40,41,44] had changes in dietary health behavior as an outcome, none described the mechanism of the relationship between the eHealth intervention and the dietary health behavior. One study [32] classified an educational interventional app in terms of behavior change techniques

(BCTs) used in behavior change interventions. Future research to investigate the underlying mechanism of BCTs will be useful in clarifying which interventions are likely to be effective.

Future Directions of eHealth Interventions

Based on user evaluations of the interventions and opinions on their future development, eHealth interventions using visual methods and interactive approaches improve user motivation to improve lifestyle and health literacy. However, 1 study [30] reported that the use of educational videos was not acceptable due to the difficulty older participants experienced in using the mobile app. Future eHealth-based interventions will be required to improve users' computer literacy and identify the mediating effects of age and sex.

Strengths and Limitations

To our knowledge, this study is the first to review the educational content of eHealth aimed at improving individuals' health literacy on lifestyle-related diseases. Our findings may provide a new perspective on the development of apps that use eHealth to address lifestyle-related diseases and improve people's health literacy.

Several limitations of our study warrant mention. One limitation is the search process: we searched only a single database and included only studies written in English and published in the decade up to April 2019. As discussed in other review papers about eHealth-based interventions [3,14], selecting from additional databases that include unpublished studies is useful for a broader review and should be a consideration for any future systematic review on the topic. Nevertheless, the search was conducted in a systematic manner using the PRISMA flow diagram [25,26]. Additionally, we excluded review papers; papers about AIDS, cancer, psychiatric conditions, odontology, and pediatrics; papers whose target populations were medical personnel and pregnant women; and papers that did not include educational videos, films, or movies. This strategy may have been too strict and some relevant papers may have been missed. Future reviews should be based on wider search criteria.

Conclusions

Our review provides an overview of the relationships between eHealth-based interventions with selected visual methods, such as videos, films, and movies, and improved health outcomes (ie, changes in measured values and dietary health behavior). Despite the necessity of self-management systems using ICT to control lifestyle-related diseases, relatively few studies have explored educational videos, films, and movies aimed at improving health outcomes.

We also found that the concept of literacy and the tools used to measure the outcome of health literacy–related interventions have not been unified. To more accurately evaluate levels of health literacy and the effects of interventions, future studies need to clarify the concept of health literacy and develop health literacy screening tools.

Acknowledgments

We thank all staff members at the Center of Innovation, University of Tokyo, as well as all members of Precision Health, University of Tokyo for their support.

This research was supported by the Center of Innovation Program from the Japan Science and Technology Agency (grant No. JPMJCE1304).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods.
[DOCX File , 23 KB - mhealth_v8i12e18316_app1.docx]

Multimedia Appendix 2
Characteristics of the included studies.
[DOCX File , 47 KB - mhealth_v8i12e18316_app2.docx]

Multimedia Appendix 3
Summary of the study design, intervention, and the results of the included randomized controlled trial (RCT), intervention study, pilot studies, and feasibility studies.
[DOCX File , 25 KB - mhealth_v8i12e18316_app3.docx]

References

1. Noncommunicable diseases. World Health Organization. URL: https://www.who.int/gho/ncd/en/ [accessed 2020-06-29]
2. The Top 10 causes of death. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2020-07-14]

3. Kankanhalli A, Shin J, Oh H. Mobile-Based Interventions for Dietary Behavior Change and Health Outcomes: Scoping Review. JMIR Mhealth Uhealth 2019 Jan 21;7(1):e11312 [FREE Full text] [doi: 10.2196/11312] [Medline: 30664461]
4. Davey J, Holden CA, Smith BJ. The correlates of chronic disease-related health literacy and its components among men: a systematic review. BMC Public Health 2015 Jun 26;15:589 [FREE Full text] [doi: 10.1186/s12889-015-1900-5] [Medline: 26112264]
5. Yamanaka-Okumura H, Tatano H, Kajiura D, Yamaguchi C, Masuda M, Taketani Y. [Key treatment of lifestyle-related diseases: nutritional education and practice]. Clin Calcium 2016 Mar;26(3):441-446. [Medline: 26923983]
6. Risk factors contributing to chronic disease. Australian Institute of Health and Wellness. 2012. URL: https://www.aihw.gov.au/getmedia/74121d1b-69ca-4a34-a08a-51b741ea26b2/12724.pdf.aspx?inline=true [accessed 2020-04-29]
7. Sørensen K, Van DBS, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
8. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. Am J Health Behav 2007;31(Suppl 1):S19-S26. [doi: 10.5555/ajhb.2007.31.supp.S19] [Medline: 17931132]
9. Baker DW, Parker RM, Williams MV, Pitkin K, Parikh NS, Coates W, et al. The health care experience of patients with low literacy. Arch Fam Med 1996 Jun;5(6):329-334. [doi: 10.1001/archfami.5.6.329] [Medline: 8640322]
10. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. Health Promot Int 2018 Oct 01;33(5):901-911. [doi: 10.1093/heapro/dax015] [Medline: 28369557]
11. Baur C, Harris L, Squire E. The U.S. National Action Plan to Improve Health Literacy: A Model for Positive Organizational Change. Stud Health Technol Inform 2017;240:186-202. [Medline: 28972518]
12. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013 Jan;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
13. Peterson PE, Kwan S. The 7th WHO Global Conference on Health Promotion - towards integration of oral health (Nairobi, Kenya 2009). Community Dent Health 2010;27(3)(suppl 1):129-136. [doi: 10.1922/CDH_2643Petersen08]
14. Kim H, Xie B. Health literacy in the eHealth era: A systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
15. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
16. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):e20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
17. Ownby RL, Acevedo A, Waldrop-Valverde D, Jacobs RJ, Caballero J, Davenport R, et al. Development and initial validation of a computer-administered health literacy assessment in Spanish and English: FLIGHT/VIDAS. Patient Relat Outcome Meas 2013;4:21-35 [FREE Full text] [doi: 10.2147/PROM.S48384] [Medline: 23990736]
18. Bindoff I, de Salas K, Peterson G, Ling T, Lewis I, Wells L, et al. Quittr: The Design of a Video Game to Support Smoking Cessation. JMIR Serious Games 2016 Dec 01;4(2):e19 [FREE Full text] [doi: 10.2196/games.6258] [Medline: 27908844]
19. Levin-Zamir D, Bertschi I. Media Health Literacy, eHealth Literacy, and the Role of the Social Environment in Context. Int J Environ Res Public Health 2018 Dec 03;15(8) [FREE Full text] [doi: 10.3390/ijerph15081643] [Medline: 30081465]
20. Paige SR, Alber JM, Stellefson ML, Krieger JL. Missing the mark for patient engagement: mHealth literacy strategies and behavior change processes in smoking cessation apps. Patient Educ Couns 2018 May;101(5):951-955. [doi: 10.1016/j.pec.2017.11.006] [Medline: 29153592]
21. 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;12(3):e36 [FREE Full text] [doi: 10.2196/jmir.1269] [Medline: 20805091]
22. Yost KJ, Webster K, Baker DW, Jacobs EA, Anderson A, Hahn EA. Acceptability of the talking touchscreen for health literacy assessment. J Health Commun 2010;15(Suppl 2):80-92 [FREE Full text] [doi: 10.1080/10810730.2010.500713] [Medline: 20845195]
23. Choi J, Bakken S. Web-based education for low-literate parents in Neonatal Intensive Care Unit: development of a website and heuristic evaluation and usability testing. Int J Med Inform 2010 Aug;79(8):565-575 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.05.001] [Medline: 20617546]
24. Lenhart A, Kahne J, Middaugh E, Rankin Macgill A, Evans C, Vitak J. Teens, video games, and civics: teens' gaming experiences are diverse and include significant social interaction and civic engagement. Pew Internet and American Life Project. 2008. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2008/09/PI_2008.09.16_teens-video-games_REPORT.pdf [accessed 2020-06-23]
25. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
26. Huhta A, Hirvonen N, Huotari M. Health literacy in web-based health information environments: systematic review of concepts, definitions, and operationalization for measurement. J Med Internet Res 2018 Dec 19;20(12):e10273 [FREE Full text] [doi: 10.2196/10273] [Medline: 30567690]

27. Sarfati D, McLeod M, Stanley J, Signal V, Stairmand J, Krebs J, et al. BetaMe: impact of a comprehensive digital health programme on HbA1c and weight at 12 months for people with diabetes and pre-diabetes: study protocol for a randomised controlled trial. Trials 2018 Mar 05;19(1):161 [FREE Full text] [doi: 10.1186/s13063-018-2528-4] [Medline: 29506562]
28. Ceasar JN, Claudel SE, Andrews MR, Tamura K, Mitchell V, Brooks AT, et al. Community Engagement in the Development of an mHealth-Enabled Physical Activity and Cardiovascular Health Intervention (Step It Up): Pilot Focus Group Study. JMIR Form Res 2019 Jan 04;3(1):e10944 [FREE Full text] [doi: 10.2196/10944] [Medline: 30684422]
29. Nguyen AD, Frensham LJ, Wong MX, Meslin SM, Martin P, Lau AY, et al. mHealth App Patient Testing and Review of Educational Materials Designed for Self-Management of Gout Patients: Descriptive Qualitative Studies. JMIR Mhealth Uhealth 2018 Oct 15;6(10):e182 [FREE Full text] [doi: 10.2196/mhealth.9811] [Medline: 30322835]
30. Choo S, Kim JY, Jung SY, Kim S, Kim JE, Han JS, et al. Development of a Weight Loss Mobile App Linked With an Accelerometer for Use in the Clinic: Usability, Acceptability, and Early Testing of its Impact on the Patient-Doctor Relationship. JMIR Mhealth Uhealth 2016;4(1):e24 [FREE Full text] [doi: 10.2196/mhealth.4546] [Medline: 27032541]
31. Wood FG, Alley E, Baer S, Johnson R. Interactive Multimedia Tailored to Improve Diabetes Self-Management. Nurs Clin North Am 2015 Sep;50(3):565-576. [doi: 10.1016/j.cnur.2015.05.009] [Medline: 26333610]
32. Boyd AD, Moores K, Shah V, Sadhu E, Shroff A, Groo V, et al. My Interventional Drug-Eluting Stent Educational App (MyIDEA): Patient-Centered Design Methodology. JMIR Mhealth Uhealth 2015;3(3):e74 [FREE Full text] [doi: 10.2196/mhealth.4021] [Medline: 26139587]
33. Materia FT, Smyth JM, Heron KE, Hillemeier M, Feinberg ME, Fonzi P, et al. Preconceptional health behavior change in women with overweight and obesity: prototype for SMART strong healthy women intervention. Mhealth 2018;4:24 [FREE Full text] [doi: 10.21037/mhealth.2018.06.06] [Medline: 30148139]
34. Finkelstein J, Cha EM. Using a Mobile App to Promote Smoking Cessation in Hospitalized Patients. JMIR Mhealth Uhealth 2016 May 06;4(2):e59 [FREE Full text] [doi: 10.2196/mhealth.5149] [Medline: 27154792]
35. Fontil V, McDermott K, Tieu L, Rios C, Gibson E, Sweet CC, et al. Adaptation and Feasibility Study of a Digital Health Program to Prevent Diabetes among Low-Income Patients: Results from a Partnership between a Digital Health Company and an Academic Research Team. J Diabetes Res 2016;2016:8472391 [FREE Full text] [doi: 10.1155/2016/8472391] [Medline: 27868070]
36. Sureshkumar K, Murthy G, Natarajan S, Naveen C, Goenka S, Kuper H. Evaluation of the feasibility and acceptability of the 'Care for Stroke' intervention in India, a smartphone-enabled, carer-supported, educational intervention for management of disability following stroke. BMJ Open 2016 Feb 02;6(2):e009243 [FREE Full text] [doi: 10.1136/bmjopen-2015-009243] [Medline: 26839011]
37. Sureshkumar K, Murthy GVS, Munuswamy S, Goenka S, Kuper H. 'Care for Stroke', a web-based, smartphone-enabled educational intervention for management of physical disabilities following stroke: feasibility in the Indian context. BMJ Innov 2015 Jul;1(3):127-136 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000056] [Medline: 26246902]
38. Burkow TM, Vognild LK, Østengen G, Johnsen E, Risberg MJ, Bratvold A, et al. Internet-enabled pulmonary rehabilitation and diabetes education in group settings at home: a preliminary study of patient acceptability. BMC Med Inform Decis Mak 2013 Mar 05;13:33 [FREE Full text] [doi: 10.1186/1472-6947-13-33] [Medline: 23496829]
39. Timmers T, Janssen L, Pronk Y, van der Zwaard BC, Koëter S, van Oostveen D, et al. Assessing the Efficacy of an Educational Smartphone or Tablet App With Subdivided and Interactive Content to Increase Patients' Medical Knowledge: Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e10742 [FREE Full text] [doi: 10.2196/10742] [Medline: 30578185]
40. Brewer LC, Jenkins S, Lackore K, Johnson J, Jones C, Cooper LA, et al. mHealth Intervention Promoting Cardiovascular Health Among African-Americans: Recruitment and Baseline Characteristics of a Pilot Study. JMIR Res Protoc 2018 Jan 31;7(1):e31 [FREE Full text] [doi: 10.2196/resprot.8842] [Medline: 29386174]
41. Lloyd T, Buck H, Foy A, Black S, Pinter A, Pogash R, et al. The Penn State Heart Assistant: A pilot study of a web-based intervention to improve self-care of heart failure patients. Health Informatics J 2019 Jun;25(2):292-303. [doi: 10.1177/1460458217704247] [Medline: 28504048]
42. Frøisland DH, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: mixed-methods study. J Med Internet Res 2012;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
43. Alanzi T, Bah S, Alzahrani S, Alshammari S, Almunsef F. Evaluation of a mobile social networking application for improving diabetes Type 2 knowledge: an intervention study using WhatsApp. J Comp Eff Res 2018 Sep;7(9):891-899. [doi: 10.2217/cer-2018-0028] [Medline: 29943647]
44. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, SUPPORT-HF Investigators. A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study. Int J Med Inform 2015 Oct;84(10):743-753. [doi: 10.1016/j.ijmedinf.2015.05.003] [Medline: 26037921]
45. Watson AM, Alber JM, Barnett TE, Mercado R, Bernhardt JM. Content Analysis of Anti-Tobacco Videogames: Characteristics, Content, and Qualities. Games Health J 2016 Jun;5(3):216-223. [doi: 10.1089/g4h.2015.0096] [Medline: 27096576]

46. BinDhim NF, Freeman B, Trevena L. Pro-smoking apps: where, how and who are most at risk. Tob Control 2015 Mar;24(2):159-161. [doi: 10.1136/tobaccocontrol-2013-051189] [Medline: 24046212]
47. Bender MS, Martinez S, Kennedy C. Designing a Culturally Appropriate Visually Enhanced Low-Text Mobile Health App Promoting Physical Activity for Latinos: A Qualitative Study. J Transcult Nurs 2016 Dec;27(4):420-428. [doi: 10.1177/1043659614565249] [Medline: 25556208]
48. Athilingam P, Osorio RE, Kaplan H, Oliver D, O'neachtain T, Rogal PJ. Embedding Patient Education in Mobile Platform for Patients With Heart Failure: Theory-Based Development and Beta Testing. Comput Inform Nurs 2016 Feb;34(2):92-98. [doi: 10.1097/CIN.0000000000000216] [Medline: 26765655]
49. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
50. Watkins I, Xie B. eHealth literacy interventions for older adults: a systematic review of the literature. J Med Internet Res 2014;16(11):e225 [FREE Full text] [doi: 10.2196/jmir.3318] [Medline: 25386719]
51. Janiak E, Rhodes E, Foster AM. Translating access into utilization: lessons from the design and evaluation of a health insurance Web site to promote reproductive health care for young women in Massachusetts. Contraception 2013 Dec;88(6):684-690. [doi: 10.1016/j.contraception.2013.09.004] [Medline: 24120249]
52. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
53. Fagnano M, Halterman JS, Conn KM, Shone LP. Health literacy and sources of health information for caregivers of urban children with asthma. Clin Pediatr (Phila) 2012 Mar;51(3):267-273 [FREE Full text] [doi: 10.1177/0009922811421413] [Medline: 21911409]
54. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med 1995 Oct;10(10):537-541. [Medline: 8576769]
55. Smith SS, Piper ME, Bolt DM, Fiore MC, Wetter DW, Cinciripini PM, et al. Development of the Brief Wisconsin Inventory of Smoking Dependence Motives. Nicotine Tob Res 2010 May;12(5):489-499 [FREE Full text] [doi: 10.1093/ntr/ntq032] [Medline: 20231242]
56. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. JAMA 1999 Feb 10;281(6):552-557. [Medline: 10022112]
57. Noblin AM, Wan TTH, Fottler M. The impact of health literacy on a patient's decision to adopt a personal health record. Perspect Health Inf Manag 2012;9:1-13 [FREE Full text] [Medline: 23209454]
58. Schprechman JP, Gathright EC, Goldstein CM, Guerini KA, Dolansky MA, Redle J, et al. Health literacy and global cognitive function predict e-mail but not internet use in heart failure patients. Nurs Res Pract 2013;2013:507910 [FREE Full text] [doi: 10.1155/2013/507910] [Medline: 24282634]
59. Hahn EA, Choi SW, Griffith JW, Yost KJ, Baker DW. Health literacy assessment using talking touchscreen technology (Health LiTT): a new item response theory-based measure of health literacy. J Health Commun 2011;16(Suppl 3):150-162 [FREE Full text] [doi: 10.1080/10810730.2011.605434] [Medline: 21951249]
60. Ownby RL, Waldrop-Valverde D, Hardigan P, Caballero J, Jacobs R, Acevedo A. Development and validation of a brief computer-administered HIV-Related Health Literacy Scale (HIV-HL). AIDS Behav 2013 Feb;17(2):710-718 [FREE Full text] [doi: 10.1007/s10461-012-0301-3] [Medline: 22961499]
61. van Deursen AJAM, van Dijk JAGM. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
62. Baker DW. The meaning and the measure of health literacy. J Gen Intern Med 2006 Aug;21(8):878-883 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00540.x] [Medline: 16881951]

Abbreviations

BCT: behavior change technique
eHEALS: eHealth Literacy Scale
ICT: information and communication technology
mHealth: mobile health
RCT: randomized controlled trial
REALM: Rapid Estimate in Adult Literacy in Medicine
S-TOFHLA: Short Test of Functional Health Literacy

Edited by CL Parra-Calderón, L Buis; submitted 21.02.20; peer-reviewed by K Schroeder, A Wickersham; comments to author 14.06.20; revised version received 07.08.20; accepted 19.11.20; published 09.12.20.

Please cite as:
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
URL: http://mhealth.jmir.org/2020/12/e18316/
doi:10.2196/18316
PMID:33295296

Original Paper

Integrated Digital Patient Education at the Bedside for Patients with Chronic Conditions: Observational Study

Benjamin Schooley[1*], MBA, PhD; Akanksha Singh[2*], MS; Neşet Hikmet[1*], PhD; Robert Brookshire[1*], MEd, PhD; Nitin Patel[3,4*], MD

[1]Health Information Technology Program, College of Engineering and Computing, University of South Carolina, Columbia, SC, United States

[2]Department of Computer Science and Engineering, College of Engineering and Computing, University of South Carolina, Columbia, SC, United States

[3]Department of Medicine, School of Medicine, University of South Carolina, Columbia, SC, United States

[4]Digital Health, Prisma Health, Columbia, SC, United States

[*]all authors contributed equally

Corresponding Author:
Benjamin Schooley, MBA, PhD
Health Information Technology Program
College of Engineering and Computing
University of South Carolina
550 Assembly St
Columbia, SC, 29208
United States
Phone: 1 8937778223
Email: bschooley@cec.sc.edu

Abstract

Background: Patient education delivered by a health care provider increases patients' understanding and adherence to medical instructions, which helps to improve patient health. Multiple challenges exist to delivering effective patient education to patients with multiple chronic conditions, including giving the necessary time, range, and types of learning materials, and assessing the level of understanding. To help overcome these challenges, it is important to study new electronic means to assist in patient education, such as the use of mobile devices, interactive media, 3-dimensional images, and multimedia educational content at the bedside.

Objective: The goal of this study was to address the need for blended learning strategies combining technical and workflow integration of digital patient education systems for patients with chronic conditions within and across the regular process of care. Studies are needed to evaluate the utility and benefits of these technologies for providers and patients alike.

Methods: A mixed-methods approach was employed including survey administration to 178 patients after they received digital patient education in person with a health care provider, and qualitative interviews with 16 nurse educators who used the mobile digital health education technology to deliver instruction to patients. Patient survey data were analyzed using chi-square statistical tests. Qualitative interviews were analyzed for user acceptance and perceived value themes.

Results: Patients who were counseled using a blended digital health education approach reported improved understanding of educational content (P=.034) and chronic health conditions (P<.001), were more motivated to care for themselves at home (P<.001), were more likely to say that they felt capable of making health care decisions with their doctors (P<.001) and on their own (P=.001), and were more likely to report their intention to follow their doctor's instructions (P<.001) than were patients whose education was not computer-based. Nurse educators felt that the digital education system and content enhanced their education efforts and could be easily integrated into the outpatient clinical workflow.

Conclusions: Patient education for individuals with chronic conditions may be more effective than traditional formats when provided in blended digital formats supervised by a health care provider.

(JMIR Mhealth Uhealth 2020;8(12):e22947) doi:10.2196/22947

KEYWORDS

patient education; patient understanding; blended learning; adherence; digital patient education; chronic condition; understanding; outcome; mHealth; education; digital health

Introduction

Background

Patient education increases patient understanding and adherence to medical instructions, which has a major impact on long-range outcomes for patients with chronic health conditions [1-3]. A recent meta-analysis of 320 articles on patient compliance with instructions and its associated impacts on physiological progress and long-range health outcomes for chronic disease management found that patient education was the most successful of all experimental effects [4]. Information technology (IT) facilitates fundamental changes to patient education delivery, enabling digital media content that is considered a time-effective and cost-effective alternative to traditional patient education [5]. The use of digital technologies for patient learning has increased drastically through the application of web-based patient education sites, video and 3-dimensional (3D) multimedia content, mobile devices (such as smartphones, MP3 players, iPods [Apple Inc], and tablets), and other devices and technology media [6]. However, many questions remain about which digital educational delivery strategies are best for a given setting.

Patients with chronic conditions require a holistic approach for care, education, and self-management to achieve patient understanding and motivation to adhere to instructions [7]. Lack of patient awareness about a condition has been identified as the most common barrier to active self-management of chronic conditions [8]. Further, despite good patient understanding about a specific treatment, adherence worsens with increasing age and complexity of treatments [9]. For patients with multiple comorbidities, there is a greater need for in-depth education that includes provider interaction [7,10,11]. Indeed, improving patient education for people with chronic conditions may be an important component to improving individual and community health delivery systems, decision making, and related outcomes [12]. Providing better education with the level of personalization that health care providers are able to deliver is an approach advocated by patients with chronic conditions [13].

Prior research has validated the use of a wide range of education and communication interventions for improving patient comprehension [14]. For example, some educational modes include using cartoon images [15], educational booklets [16], educational pamphlets [17], simple visual aids [18], and telephone-based education and reinforcement [19]. These education modes have been shown to have a positive effect on patients. Positive effects have also been demonstrated with digital and computer-enhanced education, including educational videos to supplement a traditional office consultation, web-based education programs that utilize text and voice reminders, digital multimedia and 3D interactive content [3,20], animated messaging, personalized video and flash content based on patient health status and demographic characteristics [21], tablet/PC-based education in place of nurse-led conventional patient education [22], use of multimedia presentations [21], mobile apps [22], and a wide range of other digital tools across various health care and patient contexts [23,24]. These studies have largely demonstrated improved patient education as well as higher satisfaction levels among providers [25]. Results have also indicated high levels of technology acceptance for the purpose of patient education, while also indicating that solutions need to be tailored to the needs of providers and specific patient populations [17-20].

Prior digital patient education studies have been conducted in patients undergoing or recovering from specific procedures, such as total joint arthroplasty and total hip replacement, endovascular aneurysm repair, peripheral angioplasty, catheter insertion procedures, endoscopy, and perioperative interventions, as well as in patients seeking to improve health conditions, including immunosuppression, oral health care, heart disease, lower back pain, asthma, and other conditions. It has been widely concluded that patients in specific medical contexts, receiving relevant content, have improved understanding as well as motivation to adhere to instructions when using digital learning formats, which in turn has led to higher levels of satisfaction among these patients [21,22,26-34]. Computer-aided learning has been found to have positive impacts on the knowledge, attitude, behavior, and health of patients [27].

Evaluation measures used to assess the effectiveness of computer-enhanced patient education in most studies include patient understanding, patient satisfaction, adherence to process, and improved usability. Younger, higher educated, and internet-savvy patients are found to be more tech savvy and have higher effectiveness levels when evaluated using these measures [5,35-38]. Patients have found the use of multimedia presentations during the consent process to be helpful in their understanding and report improved satisfaction. Increased regular use of multimedia presentations to help patients understand their prescribed procedures is found to be beneficial in the care of patients undergoing some interventions, including vascular and endoscopic procedures [34,35]. One study discovered that mobile device education apps support increases in adherence and patient satisfaction and help facilitate perioperative interventions [37]. Increased utilization of digital formats has shown to enhance patient education and related factors for remote, distance, online, or unsupervised patient education [39], and chronic disease management [40]. Indeed, patient education for chronic disease management is of paramount concern, as the number of individuals being diagnosed with a chronic condition in North America continues to increase, resulting in increased health care utilization [41-43].

Although there is no one-size-fits-all patient education approach, effective instruction may apply a variety of teaching tools, methodologies, and strategies that can be tailored to the individual patient's circumstances and competencies, the learning environment, and the range and types of teaching tools available to providers [44]. Common strategies today include in-person and face-to-face instruction [45] (ranging in size from one person to large groups) and digital learning, where either patients or providers control the amount and type of learning [46]. In-person education is generally believed to allow for more effective question-and-answer communications between patients and providers and better visual and auditory determination of learning competency by the provider. Health care providers involved in the development of e-learning techniques suggest several strategies, including (1) tailoring information according to patient characteristics, (2) structuring information to be

relevant and easy to find, and (3) being aware of and sensitive to patients' emotions.

For online patient education in chronic disease management settings, essential design features include patient-tailored information, interactivity, content credibility, clear presentation of content, use of multimedia, and interpretability [34]. Moreover, e-learning is likely most effective when it facilitates individual learning needs, supports feedback on competence level and improvement, and allows input from significant others (eg, experts, peers, or patients) [36]. It is widely believed that achieving all these objectives can be very challenging when applying digital education on its own without health care provider interaction or intervention, especially for patients experiencing multiple conditions and comorbidities.

Goal of the Study

This study investigated a blended learning patient education program that integrates technology-enhanced instruction with conventional in-person teaching approaches in the process of care [47]. This approach benefits from patient interactions with their provider and from mobile, digital, multimedia-enhanced education [48]. This blended learning approach for patients with multiple chronic conditions has not been extensively researched. One recent meta-analysis included just 12 studies that ranged in length from 4 weeks to 8 months and included both self-paced learning and face-to-face learning components [49]. This study focused on clinician-directed and facilitated digital education during one critical interaction in the regular course of care—a setting that represents the most common and opportune educational environment. The study evaluated differences between groups receiving traditional in-person education and in-person digital education. We evaluated differences in patient understanding, motivation for self-care, intention to adhere to instructions, confidence in care decision making, and satisfaction with clinician-patient communication by using digital technologies in the presence of a clinician in the regular flow of a patient visit (bedside). We employed digital technology in the presence of a nurse educator (NE) to augment and personalize traditional patient education methods in an outpatient clinical setting in the context of educating patients with chronic medical conditions.

Methods

Study Design

A mixed-methods approach was used to collect survey responses from patients receiving patient education and to evaluate NE perspectives on their use of the digital patient education system. The mixed-methods design used for this study follows the guidance of Morse and Niehaus [50], employing quantitative analysis of patient surveys and qualitative interviews and responses from NEs. Informed consent was obtained from all individual participants involved in the study in compliance with the approved University of South Carolina Institutional Review Board protocol (#Pro00059265) for this study.

Overview

From October 2017 to May 2018, participants were recruited from the normal flow of patients visiting one large internal medicine clinic and one hospital in the southeastern United States. Patient participants were enrolled in the study at the time that they met with their doctor, after their doctor determined that they required patient education. Doctors relied on their clinical experience to assess patient education needs. Doctors generally selected patients for instruction who fit into at least one of the following categories: (1) very recently diagnosed with a new chronic condition, (2) new to the clinic with a diagnosed chronic condition, (3) not demonstrating compliance with medical instructions, or (4) not demonstrating expected health improvements after reportedly following medical instructions and prescriptions. Patients were eligible to participate if they were aged 18 years or older, required patient education as determined by their physician, were considered "complex" patients (ie, that they were being treated for more than one chronic condition), and deemed capable of completing a survey on their own without assistance. Physicians chose patients to enroll who were over the age of 45 years—with no patients over the age of 76 years (mean age 58 years)—who were being treated for two or more of the following types of conditions: diabetes, hypertension, congestive heart failure/heart disease, obesity, chronic obstructive pulmonary disease, anxiety/depression, stroke, kidney and/or bladder failure/disease, asthma, lung disease, arthritis, chronic back pain, or osteoporosis.

After providing their informed consent, patient participants were enrolled by the NEs into one of two study arms: (1) mobile device–enabled electronic multimedia education (intervention) or (2) usual paper/pamphlet-based education (control). The NEs, all of whom worked full-time for the participating health care institution, made every effort to alternate the assignment of patients into the two study arms. Nevertheless, this was an observational cohort study in which the assignment of the intervention was not fully at the discretion of the investigator [51,52]. A total of 16 NEs (11 ambulatory, 2 inpatient, and 3 both ambulatory and inpatient), including all of the NEs at the internal medicine clinic and the hospital inpatient educators affiliated with the same clinic, provided education in one of two formats. Patient activation measures were not collected for reporting. From a clinical standpoint, all patients in this study went through the standard patient activation process at the clinic. Participants assigned to the control arm received education face-to-face with an NE in the usual format that included paper handouts, brochures, printable content from the electronic health record, paper charts, and plastic models (eg, heart, skull, etc). The intervention group also received education face-to-face with an NE but using a touchscreen laptop computer with interactive 3D images, videos, graphs, and charts. In both groups, NEs were free to choose the content based on the diagnoses, care plan, and instructions given to each patient by his/her physician. NEs used educational material in topical areas related to internal medicine, health education, general surgery, oncology, cardiology, and orthopedics.

NEs were provided with 2 hours of training on the digital patient education tool and touchscreen laptop computer prior to patient enrolment. NEs were primarily instructed on how to use the technology and were not trained on the specific educational content. This is because all of the NEs in this study had already

become proficient in providing education on all the topics noted above while using paper/pamphlet-based education. The same NEs were interviewed in person at the end of the study, after all patient surveys were completed. Researchers informed the NEs about the study and gained their consent to participate as per the approved Institutional Review Board protocol. Qualitative results are presented using guidance from the consolidated criteria for reporting qualitative research [53].

Survey Data

An online survey was sent to each study participant at the conclusion of each patient education session. Patients responded to questions on demographics, patient understanding, patient-provider communications, motivation to care for self at home, intention to adhere to instructions, and patient confidence in care decision making on one's own and together with a physician. A total of 178 patients completed the patient consent, face-to-face education, and survey instrument. The list of survey questions that were administered are provided in Textbox 1.

Textbox 1. Survey questions.

Survey questions

- Do you understand your condition better now since receiving your education today?
- Did this educator's use of a computer or handheld device make it harder or easier for you to talk with him or her?
- Did the instruction you received motivate you to care for yourself at home?
- Did the instruction you received help you feel capable of making health care decisions together with your doctor?
- Did the instruction you received help you feel capable of making your own health care decisions?
- As a result of the education you received, how likely are you to follow the instructions given to you by your doctor?

Interview Data

All 16 NEs that provided standard and digitally enhanced education were interviewed for 30 minutes each in a private room at their workplace. Semistructured questions were asked that focused on the use of the technology, the benefits and challenges of using it in the process of patient care, and the perceived benefits and challenges to patients. Two researchers (BS and NP) conducted all of the interviews. The interviews were recorded and transcribed and analyzed using the Atlas.ti [54] qualitative analysis tool for emerging themes using a grounded theory approach [55]. Three researchers (BS, NP, and NH) worked together as a committee to code responses and modify coding until consensus was reached across all three. It was determined that data saturation was met after 13 interviews. NEs were primarily women (n=16) with a mean age of 36 years. All efforts were taken to protect the identity of participants and thus additional details are not provided. Participants were provided with the opportunity to review the study's findings prior to publication.

Results

Survey Findings

Participants were asked to evaluate the effectiveness of the education received by responding to a series of questions on a 3-point Likert scale ("Yes, definitely;" "Yes, somewhat;" or "No"). A 3-point Likert scale was chosen for this study because it was quick to use by participants [56], with a low cognitive load. This was important as the survey was administered during the process of care in a busy outpatient clinic. Further, 3-point scales have demonstrated high reliability and validity and have been used effectively [57,58] in observational studies in health care [59,60]. Very few participants responded "No" to any of the questions, indicating that most participants found the education at least somewhat effective. However, there were significant differences between the group that received computer-assisted education and the group that received paper-based education in the number of responses of "Yes, definitely" as opposed to "Yes, somewhat." The chi-square statistic was used to understand differences between observed counts. Participants' demographics are shown in Table 1.

Table 1. Participants' demographics (N=178).

Demographics	Values, n (%)
Gender	
Male	77 (43.3)
Female	101 (56.7)
Ethnicity	
Black or African American	99 (55.6)
White	77 (43.3)
Hispanic or Latino	2 (1.1)
Self-rating of health status	
Poor	6 (3.4)
Fair	65 (36.5)
Good	86 (48.3)
Very good	17 (9.6)
Excellent	4 (2.2)
Level of education	
Less than high school	9 (5.1)
High school graduate	57 (32.0)
Some college or 2-year degree	45 (25.3)
Four-year college graduate	52 (29.2)
Graduate or professional education	15 (8.4)

Participants who were counseled using the tablet device and 3D/video-based education were more likely to report that the instructions they received were definitely easy to understand (109/115, 94.8%) compared with those who received paper materials (53/62, 85.5%; *P*=.034). Those receiving tablet-assisted counseling were much more likely to say they definitely understood their chronic health conditions better (99/116, 85.3%) compared with those who did not receive this type of counseling (36/61, 59.0%; *P*<.001). Similarly, those who were counseled with the tablet were much more likely to say that they were definitely motivated to care for themselves at home (90/116, 77.6%) than were those with paper-based materials (30/62, 48.4%; *P*<.001).

With regard to making health care decisions, participants who received counseling with the tablet were much more likely to say that they definitely felt capable of making health care decisions with their doctors (91/116, 78.4%) compared with those who received paper-based materials (27/62, 43.5%; *P*<.001), and they were also more likely to say that they definitely felt more capable of making their own health care decisions (90/116, 77.6%) compared with participants who received the paper-based materials (31/62, 50.0%; *P*=.001).

Finally, participants who received computer-based education said that they were very likely to follow their doctor's instructions (92/116, 79.3%) compared with those whose education was not computer-based (25/62, 40.3%; *P*<.001).

Interview Findings

Nurse educators (NEs) provided interview responses regarding the benefits and challenges of using the blended learning digital educational intervention compared with traditional paper-based education. Several themes emerged that were derived from the data. First, the NEs noted that the technology was easy enough to integrate into practice workflow. NEs gave comments such as, "It [the laptop and digital content] really wasn't more difficult to use than my normal materials [paper/pamphlet learning material]" [NE5]; "It was pretty easy to find what I needed when I needed it" [NE11]; and "Using it didn't slow me down" [NE1]. Second, patients seemed to pay better attention to 3D and video content. NEs said, "They [the patients] watch [the videos] pretty closely" [NE15]; "They [patients] like the pictures [interactive 3D], how they move around. They get interested, they actually ask questions" [NE7]; and "I'd say they pay better attention to the screen" [NE1]. Third, the system facilitates dynamic and efficient changes to and personalization of digital content delivery. Comments included, "I could make it [instruction] more about the patient" [NE9]; and "I like that I could give him [the patient] what they needed" [NE11]. Fourth, the system prompts on-demand question and answer sessions with patients. For example, one NE said, "It was pretty easy to just look up another video on a subject, like when a patient asked about surgery when I'm talking about [congestive heart failure], I could just have them watch the other video" [NE9]. Fifth, technology enables more easily accessible content for a broader range of educational categories to ensure that patients received education on topics that align with physician-directed patient instructions and prescriptions. NEs responded, "They

[patients] could ask a question and there's tons of pictures to click on to help answer it" [NE2]; and "I think it was a little easier to really look at what [the doctor] wanted for the patient to have, and then find it" [NE8]. Sixth, technology supports education under time- and resource-constrained circumstances. NEs stated, "By the time I'm sitting with the patient, they're tired and ready to go home. They already saw the doctor, so they're like, 'It's time to go!' So it helps to get it [education] done sooner" [NE15]; "I see patients back to back, so not really time to go looking for a handout, or printing one" [NE3]; and "Yeah, I like having it all in the laptop. It's just faster" [NE7].

NEs believed that the blended learning educational strategy also provided some advantages over online-only patient education. Themes that emerged included the following: (1) digital content prompts discussion, while in-person meeting allows for resolving concerns and misunderstandings by the patient and/or caregiver; (2) digital content ensures that the patient is receiving (watching/listening/reading) the educational content that pertains to his/her instructions and prescriptions; and (3) digital content allows the NE to experience verbal and visual cues to help assess patient understanding. Participating NEs explained, "There's a lot of back and forth" [NE1]; "Pictures [3D anatomy pictures], those I could move, zoom in, rotate around and the patient is like, 'cool' and starts asking questions" [NE5]; "When they [patients] ask questions I know they're listening" [NE10]; "Yeah, well they want to go home at that point, but they want to know, they're just wanting to understand what they have to do. They talk [to me]" [NE11]; and "We'll talk about their prescription and I can tell they're learning about it. It's like 'that's what my heart looks like?'" [NE9]. NEs noted several challenges with the blended learning strategy over traditional paper-based education. First, additional time and work can be required to ensure good, valid, and time-appropriate digital content. One NE commented, "There's so much that I don't need. I really just need to know where to go to get what I need, when I need it, and it took time to figure that out" [NE13]. Second, backup paper-based education is required because technology is not always working properly. For example, a few NEs explained, "The laptop updated, and I didn't have time to wait [to give instruction]" [NE3]; and "It [computer] works pretty good. But I always have my handouts with me cause what if it [computer] doesn't work. Sometimes it's [computer] just not working for me" [NE7].

Several minor themes that emerged included that agreement is needed between educators on the most effective navigation tool to access content—depending on practice specialty and educational content needed (ie, there isn't a one-size-fits-all efficient content structure). Comments included, "I had to learn where everything was" [NE14], referring to the placement of content; and "I wanted to show the stuff about stroke; it wasn't where it should be" [NE9]. Second, significant bandwidth requirements are needed for accessing online/server-side digital media (eg, interactive 3D). NEs made comments such as, "It ran really slow sometimes" [NE2]; and "The videos were ok, but the pictures [3D images] could get really choppy" [NE8]. Finally, additional time and worry is required to work with the IT department for mobile technology security adherence. NEs explained, "I learned not to forget my password!" [NE11]; and "It took two hours for him [IT staff] to make it [the computer] work on the network" [NE9].

Discussion

Principal Findings

Education delivered by a health care provider to a patient is essential for increasing patient understanding and adherence to medical instructions and to assist with increasing the overall health of the patient [2,61,62]. While there is no unified theory on patient education delivery, how the education is delivered makes a difference in how it is received. While electronic means to assist in patient education continue to advance, studies are needed to evaluate the utility and benefits of these technologies for providers and patients alike [63], in a variety of health care settings. This study represents a comparative evaluation of patient education in two settings: traditional paper-based education, and digital content with tablet/PC-assisted in-person instruction, both in the presence of a clinician to assist in the education process for patients with multiple chronic conditions. The in-person nature of education provided in this study was chosen due to the known complexities of providing education to patients with multiple chronic conditions. Digital education provided to patients outside of the care facility (eg, at home) is often not consumed by the patient in part or in whole [64,65]. Thus, for this study providers were able to maintain a desired level of control over patient education delivery. Several past studies have focused on delivering patient education content using mobile devices with multimedia digital formats for the patient to view at home or unsupervised in a health care environment. This study addresses the need for technical and workflow integration of these systems within and across the regular process of care provision in person with the patient.

We evaluated differences in patient understanding, motivation for self-care, intention to adhere to instructions, confidence in care decision making, and satisfaction with clinician-patient communication by using digital technologies in the presence of a clinician in the regular flow of a patient visit (ie, bedside). Results indicate that the interactive and 3D content, delivered using a touchscreen tablet or PC, was perceived to be more impactful to patients than the traditional paper- and pamphlet-based education that NEs typically provide to patients. That is, patients who received the blended learning approach responded that they understood their condition better, were more motivated to care for themselves at home, felt more capable of making health care decisions both on their own and with their doctor, and reported that they were more likely to follow instructions given to them by their doctor compared with patients who received traditional education. These findings provide important insights for those who are tasked with designing patient education programs. As patient engagement has become an important pursuit across all health care organizations to increase compliance with medical instructions for value-based care, these findings demonstrate that the blended learning model, delivered through the regular course of a patient visit for chronic disease management, may provide improvements upon traditional education delivery. Studies have concluded that population subgroups with chronic conditions, who incidentally

comprise a large portion of the population with multiple comorbidities, are less likely to follow programs that provide strictly digital health–only education. Rather, these patients may need more personalized and in-person education for chronic conditions management [65]. The NEs in this study likewise reported positive perceptions about integrating the digital education technologies into their workflow.

Behavior change is noted as the primary social determinant of value-based self-management of chronic conditions [12]. Remote patient education is growing with the use of patient portals used remotely, but the challenge is to efficiently and accurately measure patient understanding and patient adherence with the use of these portals. Moreover, health care providers have personal experiences and skills in patient education that are extremely useful in ensuring that patients understand the most important aspects of self-care with chronic conditions. Thus, provider-directed personalized education may be most favorable for achieving maximum patient understanding and motivation to adhere to instructions for patients with multiple chronic conditions.

In this study, a cloud-based, mobile, web-based patient education software system was utilized with patients face-to-face with NEs in an outpatient clinic. Our study is unique in at least two ways. First, we sought to deliver—and pilot—a cloud-based, mobile, web-based patient education application and mobile tablet computer within the context of bedside care for patients with multiple chronic conditions. While other studies have shown that media-enhanced instruction alone is at least as beneficial as standards-based nurse counseling [66], this study combined both strategies for a blended learning approach for the purpose of chronic disease management. Second, our study compared traditional paper/pamphlet-based education with multimedia and video-based instruction via a mobile device in the hands of NEs, providing personalized education guided by physician instructions and prescriptions in the normal process of a patient visit. The NEs were allowed the flexibility to tailor education as needed, based upon the knowledge level and questions of each patient, as well as patients' individual diagnoses, prescriptions, and discharge instructions.

Limitations

This study has certain limitations in its applicability. The assignment of the intervention was not at the discretion of the investigator and thus the study was not a randomized controlled trial [51,52]. NEs made every effort to alternate between using the intervention and using traditional paper-based educational materials for every other patient [51,52]. This effort to alternate the assignment of patients resulted in two groups with different group sizes. However, we believe that this should not be a limiting factor in our analysis. We also did not control for specific diagnoses or treatments; all of the subjects were considered "complex" patients, as defined by their physicians, and presented with chronic and/or multiple conditions. The sample size of the patient group was somewhat small and was also limited to one geographic area. We recommend future studies for evaluating patients with specific illnesses using a larger sample size.

Future Directions

Comparative studies are needed to evaluate alternatives to human facilitators in patient education. Future studies may seek to investigate how to integrate personalized digital patient education, such as what was delivered in this study, for remote patient settings, such as using synchronous two-way video education for patients and providers to communicate. Future studies should consider measuring patient activation as a way to understand how blended patient education strategies affect the knowledge, skills, beliefs, and behaviors that a patient needs to manage a chronic illness.

Conclusions

The outcome measures that were evaluated are patient understanding, patient motivation for self-care, patient confidence in care decision making, clinician-patient communication, and patient's intention to follow instructions. Our results indicate that utilizing digital methods for patient education in the presence of a clinician to assist in the process is quite effective in enhancing these outcomes. We noted significant differences between the two groups evaluated. A blended learning patient education approach that integrates mobile 3D/video technology–enhanced instruction with conventional in-person teaching at bedside may be one important component of a comprehensive patient education strategy for chronic care management. Assessing study outcomes by providers who deliver the training along with patients who receive the education provide a balanced assessment framework for digital patient education.

Acknowledgments

We are grateful to HP Inc for their support of this project via contributing hardware and technical advisement, and for Visual Health Solutions, Inc for the use of their innovative 3D/video-based education software.

Conflicts of Interest

None declared.

References

1. Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ 1995 May 01;152(9):1423-1433 [FREE Full text] [Medline: 7728691]

2. McCarthy DM, Engel KG, Buckley BA, Forth VE, Schmidt MJ, Adams JG, et al. Emergency Department Discharge Instructions: Lessons Learned through Developing New Patient Education Materials. Emerg Med Int 2012;2012:306859 [FREE Full text] [doi: 10.1155/2012/306859] [Medline: 22666597]
3. Mayeaux EJ, Murphy PW, Arnold C, Davis TC, Jackson RH, Sentell T. Improving patient education for patients with low literacy skills. Am Fam Physician 1996 Jan;53(1):205-211. [Medline: 8546047]
4. Mazzuca SA. Does patient education in chronic disease have therapeutic value? Journal of Chronic Diseases 1982 Jan;35(7):521-529. [doi: 10.1016/0021-9681(82)90071-6]
5. Dekkers T, Melles M, Groeneveld BS, de Ridder H. Web-Based Patient Education in Orthopedics: Systematic Review. J Med Internet Res 2018 Apr 23;20(4):e143. [doi: 10.2196/jmir.9013]
6. Demiris G, Afrin LB, Speedie S, Courtney KL, Sondhi M, Vimarlund V, et al. Patient-centered applications: use of information technology to promote disease management and wellness. A white paper by the AMIA knowledge in motion working group. J Am Med Inform Assoc 2008 Jan;15(1):8-13 [FREE Full text] [doi: 10.1197/jamia.M2492] [Medline: 17947617]
7. Sattoe JN, Bal MI, Roelofs PD, Bal R, Miedema HS, van Staa A. Self-management interventions for young people with chronic conditions: A systematic overview. Patient Educ Couns 2015 Jun;98(6):704-715. [doi: 10.1016/j.pec.2015.03.004] [Medline: 25819373]
8. Jerant AF, von Friederichs-Fitzwater MM, Moore M. Patients' perceived barriers to active self-management of chronic conditions. Patient Educ Couns 2005 Jun;57(3):300-307. [doi: 10.1016/j.pec.2004.08.004] [Medline: 15893212]
9. Jansà M, Hernández C, Vidal M, Nuñez M, Bertran MJ, Sanz S, et al. Multidimensional analysis of treatment adherence in patients with multiple chronic conditions. A cross-sectional study in a tertiary hospital. Patient Education and Counseling 2010 Nov;81(2):161-168. [doi: 10.1016/j.pec.2009.12.012]
10. Janke EA, Ramirez ML, Haltzman B, Fritz M, Kozak AT. Patient's experience with comorbidity management in primary care: a qualitative study of comorbid pain and obesity. Prim Health Care Res Dev 2015 Mar 17;17(01):33-41. [doi: 10.1017/s1463423615000171]
11. Coleman MT, Newton KS. Supporting self-management in patients with chronic illness. Am Fam Physician 2005 Oct 15;72(8):1503-1510 [FREE Full text] [Medline: 16273817]
12. Robert Fields. New Strategies for Condition Management Within Value-Based Care: Managing chronic conditions must involve behavior change, new data sources, and new clinical roles. Generations 2019:24-27 [FREE Full text]
13. Bomba F, Herrmann-Garitz C, Schmidt J, Schmidt S, Thyen U. An assessment of the experiences and needs of adolescents with chronic conditions in transitional care: a qualitative study to develop a patient education programme. Health Soc Care Community 2016 May 15;25(2):652-666. [doi: 10.1111/hsc.12356]
14. Schenker Y, Fernandez A, Sudore R, Schillinger D. Interventions to improve patient comprehension in informed consent for medical and surgical procedures: a systematic review. Med Decis Making 2011;31(1):151-173. [doi: 10.1177/0272989X10364247] [Medline: 20357225]
15. Tae JW, Lee JC, Hong SJ, Han JP, Lee YH, Chung JH, et al. Impact of patient education with cartoon visual aids on the quality of bowel preparation for colonoscopy. Gastrointest Endosc 2012 Oct;76(4):804-811. [doi: 10.1016/j.gie.2012.05.026] [Medline: 22840295]
16. Spiegel BMR, Talley J, Shekelle P, Agarwal N, Snyder B, Bolus R, et al. Development and Validation of a Novel Patient Educational Booklet to Enhance Colonoscopy Preparation. American Journal of Gastroenterology 2011;106(5):875-883. [doi: 10.1038/ajg.2011.75]
17. Shaikh AA, Hussain SM, Rahn S, Desilets DJ. Effect of an educational pamphlet on colon cancer screening: a randomized, prospective trial. European Journal of Gastroenterology & Hepatology 2010;22(4):444-449. [doi: 10.1097/meg.0b013e328333fca6]
18. Calderwood AH, Lai EJ, Fix OK, Jacobson BC. An endoscopist-blinded, randomized, controlled trial of a simple visual aid to improve bowel preparation for screening colonoscopy. Gastrointest Endosc 2011 Feb;73(2):307-314. [doi: 10.1016/j.gie.2010.10.013] [Medline: 21168840]
19. Liu X, Luo H, Zhang L, Leung FW, Liu Z, Wang X, et al. Telephone-based re-education on the day before colonoscopy improves the quality of bowel preparation and the polyp detection rate: a prospective, colonoscopist-blinded, randomised, controlled study. Gut 2014 Jan;63(1):125-130. [doi: 10.1136/gutjnl-2012-304292] [Medline: 23503044]
20. Jagosh J, Donald Boudreau J, Steinert Y, Macdonald ME, Ingram L. The importance of physician listening from the patients' perspective: enhancing diagnosis, healing, and the doctor-patient relationship. Patient Educ Couns 2011 Dec;85(3):369-374. [doi: 10.1016/j.pec.2011.01.028] [Medline: 21334160]
21. Jeong YW, Kim JA. Development of computer-tailored education program for patients with total hip replacement. Healthc Inform Res 2014 Oct;20(4):258-265 [FREE Full text] [doi: 10.4258/hir.2014.20.4.258] [Medline: 25405061]
22. Suhling H, Rademacher J, Zinowsky I, Fuge J, Greer M, Warnecke G, et al. Conventional vs. tablet computer-based patient education following lung transplantation--a randomized controlled trial. PLoS One 2014;9(6):e90828 [FREE Full text] [doi: 10.1371/journal.pone.0090828] [Medline: 24608864]
23. 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]

24. San Nicolas-Rocca T, Schooley B, Joo S. Design and Development of a Patient-Centered E-Health System to Improve Patient Understanding at Discharge. CAIS 2014;34. [doi: 10.17705/1cais.03424]
25. Su C, Li T, Cho D, Ma W, Chang Y, Lee T, et al. Effectiveness of a computerised system of patient education in clinical practice: a longitudinal nested cohort study. BMJ Open 2018 May 03;8(5):e020621 [FREE Full text] [doi: 10.1136/bmjopen-2017-020621] [Medline: 29724740]
26. Fasulo SM, Testa EJ, Lawler SM, Fitzgerald M, Lowe JT, Jawa A. A Preoperative Educational Video Improves Patient Satisfaction and Perceived Knowledge, but Not Patient Understanding for Total Shoulder Arthroplasty: A Randomized, Surgeon-Blinded Study. Journal of Shoulder and Elbow Arthroplasty 2018 Aug 08;2:247154921879296. [doi: 10.1177/2471549218792966]
27. Ab Malik N, Zhang J, Lam OLT, Jin L, McGrath C. Effectiveness of computer-aided learning in oral health among patients and caregivers: a systematic review. J Am Med Inform Assoc 2017 Jan;24(1):209-217 [FREE Full text] [doi: 10.1093/jamia/ocw045] [Medline: 27274013]
28. Linn AJ, van Dijk L, van Weert JCM, Gebeyehu BG, van Bodegraven AA, Smit EG. Creating a synergy effect: A cluster randomized controlled trial testing the effect of a tailored multimedia intervention on patient outcomes. Patient Educ Couns 2018 Aug;101(8):1419-1426. [doi: 10.1016/j.pec.2018.03.017] [Medline: 29609899]
29. Kanu C, Brown C, Barner J, Chapman C, Walker H. The Effect of a Tailored Patient Activation Intervention in Inflammatory Bowel Disease Patients. J Contemp Pharm Pract 2019 Sep 1;66(3):11-21. [doi: 10.37901/jcphp18-00009]
30. Groeneveld B, Melles M, Vehmeijer S, Mathijssen N, van DL, Goossens R. Tailored Patient Experiences. In: Bagnara S, Tartaglia R, Albolino S, Alexander T, Fujita Y, editors. Advances in Intelligent Systems and Computing. Cham, Switzerland: Springer; 2018.
31. Slama-Chaudhry A, Golay A, World Health Organization. Patient education and self-management support for chronic disease: methodology for implementing patient-tailored therapeutic programmes. Public health panorama 2019;5(2-3):357-361 [FREE Full text]
32. Kerr C, Murray E, Noble L, Morris R, Bottomley C, Stevenson F, et al. The potential of Web-based interventions for heart disease self-management: a mixed methods investigation. J Med Internet Res 2010;12(4):e56 [FREE Full text] [doi: 10.2196/jmir.1438] [Medline: 21156471]
33. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, et al. Improving diabetes management with a patient portal: a qualitative study of diabetes self-management portal. J Med Internet Res 2012;14(6):e158 [FREE Full text] [doi: 10.2196/jmir.2265] [Medline: 23195925]
34. Schooley B, San Nicolas-Rocca T, Burkhard R. Cloud-based multi-media systems for patient education and adherence: a pilot study to explore patient compliance with colonoscopy procedure preparation. Health Systems 2019 Sep 12:1-15. [doi: 10.1080/20476965.2019.1663974]
35. Bowers N, Eisenberg E, Montbriand J, Jaskolka J, Roche-Nagle G. Using a multimedia presentation to improve patient understanding and satisfaction with informed consent for minimally invasive vascular procedures. Surgeon 2017 Feb;15(1):7-11. [doi: 10.1016/j.surge.2015.09.001] [Medline: 26464072]
36. Stuij SM, Labrie NHM, van Dulmen S, Kersten MJ, Christoph N, Hulsman RL, et al. Developing a digital communication training tool on information-provision in oncology: uncovering learning needs and training preferences. BMC Med Educ 2018 Sep 24;18(1). [doi: 10.1186/s12909-018-1308-x]
37. Pecorelli N, Fiore JF, Kaneva P, Somasundram A, Charlebois P, Liberman AS, et al. An app for patient education and self-audit within an enhanced recovery program for bowel surgery: a pilot study assessing validity and usability. Surg Endosc 2018 May;32(5):2263-2273. [doi: 10.1007/s00464-017-5920-3] [Medline: 29098431]
38. Win KT, Hassan NM, Oinas-Kukkonen H, Probst Y. Online Patient Education for Chronic Disease Management: Consumer Perspectives. J Med Syst 2016 Apr;40(4):88. [doi: 10.1007/s10916-016-0438-0] [Medline: 26846749]
39. Marthick M, Janssen A, Cheema BS, Alison J, Shaw T, Dhillon H. Feasibility of an Interactive Patient Portal for Monitoring Physical Activity, Remote Symptom Reporting, and Patient Education in Oncology: Qualitative Study. JMIR Cancer 2019 Nov 28;5(2):e15539 [FREE Full text] [doi: 10.2196/15539] [Medline: 31778123]
40. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015 Feb 20;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
41. Raghupathi W, Raghupathi V. An Empirical Study of Chronic Diseases in the United States: A Visual Analytics Approach. Int J Environ Res Public Health 2018 Mar 01;15(3) [FREE Full text] [doi: 10.3390/ijerph15030431] [Medline: 29494555]
42. Hoffman C. Persons With Chronic Conditions. JAMA 1996 Nov 13;276(18):1473. [doi: 10.1001/jama.1996.03540180029029]
43. Public Health Agency of Canada. How healthy are Canadians? A trend analysis of the health of Canadians from a healthy living and chronic disease perspective. Ottawa: Public Health Agency of Canada; Dec 2016.
44. Bhowmik M, Banerjee B, Banerjee J. Role of Pedagogy in Effective Teaching. Basic Res J Educ Res Rev 2013 Jan;2(1):1-5 [FREE Full text]
45. Redman BK. Advances in Patient Education. New York: Springer; 2004:978-0826127358.
46. Fredericks S, Martorella G, Catallo C. A systematic review of web-based educational interventions. Clin Nurs Res 2015 Feb 26;24(1):91-113. [doi: 10.1177/1054773814522829] [Medline: 24571963]

47. Rowe M, Frantz J, Bozalek V. The role of blended learning in the clinical education of healthcare students: a systematic review. Med Teach 2012;34(4):e216-e221. [doi: 10.3109/0142159X.2012.642831] [Medline: 22455712]
48. Liu Z, Zhang MM, Li YY, Li LX, Li YQ. Enhanced education for bowel preparation before colonoscopy: A state-of-the-art review. J Dig Dis 2017 Feb;18(2):84-91. [doi: 10.1111/1751-2980.12446] [Medline: 28067463]
49. Sangrar R, Docherty-Skippen SM, Beattie K. Blended face-to-face and online/computer-based education approaches in chronic disease self-management: A critical interpretive synthesis. Patient Educ Couns 2019 Oct;102(10):1822-1832. [doi: 10.1016/j.pec.2019.05.009] [Medline: 31101427]
50. Morse J, Niehaus L. Mixed method design: principles and procedures. New York: Routledge; 2016.
51. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. N Engl J Med 2000 Jun 22;342(25):1878-1886. [doi: 10.1056/nejm200006223422506]
52. Lao KSJ, Chui CSL, Man KKC, Lau WCY, Chan EW, Wong ICK. Medication safety research by observational study design. Int J Clin Pharm 2016 Jun;38(3):676-684 [FREE Full text] [doi: 10.1007/s11096-016-0285-6] [Medline: 27003827]
53. 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]
54. Friese, Susanne. Qualitative data analysis with ATLAS.ti (3rd ed). Thousand Oaks, CA: Sage Publications; Apr 22, 2019.
55. Foley G, Timonen V. Using Grounded Theory Method to Capture and Analyze Health Care Experiences. Health Serv Res 2015 Aug;50(4):1195-1210 [FREE Full text] [doi: 10.1111/1475-6773.12275] [Medline: 25523315]
56. Preston CC, Colman AM. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychol (Amst) 2000 Mar;104(1):1-15. [doi: 10.1016/s0001-6918(99)00050-5] [Medline: 10769936]
57. Bendig AW. Reliability of short rating scales and the heterogeneity of the rated stimuli. Journal of Applied Psychology 1954;38(3):167-170. [doi: 10.1037/h0059072]
58. Bendig AW. Reliability and the number of rating-scale categories. Journal of Applied Psychology 1954;38(1):38-40. [doi: 10.1037/h0055647]
59. Joshi A, Kale S, Chandel S, Pal D. Likert Scale: Explored and Explained. BJAST 2015 Jan 10;7(4):396-403. [doi: 10.9734/bjast/2015/14975]
60. Jacoby J, Matell MS. Three-Point Likert Scales Are Good Enough. Journal of Marketing Research 1971 Nov;8(4):495. [doi: 10.2307/3150242]
61. Zolnierek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834 [FREE Full text] [doi: 10.1097/MLR.0b013e31819a5acc] [Medline: 19584762]
62. Ong LM, de HJC, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. Soc Sci Med 1995 Apr;40(7):903-918. [Medline: 7792630]
63. Peter D, Robinson P, Jordan M, Lawrence S, Casey K, Salas-Lopez D. Reducing readmissions using teach-back: enhancing patient and family education. J Nurs Adm 2015 Jan;45(1):35-42. [doi: 10.1097/NNA.0000000000000155] [Medline: 25479173]
64. Kim H, Xie B. Health literacy in the eHealth era: A systematic review of the literature. Patient Educ Couns 2017 Jun;100(6):1073-1082. [doi: 10.1016/j.pec.2017.01.015] [Medline: 28174067]
65. Gordon NP, Hornbrook MC. Older adults' readiness to engage with eHealth patient education and self-care resources: a cross-sectional survey. BMC Health Serv Res 2018 Mar 27;18(1):220 [FREE Full text] [doi: 10.1186/s12913-018-2986-0] [Medline: 29587721]
66. Park J, Kim MS, Kim H, Kim SI, Shin CH, Lee HJ, et al. A randomized controlled trial of an educational video to improve quality of bowel preparation for colonoscopy. BMC Gastroenterol 2016 Jun 17;16(1):64 [FREE Full text] [doi: 10.1186/s12876-016-0476-6] [Medline: 27317249]

Abbreviations

IT: information technology
NE: nurse educator
3D: 3-dimensional

Edited by L Buis; submitted 27.07.20; peer-reviewed by C Kruse, YH Kwan; comments to author 18.08.20; revised version received 13.10.20; accepted 30.11.20; published 22.12.20.

Please cite as:
Schooley B, Singh A, Hikmet N, Brookshire R, Patel N
Integrated Digital Patient Education at the Bedside for Patients with Chronic Conditions: Observational Study
JMIR Mhealth Uhealth 2020;8(12):e22947
URL: http://mhealth.jmir.org/2020/12/e22947/
doi:10.2196/22947
PMID:33350961

Original Paper

Accuracy of Monocular Two-Dimensional Pose Estimation Compared With a Reference Standard for Kinematic Multiview Analysis: Validation Study

Oskar Stamm[1], MSc; Anika Heimann-Steinert[1], Dr rer medic

Geriatrics Research Group, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

Corresponding Author:
Oskar Stamm, MSc
Geriatrics Research Group
Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health
Reinickendorfer Straße 61
Berlin, D-13347
Germany
Phone: 49 30 450 553 784
Email: oskar.stamm@charite.de

Abstract

Background: Expensive optoelectronic systems, considered the gold standard, require a laboratory environment and the attachment of markers, and they are therefore rarely used in everyday clinical practice. Two-dimensional (2D) human pose estimations for clinical purposes allow kinematic analyses to be carried out via a camera-based smartphone app. Since clinical specialists highly depend on the validity of information, there is a need to evaluate the accuracy of 2D pose estimation apps.

Objective: The aim of the study was to investigate the accuracy of the 2D pose estimation of a mobility analysis app (Lindera-v2), using the PanopticStudio Toolbox data set as a reference standard. The study aimed to assess the differences in joint angles obtained by 2D video information generated with the Lindera-v2 algorithm and the reference standard. The results can provide an important assessment of the adequacy of the app for clinical use.

Methods: To evaluate the accuracy of the Lindera-v2 algorithm, 10 video sequences were analyzed. Accuracy was evaluated by assessing a total of 30,000 data pairs for each joint (10 joints in total), comparing the angle data obtained from the Lindera-v2 algorithm with those of the reference standard. The mean differences of the angles were calculated for each joint, and a comparison was made between the estimated values and the reference standard values. Furthermore, the mean absolute error (MAE), root mean square error, and symmetric mean absolute percentage error of the 2D angles were calculated. Agreement between the 2 measurement methods was calculated using the intraclass correlation coefficient (ICC[A,2]). A cross-correlation was calculated for the time series to verify whether there was a temporal shift in the data.

Results: The mean difference of the Lindera-v2 data in the right hip was the closest to the reference standard, with a mean value difference of –0.05° (SD 6.06°). The greatest difference in comparison with the baseline was found in the neck, with a measurement of –3.07° (SD 6.43°). The MAE of the angle measurement closest to the baseline was observed in the pelvis (1.40°, SD 1.48°). In contrast, the largest MAE was observed in the right shoulder (6.48°, SD 8.43°). The medians of all acquired joints ranged in difference from 0.19° to 3.17° compared with the reference standard. The ICC values ranged from 0.951 (95% CI 0.914-0.969) in the neck to 0.997 (95% CI 0.997-0.997) in the left elbow joint. The cross-correlation showed that the Lindera-v2 algorithm had no temporal lag.

Conclusions: The results of the study indicate that a 2D pose estimation by means of a smartphone app can have excellent agreement compared with a validated reference standard. An assessment of kinematic variables can be performed with the analyzed algorithm, showing only minimal deviations compared with data from a massive multiview system.

(JMIR Mhealth Uhealth 2020;8(12):e19608) doi:10.2196/19608

KEYWORDS

2D human pose estimation; motion capturing; kinematics; clinical practice; mobility; smartphone app; analysis

Introduction

Traditional movement assessments, although carried out by experienced physicians, physiotherapists, and occupational therapists, can contain inaccuracies due to subjectivity, despite the clinicians' expertise. In contrast, quantitative motion measurements by motion capture systems are a valuable tool in scientific and clinical motion analysis and offer a highly accurate and reliable way of capturing kinematic data. Quantitative analyses can be used, for example, to monitor the progress of therapies and objectively evaluate the effectiveness of specific interventions. Motion capture systems are used in sports, biomechanics, and rehabilitation, and they focus on gait analysis, injury prevention, and performance improvement [1]. However, these systems are rarely used in everyday clinical practice.

There are many optoelectronic motion capture systems based on markers (eg, Vicon [Vicon Motion Systems], Motion Analysis [Motion Analysis Corp], Optitrack [NaturalPoint Inc], and Qualisys [Qualisys AB]). These systems are often regarded in the literature as the gold standard for motion capture [2]. Today, low- to high-speed experiments show that positioning errors can even be assumed in the case of measurements less than 2 mm [3]. Nevertheless, optoelectronic systems require a restricted area, such as a laboratory environment, and the attachment of markers [4], which can also be a potential source of measurement error in these systems due to skin movement artifacts [5]. Furthermore, these systems are proving to be very costly. The aforementioned factors are detrimental to the systems' practicability and everyday use in a clinical context.

Inertial sensor measurement systems can be used as a low-cost alternative. However, an inertial sensor measurement system cannot determine global position when used as a stand-alone system (by itself), although as a fusion system, in combination with a rigid body model such as the Perception Neuron (Noitom Ltd), a position in space can still be identified [2]. Nevertheless, a major disadvantage remains in clinical practice, as users must attach numerous sensors to a patient's body. Since clinical processes are usually efficiency driven and the application of several sensors is too time-consuming in most cases due to time constraints during treatment, this also contributes to the low frequency of use. Recently, studies have examined markerless and body sensor–less image-processing systems. Depth-sensing camera systems, such as the Kinect (Microsoft Corp), Intel RealSense (Intel Corp), or Zed (Stereolabs Inc) sensors, have proven to be a cost-effective solution with acceptable accuracy for some use cases [6-11]. Another low-cost motion capture method in image processing is pose estimation, which involves transformation of two-dimensional (2D) images into three-dimensional (3D) objects, for example, by using deep convolutional neural networks of monocular images [12-15], such as images on a mobile device [16]. Research into depth estimation from a monocular image is still in its infancy in the field of computer vision and is proving to be difficult in some cases, as slight inaccuracies in estimation can lead to very different depth estimates [17,18]. Major limitations include false pose estimates due to the target person moving outside the image boundaries [19] or the pose being disturbed by objects such as shadows [15].

A 2D skeleton detector enables calculation of specific joint angles for assessment and feedback in sport and rehabilitation settings. The use of 2D human pose estimations for clinical purposes, such as the Lindera-v2 app or the motion-tracking coach on the Kaia health app [20], allows mobility analyses to be carried out using, for example, a mobile device. Pose and movement analyses via a smartphone save time for medical staff, who can use the time gained as treatment time or for patient consultations. Furthermore, such a measurement of mobility represents a more objective method of measurement compared with traditional assessments, which are based on a subjective assessment. Since in clinical practice, specialist staff highly depend on the validity of information, there is a need to validate the methodology of the Lindera-v2 measuring method. In order to achieve a performance level comparable with the gold standard motion capture systems, this study aimed to evaluate the accuracy of the Lindera-v2 2D pose estimation algorithm, using the PanopticStudio Toolbox (Carnegie Mellon University) [21,22] as a reference standard.

Methods

Data Collection

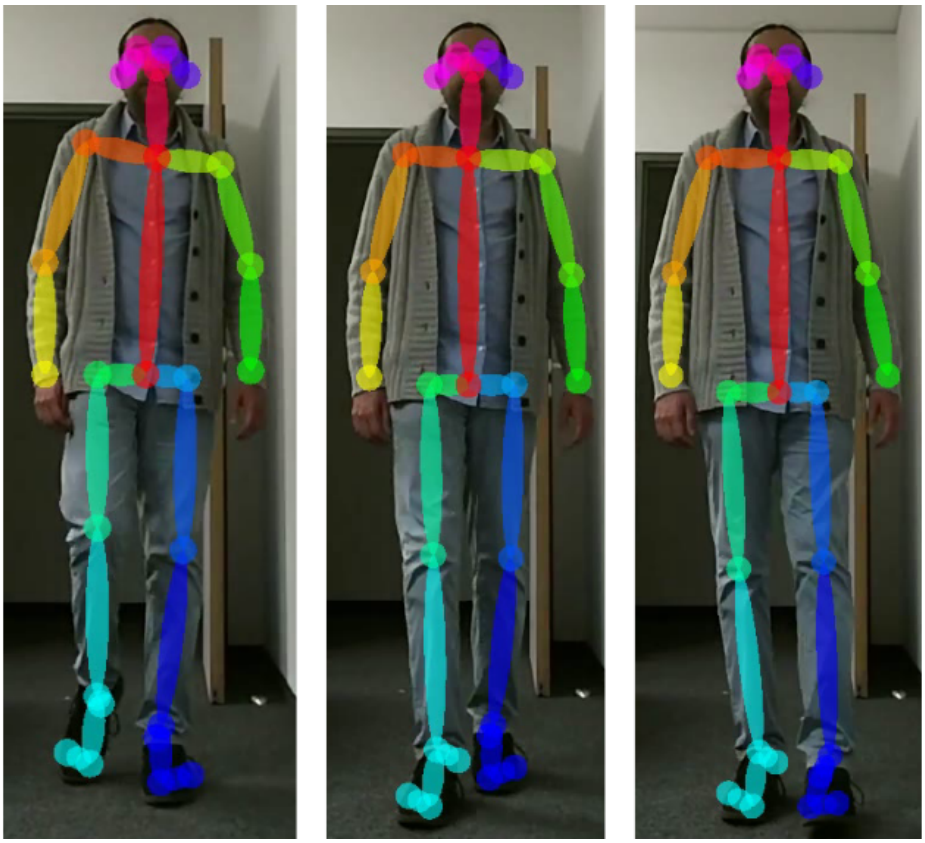
For the accuracy evaluation, 10 video sequences were generated from Panoptic Studio 3D PointCloud (data set 171204_pose1-6) [21,22]. Data shared for research purposes from Carnegie Mellon University were used as the reference standard. The data set included video sequences from 480 video graphics array (VGA) cameras, 31 high-definition (HD) cameras, and 10 Kinect cameras, as well as arrays of 2D poses of key points of body part locations, showing the range of motion of the joints. The videos were split so that only one actor appeared in each video. Subsequently, 10 videos were selected that matched the requirement for all key joint points to be visible during the movements in each frame. Within the video clips, no changes or cuts were made. The total duration of the videos selected for analysis was 24 minutes 9 seconds, with an average duration of 2 minutes 25 seconds per video. The movements were categorized by the PanopticStudio Toolbox as a range of motion.

Lindera Pose Estimation

The Lindera-v2 algorithm is a combination of a 2D and 3D skeleton-based pose estimation. For this study, we needed the output of the 2D skeleton estimator to calculate 2D joint angles [23,24]. The 2D skeleton detector module of the Lindera-v2 algorithm is based on the tf-pose-estimation repository [25]. This repository is a TensorFlow implementation of various deep learning models [13] that represent human pose estimation models based on convolutional pose machines [26]. The original repository also provides some model variants that run on mobile devices. The short version of the repository used with the Lindera-v2 algorithm has also implemented the additional Openpose Body25 model, which provides 25 body joint coordinates for each input image (x in the direction of the image width and y in the direction of the image height) instead of 18 body joint coordinates, as was the case in the original repository model variants (Figure 1). The 2D module of the Lindera-v2

algorithm produced a coordinate-time list for each joint and each frame. These time series data were then used to geometrically calculate the corresponding 2D joint angles for each frame of the processed input video. To smooth the angle-time series data, we used digital filtering with a size 11 Bartlett window.

Figure 1. Two-dimensional pose estimation by skeleton fitting, based on 25 body joint coordinates.



Reference Standard

The Panoptic Studio data set from Carnegie Mellon University [21,22] is a data set shared for research purposes. Unlike the Lindera app, the PanopticStudio Toolbox is not a monocular 2D estimator but rather a massive multiview system consisting of the following: (1) 480 VGA cameras with a resolution of 640 × 480 pixels and 25 frames per second (fps); (2) 31 HD cameras with a resolution of 1920 × 1080 pixels and 30 fps; (3) 10 Kinect II sensors of 1920 × 1080 pixels (RGB), 512 × 424 pixels (depth), and 30 fps; and (4) 5 digital light processing projectors.

The 2D skeleton of the 2D Panoptic Studio pose detector has 15 anatomical landmarks. The 2D detector uses appearance information in the interpretation and includes connectivity information.

Statistical Analysis

The value tables for the respective joint angles were clustered, and missing values were imputed using a simple moving average. The mean difference (bias) between the Lindera-v2 algorithm estimates and the reference standard values was calculated for each joint. Furthermore, the mean absolute error, the root mean squared error, and the symmetric mean absolute percent error of the 2D angles were used. The intraclass correlation coefficient (ICC[A,2]) was calculated for the data using the 2-way mixed-effects model as a measure of agreement between the 2 measurement methods. An ICC in the range of 0 indicates random evaluation behavior, and a value of 1 is regarded as an ideal reliable feature evaluation by the evaluators. We used the definition in which values greater than 0.7 are generally regarded as indicators of good agreement [27]. Values up to approximately 0.3 are regarded as a low correlation, and those of approximately 0.5 or more are regarded as a medium correlation. A further classification according to Fleiss [28] was used to assesses the ICC classification, with 0.00 to 0.40 indicating poor agreement, 0.40 to 0.75 indicating fair to good agreement, and 0.75 to 1.00 indicating excellent agreement.

A cross-correlation was calculated for the time series to verify whether there was a temporal shift in the data. To verify the stationarity of the data, which is a prerequisite for cross-correlation testing, we used the augmented Dickey-Fuller test. The data were first evaluated in IBM SPSS Statistics (version 25.0; IBM Corp) and then in the programming language R in RStudio (version 3.5.1; RStudio Inc).

Results

In order to evaluate the accuracy of the movement signals recorded, we analyzed a total of 30,000 data pairs for each joint, comparing the joint angles obtained using the Lindera-v2 algorithm with those of the PanopticStudio Toolbox data set (the reference standard). Table 1 displays the 10 joints analyzed; the key points used for calculating the joint angles; the average difference (bias) between the estimated values of the Lindera-v2

algorithm and the reference standard; and the mean absolute error (MAE), mean absolute deviation, root mean square error (RMSE), symmetric mean absolute percentage error (sMAPE), and ICC. The angles used were determined trigonometrically for both measuring methods.

Table 1. Mean angle difference and ICC of Lindera-v2 and the Panoptic Studio data set for the joints analyzed.

Joint	2D[a] key points used	Difference in 2D angles (°), mean (SD); 95% CI	MAE[b] of 2D angles (°)	MAD[c] (°)	RMSE[d] of 2D angles (°)	sMAPE[e] (%)	ICC[f] (95% CI)	SE of mean difference
Right shoulder	Right hip, shoulder, and elbow	2.71 (10.28); 2.59 to 2.83	6.48	4.10	10.63	23.33	0.978 (0.973 to 0.981)	0.06
Left shoulder	Left hip, shoulder, and elbow	–0.07 (12.11); –0.21 to 0.07	3.98	3.20	12.12	10.71	0.951 (0.950 to 0.952)	0.07
Right elbow	Right shoulder, elbow, and wrist	–1.01 (12.12); –1.15 to –0.87	6.18	4.30	12.16	6.64	0.983 (0.983 to 0.984)	0.07
Left elbow	Left shoulder, elbow, and wrist	0.24 (6.20); 0.17 to 0.31	3.15	2.84	6.21	9.17	0.997 (0.997 to 0.997)	0.04
Right hip	Right shoulder, hip, and knee	–0.05 (6.06); –0.12 to 0.02	4.45	4.68	6.06	3.01	0.983 (0.983 to 0.983)	0.04
Left hip	Left shoulder, hip, and knee	–0.61 (3.85); –0.66 to –0.57	2.29	2.29	3.90	1.74	0.992 (0.992 to 0.993)	0.02
Right knee	Right hip, knee, and ankle	–1.37 (2.97); –1.40 to –1.34	2.58	2.93	3.27	1.56	0.985 (0.974 to 0.990)	0.02
Left knee	Left hip, knee, and ankle	0.84 (4.31); 0.79 to 0.89	2.28	2.45	4.44	1.39	0.971 (0.968 to 0.974)	0.03
Neck	Pelvis, neck, and head	–3.07 (6.43); –3.14 to –2.99	4.47	3.63	7.13	3.20	0.951 (0.914 to 0.969)	0.04
Pelvis	Left knee, pelvis, and right knee	0.15 (2.03); 0.14 to 0.18	1.40	1.64	2.04	5.42	0.996 (0.996 to 0.996)	0.01

[a]2D: two-dimensional.

[b]MAE: mean absolute error.

[c]MAD: mean absolute deviation.

[d]RMSE: root mean square error.

[e]sMAPE: symmetric mean absolute percentage error.

[f]ICC: intraclass correlation coefficient ICC(A,2).

The data collected indicated both a negative and a positive bias. The mean difference of the joint angles that was nearest to the baseline was identified in the right hip (–0.05°, SD 6.06°). The joint with the highest mean difference (ie, with the greatest difference from 0) was the neck (–3.07°, SD 6.43°). The mean joint angle accuracy was used to show the average magnitude of the errors. The mean absolute error of the angle measurement closest to the baseline was observed in the pelvis (1.40°, SD 1.48°). In contrast, the highest mean absolute error was observed in the right shoulder (6.48°, SD 8.43°). The standard deviation was also lowest in the pelvis (SD 3.36°), and the highest standard deviation was found to be in the left shoulder (SD 11.45°). The root mean square error was also applied, although this tends to give weight to large errors. The RMSE indicated low accuracy in the right elbow (12.16°) and high accuracy in the pelvis (2.04°). Since the mean absolute percentage error cannot be used when values are 0 (as this would result in division by 0), we used the sMAPE, which was lowest in the left knee (1.39%) and highest in the right shoulder (23.33%).

The intraclass correlation coefficient for the joint angles is also shown in Table 1 and represents agreement between the 2 measurement methods (Lindera-v2 vs the PanopticStudio Toolbox). In accordance with the McGraw and Wong convention [29], the intraclass correlation coefficient ICC(A,2) was used (ie, a 2-way mixed type with average measures and absolute agreement). The highest ICC value was found in the left elbow joint (average measure of 0.997, 95% CI 0.997-0.997). In contrast, the lowest ICC values were in the neck (average measure of 0.951, 95% CI 0.914-0.969).

Interpretation of the measurement values based on mean values can lead to biased findings (eg, in the case of extreme outliers). Since the median is less affected by outliers, we used box plots for the differences in joint angle values measured with the

Lindera-v2 and the reference standard. Figure 2 shows a box-and-whisker plot without outliers to facilitate closer examination of the boxes. The ends of the whiskers represent 1.5 × IQR. The median with the greatest difference in comparison with the 0 value was detected in the right shoulder (3.17°), and the joint angle median nearest to the baseline was the pelvis joint (0.19°). Figure 3 shows the box-and-whisker plot with outliers. The third quartile of the right shoulder was the farthest from the baseline of all the joints, with a value of 5.87°. The lowest first quartile was in the neck, measuring –5.34°. The smallest IQR, ranging from –0.81° to 1.27°, was in the pelvis. The most extreme outliers in this plot were found in the right elbow, where the minimum was –106.00° and the maximum was 125.71°. However, the outlier with the greatest difference in comparison to 0 was in the left shoulder, with a difference of 157.10°.

Figure 2. Box plot showing differences in the Lindera-v2 and reference standard values, measured across all joints tested.

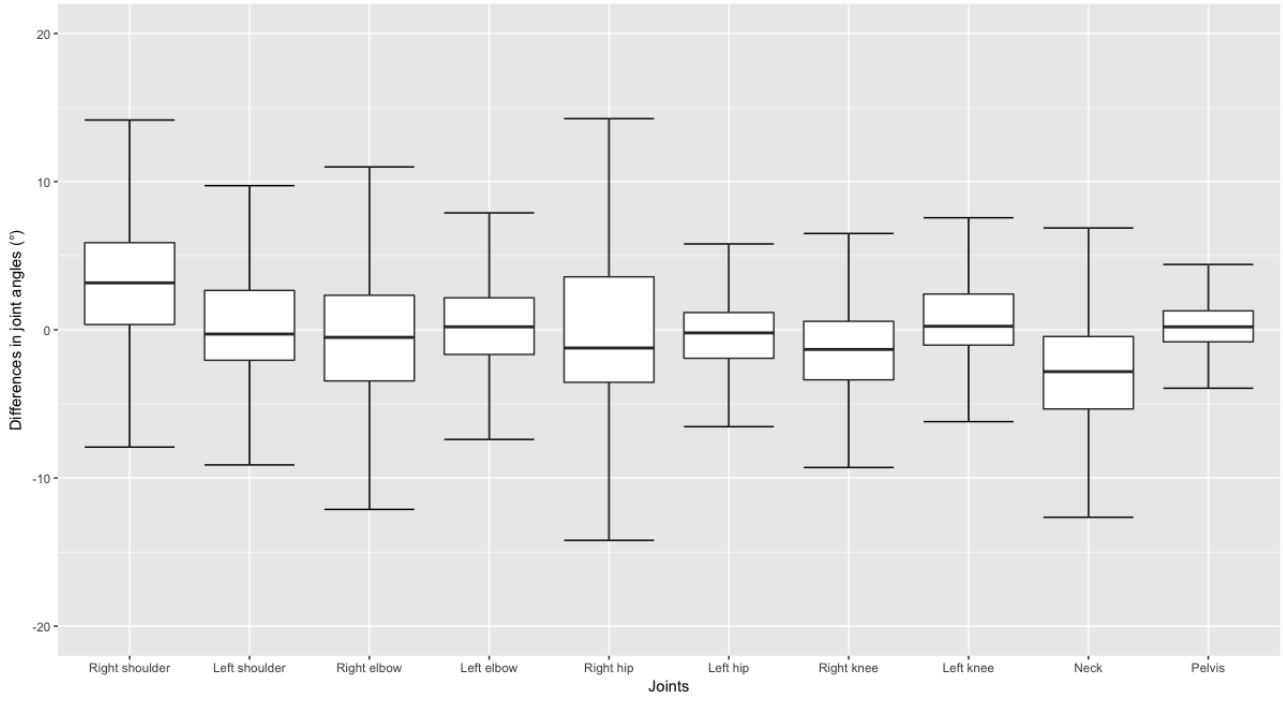
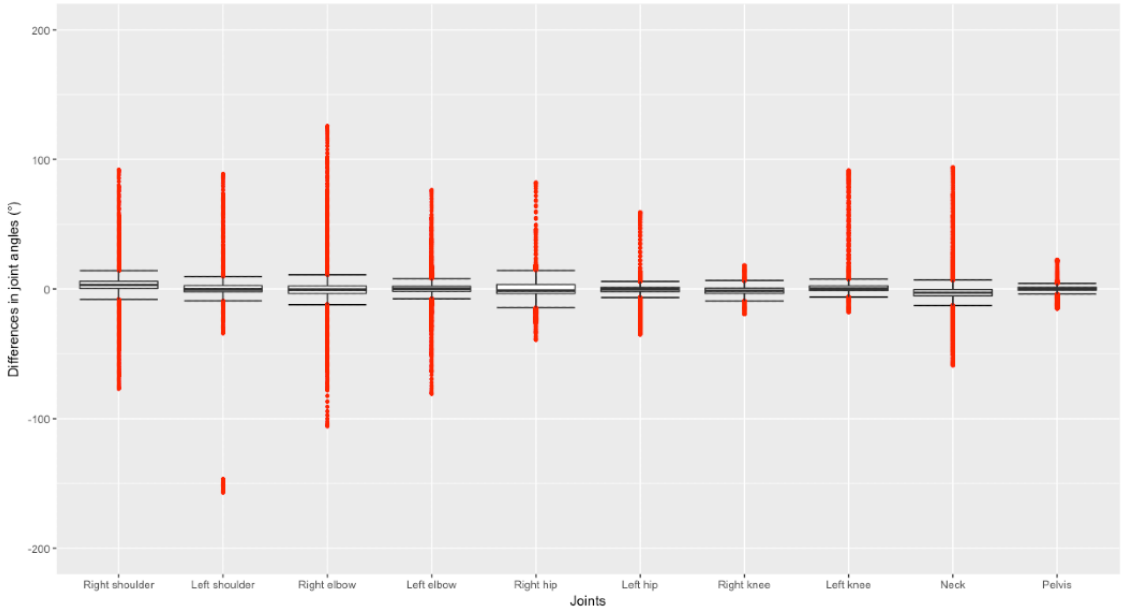


Figure 3. Box plot with outliers showing differences in the Lindera-v2 and reference standard values, measured across all joints tested.



To examine the ICC more closely and analyze the potential influence of single videos on the ICC values of the joints, our next step entailed calculating an ICC for each video. Figure 4 shows a dot plot of the ICC for the 10 videos used for the accuracy measurement in each joint. For the neck joint, 5 videos had an ICC below 0.75. However, there were no videos in which all joints had remarkably low ICC values. With the exception of the neck, all joints in all videos had an ICC value above 0.75.

The cross-correlation function was applied to the selected time series in order to examine the temporal lag. The results in Table

2 show that no time delays could be detected in the values measured. These would have been visible at an increased correlation at a time outside lag 0. However, all graphs (Figure 5) showed the highest correlation at lag 0. The dotted blue lines represent the confidence interval of the estimated correlation values. If a value was outside the range of the interval, the correlation was considered significant.

Figure 4. Dot plot of the intraclass correlation coefficient comparing Lindera-v2 and reference standard for the 10 single videos used for accuracy measurement in each joint.

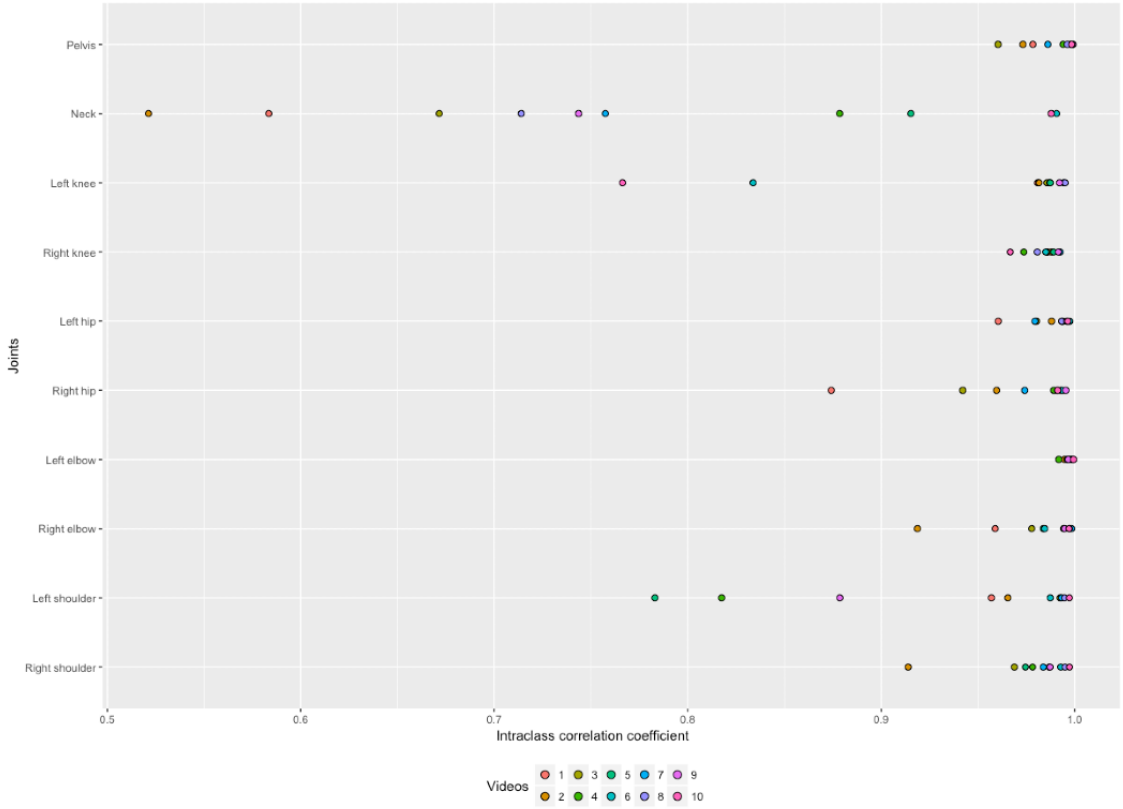


Figure 5. Cross-correlation graph of Lindera-v2 and Panoptic Studio data set values. One lag represents 1 sample (frame). ACF: autocorrelation function.

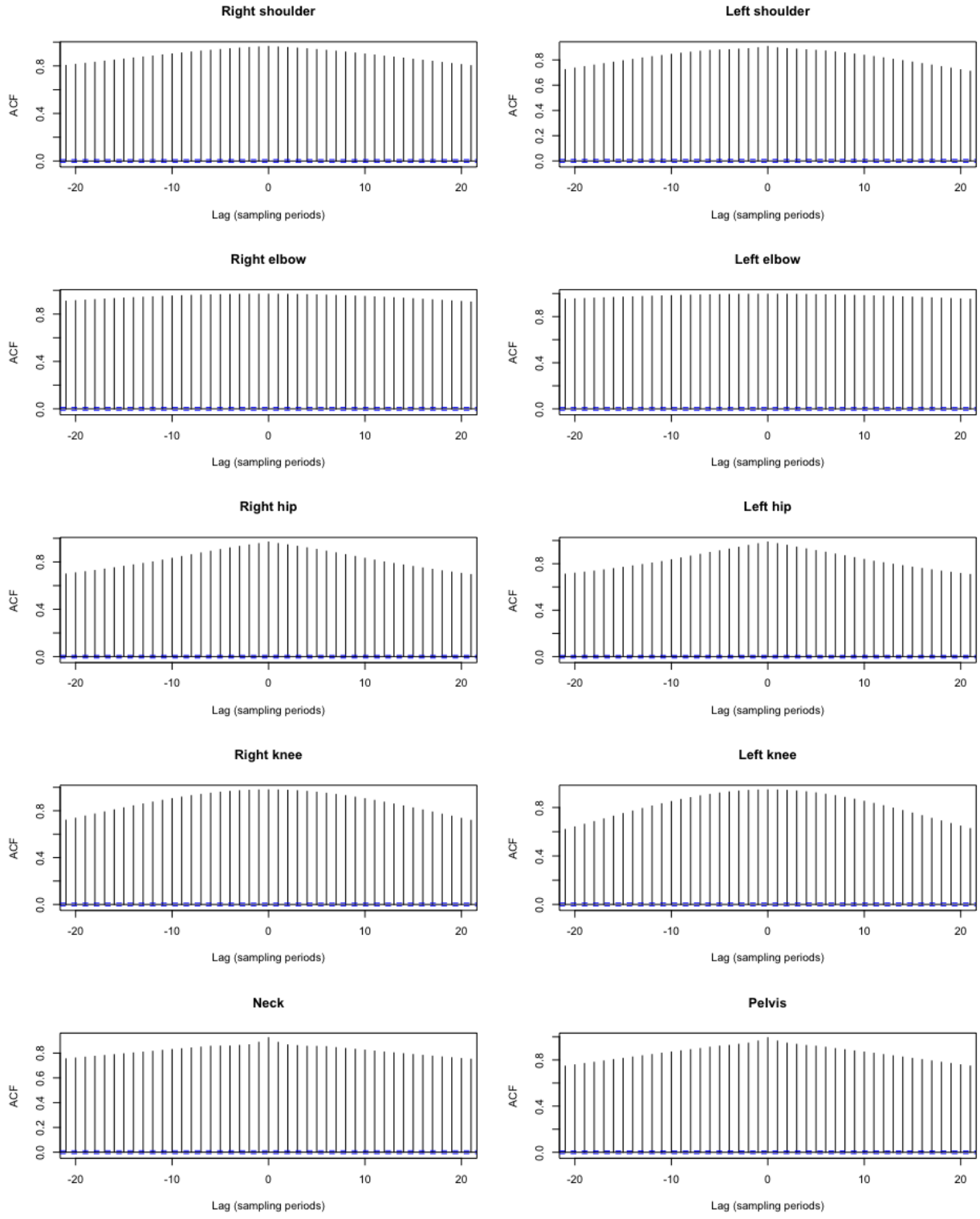


Table 2. Maximum cross-correlations of Lindera-v2 and reference standard values.

Joint	Lag value with maximum correlation	Maximum correlation coefficient
Right shoulder	0	0.96
Left shoulder	0	0.91
Right elbow	0	0.97
Left elbow	0	0.99
Right hip	0	0.97
Left hip	0	0.99
Right knee	0	0.98
Left knee	0	0.95
Neck	0	0.92
Pelvis	0	0.99

Discussion

Principal Results

The goal of this study was to validate the accuracy of the 2D pose estimation of joint angles obtained from the Lindera-v2 algorithm, using the PanopticStudio Toolbox, which served as the reference standard. Therefore, we analyzed 30,000 data pairs for each joint angle during diverse total body motion activity. First, the mean difference and error measures were compared for each joint. Second, the ICC was compared for each joint. In order to verify agreement between the 2 measurement methods (the Lindera-v2 and the PanopticStudio Toolbox data set), we analyzed the ICC values for each of the 10 videos. Finally, we examined the potential temporal lag through cross-correlation. The results of the study indicate that the 2D pose estimation method used had excellent agreement with the reference standard. Furthermore, the Lindera-v2 algorithm had no temporal lag.

The mean angle generated for the right hip by the Lindera-v2 algorithm was the closest to the reference standard. Even the value with the greatest difference from 0 (found in the neck) was acceptable. However, these values should be treated with caution because mean values can lead to biased results. Therefore, we displayed the median values in box plots. The medians of all joints compared with the reference standard ranged from a difference of 0.19° (pelvis joint) to 3.17° (right shoulder). In all joints, the IQR was within 6° and –6°, which means that 50% of the values were within this range. These acquired values provide a promising starting point upon which to base mobility assessments and 3D pose estimation. A further reason why box plots were used was to identify outliers because the RMSE used gives greater weight to large errors. Figure 3 shows that in the box-and-whisker plots, the data recording the difference between the Lindera-v2 algorithm and the reference standard had few very high outliers. In the right elbow, for example, the outliers were particularly high compared with the other joints. Since the RMSE squares the errors before averaging, the RMSE in this joint was the highest. Since large outliers can be quickly identified as such by an experienced user, the weighting of large outliers by the RMSE does not seem appropriate. Hence, the MAE might be the more appropriate measure. The sMAPE is more resistant to outliers due to defined error limits because it gives less weight to outliers than other measures that have no error limits; it was applied additionally for this reason [30]. An advantage over the mean absolute percentage error is that the sMAPE cannot be extremely large or infinite [30]. The sMAPE in our evaluation was particularly high in the right shoulder (23.33%). Possible joint “losses” could be an explanation due to bad visibility of the joints in the videos, perhaps due to the clothing of the participant or the lighting used.

The ICC agreement between the 2 measurement methods can be interpreted as excellent (according to the classification presented by Fleiss [28]). All joints had an ICC value of at least 0.951. Furthermore, the 95% confidence interval of the ICC for all joints can be classified as excellent. Our analysis of the individual videos showed why the neck had the lowest overall agreement in comparison with the other joints. In several videos, the agreement could be interpreted as fair to good. We assume that this relates to the approximations of neck positions, since these were calculated from the key point of the nose in the Lindera algorithm. By applying cross-correlation, angles estimated through the Lindera-v2 algorithm showed no temporal lag.

Early research and reviews published in 2016 reported that the Kinect skeleton-tracking algorithm indicated poor validity and large errors with respect to most kinematic variables [7]. Clark et al [10] recommended that the Kinect system be carefully chosen for specific use cases (eg, trunk angles can be highly accurate). An extensive recent review by Poitras et al [31] on the validity and reliability of wearable sensors for joint angle estimation revealed mixed results. The results presented in this study are therefore very promising, not only because of the acceptable accuracy of the angles but also because the usability of smartphone apps (compared with the Kinect system or wearable sensors) offers major advantages. Schurr et al [32] showed a moderate to strong relationship between a 2D video camera and 3D motion capture analyses. From this point of view, 2D pose estimations are applicable in clinical practice. Even though 2D cameras offer clinicians a valuable kinematic measurement tool, the use of a smartphone would be far less

complicated and would make the technology available to a wide user group.

Valid and reliable 2D joint angles are an important first step on the way to valid and reliable 3D joint angles. Therefore, in the next step, the 2D data from the evaluation will be transformed into 3D pose estimation angles using deep convolutional neural networks. A validation of the 3D joint angle accuracy of the resulting data will show whether the requirements for clinical practice are met.

Limitations

Although this study showed excellent agreement to a reference standard, a validity study using a state-of-the-art marker-based motion capture system as a ground truth is necessary for a thorough validation. The comparison to the reference standard is an important step toward accuracy assurance but does not replace a proof of validity.

To determine a systematic error in the algorithm by an offset, a static setup would be needed. From this, a Euclidean distance could be calculated to identify a precise source of error. The mean joint position error is the most frequently used method for verification of the accuracy of a pose estimation. However, since determination of the coordinates in millimeters in space was not possible in these data sets, accuracy verification was carried out for the joint angles. Verification of the precision showing the repeatability of the data was not planned in this project, since measurement using the Lindera-v2 was carried out once and the movements were not repeated in a standardized manner. However, the precision of the time stamps within the measurement of the evaluated movements can be seen from the standard errors of the mean difference. A validation of the precision will be the subject of further studies.

Perspectives

In geriatrics, orthopedics, and neurology in particular, accurate and validated mobility analyses such as the Lindera-v2 could help medical professionals confirm diagnoses and track the success of treatments. Mobility assessments have very high relevance for a multitude of clinical uses (eg, older adults and patients with more severe diseases who have a higher risk of falling) [33,34]; in this case, fall risk assessments could be of high value. Assessment of kinematic variables, such as specific joint angles, can be accessed via 2D skeleton data if viewed from specific angles, and such data can also be used for rehabilitative purposes in physical therapy or sports science. Furthermore, the 2D values analyzed in this study constitute an encouraging basis for 3D pose estimation, which will be the next step in accuracy validation.

Conclusions

The results of the study indicate that 2D pose estimation by means of a camera-based smartphone app can have excellent agreement with a validated reference standard. Furthermore, the Lindera-v2 algorithm was found to have no temporal lag. An assessment of kinematic variables, such as specific joint angles, can be performed with the algorithm, and these data showed only minimal deviations compared with data from a massive multiview system. In future studies, it will be important to test the app in a clinical context with participants with physical limitations.

Acknowledgments

We would like to thank Lindera GmbH for providing the raw data for the study. We acknowledge support from the German Research Foundation and the Open Access Publication Fund of Charité – Universitätsmedizin Berlin.

Conflicts of Interest

None declared.

References

1. Colyer SL, Evans M, Cosker DP, Salo AIT. A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. Sports Med Open 2018 Jun 05;4(1):24 [FREE Full text] [doi: 10.1186/s40798-018-0139-y] [Medline: 29869300]
2. van der Kruk E, Reijne MM. Accuracy of human motion capture systems for sport applications; state-of-the-art review. Eur J Sport Sci 2018 Jul 09;18(6):806-819. [doi: 10.1080/17461391.2018.1463397] [Medline: 29741985]
3. Merriaux P, Dupuis Y, Boutteau R, Vasseur P, Savatier X. A Study of Vicon System Positioning Performance. Sensors (Basel) 2017 Jul 07;17(7) [FREE Full text] [doi: 10.3390/s17071591] [Medline: 28686213]
4. Corazza S, Mündermann L, Gambaretto E, Ferrigno G, Andriacchi TP. Markerless Motion Capture through Visual Hull, Articulated ICP and Subject Specific Model Generation. Int J Comput Vis 2009 Sep 2;87(1-2):156-169. [doi: 10.1007/s11263-009-0284-3]
5. Leardini A, Chiari L, Della Croce U, Cappozzo A. Human movement analysis using stereophotogrammetry. Part 3. Soft tissue artifact assessment and compensation. Gait Posture 2005 Feb;21(2):212-225. [doi: 10.1016/j.gaitpost.2004.05.002] [Medline: 15639400]
6. Otte K, Kayser B, Mansow-Model S, Verrel J, Paul F, Brandt AU, et al. Accuracy and Reliability of the Kinect Version 2 for Clinical Measurement of Motor Function. PLoS One 2016 Nov 18;11(11):e0166532 [FREE Full text] [doi: 10.1371/journal.pone.0166532] [Medline: 27861541]

7. Springer S, Yogev Seligmann G. Validity of the Kinect for Gait Assessment: A Focused Review. Sensors (Basel) 2016 Feb 04;16(2):194 [FREE Full text] [doi: 10.3390/s16020194] [Medline: 26861323]
8. Bonnechère B, Jansen B, Salvia P, Bouzahouene H, Omelina L, Moiseev F, et al. Validity and reliability of the Kinect within functional assessment activities: comparison with standard stereophotogrammetry. Gait Posture 2014;39(1):593-598. [doi: 10.1016/j.gaitpost.2013.09.018] [Medline: 24269523]
9. Mentiplay BF, Perraton LG, Bower KJ, Pua Y, McGaw R, Heywood S, et al. Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. J Biomech 2015 Jul 16;48(10):2166-2170. [doi: 10.1016/j.jbiomech.2015.05.021] [Medline: 26065332]
10. Clark RA, Mentiplay BF, Hough E, Pua YH. Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives. Gait Posture 2019 Feb;68:193-200. [doi: 10.1016/j.gaitpost.2018.11.029] [Medline: 30500731]
11. Fernandez-Baena A, Susin A, Lligadas X. Biomechanical Validation of Upper-Body and Lower-Body Joint Movements of Kinect Motion Capture Data for Rehabilitation Treatments. 2012 Presented at: 2012 Fourth International Conference on Intelligent Networking and Collaborative Systems; Sep 19-21, 2012; Bucharest, Romania. [doi: 10.1109/incos.2012.66]
12. Brau E, Jiang H. 3D Human Pose Estimation via Deep Learning from 2D Annotations. 2016 Presented at: Fourth International Conference on 3D Vision (3DV); Oct 25-28, 2016; Stanford, CA. [doi: 10.1109/3dv.2016.84]
13. Cao Z, Hidalgo G, Simon T, Wei S, Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. arXiv. 2019 May 30. URL: http://arxiv.org/abs/1812.08008 [accessed 2019-11-10]
14. Bogo F, Kanazawa A, Lassner C, Gehler P, Romero J, Black M. Keep It SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image. In: Leibe B, Matas J, Sebe N, Welling M, editors. Computer Vision ECCV 2016. Cham, Switzerland: Springer International Publishing; 2016:561-578.
15. Liu S, Li Y, Hua G. Human Pose Estimation in Video via Structured Space Learning and Halfway Temporal Evaluation. IEEE Trans Circuits Syst Video Technol 2019 Jul;29(7):2029-2038. [doi: 10.1109/tcsvt.2018.2858828]
16. Moder M, Schellenbach M. Deep Learning Approach for Estimating a Human Pose on a Mobile Device. In: Kameas A, Stathis K, editors. Ambient Intelligence. Cham, Switzerland: Springer International Publishing; 2018:100-105.
17. Mehta D, Sridhar S, Sotnychenko O, Rhodin H, Shafiei M, Seidel H, et al. VNect. ACM Trans Graph 2017 Jul 20;36(4):1-14. [doi: 10.1145/3072959.3073596]
18. Qiang B, Zhang S, Zhan Y, Xie W, Zhao T. Improved Convolutional Pose Machines for Human Pose Estimation Using Image Sensor Data. Sensors (Basel) 2019 Feb 10;19(3):718 [FREE Full text] [doi: 10.3390/s19030718] [Medline: 30744191]
19. Hidalgo G, Raaj Y, Idrees H, Xiang D, Joo H, Simon T, et al. Single-Network Whole-Body Pose Estimation. 2019 Presented at: 2019 International Conference on Computer Vision (ICCV); Oct 27-Nov 2, 2019; Seoul, Korea. [doi: 10.1109/iccv.2019.00708]
20. Huber S, Priebe JA, Baumann K, Plidschun A, Schiessl C, Tölle TR. Treatment of Low Back Pain with a Digital Multidisciplinary Pain Treatment App: Short-Term Results. JMIR Rehabil Assist Technol 2017 Dec 04;4(2):e11 [FREE Full text] [doi: 10.2196/rehab.9032] [Medline: 29203460]
21. Joo H, Simon T, Li X, Liu H, Tan L, Gui L, et al. Panoptic Studio: A Massively Multiview System for Social Interaction Capture. IEEE Trans Pattern Anal Mach Intell 2019 Jan 1;41(1):190-204. [doi: 10.1109/tpami.2017.2782743]
22. Simon T, Joo H, Matthews I, Sheikh Y. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. arXiv. 2017 Apr 25. URL: http://arxiv.org/abs/1704.07809 [accessed 2019-11-11]
23. Azhand D, Rabe D, Müller D, Sattler I, Steinert D. Algorithm Based on One Monocular Video Delivers Highly Valid and Reliable Gait Parameters. arXiv. URL: http://arxiv.org/abs/2008.08045 [accessed 2020-09-08]
24. Rabe S, Azhand A, Pommer W, Müller S, Steinert A. Descriptive Evaluation and Accuracy of a Mobile App to Assess Fall Risk in Seniors: Retrospective Case-Control Study. JMIR Aging 2020 Feb 14;3(1):e16131 [FREE Full text] [doi: 10.2196/16131] [Medline: 32130111]
25. Kim I. Deep Pose Estimation implemented using Tensorflow with Custom Architectures for fast inference. GitHub. URL: https://github.com/ildoonet/tf-pose-estimation [accessed 2020-12-12]
26. Wei S, Ramakrishna V, Kanade T, Sheikh Y. Convolutional Pose Machines. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); June 27-30, 2016; Las Vegas, NV. [doi: 10.1109/cvpr.2016.511]
27. Wirtz M, Caspar F. Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen, Germany: Hogrefe Verlag; 2002:978-973.
28. Fleiss J. The Design and Analysis of Clinical Experiments. New York, NY: John Wiley & Sons; 1986.
29. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1(1):30-46. [doi: 10.1037/1082-989X.1.1.30]
30. Chen C, Twycross J, Garibaldi JM. A new accuracy measure based on bounded relative error for time series forecasting. PLoS One 2017 Mar 24;12(3):e0174202 [FREE Full text] [doi: 10.1371/journal.pone.0174202] [Medline: 28339480]
31. Poitras I, Dupuis F, Bielmann M, Campeau-Lecours A, Mercier C, Bouyer L, et al. Validity and Reliability of Wearable Sensors for Joint Angle Estimation: A Systematic Review. Sensors (Basel) 2019 Mar 31;19(7):1555 [FREE Full text] [doi: 10.3390/s19071555] [Medline: 30935116]

32. Schurr S, Marshall AN, Resch JE, Saliba SA. Two-Dimensional Video Analysis Is Comparable To 3D Motion Capture In Lower Extremity Movement Assessment. Int J Sports Phys Ther 2017 Apr;12(2):163-172 [FREE Full text] [Medline: 28515970]
33. Jia H, Lubetkin EI, DeMichele K, Stark DS, Zack MM, Thompson WW. Prevalence, risk factors, and burden of disease for falls and balance or walking problems among older adults in the U.S. Prev Med 2019 Sep;126:105737. [doi: 10.1016/j.ypmed.2019.05.025] [Medline: 31150739]
34. Jacobi L, Petzold T, Hanel A, Albrecht M, Eberlein-Gonska M, Schmitt J. [Epidemiology and prediction of the risk of falling in patients in acute care settings: Analysis of routine data from a university hospital]. Z Evid Fortbild Qual Gesundhwes 2017 Feb;120:9-15. [doi: 10.1016/j.zefq.2016.12.006] [Medline: 28284369]

Abbreviations

2D: two-dimensional
3D: three-dimensional
fps: frames per second
HD: high definition
ICC: intraclass correlation coefficient
MAE: mean absolute error
RMSE: root mean square error
sMAPE: symmetric mean absolute percentage error
VGA: video graphics array

Edited by L Buis; submitted 24.04.20; peer-reviewed by M Schmid; comments to author 20.07.20; revised version received 14.09.20; accepted 17.11.20; published 21.12.20.

Please cite as:
Stamm O, Heimann-Steinert A
Accuracy of Monocular Two-Dimensional Pose Estimation Compared With a Reference Standard for Kinematic Multiview Analysis: Validation Study
JMIR Mhealth Uhealth 2020;8(12):e19608
URL: http://mhealth.jmir.org/2020/12/e19608/
doi:10.2196/19608
PMID:33346739

Original Paper

Wearability Testing of Ambulatory Vital Sign Monitoring Devices: Prospective Observational Cohort Study

Carlos Areia[1,2*], PT, MSc; Louise Young[1,2*], RN; Sarah Vollam[1,2], RN, MSc; Jody Ede[1,2], RN, MSc; Mauro Santos[2,3], BSc, DPhil; Lionel Tarassenko[2,3], MA, DPhil; Peter Watkinson[1,2], MD

[1]Critical Care Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

[2]National Institute for Health Research, Biomedical Research Centre, Oxford, United Kingdom

[3]Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Carlos Areia, PT, MSc
Critical Care Research Group
Nuffield Department of Clinical Neurosciences
University of Oxford
Kadoorie Research Centre, John Radcliffe Hospital
Oxford
United Kingdom
Phone: 44 1865 231440
Email: carlos.morgadoareia@ndcn.ox.ac.uk

Abstract

Background: Timely recognition of patient deterioration remains challenging. Ambulatory monitoring systems (AMSs) may provide support to current monitoring practices; however, they need to be thoroughly tested before implementation in the clinical environment for early detection of deterioration.

Objective: The objective of this study was to assess the wearability of a selection of commercially available AMSs to inform a future prospective study of ambulatory vital sign monitors in an acute hospital ward.

Methods: Five pulse oximeters (4 with finger probes and 1 wrist-worn only, collecting pulse rates and oxygen saturation) and 2 chest patches (collecting heart rates and respiratory rates) were selected to be part of this study: The 2 chest-worn patches were VitalPatch (VitalConnect) and Peerbridge Cor (Peerbridge); the 4 wrist-worn devices with finger probe were Nonin WristOx2 3150 (Nonin), Checkme O2+ (Viatom Technology), PC-68B, and AP-20 (both from Creative Medical); and the 1 solely wrist-worn device was Wavelet (Wavelet Health). Adult participants wore each device for up to 72 hours while performing usual "activities of daily living" and were asked to score the perceived exertion and perception of pain or discomfort by using the Borg CR-10 scale; thoughts and feelings caused by the AMS using the Comfort Rating Scale (CRS); and to provide general free text feedback. Median and IQRs were reported and nonparametric tests were used to assess differences between the devices' CRS scores.

Results: Quantitative scores and feedback were collected in 70 completed questionnaires from 20 healthy volunteers, with each device tested approximately 10 times. The Wavelet seemed to be the most wearable device (P<.001) with an overall median (IQR) CRS score of 1.00 (0.88). There were no statistically significant differences in wearability between the chest patches in the CRS total score; however, the VitalPatch was superior in the Attachment section (P=.04) with a median (IQR) score of 3.00 (1.00). General pain and discomfort scores and total percentage of time worn are also reflective of this.

Conclusions: Our results suggest that adult participants prefer to wear wrist-worn pulse oximeters without a probe compressing the fingertip and they prefer to wear a smaller chest patch. A compromise between wearability, reliability, and accuracy should be made for successful and practical integration of AMSs within the hospital environment.

(JMIR Mhealth Uhealth 2020;8(12):e20214) doi:10.2196/20214

KEYWORDS

wearables; pulse oximeter; chest patch; wearability; vital signs; ambulatory monitoring

Introduction

Background

Failure to recognize and act on deteriorating signs of acute illness has been documented previously [1,2]. The National Institute for Health and Care Excellence [3] recommends the use of an early warning score, which is designed to quantitatively assess the severity of abnormal vital signs triggering the appropriately graded clinical response. A limitation of early warning score systems is the requirement for clinical staff to measure vital signs at the correct frequency. There are several factors that can affect monitoring frequency such as clinical shift duration [4], ward staff levels [5] and workload associated with the vital sign measurements [6]; hence, the ideal frequency is often not achieved [7-9].

Research has shown that the current ward-based vital-sign monitoring of patients is time consuming, as there can be several processes involved in addition to manual measurement, for example, explaining the process and obtaining consent from patients, documenting vital signs in patient records, calculating the early warning score, among others [6,10]. Additionally, even if the ideal frequency of measurement is achieved, patients might deteriorate between observation sets [11].

To address this, patients could be continuously monitored with the aim of increasing early detection of deterioration [12]. In the United Kingdom, continuous monitoring is undertaken in clinical practice but is not commonly performed in wards [13]. It has also been suggested that the most frequent reason for nonuse of continuous monitoring systems is restriction of patient movement and that, to maximize clinical integration, continuous monitoring should be comfortable and less restrictive [13,14].

Ambulatory monitoring systems (AMSs) may provide an alternative to either intermittent measurement of manual vital signs or wired continuous monitoring, affording the patients more mobility and comfort while supporting clinical staff by providing regular vital signs data [15]. There is an increased focus on the development of wireless vital sign monitors for use in the health care setting; however, their reliability and efficiency are still uncertain and need to be tested [16]. Additionally, it has been suggested that introducing AMSs may have physical or psychological effects that should be assessed [17-19] in order to maximize patient compliance and data retrieval. Previous studies have shown that AMSs are removed prematurely owing to patient irritation, discomfort, feeling unwell, or equipment failure [20].

This is the first study of our virtual high dependency unit project, with the overall aim of testing the feasibility of deploying ambulatory vital sign monitoring in the hospital environment. To achieve this, and considering the nonadoption, abandonment, scale-up, spread, and sustainability framework [21], this study will support the initial AMS selection to move into further testing within the virtual high dependency unit project. This will include ward locational testing as well as tests of device accuracy during patient movement and in the detection of hypoxia [22]. Once the final devices have been selected, we will integrate these within our user interface and test its clinical deployment.

Objective

The aim of this study was to assess the wearability of a selection of commercially available AMSs to inform a future prospective study of wearable vital sign monitors in an acute hospital ward.

Methods

Device Selection

The research team conducted a preliminary review of commercially available ambulatory vital sign monitors. To be considered for this study, the devices were required to have wireless connectivity, to measure at least two of the target parameters (ie, heart rate, oxygen saturation, respiratory rate) and to provide third-party permission to access raw data. Based on these requirements, we selected the following monitors: 2 chest-worn patches, that is, VitalPatch (VitalConnect) and Peerbridge Cor (Peerbridge); 4 wrist-worn devices with finger probe, that is, Nonin WristOx2 3150 (Nonin), Checkme O2+ (Viatom Technology), PC-68B, and AP-20 (both from Creative Medical); and 1 solely wrist-worn device, that is, Wavelet (Wavelet Health). Nonin WristOx2 3150 is named Nonin hereafter (Figure 1).

Figure 1. Devices included in this study.

Study Design and Participants

This study used a prospective observational cohort design. It was reviewed and approved by the Oxford University Research and Ethics Committee and Clinical Trials and Research Governance teams (R55430/RE003). This study is compliant with the cohort checklist of the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) statement Multimedia Appendix 1.

Adult participants were in-house research staff based in the Kadoorie Research Centre (Level 3, John Radcliffe Hospital) and the Oxford Institute of Biomedical Engineering (University of Oxford, Old Road Campus Research Building), who were recruited through posters placed in target locations such as office common spaces. We also used the internal departmental newsletter and distributed the participant information sheet within the departments. Once a volunteer expressed interest in the study via email/telephone, a study session was organized. All participants were healthy adult volunteers, with the only exclusion criterion being known allergies to adhesive stickers or intracardiac devices (permanent pacemaker).

Following informed consent, participants had at least one of the AMSs fitted and were guided through the device practicalities (eg, how to remove and reattach, waterproofing advice). Participants were required to wear up to 4 different AMSs for up to 72 hours each to mimic in-hospital use. They were advised that they could remove the device if desired; they were then requested to log the time and reason for removal (eg, not being able to wear the finger probe while cooking). Participants were also asked to score various activities that patients are likely to perform during their hospital stay by using a validated questionnaire. No incentives (monetary or otherwise) were given to the participants.

Measurements

All participants completed 1 “Ambulatory Monitoring Wearability Assessment Questionnaire” for each device tested. Data collected included participant demographics and device details (eg, sex, age, device used), the perceived exertion while performing “activities of daily living” (ADLs) using the Borg CR-10 scale [23], the perception of pain or discomfort in specific body areas using body maps with the Borg CR-10 scale [23], and thoughts and feelings about emotions, anxiety, harm, etc, caused by the AMS by using the Comfort Rating Scale (CRS), as described in Multimedia Appendix 2 (CRS information [19] and an open comment section for participants to share general feedback).

The CRS [19] uses a 21-point scale throughout 14 statements, split into 6 categories; 3 statements for emotion, 4 for attachment, 1 for harm, 2 for perceived change, 1 for movement, and 3 for anxiety. All but one of the 14 statements are negatively worded such that, to strongly disagree with a statement (lower score), is a positive outcome [19]. In the case of the 1 positive statement, the answers were further preprocessed (ie, inverted) to make them homogeneous with the other answers, as previously described in another study [17]. For each participant, the median score was first determined for each questionnaire section. For better interpretation, we have also calculated the percentage of responses within each question/category and colored it according to positive or negative outcome (further information in Multimedia Appendix 2). The median score of all the sections was then computed to determine the participant’s overall median CRS score.

To minimize the risk of missing data, the clinical researchers double-checked all the received questionnaires with the participants when collecting them. To minimize wearability bias between devices, we mixed them and documented the order/combination they were used by participants, introduced a washout period of at least one week before testing another device and checked for any clear bias in the free-text feedback section of the questionnaire.

Data Analysis

Sample Size

Owing to the exploratory pragmatic nature of this study, no sample size calculation was performed. We recruited a convenience sample of 20 healthy volunteers to offer a wide range of experiences with wearability of the test devices.

Data Preprocessing

For comparisons, we grouped the chest patches (Peerbridge Cor and VitalPatch) and the pulse oximeters (AP-20, Checkme O2+, Nonin, PC-68B, and Wavelet). This grouping allowed us to conduct separate comparisons, as the selected main measurements from the chest patches are heart rate and respiratory rate, while the pulse oximeters include pulse rate and peripheral capillary oxygen saturations (SpO_2). It is expected that these 2 types of monitors will be part of the same AMS.

Statistical Analysis

Due to the limited sample size and data skewness (normality was assessed using the Shapiro-Wilk test), median and IQRs were reported. Nonparametric tests were used to assess differences between the devices' CRS scores. The Wilcoxon test was used to compare the median CRS scores of the chest patches. Finally, the Kruskal-Wallis test, followed by post hoc Dunn tests [24] (with Bonferroni correction [25]), was used to compare the median CRS scores of the pulse oximeters. All statistical tests were conducted using R v3.6.1 [26] and the tidyverse package [27].

Free Text Analysis

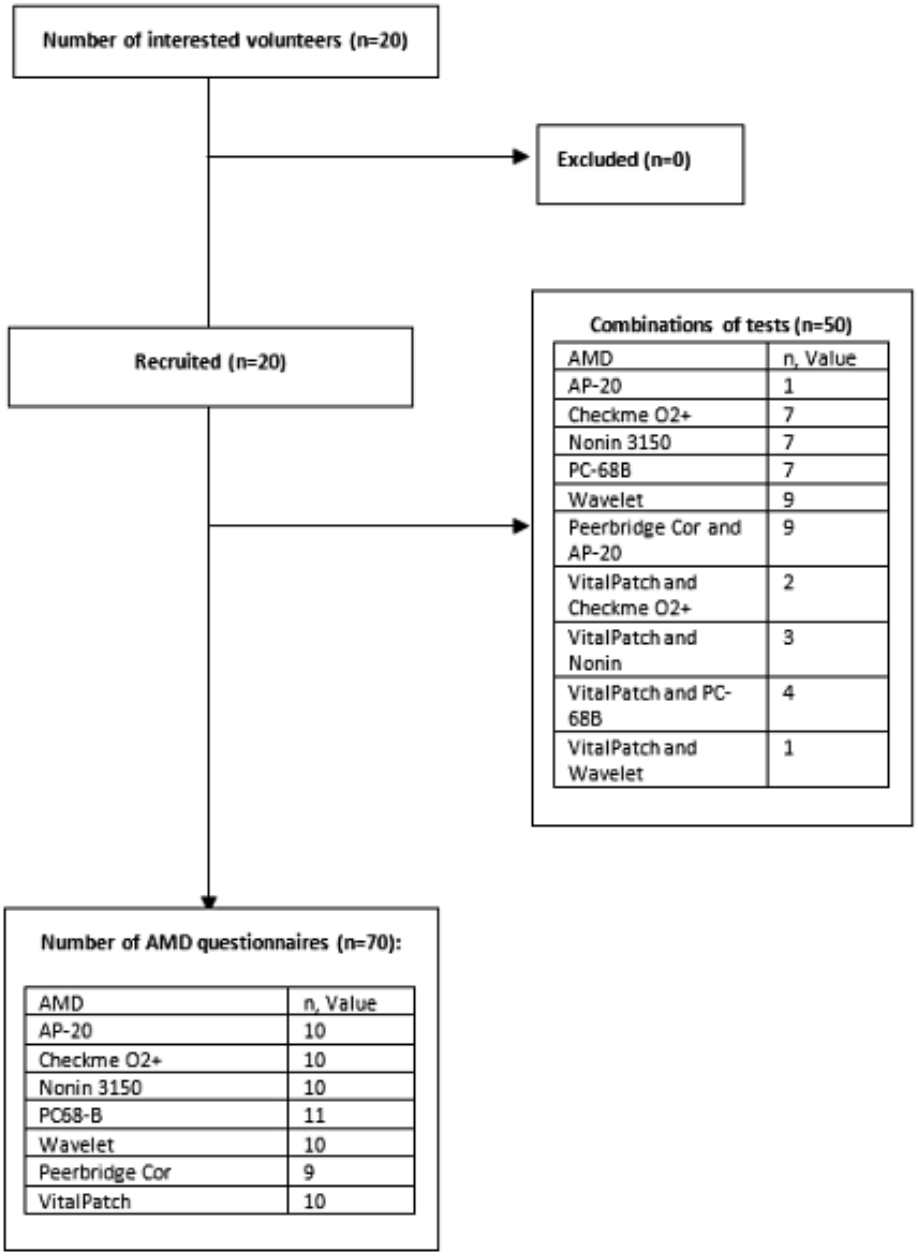
Participants were also encouraged to write free text to describe the problems and challenges they encountered with each device. NVivo 12 (NVivo qualitative data analysis software, QSR International Pty Ltd Version 12, 2018) was used to analyze and collate feedback into common categories. For each category, we included the number of participants with at least one negative comment (eg, for disrupted ADLs, the number of participants who took off the device for ADLs or mentioned that these were disruptive when performing daily tasks is reported). This number was reviewed and agreed by 2 researchers from the study team.

Results

Participant Characteristics

Twenty in-house volunteers (13 women and 7 men) were recruited between May 4, 2018 and October 30, 2018, with a median (IQR) age of 34 (32-40) years. For each session, participants wore either a pulse oximeter, a chest patch, or both for up to 72 hours before completing the wearability questionnaire. All participants wore at least one device, with a median (IQR) washout period of 31 (13-58) days before wearing another one, and they all completed 1 questionnaire per device (Figure 2).

Figure 2. The study participant flowchart. AMD: ambulatory monitoring device.



Wearability Questionnaire Outcomes

The total wear duration and total number of removals per device are presented in Table 1. Wavelet and Checkme O2+ were the most used pulse oximeters (644/720, 89.4%) and the VitalPatch was the most used chest patch (663/720, 92.1%). These devices also had the lowest pain/discomfort and median exertion scores in the included activities as described in Table 2.

Table 1. Device testing duration, removal duration, and reasons for removal.

Measurements	Pulse oximeters					Chest patches	
	AP-20 (n=10)	Checkme O2+ (n=10)	Nonin (n=10)	PC68-B (n=11)	Wavelet (n=10)	Peerbridge Cor (n=9)	VitalPatch (n=10)
Testing duration (hours:minutes)							
Total planned duration[a]	720:00	720:00	720:00	792:00	720:00	648:00	720:00
Used duration[b] (% of total planned duration)	522:31 (72.6)	643:35 (89.4)	590:11 (82.0)	558:43 (70.6)	640:45 (90.0)	491:19 (75.8)	662:52 (92.1)
Median session duration (IQR)	70:45 (51:50-72:00)	70:40 (58:17-72:00)	64:42 (53:46-69:55)	67:50 (36:58-70:44)	72:00 (55:11-72:00)	70:45 (45:10-71:55)	69:58 (68:12-72:00)
Removals							
Number of participants (total removals)	8 (36)	10 (45)	9 (49)	8 (31)	3 (4)	4 (8)	0
Median removal duration[c] (IQR)	1:06 (0:30-5:37)	1:21 (0:50-3:30)	2:01 (1:04-5:05)	01:20 (0:35-3:15)	1:50 (1:33-3:00)	0:10 (0:10-0:45)	0:00 (0:00-0:00)
Reasons for removal (n)							
Hygiene	10	6	6	1	0	6	0
Discomfort	3	3	7	6	0	0	0
Cooking or eating	6	6	11	10	0	0	0
Exercise	0	3	3	4	0	0	0
Work/social	2	5	4	3	0	0	0
Battery/hardware failure	10	17	2	2	3	0	0
Other/unknown	5	5	16	5	1	2	0

[a]Total planned duration refers to the total amount of time if all participants wore the respective device for the full 72 hours (Total=72 × n). Values are shown as hours:minutes.

[b]Used duration: reflects the actual time that the devices were worn by the participants, with missing times representing a combination of device removal periods and differences between the actual end of the session and 72 hours (ie, when the full 72 hours are not achieved). Values are shown as hours:minutes.

[c]Values are shown as hours:minutes.

Table 2. Pain/discomfort and exertion scores per device, body part, and activity by using Borg CR-10 scale [23].

Device	Pulse oximeters					Chest patches	
	AP-20	Checkme O2+	Nonin	PC68-B	Wavelet	Peerbridge Cor	VitalPatch
Median pain/discomfort score per body part[a] (IQR)	3.00 (5.00)	1.50 (2.38)	3.50 (2.50)	3.00 (2.50)	0.75 (0.50)	2.00 (6.00)	1.00 (0.88)
Median exertion score per activity							
Walking	0	0	0	0	0	0	0
Eating	3	2	3	4	0	0	0
Drinking	1	0	1	3	0	0	0
Dressing	5	4	5	8	1	0	0
Writing	1	0	3	5	0	0	0
Using phone/tablet	0	0	3	3	0	0	0
Reading	0	0	0	0	0	0	0
Hand washing	9	6	7	10	0	0	0
Sleeping	4	2	3	4	0	3	0

[a]For the pulse oximeters, the body part of interest was the nondominant wrist and for the chest patches, it was the chest.

Common problems identified in free text sections of the questionnaires are presented in Table 3. We grouped free text comments into 5 categories: (1) device size (eg, device being too big or bulky), (2) disrupted ADLs (eg, limited daily tasks, so needed to be removed), (3) skin irritation (eg, some concerns of wrist strap/finger-probe becoming itchy), (4) finger probe uncomfortable (eg, sweaty and annoying), and (5) affected sleep (eg, participants kept waking up and unable to sleep with it on).

Significant differences for most sections were found between the CRS scores of the pulse oximeters as well as between the overall median CRS scores (Table 4 and Table 5). Figure 3 and Figure 4 represent the percentage of positive/negative CRS score outcomes per section and per device.

Table 3. Device specifications and participant-identified problems.

Measurements and problems	Pulse oximeters					Chest patches	
	AP-20	Checkme O2+	Nonin	PC-68B	Wavelet	Peerbridge Cor	VitalPatch
Device measurements	SpO_2[a], PR[b], RR[c], perfusion index	SpO_2, PR, steps	SpO_2, PR	SpO_2, PR, perfusion index	SpO_2, PR[b]	HR[d], RR, ECG[e]	HR, RR, ECG, body position
Participant-identified problems (n)[f]							
Device size	2	0	3	10	0	16	0
Disrupted ADLs[g]	8	6	5	6	0	0	0
Skin irritation	0	2	1	0	2	3	3
Finger probe uncomfortable	6	0	5	8	N/A[h]	N/A	N/A
Affected sleep	1	0	3	5	0	5	1

[a]SpO_2: oxygen saturation.

[b]PR: pulse rate.

[c]RR: respiratory rate. For the AP-20, respiratory rate is only possible using a nasal cannula.

[d]HR: heart rate.

[e]ECG: electrocardiogram. Peerbridge Cor uses a 2-lead ECG and VitalPatch a single-lead ECG.

[f]The cells show the number of participants identifying at least one problem in each category.

[g]ADLs: activities of daily living.

[h]N/A: not applicable.

Table 4. Comparison of the Comfort Rating Scale scores across different pulse oximeters.

CRS[a] section	Pulse oximeters, median (IQR) score					
	AP-20 (n=10)	Checkme O2+ (n=10)	Nonin (n=10)	PC68-B (n=11)	Wavelet (n=10)	*P* value
Overall score	7.50 (4.00)[b]	3.25 (8.63)[b]	6.00 (6.38)[b]	9.00 (4.75)[b]	1.00 (0.88)[c,d,e,f]	<.001
Emotion	4.00 (6.00)	4.00 (7.25)	6.50 (12.25)[b]	7.00 (5.5)[b]	1.00 (1.00)[e,f]	.02
Attachment	10.75 (5.13)[b]	8.50 (9.75)	10.75 (2.00)[b]	14.00 (3.25)[b]	2.5 (3.00)[c,e,f]	.001
Harm	1.00 (1.75)	1.50 (6.75)	1.00 (0.00)	2.00 (5.00)	1.00 (1.50)	.30
Perceived change	13.75 (4.25)[b]	6.75 (9.13)	10.50 (3.88)	15.00 (6.50)[b]	1.25 (1.38) [c,f]	<.001
Movement	17.00 (6.00)[b,d]	7.00 (6.75)[g]	10.50 (8.00)[b]	16.00 (7.00)[b]	1.00 (1.00)[c,e,f]	<.001
Anxiety	2.50 (7.75)	1.50 (2.75)	5.00 (6.50)	4.00 (5.50)	1.00 (1.00)	.25

[a]CRS: Comfort Rating Scale.

[b]Statistically significant difference from Wavelet.

[c]Statistically significant difference from AP-20.

[d]Statistically significant difference from Checkme O2+.

[e]Statistically significant difference from Nonin.

[f]Statistically significant difference from PC-68B.

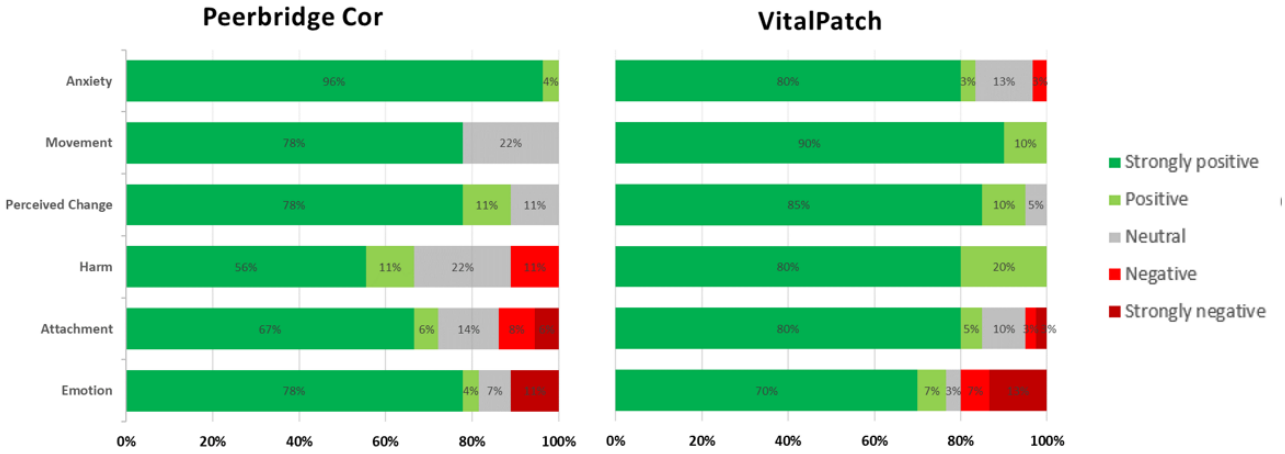
Table 5. Comparison of the Comfort Rating Scale scores between the chest patches.

CRS[a] section	Chest patches, median (IQR)		
	Peerbridge Cor (n=9)	VitalPatch (n=10)	*P* value
Overall score	1.50 (1.50)	1.25 (2.00)	.60
Emotion	1.00 (2.00)	1.00 (1.00)	.92
Attachment	4.50 (2.50)	3.00 (1.00)	.04
Harm	3.00 (10.00)	2.50 (3.50)	.48
Perceived change	2.50 (3.00)	1.00 (3.13)	.17
Movement	3.00 (3.00)	1.00 (1.50)	.06
Anxiety	1.00 (0.00)	1.50 (3.50)	.22

[a]CRS: Comfort Rating Scale.

Figure 3. Comfort Rating Scale scores for each pulse oximeter. Green represents the percentage of positive outcomes and red represents the percentage of negative outcomes from the Comfort Rating Scale scores (Multimedia Appendix 2).

Figure 4. Comfort Rating Scale scores for each chest patch device. Green represents the percentage of positive outcomes and red represents the percentage of negative outcome from the Comfort Rating Scale scores (Multimedia Appendix 2).



A post hoc analysis showed that the Wavelet had significantly better scores than other pulse oximeters in most sections of the CRS (P<.05), for example, in the CRS total score, the Wavelet was statistically superior to AP-20 (P=.004), Checkme O2+ (P=.048), Nonin (P=.02), and PC-68B (P<.001). The Checkme O2+ also showed a significantly better score for Movement against the PC-68B (P=.048) and close to statistical significance against the AP-20 (P=.05) (Multimedia Appendix 3).

For the chest patches, although there was a statistically significant difference in the Attachment section, no difference was found in the overall median CRS scores between the VitalPatch and Peerbridge Cor (Table 5). In the open feedback, the main difference reported by participants was related to sleep, with the Peerbridge Cor more frequently reported to be uncomfortable.

Discussion

Main Results

Twenty participants were recruited for this study, with 70 questionnaires completed (approximately 10 per device) to assess AMS wearability. The Wavelet was found to be the most comfortable wearable pulse oximeter, with the absence of a finger probe, thereby contributing to its better wearability score. Having a probe on the fingertip seemed to negatively impact a device's wearability, with participants reporting a feeling of tightness and sweatiness after prolonged and continuous use. The finger probe was also often reported as hindering function, requiring its removal to perform activities and affecting the total time the device was worn. Amongst the pulse oximeters with a finger probe, the Checkme O2+ was preferred, probably due to the smaller and ring-shaped finger probe, with placement away from the fingertip. For the remaining pulse oximeters (AP-20, PC-68B, and Nonin), no significant differences were found; however, PC-68B was the most negatively commented on, as participants consistently described this device as bulky and disruptive to ADLs. In addition, it had the most negative feedback for a finger probe, in comparison with AP-20 and Nonin.

For the chest patches, there were only significant differences between the 2 selected devices in favor of the VitalPatch in the Attachment section of the CRS scale. Participants noted that they found it difficult to sleep on their front or side with the Peerbridge Cor.

Preference for the Wavelet, Checkme O2+, and VitalPatch was also reflected in the pain/discomfort score, activity exertion score, and free text feedback given. Participants also reported preference for smaller devices, both for the pulse oximeters and the chest patches. To our knowledge, this is the first study comparing wearability for a number of wearable devices.

Study Limitations

A clear limitation of this study was the recruitment of in-house healthy volunteers; thus, our results may not reflect the hospitalized population. However, this was the first step within our project to provide evidence on the wearability and to select feasible devices to be further tested. As there was a limited number of devices available, the order was not randomly assigned; these were allocated to participants as they were available, with the order being documented.

Another limitation was that not all participants were able to test all 7 devices, with an average of 3.5 devices used per participant (1 being a chest patch) between all sessions. To avoid bias, participants were encouraged to assess each device individually and not by comparison with previous devices worn. Additionally, although a log of temporary removals was provided, not all participants were fully compliant with it, which may explain the variability in the device removal numbers.

Furthermore, no specific instruction was given to participants regarding the finger probe placement for any of the pulse oximeters (we note that Viatom Technology recommends using the Checkme O2+ on the thumb [28]), and participants were therefore not asked to indicate which finger they used. This could provide an additional comparison between pulse oximeters that use finger probes and such a comparison will be included in future wearability studies.

We have used the CRS scale [18,19] for the main assessment of device wearability in healthy volunteers; however, the 6 domains evaluated and their distribution might not be the most appropriate method to evaluate wearability for clinical

monitoring devices, as it does not necessarily ensure a correct representation of and applicability to the clinical environment.

Comparison With Prior Work

Our questionnaire methodology for wearability evaluation has been described in previous studies [17,18] and adapted here to compare our selected devices using similar outcomes such as the CRS scale, which is designed to assess comfort over a range of dimensions [29]. Wearability has a direct impact on system usability and its clinical implementation, as patients will be more likely to wear the AMS if they feel comfortable, thus improving data availability and quality [30]. A recent review analyzed the validation, feasibility, clinical outcomes, and costs of 13 different wearable devices and concluded that these were predominantly in the validation and feasibility testing phases [31]. Despite the exponential growth in wearable technology, little evidence is available regarding wearability and acceptance in the clinical setting [30,32,33]. We note that, for all the devices under test, only the VitalPatch had indexed wearability studies available [34,35].

This exploratory study is embedded in a comprehensive research project, which aims to test, refine, and deploy these devices in clinical practice. This project follows a human-centered design process that requires a full exploration of the environment into which the technology is to be placed and understanding the eventual end users of the technology [21]. Although this was not tested on end users, our study was the most ethically sound surrogate, as it did not expose patients to equipment that does not work or that they would find intolerable because of discomfort.

The findings of this study will support the initial selection of wearable devices for the next phases of our project for testing the reliability, accuracy, and functionality of the selected devices [22], as it is not known the AMSs that are the most reliable for use in the hospital environment. Once the initial devices have been selected, tested, and refined, patients will also have the opportunity to provide both qualitative and quantitative data on the wearability of the devices.

Conclusions and Future Research

Our results suggest that traditional pulse oximeter finger probes hinder function, as participants preferred the wrist-worn (Wavelet) and ring-style pulse oximeters (Checkme O2+). The smaller chest patch (VitalPatch) was found to be less noticeable and more comfortable. These preferences were reflected in the total time participants wore the device. These results help to inform which wearable device designs are more likely to be deployed successfully within the hospital environment.

Acknowledgments

The research was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. PW and LT are supported by the NIHR Biomedical Research Centre, Oxford. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. We thank Delaram Jarchi for performing the initial review that supported our device selection as part of this study. We would also like to thank Philippa Piper, Dario Salvi, and Marco Pimentel for their support in the design of the study.

Authors' Contributions

CA, LY, SV, JE, LT and PW made significant contributions to the conception and design or acquisition of data. CA, LY, and MS analyzed the data. CA and LY wrote the first draft of the manuscript, which was reviewed and approved by all authors.

Conflicts of Interest

PW and LT report significant grants from the NIHR, UK, and the NIHR Biomedical Research Centre, Oxford, during the conduct of the study. PW and LT report modest grants and personal fees from Sensyne Health, outside the submitted work. LT works part-time for Sensyne Health and has share options in the company. PW holds shares in the company. All other authors have no conflicts to declare.

Multimedia Appendix 1
STROBE checklist.
[PDF File (Adobe PDF File), 136 KB - mhealth_v8i12e20214_app1.pdf]

Multimedia Appendix 2
Comfort Rating Scale information.
[DOCX File , 21 KB - mhealth_v8i12e20214_app2.docx]

Multimedia Appendix 3
Post hoc Dunn tests with Bonferroni Correction for Pulse Oximeters.
[DOCX File , 23 KB - mhealth_v8i12e20214_app3.docx]

References

1. McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, et al. Confidential inquiry into quality of care before admission to intensive care. BMJ 1998 Jun 20;316(7148):1853-1858 [FREE Full text] [doi: 10.1136/bmj.316.7148.1853] [Medline: 9632403]
2. Watkinson PJ, Barber VS, Price JD, Hann A, Tarassenko L, Young JD. A randomised controlled trial of the effect of continuous electronic physiological monitoring on the adverse event rate in high risk medical and surgical patients. Anaesthesia 2006 Nov;61(11):1031-1039 [FREE Full text] [doi: 10.1111/j.1365-2044.2006.04818.x] [Medline: 17042839]
3. Acutely ill adults in hospital: recognising and responding to deterioration. The National Institute for Health and Care Excellence. 2007 Jul 25. URL: https://www.nice.org.uk/guidance/cg50 [accessed 2019-12-17]
4. Dall'Ora C, Griffiths P, Redfern O, Recio-Saucedo A, Meredith P, Ball J, Missed Care Study Group. Nurses' 12-hour shifts and missed or delayed vital signs observations on hospital wards: retrospective observational study. BMJ Open 2019 Feb 01;9(1):e024778 [FREE Full text] [doi: 10.1136/bmjopen-2018-024778] [Medline: 30782743]
5. Griffiths P, Recio-Saucedo A, Dall'Ora C, Briggs J, Maruotti A, Meredith P, Missed Care Study Group. The association between nurse staffing and omissions in nursing care: A systematic review. J Adv Nurs 2018 Jul;74(7):1474-1487 [FREE Full text] [doi: 10.1111/jan.13564] [Medline: 29517813]
6. Dall'Ora C, Griffiths P, Hope J, Barker H, Smith GB. What is the nursing time and workload involved in taking and recording patients' vital signs? A systematic review. J Clin Nurs 2020 Jul;29(13-14):2053-2068. [doi: 10.1111/jocn.15202] [Medline: 32017272]
7. Jansen JO, Cuthbertson BH. Detecting critical illness outside the ICU: the role of track and trigger systems. Curr Opin Crit Care 2010 Jun;16(3):184-190. [doi: 10.1097/MCC.0b013e328338844e] [Medline: 20305556]
8. Clifton DA, Clifton L, Sandu D, Smith GB, Tarassenko L, Vollam SA, et al. 'Errors' and omissions in paper-based early warning scores: the association with changes in vital signs--a database analysis. BMJ Open 2015 Jul 03;5(7):e007376 [FREE Full text] [doi: 10.1136/bmjopen-2014-007376] [Medline: 26141302]
9. Downey C, Tahir W, Randell R, Brown J, Jayne D. Strengths and limitations of early warning scores: A systematic review and narrative synthesis. Int J Nurs Stud 2017 Nov;76:106-119. [doi: 10.1016/j.ijnurstu.2017.09.003] [Medline: 28950188]
10. Cardona-Morrell M, Prgomet M, Turner RM, Nicholson M, Hillman K. Effectiveness of continuous or intermittent vital signs monitoring in preventing adverse events on general wards: a systematic review and meta-analysis. Int J Clin Pract 2016 Oct;70(10):806-824. [doi: 10.1111/ijcp.12846] [Medline: 27582503]
11. Tarassenko L, Hann A, Young D. Integrated monitoring and analysis for early warning of patient deterioration. Br J Anaesth 2006 Jul;97(1):64-68 [FREE Full text] [doi: 10.1093/bja/ael113] [Medline: 16707529]
12. Prgomet M, Cardona-Morrell M, Nicholson M, Lake R, Long J, Westbrook J, et al. Vital signs monitoring on general wards: clinical staff perceptions of current practices and the planned introduction of continuous monitoring technology. Int J Qual Health Care 2016 Sep;28(4):515-521. [doi: 10.1093/intqhc/mzw062] [Medline: 27317251]
13. Bonnici T, Tarassenko L, Clifton DA, Watkinson P. The digital patient. Clin Med (Lond) 2013 Jun;13(3):252-257 [FREE Full text] [doi: 10.7861/clinmedicine.13-3-252] [Medline: 23760698]
14. Downey C, Chapman S, Randell R, Brown J, Jayne D. The impact of continuous versus intermittent vital signs monitoring in hospitals: A systematic review and narrative synthesis. Int J Nurs Stud 2018 Aug;84:19-27. [doi: 10.1016/j.ijnurstu.2018.04.013] [Medline: 29729558]
15. Weenk M, van Goor H, Frietman B, Engelen L, van Laarhoven CJ, Smit J, et al. Continuous Monitoring of Vital Signs Using Wearable Devices on the General Ward: Pilot Study. JMIR Mhealth Uhealth 2017 Jul 05;5(7):e91 [FREE Full text] [doi: 10.2196/mhealth.7208] [Medline: 28679490]
16. Appelboom G, Camacho E, Abraham ME, Bruce SS, Dumont EL, Zacharia BE, et al. Smart wearable body sensors for patient self-assessment and monitoring. Arch Public Health 2014;72(1):28 [FREE Full text] [doi: 10.1186/2049-3258-72-28] [Medline: 25232478]
17. Cancela J, Pastorino M, Tzallas AT, Tsipouras MG, Rigas G, Arredondo MT, et al. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. Sensors (Basel) 2014 Sep 16;14(9):17235-17255 [FREE Full text] [doi: 10.3390/s140917235] [Medline: 25230307]
18. Knight J, Deen-Williams D, Arvanitis T, Baber C, Sotiriou S, Anastopoulou S, et al. Assessing the wearability of wearable computers. : IEEE; 2007 Presented at: Proceedings - International Symposium on Wearable Computers, ISWC; 11-14 Oct 2006; Montreux, Switzerland p. 75-82. [doi: 10.1109/ISWC.2006.286347]
19. Knight JF, Baber C. A tool to assess the comfort of wearable computers. Hum Factors 2005;47(1):77-91. [doi: 10.1518/0018720053653875] [Medline: 15960088]
20. Jeffs E, Vollam S, Young JD, Horsington L, Lynch B, Watkinson PJ. Wearable monitors for patients following discharge from an intensive care unit: practical lessons learnt from an observational study. J Adv Nurs 2016 Aug;72(8):1851-1862. [doi: 10.1111/jan.12959] [Medline: 26990704]
21. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]

22. Areia C, Vollam S, Piper P, King E, Ede J, Young L, et al. Protocol for a prospective, controlled, cross-sectional, diagnostic accuracy study to evaluate the specificity and sensitivity of ambulatory monitoring systems in the prompt detection of hypoxia and during movement. BMJ Open 2020 Jan 12;10(1):e034404 [FREE Full text] [doi: 10.1136/bmjopen-2019-034404] [Medline: 31932393]
23. Borg G. Borg's perceived exertion and pain scales. Illinois: Human Kinetics, 1998; 1998.
24. Daniel W. Applied nonparametric statistics. In: Applied nonparametric statistics. Boston: Cengage Learning; 2nd Revised edition (1 Jan. 1980); 1990.
25. Neyman J, Pearson ES. On The Use And Interpretation Of Certain Test Criteria For Purposes Of Statistical Inference Part I. Biometrika 1928;20A(1-2):175-240. [doi: 10.1093/biomet/20a.1-2.175]
26. A Language and Environment for Statistical Computing Internet. Vienna, Austria. 2019. URL: https://www.r-project.org/ [accessed 2020-03-01]
27. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. JOSS 2019 Nov;4(43):1686. [doi: 10.21105/joss.01686]
28. Shenzhen VTC. O2 Vibe Wrist Pulse Oximeter Quick Start Guide. 2016. URL: https://fccid.io/2ADXK-1600/User-Manual/Users-Manual-3119531.pdf [accessed 2019-08-01]
29. Knight J, Baber C, Schwirtz A, Bristow H. The comfort assessment of wearable computers. : IEEE; 2003 Presented at: Proceedings Sixth Int Symp Wearable Comput Internet IEEE; 10 Oct. 2002; Seattle, WA, USA p. 65-72. [doi: 10.1109/ISWC.2002.1167220]
30. 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 Jul;41(7):115. [doi: 10.1007/s10916-017-0760-1] [Medline: 28631139]
31. Leenen J, Leerentveld C, van Dijk JD, van Westreenen HL, Schoonhoven L, Patijn G. 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]
32. Botros A, Schütz N, Camenzind M, Urwyler P, Bolliger D, Vanbellingen T, et al. Long-Term Home-Monitoring Sensor Technology in Patients with Parkinson's Disease—Acceptance and Adherence. Sensors 2019 Nov 26;19(23):5169. [doi: 10.3390/s19235169]
33. Pavic M, Klaas V, Theile G, Kraft J, Tröster G, Guckenberger M. Feasibility and Usability Aspects of Continuous Remote Monitoring of Health Status in Palliative Cancer Patients Using Wearables. Oncology 2020;98(6):386-395 [FREE Full text] [doi: 10.1159/000501433] [Medline: 31336377]
34. Tonino R, Larimer K, Eissen O, Schipperus M. Remote Patient Monitoring in Adults Receiving Transfusion or Infusion for Hematological Disorders Using the VitalPatch and accelerateIQ Monitoring System: Quantitative Feasibility Study. JMIR Hum Factors 2019 Dec 02;6(4):e15103 [FREE Full text] [doi: 10.2196/15103] [Medline: 31789596]
35. Selvaraj N. Long-term remote monitoring of vital signs using a wireless patch sensor. 2015 Feb 12 Presented at: IEEE Healthc Innov Conf Internet IEEE; 8-10 Oct 2014; Seattle, WA, USA p. 83-86. [doi: 10.1109/HIC.2014.7038880]

Abbreviations

ADLs: activities of daily living
AMS: ambulatory monitoring system
CRS: comfort rating scale
NIHR: National Institute for Health Research
SpO_2: oxygen saturation

Edited by G Eysenbach, L Buis; submitted 13.05.20; peer-reviewed by C King, K Ng; comments to author 16.07.20; revised version received 28.08.20; accepted 21.10.20; published 16.12.20.

Please cite as:
Areia C, Young L, Vollam S, Ede J, Santos M, Tarassenko L, Watkinson P
Wearability Testing of Ambulatory Vital Sign Monitoring Devices: Prospective Observational Cohort Study
JMIR Mhealth Uhealth 2020;8(12):e20214
URL: http://mhealth.jmir.org/2020/12/e20214/
doi:10.2196/20214
PMID:33325827

Original Paper

Comparison of the Physical Activity Measured by a Consumer Wearable Activity Tracker and That Measured by Self-Report: Cross-Sectional Analysis of the Health eHeart Study

Alexander J Beagle[1], MD; Geoffrey H Tison[2], MPH, MD; Kirstin Aschbacher[2], PhD; Jeffrey E Olgin[2], MD; Gregory M Marcus[2], MAS, MD; Mark J Pletcher[3], MPH, MD

[1]Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[2]Division of Cardiology, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[3]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, United States

Corresponding Author:
Alexander J Beagle, MD
Department of Medicine
University of California San Francisco
505 Parnassus Ave
San Francisco, CA, 94143
United States
Phone: 1 9098165831
Email: alexanderjbeagle@gmail.com

Abstract

Background: Commercially acquired wearable activity trackers such as the Fitbit provide objective, accurate measurements of physically active time and step counts, but it is unclear whether these measurements are more clinically meaningful than self-reported physical activity.

Objective: The aim of this study was to compare self-reported physical activity to Fitbit-measured step counts and then determine which is a stronger predictor of BMI by using data collected over the same period reflecting comparable physical activities.

Methods: We performed a cross-sectional analysis of data collected by the Health eHeart Study, a large mobile health study of cardiovascular health and disease. Adults who linked commercially acquired Fitbits used in free-living conditions with the Health eHeart Study and completed an International Physical Activity Questionnaire (IPAQ) between 2013 and 2019 were enrolled (N=1498). Fitbit step counts were used to quantify time by activity intensity in a manner comparable to the IPAQ classifications of total active time and time spent being sedentary, walking, or doing moderate activities or vigorous activities. Fitbit steps per day were computed as a measure of the overall activity for exploratory comparisons with IPAQ-measured overall activity (metabolic equivalent of task [MET]-h/wk). Measurements of physical activity were directly compared by Spearman rank correlation. Strengths of associations with BMI for Fitbit versus IPAQ measurements were compared using multivariable robust regression in the subset of participants with BMI and covariates measured.

Results: Correlations between synchronous paired measurements from Fitbits and the IPAQ ranged in strength from weak to moderate (0.09-0.48). In the subset with BMI and covariates measured (n=586), Fitbit-derived predictors were generally stronger predictors of BMI than self-reported predictors. For example, an additional hour of Fitbit-measured vigorous activity per week was associated with nearly a full point reduction in BMI (–0.84 kg/m^2, 95% CI –1.35 to –0.32) in adjusted analyses, whereas the association between self-reported vigorous activity measured by IPAQ and BMI was substantially smaller in magnitude (–0.17 kg/m^2, 95% CI –0.34 to –0.00; P<.001 versus Fitbit) and was dominated by the Fitbit-derived predictor when compared head-to-head in a single adjusted multivariable model. Similar patterns of associations with BMI, with Fitbit dominating self-report, were seen for moderate activity and total active time and in comparisons between overall Fitbit steps per day and IPAQ MET-h/wk on standardized scales.

Conclusions: Fitbit-measured physical activity was more strongly associated with BMI than self-reported physical activity, particularly for moderate activity, vigorous activity, and summary measures of total activity. Consumer-marketed wearable activity trackers such as the Fitbit may be useful for measuring health-relevant physical activity in clinical practice and research.

(JMIR Mhealth Uhealth 2020;8(12):e22090) doi:10.2196/22090

KEYWORDS

exercise; body mass index; overweight; obesity; fitness trackers; self-report; adult; mHealth; public health; cardiovascular diseases

Introduction

The prevalence of overweight and obesity has steadily increased worldwide since 1980 [1], with concomitant increases in all-cause mortality [2] and morbidity from diseases such as cardiovascular disease [3,4]. Increased physical activity in adults is associated with reductions in BMI, cardiovascular disease, and all-cause mortality [5-8], thereby making it a crucial target for efforts to improve health [9]. Measuring physical activity is challenging but essential for clinical care and research [10].

Data from self-report instruments such as the International Physical Activity Questionnaire (IPAQ) [11] are associated with BMI [12], but only modestly. Moreover, self-reported physical activity is less accurate than measurements from medical-grade accelerometers such as the ActiGraph [13-16]. Commercial wearable activity monitors or smartwatches such as the Fitbit are growing in popularity and could be used by clinicians and researchers to obtain objective physical activity measurements. They provide step counts that are similar to hand-counted steps [17] and medical-grade devices [16,18,19] and can be used to define minute-to-minute activity intensity [20,21]. Such devices could supplant self-report for physical activity measurement, but they are not designed to capture activities such as biking or swimming that can be self-reported. Several studies have highlighted that compared to activity measurement by self-report, associations between physical activity and BMI are likely stronger when physical activity is measured by objective means by using medical grade accelerometers provided by studies [22-25]. However, no study has explored whether this extends to consumer wearable activity trackers such as Fitbits. This is of particular interest for activities of various intensities where some may not be well-represented on an activity tracker, especially when measuring the same 7 days by activity tracker and self-report.

The Health eHeart Study, which has collected BMI and physical activity information via self-report and Fitbit, provides an opportunity to address this question. We conducted a cross-sectional study to (1) assess the agreement between objective Fitbit-measured physical activity and self-report (IPAQ) and (2) compare the strength of the associations with BMI between Fitbit and IPAQ, utilizing data from the Health eHeart Study [26], a large-scale mobile health study coordinated by the University of California San Francisco. Minute-by-minute Fitbit step counts were used to categorize time by activity intensity for comparisons with analogous self-reported IPAQ measures and calculated measures of overall physical activity from Fitbit (steps/d) and IPAQ (metabolic equivalent of task [MET]/wk). We then compared Fitbit to self-report in terms of cross-sectional associations with BMI. We hypothesized that Fitbit-measured activity would be more strongly associated with BMI than self-reported activity but this would not be consistent across various activity intensities.

Methods

Participants in This Study

Health eHeart is a prospective longitudinal cohort study that has enrolled all interested adult volunteers (age≥18 years) who have an email address and have been residing within the United States since March 2013. Participation in the study is entirely web-based and recruitment was conducted on the internet through social media advertisements, email campaigns with advocacy and research organizations, and in person at clinics at the University of California San Francisco. We included participants who completed at least one IPAQ and opted before June 2019 to share data from commercially acquired Fitbits with Health eHeart. Participants were excluded only for missing or invalid data. Health eHeart was approved by the University of California San Francisco Committee on Human Research. Consent was obtained electronically on enrollment. We obtained deidentified data for this study.

Physical Activity Measurement

Self-reported activity was assessed with the IPAQ short form, which assesses activity at discrete intensities (sedentary, walking, moderate activity, or vigorous activity) over 7 consecutive days [11,12]. Data were cleaned according to the IPAQ protocol with additional criteria developed to address erroneous values (Multimedia Appendix 1). Total active time (sum of time spent walking, moderately active, and vigorously active) and overall activity in weekly MET hours were computed [12].

Physical activity was measured objectively on the same 7 days with Fitbit-brand wearable activity trackers, which are triaxial accelerometers using proprietary algorithms to convert accelerometer counts to minute-by-minute step counts. The Fitbit devices that contributed study data were owned by study participants who obtained their devices outside of the Health eHeart study; the study did not request that participants acquire a Fitbit for this study and did not exclude specific Fitbit device types. Participants received no specific instructions on Fitbit use from Health eHeart. No specific devices were excluded and all were Fitbit brand (Tables S1 and S2 in Multimedia Appendix 1). We defined nonwear time as ≥120 contiguous minutes with 0 steps counted; all other time was considered wear time [27]. We recoded minutes with more than 200 steps as missing, as described previously in accelerometer studies [28,29]. Fitbit data were considered adequate for a given day when there were at least 10 hours of Fitbit wear time during that day. An activity observation was defined as a period of 7 days recalled by a participant on the IPAQ for which the same 7 days had adequate Fitbit data.

We then computed a series of Fitbit-derived physical activity estimates corresponding to standard estimates derived from the IPAQ [30]. Fitbit overall activity was defined as a within-participant average of steps/day during the week. We

calculated time spent walking (10-100 steps/min, eg, slow walking), moderately active (101-130 steps/min, eg, brisk walking), vigorously active (131 to 200 steps/min, eg, vigorous walking, jogging, running), and sedentary time, defined as bouts shorter than 120 consecutive minutes with 0-9 steps, to allow for incidental movement [31] and typical sedentary bout lengths [32]. Minutes with at least 10 steps counted were considered active time.

This intensity-defined approach was designed to reflect energy expenditure classifications for the IPAQ [31,33,34]. In a review of literature correlating step rates and METs expended in adults, 64-96 steps/min ranged from 2.0 to 3.1 METs (5 studies), 102-129 steps/min ranged from 2.9 to 5.5 METs (8 studies), and 134 to 170 steps/min ranged from 6.8 to 13.0 METs (7 studies) [35], which are closely aligned with IPAQ definitions of walking (<3 METs), moderate activity (3-6 METs), and vigorous (>6 METs) activity [36,37].

BMI Measurement

Height, weight, and BMI were self-reported through Health eHeart. Values reported within 90 days of the activity observation were averaged for weight and BMI, and values within 365 days were averaged for height. When both were available, height and weight data were combined to generate a calculated BMI value associated with each activity observation. Calculated BMI values were merged with self-reported BMI values using the median when both were available for an activity observation (supplementary methods in Multimedia Appendix 1).

Other Measurements

Demographic characteristics, health-related behaviors (ie, alcohol and smoking), and clinical characteristics (ie, history of coronary artery disease, hypertension, hyperlipidemia, and diabetes) were obtained through questionnaires. We used responses within 365 days of and closest to each activity observation.

Statistical Analyses

Some participants contributed multiple observations over time (Figure S1 in Multimedia Appendix 1). Comparative analyses were restricted to the first activity observation (N=1498). We investigated agreement between IPAQ-measured and Fitbit-measured activity by using Wilcoxon matched-pairs signed-rank test and Spearman's rank correlations on untransformed data. For the BMI analysis, we excluded participants lacking BMI or covariate data to generate a subsample of 697 observations from 586 participants (Multimedia Appendix 1). We compared normally distributed continuous variables by two-sided *t* test, nonnormal variables by Wilcoxon rank sum test and categorical variables by chi-squared test. Bivariate associations between activity and BMI were examined by Spearman's rank correlation and nonparametric test-of-trend across heuristically defined physical activity categories at participant level restricted to the first activity observation. We next used multivariable regressions of BMI on Fitbit-measured and IPAQ-measured activity variables by using robust errors clustered by participants to account for activity outliers and nonindependence for participants with multiple observations. Activity distributions and BMI were not transformed. Twelve separate predictor models were fit—one for each of the 6 analogous activity variables measured by either Fitbit or IPAQ. Next, 6 head-to-head combined models were fit with both analogous measurements as predictors (eg, vigorous activity measured by IPAQ and Fitbit in the same model) to assess BMI association strength while controlling for the analogous measurement. All models were adjusted for potential confounding by Fitbit wear-time, device wear location (torso or wrist), and season of data collection and estimated in natural units and in standardized units to account for differences in self-reported and Fitbit-measured activity distributions. We then sequentially adjusted the models for demographic characteristics, health-related behaviors, and clinical characteristics.

We conducted 2 sensitivity analyses to investigate whether data sources affected the main findings of the study. The first examined whether the main findings were sensitive to the inclusion of participants whose data were supplied by MobileTrack (81/1392, 5.8% of comparative analysis cohort; 32/586, 5.5% of BMI analysis subset), which was derived from smartphone measurements and not a wearable activity tracker. The second examined whether the main findings were sensitive to the inclusion of participants whose activity trackers were torso-worn (397/1392, 28.5% of the comparative analysis cohort; 166/586, 28.3% of the BMI analysis subset). For both sensitivity analyses, we excluded the relevant cohort and repeated the study analysis.

Data preparation and analyses were performed using STATA 14 (StataCorp). Two-tailed *P* values less than .05 were considered significant. We report this study according to STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines [38].

Results

Comparative Analysis

The age of the participants (N=1498) ranged from 19 to 92 years and they were predominantly Whites (1299/1498, 86.7%), females (895/1498, 59.7%), with at least a bachelor's degree (1067/1415, 75.4%), and an annual income exceeding US $100,000 (695/1282, 54.2%) (Table 1). All characteristics in Table 1 are summarized at participant level at the first activity observation by using mean (SD) for normal continuous data, median (IQR) for nonnormal continuous, and not available for frequency data.

Table 1. Demographic and clinical characteristics of the participants in the comparative analysis and BMI analysis subsets.

Characteristics	Comparative analysis	BMI analysis
Demographic characteristics		
Age (years)[a], mean (SD)	52.9 (13.9)	51.9 (13.4)
Female gender[a], n (%)	895 (59.8)	370 (63.1)
Education[b], n (%)		
<Bachelor's degree	348 (24.6)	151 (25.7)
Bachelor's degree	492 (34.8)	183 (31.2)
Postgraduate degree	575 (40.6)	252 (43.0)
Annual income (US $)[c], n (%)		
<$50,000	217 (16.9)	99 (16.9)
$50,000-$100,000	370 (28.9)	175 (29.9)
>$100,000	695 (54.2)	312 (53.2)
Race[d], n (%)		
White	1299 (87.1)	514 (87.7)
Asian	64 (4.3)	19 (3.2)
Black/African-American	56 (3.8)	23 (3.9)
Multiracial	41 (2.8)	17 (2.9)
Other	32 (2.1)	13 (2.2)
Hispanic or Latino	82 (5.5)	32 (5.5)
Health-related behaviors		
Smoking[e], n (%)		
Never	990 (68.3)	392 (66.9)
Past	420 (28.9)	17 (29.5)
Current	39 (2.7)	21 (3.6)
Alcoholic drinks/wk[f], median (IQR)	3 (1-7)	3 (1-7)
Clinical characteristics		
Height[g], (m), median (IQR)	1.70 (1.63-1.78)	1.68 (1.63-1.75)
Weight[h], (kg), median (IQR)	77.6 (66.5-89.4)	77.6 (66.4-8.9)
BMI[i], (kg/m^2), median (IQR)	26.5 (23.5-30.2)	26.7 (23.5-30.2)
Coronary artery disease[j], n (%)	64 (6.6)	38 (6.5)
Diabetes[j], n (%)	61 (6.3)	36 (6.1)
Hyperlipidemia[k], n (%)	392 (40.7)	240 (41.0)
Hypertension[l], n (%)	347 (35.8)	215 (36.7)

[a]Calculated for all 1498 comparative analysis participants and 586 BMI analysis participants.
[b]Calculated for 1415 comparative analysis participants and 586 BMI analysis participants.
[c]Calculated for 1282 comparative analysis participants and 586 BMI analysis participants.
[d]Calculated for 1492 comparative analysis participants and 586 BMI analysis participants.
[e]Calculated for 1449 comparative analysis participants and 586 BMI analysis participants.
[f]Available for 1288 comparative analysis participants and 586 BMI analysis participants.
[g]Available for 1076 comparative analysis participants and 533 BMI analysis participants.
[h]Available for 1095 comparative analysis participants and 546 BMI analysis participants.
[i]Available for 1196 comparative analysis participants and 586 BMI analysis participants.

[j]Available for 968 comparative analysis participants and 586 BMI analysis participants.

[k]Available for 963 comparative analysis participants and 586 BMI analysis participants.

[l]Available for 967 comparative analysis participants and 586 BMI analysis participants.

Physical activity observations were not evenly distributed among seasons (Wald test, $P<.001$) and 995 (71.5%) of 1392 Fitbit devices were wrist-worn (Table 2). In Table 2, all characteristics are summarized at participant level at the first activity observation by using mean (SD) for normal continuous data, median (IQR) for nonnormal continuous, and not available for frequency data. There were no missing data for the BMI analysis. The median overall activity at the first activity observation was 35.5 MET-h/wk (IQR 18.1-62.2) by IPAQ. Fitbits counted a median of 8622 steps/d (IQR 6191-11061) over the same week. Compared to Fitbit measurements, participants self-reported more moderate activity (median of 90 min/wk vs 75 min/wk, Wilcoxon matched-pairs sign-rank test $P<.001$) and vigorous activity (80 min/wk vs 0 min/wk, $P<.001$), but self-reported less time spent sedentary (42.0 h/wk vs 87.2 h/wk; $P<.001$) and walking (3.5 h/wk vs 20.2 h/wk, $P<.001$).

Table 2. Physical activity characteristics of the participants in the comparative analysis and BMI analysis subset.

Characteristics	Comparative analysis	BMI analysis subset
Season of data collection[a], n (%)		
Spring	340 (22.7)	161 (27.5)
Summer	445 (29.7)	149 (25.4)
Fall	323 (21.6)	141 (24.1)
Winter	390 (26.0)	135 (23.0)
Fitbit		
Wearing on wrist[b], n (%)	995 (71.5)	420 (71.7)
Wearing on torso[b], n (%)	397 (28.5)	166 (28.3)
Wear time (h/d)[a], median (IQR)	17.1 (16.0-18.5)	17.1 (15.9-18.5)
Overall activity (steps/d)[a], median (IQR)	8622 (6191-11,061)	8778 (6385-11,115)
Active time (h/wk)[a], median (IQR)	22.9 (18.0-28.0)	23.1 (18.5-28.1)
Sedentary (h/wk)[a], median (IQR)	87.2 (80.5-94.2)	87.3 (80.7-94.9)
Walking (h/wk)[a], median (IQR)	20.2 (15.9-25.0)	20.8 (16.6-25.5)
Moderate activity (min/wk)[a], median (IQR)	75 (26-163)	75 (28-166)
Vigorous activity (min/wk)[a], median (IQR)	0 (0-11)	1 (0-12)
IPAQ[a,c]		
Overall activity (metabolic equivalent of tasks-h/wk), median (IQR)	35.5 (18.1-62.2)	34.9 (19.3-60.0)
Active time (h/wk), median (IQR)	8.0 (4.3-14.0)	8.2 (4.5-13.2)
Sedentary (h/wk), median (IQR)	42.0 (21.0-56.0)	42.0 (26.8-56.0)
Walking (h/wk), median (IQR)	3.5 (2.0-7.0)	3.5 (1.7-7.0)
Moderate activity (min/wk), median (IQR)	90 (20-200)	90 (30-200)
Vigorous activity (min/wk), median (IQR)	80 (0-180)	80 (0-180)

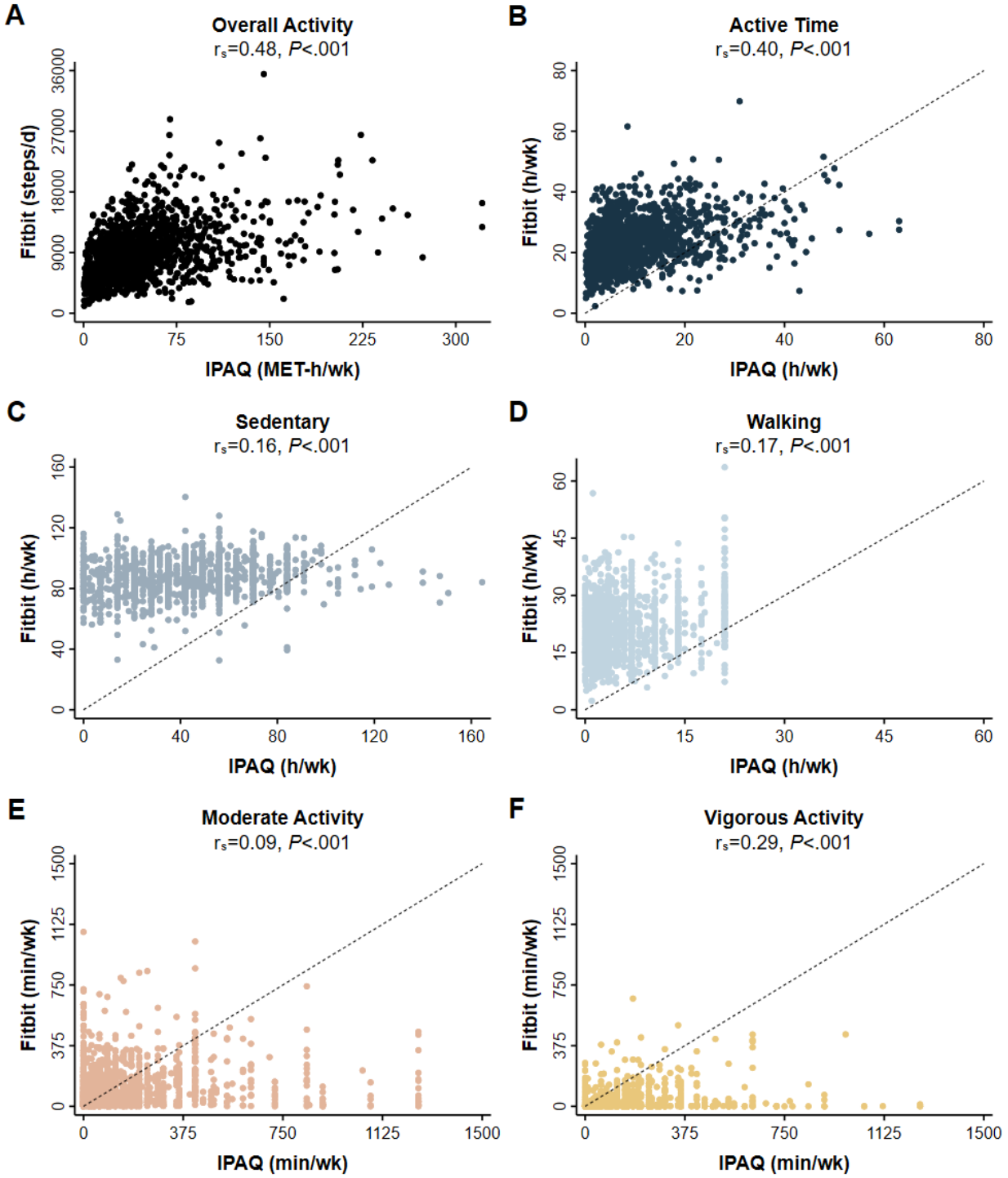
[a]Calculated for 1498 comparative analysis participants and 586 BMI analysis participants.

[b]Calculated for 1345 comparative analysis participants and 586 BMI analysis participants.

[c]IPAQ: International Physical Activity Questionnaire.

Overall activity measurements by Fitbit (steps/d) and IPAQ (MET-h/wk) were moderately correlated (Spearman's rank $r_s=0.48$), as was total active time ($r_s=0.40$) and vigorous activity ($r_s=0.29$) (Figure 1, Figure S2 in Multimedia Appendix 1). By contrast, time spent sedentary, walking, or moderately active were weakly correlated (walking $r_s=0.17$, sedentary $r_s=0.16$, moderate activity $r_s=0.09$). Findings were very similar when we excluded participants whose data source was the MobileTrack app, and when we excluded participants whose data were obtained from an activity tracker worn on the torso.

Figure 1. Fitbit-measured and self-reported physical activity of 1498 participants. Each symbol represents a participant's first paired activity observation. Spearman's rank correlations (rs) and *P* values are provided for the following: A. Overall activity and intensity-defined time spent. B. Active in total (sum of walking, moderate activity, and vigorous activity). C. Sedentary D. Walking E. Moderately active, and F. Vigorously active. Fitbit measurements were acquired during the same period queried by self-report (IPAQ). Dashed lines represent identical measurements. Overall activity was measured on different scales (Fitbit, steps/d; IPAQ, MET-h/wk) and no dashed line is provided. IPAQ: International Physical Activity Questionnaire; MET: metabolic equivalent of task.



BMI Analysis

The subset of participants with a BMI measurement was included in the BMI analysis sample (n=586). In bivariate analyses using categorical physical activity, Fitbit measurements of activity were significantly associated with BMI for all intensity-defined (Table 3) and overall activity measurements (Table 4) (*P* values range from <.001 to .01), whereas overall activity, total active time, sedentary time, and walking time measured by IPAQ were significantly associated with BMI (*P* values range from .001 to .01).

Table 3. Activity distributions and bivariate associations with BMI by activity intensity in the BMI analysis subset (n=586)[a].

Fitbit				IPAQ[b]			
Time (h/wk)	n (%)	BMI, mean (SD)	*P* value	Time (h/wk)	n (%)	BMI, mean (SD)	*P* value
Sedentary			.01				.001
0-19	0	N/A[c]		0-19	101 (17.2)	26.4 (5.2)	
20-39	0	N/A		20-39	159 (27.1)	26.9 (5.0)	
40-59	4 (0.7)	28.2 (7.8)		40-59	197 (33.6)	28.0 (6.2)	
60-79	133 (22.7)	26.5 (5.7)		60-79	93 (15.9)	28.1 (6.0)	
80+	449 (76.6)	28.0 (6.0)		80+	36 (6.1)	31.0 (8.6)	
Walking			.001				.01
0-4	0	N/A		0-4	357 (60.9)	28.0 (6.0)	
5-9	19 (3.2)	30.4 (7.8)		5-9	136 (23.2)	27.3 (5.5)	
10-14	79 (13.5)	28.8 (6.8)		10-14	49 (8.4)	27.6 (7.6)	
15-19	170 (29.0)	28.4 (6.6)		15-19	12 (2.0)	26.6 (5.3)	
20-24	158 (27.0)	27.3 (6.0)		20-24	32 (5.5)	25.4 (4.8)	
25+	160 (27.3)	26.3 (4.0)		25+	0	N/A	
Moderate activity			<.001				.19
0-1	241 (41.1)	28.7 (6.4)		0-1	194 (33.1)	28.3 (7.0)	
1-2	134 (22.9)	27.5 (5.1)		1-2	113 (19.3)	27.3 (5.9)	
2-3	82 (14.0)	26.8 (5.4)		2-3	90 (15.4)	27.8 (5.2)	
3-4	53 (9.0)	26.8 (5.5)		3-4	64 (10.9)	27.3 (5.2)	
4-5	26 (4.4)	25.5 (4.4)		4-5	33 (5.6)	27.3 (4.9)	
5+	50 (8.5)	26.0 (7.1)		5+	92 (15.7)	26.8 (5.0)	
Vigorous activity			<.001				.13
0-1	534 (91.1)	27.9 (5.9)		0-1	249 (42.5)	28.4 (6.5)	
1-2	29 (4.9)	26.5 (8.2)		1-2	99 (16.9)	27.4 (6.5)	
2-3	13 (2.2)	22.8 (1.9)		2-3	84 (14.3)	26.9 (5.7)	
3-4	5 (0.9)	23.7 (2.5)		3-4	43 (7.3)	26.4 (3.6)	
4-5	3 (0.5)	25.8 (7.7)		4-5	44 (7.5)	28.2 (5.6)	
5+	2 (0.3)	24.0 (1.9)		5+	67 (11.4)	26.6 (4.2)	

[a]Data were restricted to the first activity observation (586 observations). *P* values are for nonparametric test-of-trend across heuristically defined categories of time spent at each level of activity.

[b]IPAQ: International Physical Activity Questionnaire.

[c]N/A: not applicable.

Table 4. Overall activity distributions and bivariate associations with BMI in the BMI analysis subset (n=586)[a].

Fitbit				IPAQ[b]			
Time (h/wk)	n (%)	BMI, mean (SD)	*P* value	Time (h/wk)	n (%)	BMI, mean (SD)	*P* value
Total active time			<.001				.008
0-9	8 (1.4)	32.7 (8.2)		0-9	366 (62.5)	28.2 (6.2)	
10-19	184 (31.4)	29.1 (7.1)		10-19	153 (26.1)	26.8 (5.5)	
20-29	285 (48.6)	27.2 (5.0)		20-29	46 (7.8)	26.9 (6.1)	
30-39	99 (16.9)	26.1 (5.7)		30-39	12 (2.0)	26.8 (5.3)	
40+	10 (1.7)	25.3 (3.1)		40+	9 (1.5)	25.1 (3.0)	
Overall activity, (steps/d)			<.001				c
0-2999	14 (2.4)	31.0 (8.1)		c	c	c	
3000-5999	107 (18.3)	30.6 (7.8)		c	c	c	
6000-8999	191 (32.6)	27.6 (5.0)		c	c	c	
9000-11,999	166 (28.3)	26.9 (4.7)		c	c	c	
12,000-14,999	70 (11.9)	25.5 (4.8)		c	c	c	
15,000+	38 (6.5)	25.4 (7.2)		c	c	c	
Overall activity, (MET-h/wk)			c				.006
c	c	c		0-19	155 (26.4)	29.2 (7.4)	
c	c	c		20-39	181 (30.9)	27.3 (5.1)	
c	c	c		40-59	103 (17.6)	27.2 (5.0)	
c	c	c		60-79	68 (11.6)	26.5 (6.3)	
c	c	c		80-99	32 (5.5)	26.0 (4.1)	
c	c	c		100+	47 (8.0)	27.3 (5.5)	

[a]Data are restricted to the first activity observation (586 observations). *P* values are for nonparametric test-of-trend across heuristically defined categories of time spent at each level of activity.

[b]IPAQ: International Physical Activity Questionnaire.

[c]Cells omitted because overall activity was measured on different scales (Fitbit, steps/d; IPAQ, MET-h/wk).

All Fitbit measurements, including those of lower intensity activities (Figure 2), higher intensity activities (Figure 3), and overall activity (Figure 4) were correlated with BMI (Spearman's rank test, 6 comparisons, all $P<.001$). By contrast, IPAQ measurements of lower intensity activity (Figure 2) and overall activity (Figure 4) were correlated with BMI (P values range from .001 to .005) but not the higher intensity moderate activity ($P=.26$) or vigorous activity ($P=.06$) (Figure 3).

Figure 2. BMI correlations for Fitbit and IPAQ-measured sedentary and walking time of 586 participants. Each symbol represents a participant's first paired activity observation. Spearman's rank correlations (rs) and *P* values are provided for time spent in A. Sedentary and B. Walking. Fitbit measurements were acquired during the same period queried by self-report (IPAQ). Solid red lines are results of locally weighted regressions. Overall activity was measured on different scales (Fitbit, steps/d; IPAQ, MET-h/wk). IPAQ: International Physical Activity Questionnaire; MET: metabolic equivalent of task.

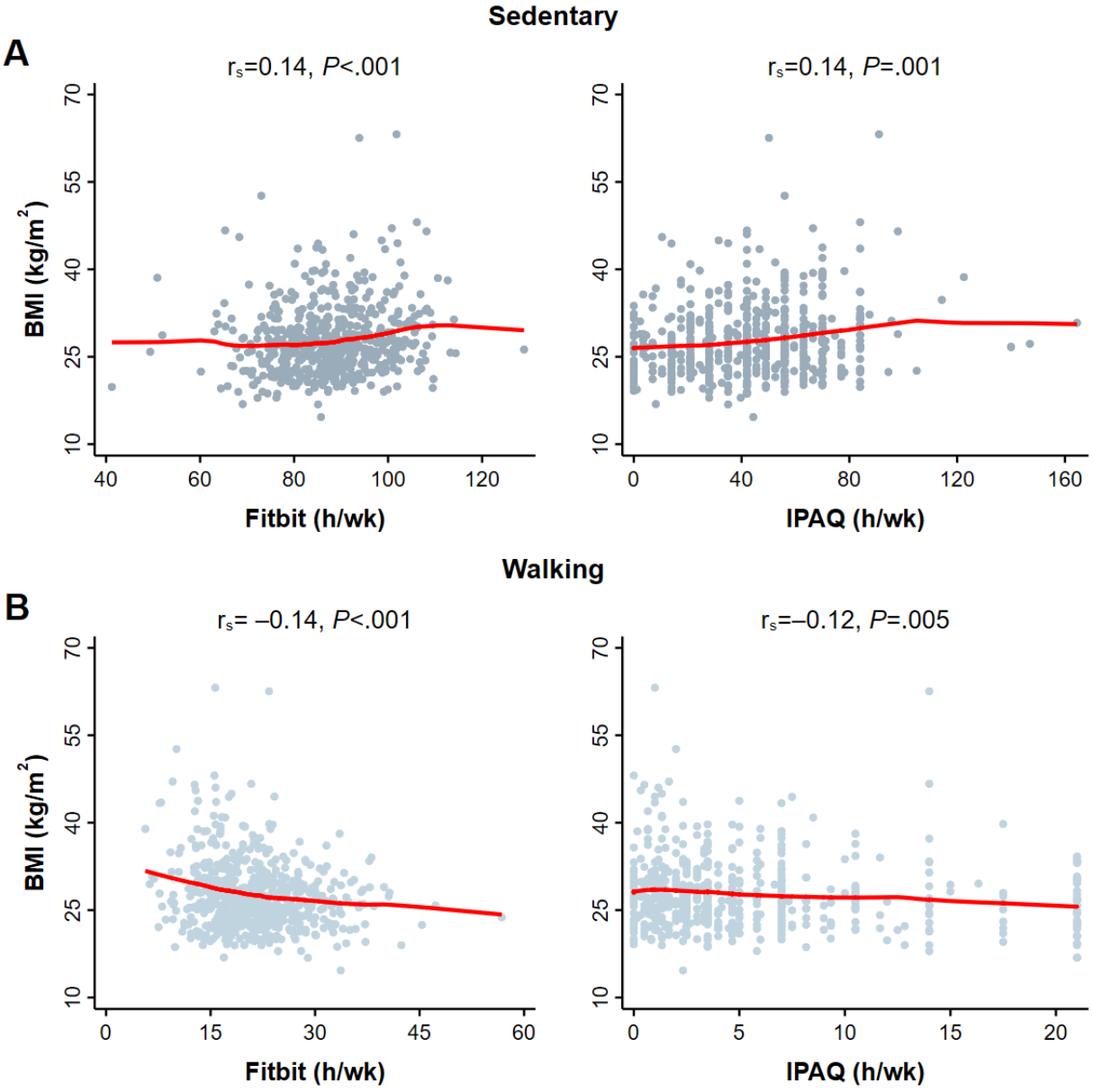


Figure 3. BMI correlations for Fitbit and IPAQ-measured moderate and vigorous activity of 586 participants. Each symbol represents a participant's first paired activity observation. Spearman's rank correlations (rs) and *P* values are provided for A. Moderate activity and B. Vigorous activity. Fitbit measurements were acquired during the same period queried by self-report (IPAQ). Solid red lines are results of locally weighted regressions. Overall activity was measured on different scales (Fitbit, steps/d; IPAQ, MET-h/wk). IPAQ: International Physical Activity Questionnaire; MET: metabolic equivalent of task.

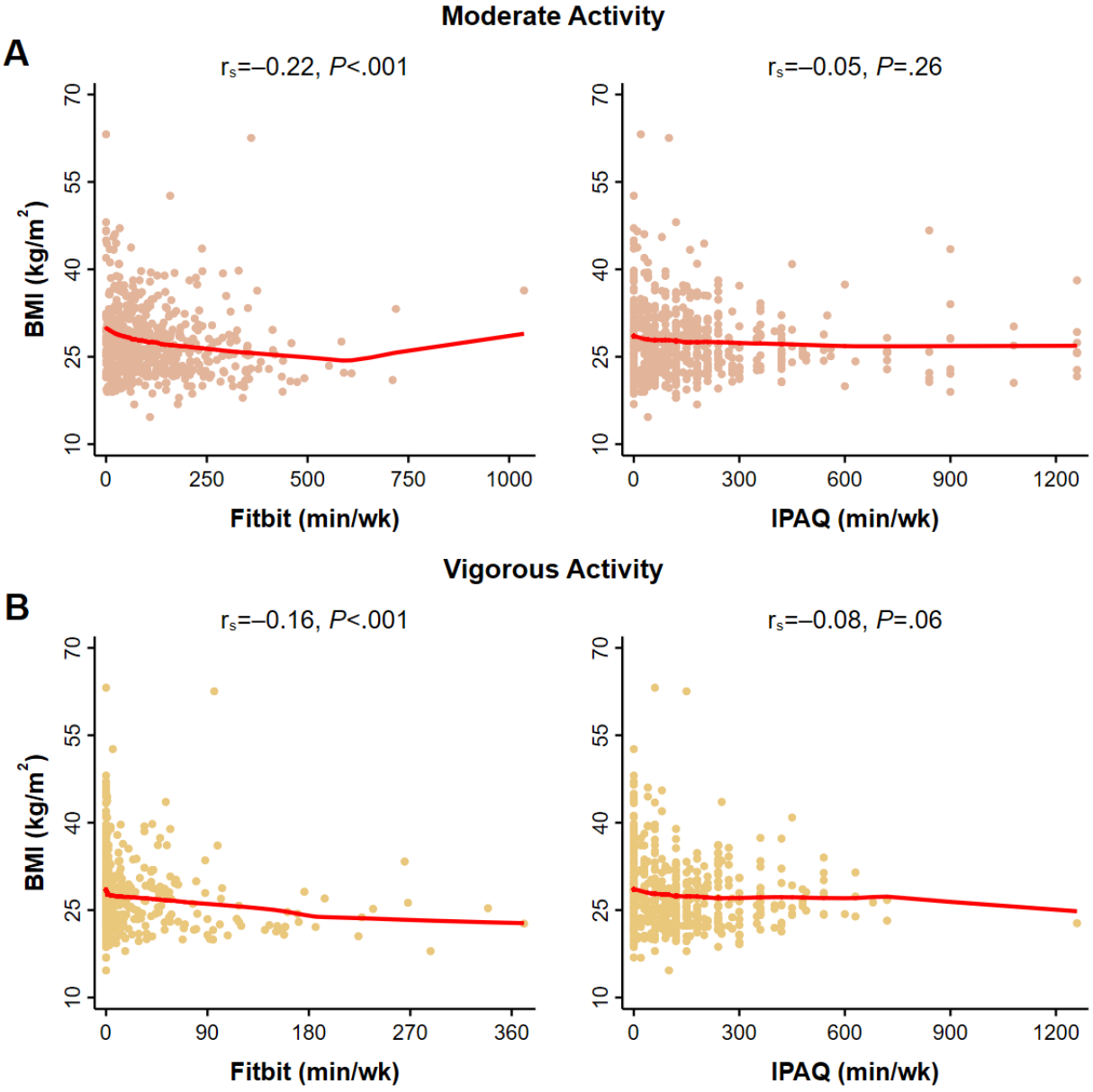
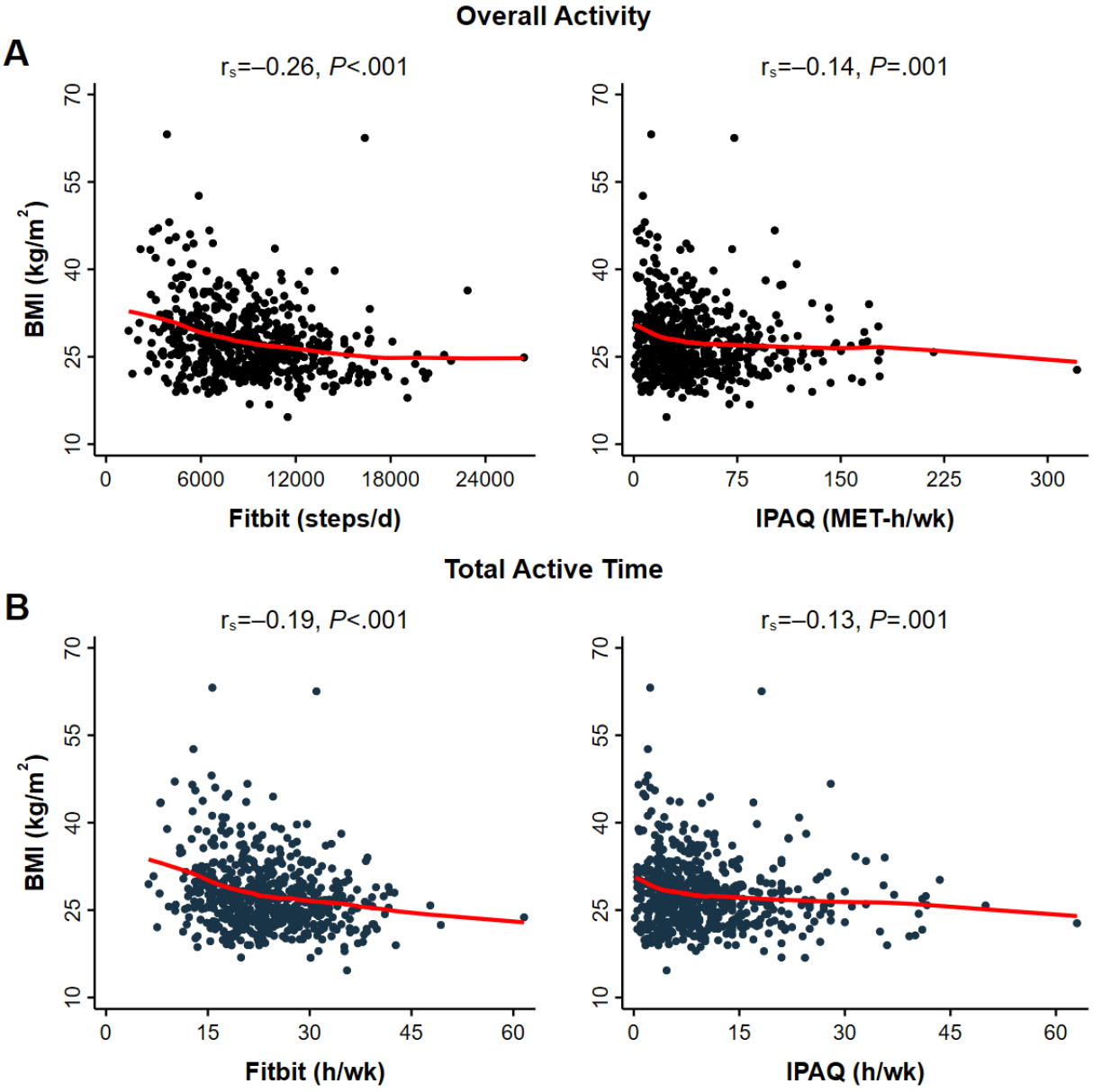


Figure 4. BMI correlations for Fitbit and IPAQ-measured overall activity and total active time of 586 participants. Each symbol represents a participant's first paired activity observation. Spearman's rank correlations (rs) and *P* values are provided for A. Overall activity and B. Active in total (sum of walking, moderate activity, and vigorous activity). Fitbit measurements were acquired during the same period queried by self-report (IPAQ). Solid red lines are results of locally weighted regressions. Overall activity was measured on different scales (Fitbit, steps/d; IPAQ, MET-h/wk). IPAQ: International Physical Activity Questionnaire; MET: metabolic equivalent of task.



After adjustment for demographics, health-related behaviors, and clinical characteristics, an increase of 1 h/wk in Fitbit-measured vigorous activity was associated with a BMI that was 0.84 kg/m^2 lower (95% CI –1.35 to –0.32, *P*<.001) (Table 5). By contrast, an increase of 1 h/wk in self-reported vigorous activity on the IPAQ was associated with a BMI that was 0.17 kg/m^2 lower (95% CI –0.34 to –0.00, *P*=.045), which was significantly smaller than the coefficient for Fitbit-measured vigorous activity (*P*=.01).

Table 5. BMI associations for Fitbit versus International Physical Activity Questionnaire activity by intensity in multivariable logistic regressions[a].

Activity predictor	Unadjusted		Adjusted for demographics		Adjusted for demographics, health-related behaviors, and clinical characteristics	
	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value
Sedentary (h/wk)						
Separate predictor models		.11		.16		.16
Fitbit	0.09[b] (0.02 to 0.15)		0.09[b] (0.03 to 0.15)		0.08[b] (0.02 to 0.14)	
IPAQ[c]	0.04[d] (0.02 to 0.06)		0.05[d] (0.03 to 0.07)		0.04[d] (0.02 to 0.06)	
Combined predictor model		.35		.61		.58
Fitbit	0.07[e] (0.00 to 0.13)		0.06 (–0.00 to 0.12)		0.06 (–0.00 to 0.12)	
IPAQ	0.03[b] (0.01 to 0.06)		0.04[d] (0.02 to 0.06)		0.04[d] (0.02 to 0.06)	
Walking (h/wk)						
Separate predictor models		.30		.35		.30
Fitbit	–0.20[d] (–0.27 to –0.13)		–0.19[d] (–0.27 to –0.12)		–0.18[c] (–0.24 to –0.11)	
IPAQ	–0.15[b] (–0.23 to –0.06)		–0.15[d] (–0.23 to –0.07)		–0.12[b] (–0.20 to –0.04)	
Combined predictor model		.20		.24		.22
Fitbit	–0.18[d] (–0.26 to –0.11)		–0.18[d] (–0.25 to –0.10)		–0.16[d] (–0.23 to –0.09)	
IPAQ	–0.10[e] (–0.19 to –0.01)		–0.10[e] (–0.19 to –0.02)		–0.08[e] (–0.17 to 0.00)	
Moderate activity (h/wk)						
Separate predictor models		.01		.04		.06
Fitbit	–0.50[d] (–0.79 to –0.22)		–0.44[d] (–0.71 to –0.17)		–0.41[b] (–0.68 to –0.15)	
IPAQ	–0.11 (–0.23 to 0.01)		–0.13[e] (–0.25 to –0.01)		–0.13[e] (–0.25 to –0.02)	
Combined predictor model		.02		.06		.08
Fitbit	–0.49[b] (–0.77 to –0.21)		–0.42[b] (–0.70 to –0.15)		–0.40[b] (–0.66 to –0.13)	
IPAQ	–0.08 (–0.20 to 0.03)		–0.11 (–0.23 to 0.01)		–0.11 (–0.23 to 0.00)	
Vigorous activity (h/wk)						
Separate predictor models		<.001		<.001		.01
Fitbit	–1.19[d] (–1.70 to –0.68)		–1.04[d] (–1.52 to –0.56)		–0.84[b] (–1.35 to –0.32)	
IPAQ	–0.21[e] (–0.38 to –0.05)		–0.21[e] (–0.37 to –0.04)		–0.17[e] (–0.34 to –0.00)	
Combined predictor model		<.001		.003		.03
Fitbit	–1.10[c] (–1.61 to –0.59)		–0.95[d] (–1.43 to –0.47)		–0.76[b] (–1.28 to –0.25)	
IPAQ	–0.16 (–0.32 to 0.00)		–0.16 (–0.32 to 0.00)		–0.14 (–0.30 to 0.03)	

[a]All models were adjusted for Fitbit wear-time, data collection season, and Fitbit device wear location. Demographics include age, gender, education, income, race, and ethnicity. Health-related behaviors include smoking and alcohol use. Clinical characteristics include coronary artery disease, diabetes, hyperlipidemia, and hypertension. This table shows the BMI difference (kg/m^2) per predictor increase (95% CI). Regressions accounted for multiple observations with robust standard errors clustered by participant. *P* values shown are for Wald tests comparing Fitbit versus IPAQ coefficients.

[b]These values were significant at $P<.01$ for cluster robust association with BMI (586 participant clusters).

[c]IPAQ: International Physical Activity Questionnaire.

[d]These values were significant at $P<.001$ for cluster robust association with BMI (586 participant clusters).

[e]These values were significant at $P<.05$ for cluster robust association with BMI (586 participant clusters).

In fully adjusted combined models with both analogous Fitbit and IPAQ measurements in the same model (eg, vigorous activity measured by Fitbit and IPAQ included as predictors in the same model), coefficients displayed a similar pattern of magnitude to the fully-adjusted single predictor models (Table 5). However, only Fitbit measurements were significantly associated with BMI in combined head-to-head models for

moderate and vigorous activity, whereas the reverse was the case for sedentary time.

When using standardized units to account for differences in the typical ranges between self-reported and Fitbit-measured activity using fully-adjusted single predictor models, magnitudes of association with BMI were larger for Fitbit versus IPAQ-measured time spent walking, moderately active, or vigorously active, but these did not reach statistical significance (Table S5 in Multimedia Appendix 1).

In separate regressions, an increase in Fitbit-measured overall activity of 1 SD (3663 steps/d) was associated with a BMI that was 1.37 kg/m^2 lower after adjusting for demographic factors, health-related behaviors, and clinical characteristics (β=–1.37, 95% CI –1.84 to –0.90; *P*<.001), whereas 1 SD increase in IPAQ-measured overall activity (37.2 MET-h/wk) was associated with a BMI that was 0.72 kg/m^2 lower (β=–.72, 95% CI –1.14 to –0.29; *P*=.002) (Table 6, complete regression results in Table S3 in Multimedia Appendix 1). In Table 6, overall activity units were standardized due to their different scales (natural unit analyses of overall activity are presented in Table S4 in Multimedia Appendix 1). Findings were very similar when we excluded participants whose data source was the MobileTrack app and when we excluded participants whose data were obtained from an activity tracker worn on the torso.

Table 6. BMI associations for Fitbit versus IPAQ by overall activity and total active time in multivariable logistic regressions[a].

Activity predictor	Unadjusted		Adjusted for demographics		Adjusted for demographics, health-related behaviors, and clinical characteristics	
	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value
Overall activity per 1 SD/wk[b]						
Separate predictor models		.001		.007		.01
Fitbit	–1.65[c] (–2.15 to –1.16)		–1.51[c] (–1.99 to –1.03)		–1.37[c] (–1.84 to –0.90)	
IPAQ[d]	–0.80[e] (–1.23 to –0.35)		–0.82[c] (–1.24 to –0.39)		–0.72[e] (–1.14 to –0.29)	
Combined predictor model		.001		.005		.01
Fitbit	–1.57[c] (–2.09 to –1.05)		–1.40[c] (–1.90 to –0.89)		–1.28[c] (–1.78 to –0.77)	
IPAQ	–0.24 (–0.66 to 0.19)		–0.30 (–0.72 to 0.12)		–0.25 (–0.69 to 0.19)	
Total active time (h/wk)						
Separate predictor models		<.001		<.001		.002
Fitbit	–0.23[c] (–0.30 to –0.16)		–0.22[c] (–0.28 to –0.15)		–0.20[c] (–0.26 to –0.13)	
IPAQ	–0.10[c] (–0.15 to –0.04)		–0.10[c] (–0.15 to –0.05)		–0.09[e] (–0.14 to –0.04)	
Combined predictor model		.001		.002		.006
Fitbit	–0.21[c] (–0.28 to –0.14)		–0.20[c] (–0.27 to –0.13)		–0.18[c] (–0.25 to –0.11)	
IPAQ	–0.04 (–0.10 to 0.01)		–0.05 (–0.10 to 0.01)		–0.04 (–0.09 to 0.01)	

[a]All models were adjusted for Fitbit wear-time, data collection season, and Fitbit device wear location. Demographics include age, gender, education, income, race, and ethnicity. This table shows the BMI difference (kg/m^2) per predictor increase (95% CI). Health-related behaviors include smoking and alcohol use. Clinical characteristics include coronary artery disease, diabetes, hyperlipidemia, and hypertension. Regressions accounted for multiple observations with robust standard errors clustered by participant. *P* values shown are for Wald tests comparing Fitbit versus IPAQ coefficients.

[b]1 SD for overall activity was 3663 steps/d for Fitbit and 37.2 MET-h/wk for International Physical Activity Questionnaire.

[c]These values were significant at *P*<.001 for cluster robust association with BMI (586 participant clusters).

[d]IPAQ: International Physical Activity Questionnaire.

[e]These values were significant at *P*<.01 for cluster robust association with BMI (586 participant clusters).

Discussion

In this study of adults who owned a Fitbit and enrolled in the web-based Health eHeart study, we found that physical activity measured by a wearable activity tracker (Fitbit) was not strongly correlated with self-reported activity but generally proved to be more strongly associated than self-reported activity with a key marker of cardiometabolic health—BMI. Fitbit and self-reported activity measurements were moderately correlated for overall activity, total active time, and nearly so for vigorous activity, but only weakly correlated for sedentary time, walking, and moderate activity. Participants self-reported more vigorous activity, slightly more moderate activity, and less sedentary and walking time than was measured by Fitbit. Associations with BMI were typically stronger for Fitbit compared to self-reported activity from the IPAQ, particularly for higher intensity activities (moderate and vigorous activity) and summary measures (time spent actively and steps/d), both in separate regression models and head-to-head combined models.

Our study found only modest correlations between self-report and device data, a phenomenon that has been previously described [39,40]. The stronger moderate correlation observed for overall activity between Fitbit and IPAQ (r_s=0.48) agrees with previously reported values of correlations between self-report and device data for general physical activity measurements [13].

Our models estimated that Fitbit-measured steps/d, a common measure of overall physical activity derived from wearable activity trackers, was associated with about twice as much difference in BMI as self-reported MET-h/wk, the IPAQ summary measure encapsulating overall physical activity. This is consistent with 3 recent studies using medical-grade accelerometers that found stronger associations with BMI compared to self-report—one in a representative sample of over 4700 adults in the United States [22], one of nearly 500 adults in Malaysia [23], and a third of 317 in Chile that also demonstrated stronger associations with other cardiometabolic risk markers such as fasting glucose and lipid levels [25]. Our twofold larger association also approximates that found in a study of over 80,000 participants in the United Kingdom, in which there was typically a 5-6-year discrepancy between self-report and accelerometer data collection [24]. Our study is the first such analysis to estimate comparative associations with BMI by using activity measurements acquired from consumer wearable activity trackers under free-living conditions that were collected during the same period of time, providing a within-subject control for activities performed. The larger coefficient for the activity tracker measures in this study thus cannot by explained by differences in the actual activity levels between data collection for the wearable devices and self-report. The effect magnitudes estimated in this study are also consistent with studies estimating the activity-BMI relationship separately using objective and self-report methods [12,41].

Across all standardized unit activity measurements from both Fitbit and self-report, Fitbit steps/d was associated with a substantial difference in BMI per SD increase (β=–1.37, *P*<.001). Measurements from Fitbit but not IPAQ remained significantly associated with BMI in regression models that were specified with both for overall activity, total active time, moderate activity, and vigorous activity, thereby suggesting that Fitbit data captures BMI-relevant variability in physical activity better than self-report. Taken together, these findings suggest that Fitbits are advantageous in comparison to self-report primarily because they provide more informative measurements of moderate-to-vigorous physical activity and overall physical activity (steps/d or total active time).

If we assume that these Fitbit associations reflect entirely causal effects of physical activity on BMI (ie, no residual confounding), the impact of Fitbit-measured physical activity can be described in terms that could be useful for clinician-patient interactions and other purposes. For example, 1 SD increase in steps/d measured by Fitbit (3663 steps/d) was associated with a BMI that was 1.37 kg/m^2 lower in this study. Such step increases are reasonably attainable given that prospective interventions with activity trackers have demonstrated typical step/d increases of about 2500 [42,43]. It is worth noting, though, that more modest decreases in BMI are typically observed in prospective observational studies and trials using activity tracker interventions [43,44], and the relationship between BMI and physical activity when using a wearable activity tracker is still unclear. For example, a recent randomized trial providing obese participants with a diet and exercise intervention along with a wearable activity tracker found that the group receiving the device lost significantly less weight over 2 years [45]. However, the techniques to increase physical activity most effectively and otherwise modify behavior to induce healthy weight loss remains unclear [46].

This study has several limitations. Although our participants self-reported physical activity levels that were similar to those mentioned in previous multinational studies (about 36 MET-h/wk) [11], our Fitbit-measured levels of activity (median 8622 steps/d) were higher than previous estimates from nationally representative samples (about 6500 steps/d) [29]. This likely reflects healthy volunteer bias in the Health eHeart Study, but associations with BMI should still be valid given the range of activity levels reflected in the study. The high step counts in this study might also reflect a bias toward increased device adherence when exercising more often than usual, although such systematic overestimation might not affect associations with BMI. Generalizability is limited by the sample's demographic homogeneity and intrinsic self-selection biases. For example, individuals who self-select to use an activity tracker consistently over time such as our participants likely differ in unclear ways from those who abandon the device rapidly or use it infrequently. Participants could receive feedback on activity via their Fitbit, which could influence recall for the IPAQ and bias estimate comparisons. We conducted a large number of analyses and presented *P* values for many of them to illustrate the potential role of chance; it is important for readers to note that we did not correct for multiple comparisons, on the grounds that our hypotheses are separable and not merely equivalent paths to achieving statistical equivalence for a single global hypothesis test. However, it is certainly possible that one or more of the statistically significant findings that we observed occurred by chance. Lastly, we recognize that step-counting algorithms have evolved over the study period, but the predictive advantages of wearables will likely increase over time as algorithms improve and become more accurate.

Objective measurements of moderate and vigorous activity that were less than self-reported values could reflect a wearable activity tracker's inability to detect activities such as biking or user behavior such as removing a device before swimming, but failing to capture these activities could be expected to weaken BMI associations for Fitbit-derived measures only and would not account for the larger BMI associations for Fitbit in this study. Future studies comparing objective and subjective measurements of activity using commercially available accelerometers designed to also capture such activities would help clarify this point.

In summary, our study provides evidence that wearable activity trackers such as Fitbits provide information about physical activity that is likely more clinically meaningful than self-report. The health benefits of physical activity span numerous disease processes responsible for significant morbidity and mortality

[47] such that improving physical activity measurement could significantly impact public health. As the risk attributable to physical inactivity and obesity from noncommunicable diseases such as coronary heart disease continues to grow worldwide [48,49], objective measurements of physical activity from Fitbit and similar wearable activity trackers may play an increasingly important role in improving health through physical activity measurement [50].

In conclusion, in this cross-sectional analysis of Health eHeart Study participants using their Fitbit devices in free-living conditions, step measurements from Fitbits were more strongly associated with BMI than self-reported activity from the same period, particularly for higher intensity activity and summary measures of activity. Fitbit measurements were only moderately or weakly correlated with self-reported physical activity. Wearable activity trackers such as Fitbit likely provide more meaningful data than self-report with respect to weight for clinicians, researchers, and patients.

Acknowledgments

This study was supported by the Patient Centered Outcomes Research Institute, PPRN-1306-04709. The funding organization 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. We thank David Wen and Hannah Gittleman for their support in accessing study data as well as the participants of the Health eHeart Study for their generosity and dedication to advancing our understanding of cardiovascular health and disease. The data sets used and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AJB and MJP conceptualized the study, developed the methodology, performed the analyses, and drafted the original manuscript. All coauthors contributed to the interpretation of results and were involved in manuscript review and editing. All coauthors contributed to the final approval of the manuscript and fulfil the ICMJE criteria for authorship. MJP provided supervision. MJP additionally had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

GMM has received research funding from Jawbone Health.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 674 KB - mhealth_v8i12e22090_app1.docx]

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug 30;384(9945):766-781. [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]
2. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. JAMA 2013 Jan 02;309(1):71-82 [FREE Full text] [doi: 10.1001/jama.2012.113905] [Medline: 23280227]
3. Wilson PWF, D'Agostino RB, Sullivan L, Parise H, Kannel WB. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. Arch Intern Med 2002 Sep 09;162(16):1867-1872. [doi: 10.1001/archinte.162.16.1867] [Medline: 12196085]
4. Barquera S, Pedroza-Tobías A, Medina C, Hernández-Barrera L, Bibbins-Domingo K, Lozano R, et al. Global Overview of the Epidemiology of Atherosclerotic Cardiovascular Disease. Arch Med Res 2015 Jul;46(5):328-338. [doi: 10.1016/j.arcmed.2015.06.006] [Medline: 26135634]
5. Ekelund U, Ward HA, Norat T, Luan J, May AM, Weiderpass E, et al. Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). Am J Clin Nutr 2015 Mar;101(3):613-621 [FREE Full text] [doi: 10.3945/ajcn.114.100065] [Medline: 25733647]
6. Moore SC, Patel AV, Matthews CE, Berrington de Gonzalez A, Park Y, Katki HA, et al. Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. PLoS Med 2012;9(11):e1001335 [FREE Full text] [doi: 10.1371/journal.pmed.1001335] [Medline: 23139642]
7. Physical Activity Guidelines Advisory Committee Report, 2008. Washington, DC, USA: US Department of Health and Human Services; 2008. URL: https://health.gov/sites/default/files/2019-10/CommitteeReport_7.pdf [accessed 2019-08-02]

8. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I, American College of Sports Medicine. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011 Jul;43(7):1334-1359. [doi: 10.1249/MSS.0b013e318213fefb] [Medline: 21694556]
9. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2019 Sep 10;74(10):1376-1414 [FREE Full text] [doi: 10.1016/j.jacc.2019.03.009] [Medline: 30894319]
10. Kelly P, Fitzsimons C, Baker G. Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. Int J Behav Nutr Phys Act 2016 Mar 01;13:32 [FREE Full text] [doi: 10.1186/s12966-016-0351-4] [Medline: 26931142]
11. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
12. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006 Sep;9(6):755-762. [doi: 10.1079/phn2005898] [Medline: 16925881]
13. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8:115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
14. Skender S, Ose J, Chang-Claude J, Paskow M, Brühmann B, Siegel EM, et al. Accelerometry and physical activity questionnaires - a systematic review. BMC Public Health 2016 Jun 16;16:515 [FREE Full text] [doi: 10.1186/s12889-016-3172-0] [Medline: 27306667]
15. Sabia S, van Hees VT, Shipley MJ, Trenell MI, Hagger-Johnson G, Elbaz A, et al. Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors. Am J Epidemiol 2014 Mar 15;179(6):781-790 [FREE Full text] [doi: 10.1093/aje/kwt330] [Medline: 24500862]
16. Farina N, Lowry RG. The Validity of Consumer-Level Activity Monitors in Healthy Older Adults in Free-Living Conditions. J Aging Phys Act 2018 Jan 01;26(1):128-135. [doi: 10.1123/japa.2016-0344] [Medline: 28595019]
17. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
18. Chu AHY, Ng SHX, Paknezhad M, Gauterin A, Koh D, Brown MS, et al. Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults. PLoS One 2017;12(2):e0172535 [FREE Full text] [doi: 10.1371/journal.pone.0172535] [Medline: 28234953]
19. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fibit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Res Notes 2014 Dec 23;7:952 [FREE Full text] [doi: 10.1186/1756-0500-7-952] [Medline: 25539733]
20. Reid RE, Insogna JA, Carver TE, Comptour AM, Bewski NA, Sciortino C, et al. Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. J Sci Med Sport 2017 Jun;20(6):578-582. [doi: 10.1016/j.jsams.2016.10.015] [Medline: 27887786]
21. Bai Y, Welk GJ, Nam YH, Lee JA, Lee J, Kim Y, et al. Comparison of Consumer and Research Monitors under Semistructured Settings. Med Sci Sports Exerc 2016 Jan;48(1):151-158. [doi: 10.1249/MSS.0000000000000727] [Medline: 26154336]
22. Wanner M, Richard A, Martin B, Faeh D, Rohrmann S. Associations between self-reported and objectively measured physical activity, sedentary behavior and overweight/obesity in NHANES 2003-2006. Int J Obes (Lond) 2017 Jan;41(1):186-193. [doi: 10.1038/ijo.2016.168] [Medline: 27677618]
23. Lee YY, Kamarudin KS, Wan Muda WAM. Associations between self-reported and objectively measured physical activity and overweight/obesity among adults in Kota Bharu and Penang, Malaysia. BMC Public Health 2019 May 22;19(1):621 [FREE Full text] [doi: 10.1186/s12889-019-6971-2] [Medline: 31118019]
24. Guo W, Key TJ, Reeves GK. Accelerometer compared with questionnaire measures of physical activity in relation to body size and composition: a large cross-sectional analysis of UK Biobank. BMJ Open 2019 Jan 29;9(1):e024206 [FREE Full text] [doi: 10.1136/bmjopen-2018-024206] [Medline: 30700478]
25. Celis-Morales CA, Perez-Bravo F, Ibañez L, Salas C, Bailey MES, Gill JMR. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. PLoS One 2012;7(5):e36345 [FREE Full text] [doi: 10.1371/journal.pone.0036345] [Medline: 22590532]
26. Health eHeart Study. URL: http://www.health-eheartstudy.org [accessed 2020-01-01]
27. Hutto B, Howard VJ, Blair SN, Colabianchi N, Vena JE, Rhodes D, et al. Identifying accelerometer nonwear and wear time in older adults. Int J Behav Nutr Phys Act 2013 Oct 25;10:120 [FREE Full text] [doi: 10.1186/1479-5868-10-120] [Medline: 24156309]
28. Beddhu S, Wei G, Marcus RL, Chonchol M, Greene T. Light-Intensity Physical Activities and Mortality in the United States General Population and CKD Subpopulation. CJASN 2015 Apr 30;10(7):1145-1153. [doi: 10.2215/cjn.08410814]

29. Tudor-Locke C, Johnson WD, Katzmarzyk PT. Accelerometer-determined steps per day in US adults. Med Sci Sports Exerc 2009 Jul;41(7):1384-1391. [doi: 10.1249/MSS.0b013e318199885c] [Medline: 19516163]
30. Matthews CE, Ainsworth BE, Thompson RW, Bassett DR. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc 2002 Aug;34(8):1376-1381. [doi: 10.1097/00005768-200208000-00021] [Medline: 12165695]
31. Tudor-Locke C, Camhi SM, Leonardi C, Johnson WD, Katzmarzyk PT, Earnest CP, et al. Patterns of adult stepping cadence in the 2005-2006 NHANES. Prev Med 2011 Sep;53(3):178-181. [doi: 10.1016/j.ypmed.2011.06.004] [Medline: 21708187]
32. Diaz KM, Howard VJ, Hutto B, Colabianchi N, Vena JE, Blair SN, et al. Patterns of Sedentary Behavior in US Middle-Age and Older Adults: The REGARDS Study. Med Sci Sports Exerc 2016 Mar;48(3):430-438 [FREE Full text] [doi: 10.1249/MSS.0000000000000792] [Medline: 26460633]
33. Tudor-Locke C, Aguiar EJ, Han H, Ducharme SW, Schuna JM, Barreira TV, et al. Walking cadence (steps/min) and intensity in 21-40 year olds: CADENCE-adults. Int J Behav Nutr Phys Act 2019 Jan 17;16(1):8 [FREE Full text] [doi: 10.1186/s12966-019-0769-6] [Medline: 30654810]
34. Tudor-Locke C, Han H, Aguiar EJ, Barreira TV, Schuna JM, Kang M, et al. How fast is fast enough? Walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. Br J Sports Med 2018 Jun;52(12):776-788 [FREE Full text] [doi: 10.1136/bjsports-2017-097628] [Medline: 29858465]
35. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De Cocker K, Giles-Corti B, et al. How many steps/day are enough? For adults. Int J Behav Nutr Phys Act 2011 Jul 28;8:79 [FREE Full text] [doi: 10.1186/1479-5868-8-79] [Medline: 21798015]
36. Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995 Feb 01;273(5):402-407. [doi: 10.1001/jama.273.5.402] [Medline: 7823386]
37. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, American College of Sports Medicine, American Heart Association. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Circulation 2007 Aug 28;116(9):1094-1105. [doi: 10.1161/CIRCULATIONAHA.107.185650] [Medline: 17671236]
38. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
39. Brewer W, Swanson BT, Ortiz A. Validity of Fitbit's active minutes as compared with a research-grade accelerometer and self-reported measures. BMJ Open Sport Exerc Med 2017;3(1):e000254 [FREE Full text] [doi: 10.1136/bmjsem-2017-000254] [Medline: 29018543]
40. Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exerc 2014 Jan;46(1):99-106. [doi: 10.1249/MSS.0b013e3182a0595f] [Medline: 23793232]
41. Tudor-Locke C, Schuna JM, Han HO, Aguiar EJ, Green MA, Busa MA, et al. Step-Based Physical Activity Metrics and Cardiometabolic Risk: NHANES 2005-2006. Med Sci Sports Exerc 2017 Feb;49(2):283-291 [FREE Full text] [doi: 10.1249/MSS.0000000000001100] [Medline: 27669450]
42. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step Counting: A Review of Measurement Considerations and Health-Related Applications. Sports Med 2017 Jul;47(7):1303-1315 [FREE Full text] [doi: 10.1007/s40279-016-0663-1] [Medline: 28005190]
43. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
44. Richardson CR, Newton TL, Abraham JJ, Sen A, Jimbo M, Swartz AM. A meta-analysis of pedometer-based walking interventions and weight loss. Ann Fam Med 2008;6(1):69-77 [FREE Full text] [doi: 10.1370/afm.761] [Medline: 18195317]
45. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. JAMA 2016 Sep 20;316(11):1161-1171 [FREE Full text] [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
46. Bray GA, Heisel WE, Afshin A, Jensen MD, Dietz WH, Long M, et al. The Science of Obesity Management: An Endocrine Society Scientific Statement. Endocr Rev 2018 Apr 01;39(2):79-132 [FREE Full text] [doi: 10.1210/er.2017-00253] [Medline: 29518206]
47. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
48. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
49. GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, et al. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. N Engl J Med 2017 Jul 06;377(1):13-27 [FREE Full text] [doi: 10.1056/NEJMoa1614362] [Medline: 28604169]

50. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 03;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]

Abbreviations

IPAQ: International Physical Activity Questionnaire
MET: metabolic equivalent of task

Edited by G Eysenbach; submitted 03.07.20; peer-reviewed by C Friel, T Fazzino; comments to author 26.08.20; revised version received 20.10.20; accepted 12.11.20; published 29.12.20.

Please cite as:
Beagle AJ, Tison GH, Aschbacher K, Olgin JE, Marcus GM, Pletcher MJ
Comparison of the Physical Activity Measured by a Consumer Wearable Activity Tracker and That Measured by Self-Report: Cross-Sectional Analysis of the Health eHeart Study
JMIR Mhealth Uhealth 2020;8(12):e22090
URL: http://mhealth.jmir.org/2020/12/e22090/
doi:10.2196/22090
PMID:33372896

Original Paper

Augmented Reality for Smoking Cessation: Development and Usability Study

Christine Vinci[1,2*], PhD; Karen O Brandon[1,3*], PhD; Marloes Kleinjan[4,5], PhD; Laura M Hernandez[1], BA; Leslie E Sawyer[1,3], BS; Jody Haneke[6], BGD; Steven K Sutton[2,3,7], PhD; Thomas H Brandon[1,2,3], PhD

[1]Department of Health Outcomes and Behavior, Moffitt Cancer Center, Tampa, FL, United States

[2]Department of Oncologic Sciences, University of South Florida, Tampa, FL, United States

[3]Department of Psychology, University of South Florida, Tampa, FL, United States

[4]Trimbos Institute, Utrecht, Netherlands

[5]Department of Interdisciplinary Social Science, Utrecht University, Utrecht, Netherlands

[6]Haneke Design, Tampa, FL, United States

[7]Department of Biostatistics and Bioinformatics, Moffitt Cancer Center, Tampa, FL, United States

[*]these authors contributed equally

Corresponding Author:
Christine Vinci, PhD
Department of Health Outcomes and Behavior
Moffitt Cancer Center
Tobacco Research and Intervention Program
4115 East Fowler Avenue
Tampa, FL, 33617
United States
Phone: 1 1 813 745 5421
Email: Christine.Vinci@moffitt.org

Abstract

Background: The recent widespread availability of augmented reality via smartphone offers an opportunity to translate cue exposure therapy for smoking cessation from the laboratory to the real world. Despite significant reductions in the smoking rates in the last decade, approximately 13.7% of the adults in the United States continue to smoke. Smoking-related cue exposure has demonstrated promise as an adjuvant therapy in the laboratory, but practical limitations have prevented its success in the real world. Augmented reality technology presents an innovative approach to overcome these limitations.

Objective: The aim of this study was to develop a smartphone app that presents smoking-related augmented reality images for cue exposure. Smokers provided feedback on the images and reported on the perceived urge to smoke, qualities of reality/coexistence, and general feedback about quality and functioning. The feedback was used to refine the augmented reality images within the app.

Methods: In collaboration with an augmented reality design company, we developed 6 smoking-related images (cigarette, lighter, ashtray, lit cigarette in ashtray, etc) and 6 neutral images similar in size or complexity for comparison (pen, eraser, notebook, soda bottle with droplets, etc). Ten smokers completed a survey of demographic characteristics, smoking history and behavior, dependence on nicotine, motivation to quit smoking, and familiarity with augmented reality technology. Then, participants viewed each augmented reality image and provided ratings on 10-point Likert scales for urge to smoke and reality/coexistence of the image into the scene. Participants were also queried with open-ended questions regarding the features of the images.

Results: Of the 10 participants, 5 (50%) had experienced augmented reality prior to the laboratory visit, but only 4 of those 5 participants used augmented reality at least weekly. Although the sample was small (N=10), smokers reported significantly higher urge to smoke after viewing the smoking-related augmented reality images (median 4.58, SD 3.49) versus the neutral images (median 1.42, SD 3.01) (Z=–2.14, P=.03; d=0.70). The average reality and coexistence ratings of the images did not differ between smoking-related and neutral images (all P>.29). Augmented reality images were found on average to be realistic (mean [SD] score 6.49 [3.11]) and have good environmental coexistence (mean [SD] score 6.93 [3.04]) and user coexistence (mean [SD] score 6.38 [3.27]) on the 10-point scale. Participant interviews revealed some areas of excellence (eg, details of the lit cigarette) and areas for improvement (eg, stability of images, lighting).

Conclusions: All images were generally perceived as being realistic and well-integrated into the environment. However, the smoking augmented reality images produced higher urge to smoke than the neutral augmented reality images. In total, our findings support the potential utility of augmented reality for cue exposure therapy. Future directions and next steps are discussed.

(JMIR Mhealth Uhealth 2020;8(12):e21643) doi:10.2196/21643

KEYWORDS

augmented reality; smoking cessation; cue exposure therapy; cue reactivity; behavior change; smoking; smartphone app; mobile phone

Introduction

Augmented reality is an emerging technology that could provide a novel and exciting treatment strategy for substance use disorders [1]. Augmented reality superimposes virtual, digital objects into the real world environment, often via a smartphone or a tablet [2]. The advantages of augmented reality as a treatment modality include its scalability and application in the real world environment (vs treatment in a clinic). To date, augmented reality has been used primarily in gaming and retail phone apps as well as for training purposes in the medical field [3,4]. With respect to behavioral treatments, augmented reality has mainly been tested for the treatment of specific phobias via cue exposure, with promising outcomes [5]. We propose that augmented reality might be particularly well suited for the treatment of substance use disorders, and we begin by focusing on tobacco dependence.

Cigarette smoking is the leading preventable cause of morbidity and mortality in the United States [6]. Although tobacco use has declined over the past 50 years, smoking prevalence was reported to be 13.7% among adults in the United States in 2018 [7]. As such, developing more effective cessation interventions remains a public health priority.

There is a specific need for cessation interventions that reduce barriers to dissemination and implementation, and mobile health holds particular promise in this regard [8-11]. Moreover, harnessing recent advances in technology may increase the uptake of cessation treatment, while also targeting identified roadblocks to cessation. For example, in 2019, 81% of Americans reported owning a smartphone, which was up from 35% from only 7 years earlier, with 70% ownership even among low-income populations [12]. Further, among all smartphone owners in 2018, almost half reported using an app to monitor their health [13]. Thus, novel interventions that leverage advances in technology may be an ideal way to increase the reach and efficacy of smoking cessation treatments.

Through the display of virtual smoking cues superimposed upon smokers' natural environment, augmented reality appears ideal for improving the long-term effects of cue exposure. Cue exposure therapy is based on the principles of Pavlovian conditioning [14,15], which when applied to drug behavior consists of the repeated presentation of smoking-related cues (ie, conditioned stimuli) without access to nicotine (ie, the unconditioned stimulus), resulting in the extinction of craving (ie, the conditioned response). Thus, cue exposure therapy typically consists of presenting a drug cue (eg, cigarette) to individuals multiple times while not allowing them to engage in the typical drug behavior (in this case, smoking), with the goal of extinguishing the urge to smoke in response to these cues. The ultimate clinical goal is to reduce the risk of smoking relapse when individuals encounter similar cues in their environment [16,17]. Cue exposure treatments have demonstrated efficacy in decreasing craving in the laboratory, but these effects are often short-lived as they do not generalize well beyond the extinction setting (eg, laboratory or clinic) into the real world [18,19]. Augmented reality allows the entire cue exposure session to take place in the real world, which could greatly enhance the efficacy of cue exposure treatment. Additionally, the COVID-19 pandemic has elevated the utility of mobile health interventions that move treatment from a clinic to a patient's own environment.

To our knowledge, there have been no previous studies examining the potential of augmented reality for cue exposure treatment for substance use disorders. Thus, this paper presents the initial phase of designing and testing the usability of augmented reality stimuli in a small sample of smokers. The primary aim was to receive feedback on the stimuli and modify as needed.

Methods

Study Participants

Ten participants were recruited in October and November of 2019. Inclusion criteria were as follows: (1) ≥18 years of age; (2) currently smoking ≥3 cigarettes per day during the past year; (3) having a breath carbon monoxide (CO) level ≥5 ppm to verify smoking status; (4) motivated to quit smoking; (5) having a valid home address in the local area; (6) having a functioning telephone number; and (7) the ability to speak, read, and write in English. Exclusion criteria were regular use (eg, on more than one-third of tobacco use occasions) of other tobacco products or a household member already enrolled in the study.

Procedures

Smartphone App and Stimuli Development

A smartphone app was developed as a platform for presenting cue-reactivity stimuli as augmented reality images. The stimuli were consistent with the existing cue-reactivity literature and commonly used cues within the field [20-24]. Proximal smoking cues (eg, cigarettes, lighters, ashtrays) were developed, and for comparison purposes, we selected neutral cues that were similar in size and complexity to the smoking cues whenever feasible (eg, pencil, eraser, notebook). Single and compound cues, in conjunction with one motion cue, were used for smoking-related and neutral images. Neutral items represented the general

category of common school or office supplies, apart from a moving image (a soda bottle with condensation). Table 1 lists the type of cue (single, compound, motion item) for each category of cue (smoking-related or neutral) that was used.

The augmented reality design company, Haneke Design, created 12 initial images by using the above guidelines (Table 1). Some images (assets) were provided by the Unity [25] platform software, while others were purchased (ie, animated cigarette smoke). Image rendering was conducted to create images not available in the Unity library, including a pack of cigarettes and a soda bottle. In such instances, rendering was done using extrusion modeling and an additional app, Blender v.2.80 [26], which is a free and open-source 3D computer graphics software toolset for visual effects, motion graphics, and interactive 3D apps. The developer also used photos of a plate with a cigarette to render an ashtray since the asset was not available in the Unity library. Other assets, including the soda bottle and cigarette pack, were also captured using digital photography and by using extrusion modeling and the Blender software app to provide realistic augmented reality images. Initial steps by the research team included contacting the developers to discuss features, images, and overall design of the app. After all the features and images were determined, the company developed and deployed the initial app to the study team for review and staff testing. The research team tested the images and provided feedback regarding item features (color, orientation, size, texture, etc) through an iterative process until the images were deemed satisfactory. Next, the developers incorporated the images in a smartphone compatible app created to deliver the images to the participants. For example, features such as participant identification number, timed presentation of the augmented reality images, user-friendly urge and reality/coexistence ratings, and data (ID, order of viewing, ratings) storage and export capabilities were included. Figure 1 and Figure 2 are pictures of several of the final images. The video of an image placement and rating can be found in Multimedia Appendix 1.

Table 1. Types of cues for smoking and neutral augmented reality images.

Type of cue, subtype	Smoking images	Neutral images
Single cues		
Single cue 1	Cigarette	Pen
Single cue 2	Pack of cigarettes	Notepad
Compound cues		
Compound cue 1	Pack of cigarettes with ashtray	Pencil and eraser
Compound cue 2	Cigarette and lighter	Pencil with notepad
Compound cue 3	Pack of cigarettes with lighter	Sticky notes and pen
Motion cue	Ashtray with lit cigarette emitting smoke	Soda bottle with condensation droplets and effervescence

Figure 1. A subset of digitally created augmented reality smoking-related and neutral cues.

Figure 2. Digitally created augmented reality smoking image superimposed on a user's table in real time. This image includes the motion feature of smoke rising from the cigarette tip.

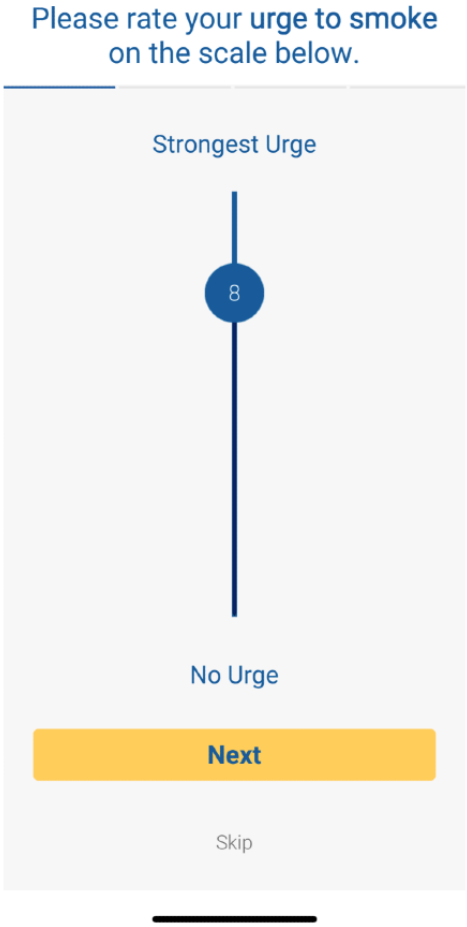
Participant Recruitment and In-Person Session

Participants were recruited through referrals from previous studies within our laboratory. Individuals were contacted by phone and they completed a brief phone screen for eligibility assessment. Eligible participants were invited to attend an in-person session, where the informed consent process was completed. After participants provided consent, a breath CO sample was collected to confirm smoking status. Participants completed a brief survey of demographics, smoking history, self-reported nicotine dependence, motivation to quit smoking, and familiarity with augmented reality. Participants were instructed on app use through an experimenter-guided user manual. The manual detailed how to position the images on the table and showed examples of sequential screenshots. Participants were provided with an Apple 10XR iPhone with the augmented reality app to view each of the 12 images. Each participant was assigned a random order for viewing the images.

After the experimenter selected the first image to be viewed, the app guided the participant on image placement upon a table that was empty apart from a tissue box placed to provide context for the augmented reality image. Participants were instructed to place an augmented reality–generated blue arrow (Figure 3) anywhere on the table. Once the blue arrow was placed where they liked, participants tapped on the blue arrow, which allowed the selected image to appear in the center of the arrow. Once the participants were satisfied with the placement, the image was locked into place by pressing the "Start" button that appeared at the bottom of the screen. The augmented reality image was visible for 1 minute, during which the participant was free to view the augmented reality image from multiple directions and distances. After 1 minute, the image disappeared from the screen and the app displayed Likert scale questions to obtain ratings on smoking urge and image reality/coexistence (Figure 4). Next, the experimenter asked open-ended questions to obtain general feedback on the image (discussed in further detail below). Once completed, the next augmented reality image was selected for viewing, and the process was repeated for each of the 12 augmented reality images. Sessions lasted approximately 1 hour, and participants were compensated US $35 for their time. All procedures were approved by the Advarra Institutional Review Board.

Figure 3. Screenshot of the digitally created blue placement circle.

Figure 4. Screenshot of a smoking urge rating after viewing the augmented reality image.



Measures

Baseline Variables

Self-report items assessed demographic variables (age, race, ethnicity, sex, and education with 8 categories from no high school to professional degree, and household income). Smoking and other tobacco use in the past month was also assessed (eg, “In the past month, how often have you smoked cigarettes on average?” with response categories of “no use in the past month,” “1-5 times per month,” “1-3 days a week,” “4-6 days a week,” and “7 days a week”). The number of previous cessation attempts (“In the past month, how many cigarettes per day have you smoked on average?”), motivation to quit smoking, and nicotine dependence were also assessed. Motivation to quit was assessed by 2 Yes/No items asking about plans to quit smoking in the next 3 and 6 months, 1 Likert scale item assessing confidence that they could stay off cigarettes for good if they tried to quit (1=“Not at all” to 5=“Very”), 1 Likert item rating how strongly they agree with the statement, “I am committed to quitting smoking” (1=“Not at all” to 5=“Very”), and the Contemplation Ladder, [27] a 10-point scale (1=“No thoughts of quitting smoking cigarettes” to 10=“Taking action to quit smoking cigarettes”). Tobacco dependence was measured by the 6-item Fagerström Test for Nicotine Dependence scale (FTND) [28]. The FTND items assess the number of cigarettes smoked per day, time to first smoked cigarette after waking, difficulty not smoking in places where smoking is prohibited, most difficult cigarette to give up (first one in the morning or any other), smoking more within the first 2 hours of the day, and smoking even when ill. We also assessed participants’ familiarity and past use of augmented reality by asking “Have you ever used any kind of augmented reality (AR) app before? (eg, augmented reality feature in Pokémon Go, Ikea or other furniture app, Snapchat filters, Google Sky Map, etc)"; the answer options were Yes/No/Don’t know and by asking “How frequently do you use augmented reality (AR) apps?” (ie, answer options were 7 days a week/4-6 days a week/1-3 days a week/1-5 times a month/less than once a month/Never/Don’t know).

Ratings of Augmented Reality Images

Urge (craving) to smoke was rated on a Likert scale (1=absolutely no urge to smoke to 10=strongest urge to smoke) (Figure 4). Three items assessed the quality of the augmented reality experience [29] using 10-point Likert scales: (1) a reality item captured the perceived realism of the augmented reality object, “How real did the object seem to you?” from 1=Not all to 10=Very Real; (2) an environment coexistence item assessed the integration of the object into the surrounding environment, as seen on the smartphone screen, “How well did the object appear to be part of the scene?” from 1=Not at all to 10=Very Well; and (3) user coexistence captured the degree to which the user felt in the presence of the object, “How much did you feel the object was right there in front of you?” from 1=Not at all to 10=Very Much.

Additional Feedback on Images

After viewing each image, participants were asked for their open-ended feedback with the following questions: "What were your initial thoughts about the item?", "Is there anything else that seems wrong, off, or not quite right about the item?", "Do you have any other feedback that you think would improve the item?" Additional specific questions were asked about the 2 images with motion aspects (ie, lit cigarette in ashtray and soda bottle with droplets). Finally, we asked participants if they could recommend additional images that might elicit urges to smoke.

Statistical Analysis

Although this development study was not powered for inferential statistics, we nevertheless compared the median urge scores of smoking versus neutral augmented reality images by using the Wilcoxon signed-rank test after calculating each participant's mean rating across smoking and neutral images. Cue reactivity would be indicated by higher urges elicited by the smoking images compared to the neutral images. Evidence of such cue reactivity is a necessary condition for subsequent cue exposure therapy designed to reduce conditioned urge. We also compared smoking and neutral images with respect to realism/coexistence.

Results

Sample Characteristics

Thirty-nine people were screened for eligibility. Seventeen individuals met the eligibility criteria, 15 were scheduled for a session, and 11 consented. One participant who gave consent did not meet the CO eligibility and thus did not participate in the session; therefore, only 10 participants were included in the analyses. The descriptive characteristics of the participants are summarized in Table 2. Of the 10 participants, 6 (60%) were males and the majority was white, non-Hispanic; 8 (80%) of them had an annual household income of less than US $40,000. All participants were daily smokers and smoked an average of 18 cigarettes per day. Participants' nicotine dependence as indexed by the FTND demonstrated a full range of dependence levels in our sample. The average participant motivation to quit was moderately high, with all but 1 participant planning a quit attempt in the next 6 months.

Table 2. Statistical summary of the descriptive characteristics of the participants (N=10).

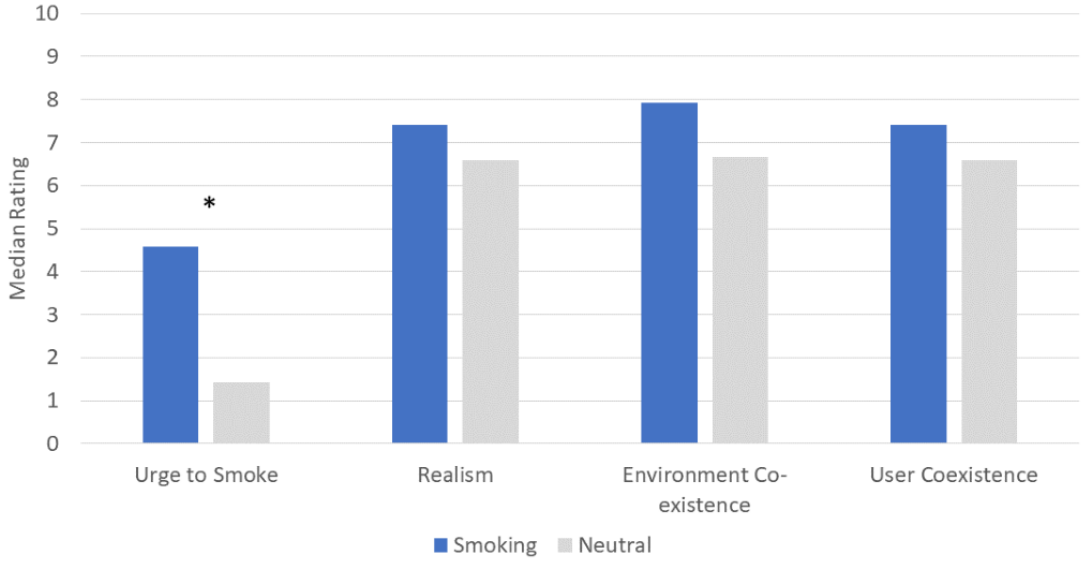
Descriptive characteristics, variable, subvariable	Values
Demographics	
Age (years), mean (SD)	50.70 (9.84)
Male, n (%)	6 (60)
Latino, n (%)	3 (30)
Income <US $40,000, n (%)	8 (80)
Education < High school, n (%)	3 (30)
Race, n (%)	
White	8 (80)
Black	2 (20)
Smoking characteristics	
Cigarettes per day, mean (SD)	18.30 (6.57)
FTND[a] total, mean (SD)	4.80 (2.39)
Nicotine dependence, n (%)	
Low dependence	1 (10)
Low-to-moderate dependence	4 (40)
Moderate dependence	3 (30)
High dependence	2 (20)
Other household members smoke, n (%)	3 (30)
Plans to quit smoking, n (%)	
Within 6 months	9 (90)
Within 3 months	7 (70)
Quitting motivation, mean (SD) scores	
Confidence in quitting (scale 1-5)	2.90 (1.37)
Commitment to quitting smoking (scale 1-5)	3.70 (1.06)
Contemplation ladder (scale 1-10)	6.10 (2.38)
Augmented reality ever used, n (%)	5 (50)
Frequency of use of those who had used augmented reality, n (%)	
4-6 days/week	1 (20)
1-5 days/week	1 (20)
Less than once/month	3 (60)

[a]FTND: Fagerström Test for Nicotine Dependence.

Ratings of Augmented Reality Images

Figure 5 shows the median ratings of urge to smoke and reality/coexistence across the 6 smoking-related images when compared to the median ratings across the 6 neutral images. As expected, participants reported higher ratings of smoking urge in response to the smoking stimuli (median 4.58) compared to neutral stimuli (median 1.42) (Z=–2.14, *P*=.03), with a moderate effect size (Cohen *d*=0.70). No significant differences were found in the average ratings of reality/coexistence between the categories of augmented reality images (*P*>.29). The average ratings of augmented reality images indicated that these images were realistic (mean [SD] score 6.49 [3.11]) and have both good environmental coexistence (mean [SD] score 6.93 [3.04]) and user coexistence (mean [SD] score 6.38 [3.27]). Multimedia Appendix 2 and Multimedia Appendix 3 show the mean ratings of each individual smoking-related and neutral cue (respectively) urge to smoke and reality/coexistence.

Figure 5. Median ratings of urge to smoke and reality/coexistence measures across image categories. *P=.03.



Participant Usability Feedback

Across the 12 items and 10 participants (120 reports), there were 18 mentions of the image moving, rolling, or floating above or off the table, 16 reports of the image items being either too big or too small, 16 comments regarding shadow/lighting oddities, 20 mentions of color inconsistencies or dislikes, 8 reports of pixilation problems, 16 mentions that the image or details of the image were not the same as their brand (smoking-related images only), and 21 comments regarding details that were missing (eg, no. 2 on the pencil, ashtray too dirty or too clean, lack of brand name on the cigarette filter). Although we were not focused on capturing the positive aspects of the stimuli per se, many participants offered comments referencing item realism (eg, “seemed realistic”), level of detail (eg, “pretty good graphics,” “really good, seemed detailed”), and urge response (eg, “made me want to have a cigarette”).

Regarding the 2 images with movement, all participants thought the movement element (smoke) of the lit cigarette looked realistic and moved in a realistic way. All but 2 participants thought the color and density of the smoke were realistic. One participant thought the smoke was too dark at the top, while the other thought it was too light. For the soda bottle with condensation dripping and effervescence, 7 of the 10 participants thought the movement elements looked realistic, while 3 did not really notice the movement. Among the participants who noticed the condensation movement, 6 said the droplets were perfect in size and transparency.

Participants’ recommendations for additional smoking-related images to include in the future were largely secondary or distal cues, that is, they suggested items that often co-occur with smoking but that are not necessary for smoking itself (coffee, alcohol, a joint, food, phones, car, crowds, TV, and other smokers).

Discussion

Principal Results

Twelve smoking and nonsmoking augmented reality images were created using the Unity platform, Blender, and a combination of digital photography and extrusion modeling. Average participant ratings were all above 6 on a 10-point scale, suggesting that the images were reasonably realistic and coexisted in the overall environment well. Although this development and usability study was not powered for inferential comparisons, the urge to smoke following the visualization of smoking-related images was rated significantly higher than that following the visualization of non–smoking-related images, with a robust effect size. This finding is consistent with the existing literature on cue reactivity using in vivo items [30], pictures [21], and virtual reality cues [31].

In the open-ended usability interviews, the participants found the images to be generally realistic and detailed. However, there were numerous issues with image stability, lighting, shadows, and sizing. A number of participants reported that image details were either missing or different from their brand, which may be a consequence of our design choice for items to not look like any specific brand but rather be representative of the class of object (to be generalizable to a range of individuals). Based on participant feedback and experimenter observation, we developed a set of “best practice” recommendations to maximize the quality of the augmented reality experience (Table 3).

Table 3. Best practices for placement and viewing of augmented reality images.

Problem	Description	Solution(s)
Image placement	Difficulty for the app in locking an image into place on a surface	Use a surface with some variation or texture (eg, woodgrain) rather than glossy or uniformly white or black.
		When choosing where to place the image, try to have the edge of the surface (eg, table) in the screen.
		Do not have the smartphone camera lens close to and perpendicular to the plane onto which the image is to be placed.
		Do not place the object with the camera lens facing directly down (phone looking directly down on the object).
		If there are other objects on the table, do not attempt to place the augmented reality object too near the real object or the app may have difficulty identifying the correct plane and the augmented reality object may "jump on" the item.
Image stability	Shaking, moving, floating away	Do not have the smartphone camera lens close to and perpendicular to the plane onto which the image has been placed.
		Do not move the smartphone lens too close to the image (ie, so that it appears one is millimeters away from the object).
Light and shadow	Light reflection looking unnatural or static	Lighting conditions in the room in which the image was viewed appeared to affect the realism of the light and shadow on the object. Very bright rooms highlighted any deficiencies in shadowing of the object.
	Lack of shadow or shadow that does not move naturally as the object is viewed from different angles.	
		Items that often reflect light (eg, metal pen) were more likely to have pixilation problems in the areas that were meant to reflect light.
Sizing	Image, once placed, is either much too large or small.	Scale appears to be affected by identifying the size of the plane through edges, intersecting walls, etc. Sizing problems were more common on large uniform surfaces where edges or walls were not in the screen view.

Many of the difficulties mentioned above may be resolved as augmented reality hardware and software advances, and some may already be addressable with more extensive programming and budgets to support it. Although other platforms, including Autodesk Maya and Autodesk 3ds Max, do provide additional capabilities, the developers used Blender, as it was the most suitable low-cost option.

Limitations

There are several limitations to this development and usability study. First, although the augmented reality app can be used with any iPhone operating system device with augmented reality capabilities, a very recent model with an oversized screen was used in this study. Older models, smaller screens, or Android devices may have a different user experience in image quality or stability. Second, the smoking-related images chosen were a small sample of potentially relevant smoking stimuli. Although we selected items that are common to most smokers, there may be more potent triggers for some individuals, such as coffee and cigarette rolling papers. Likewise, the images that we chose were designed to be generic (ie, not tied to any specific brand, so as to be equally familiar to all participants). Brand-specific items may better elicit urges to smoke and appear more realistic to participants. Third, it is possible that the images intended to be neutral might nevertheless have elicited smoking urges in some participants. For example, if soda consumption had been reliably paired with smoking, the soda bottle could have been a conditioned stimulus rather than a neutral cue. Further, the laboratory setting in which participants viewed the augmented reality images was most likely dissimilar to their naturalistic smoking environments, which could have affected the ratings of smoking urges and reality/coexistence. Of course, the advantage of augmented reality is that the stimuli ultimately can be presented in smokers' true smoking settings. Fourth, the results of this study may be limited by the small sample size of the participants.

Conclusions

To our knowledge, this is the first application of augmented reality technology to create smoking stimuli that can be further evaluated for cue exposure therapy for smoking cessation. Smoking-related augmented reality images were successfully created and demonstrated to elicit urges to smoke, which were very similar to previous findings with in vivo cues [20-24]. Thus, it should be possible to use these images to extinguish smoking urges during quit attempts in locations associated with smoking in the real world. This has the potential to increase the long-term efficacy of cue exposure therapy. The next step toward the development of a functional smoking cessation augmented reality app is to test both cue reactivity and extinction of urge to smoke (craving) in a carefully controlled, fully powered experimental setting. This would be followed by continued testing in smokers' naturalistic environments and by the addition of other user-friendly app features designed to improve ease of use and treatment adherence. The inclusion of an app readily accessible to smokers may improve the reach of cue exposure therapy. Ultimately, augmented reality technology may prove to facilitate cue exposure treatment and to be an effective, portable, adjuvant to treatments for dependence on tobacco as well as other substances.

Acknowledgments

Research reported in this publication was supported by an award from the National Institute on Drug Abuse (R34DA047598) and by the Participant Research, Interventions, and Measurement Core at the H. Lee Moffitt Cancer Center & Research Institute, a comprehensive cancer center designated by the National Cancer Institute and funded in part by Moffitt's Cancer Center Support Grant (P30-CA076292).

Authors' Contributions

CV, KB, MK, LH, LS, SS, and TB contributed to the study design and interpretation of data. LH collected the data. KB conducted the statistical analyses. CV and KB wrote the first draft of the manuscript. JH is responsible for the technical creation of the augmented reality app. All authors have contributed to the final version of the manuscript and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

TB has received research support from Pfizer Inc and is on the advisory board for Hava Health, Inc.

Multimedia Appendix 1
Video of the augmented reality app in action.
[MOV File , 10556 KB - mhealth_v8i12e21643_app1.mov]

Multimedia Appendix 2
Mean urge and reality/coexistence ratings of smoking-related augmented reality images.
[PNG File , 25 KB - mhealth_v8i12e21643_app2.png]

Multimedia Appendix 3
Mean urge and reality/coexistence ratings of neutral augmented reality images.
[PNG File , 53 KB - mhealth_v8i12e21643_app3.png]

References

1. Vinci C, Brandon KO, Kleinjan M, Brandon TH. The clinical potential of augmented reality. Clin Psychol (New York) 2020 Sep;27(3). [doi: 10.1111/cpsp.12357] [Medline: 33223628]
2. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput. Grap. Appl 2001 Nov;21(6):34-47. [doi: 10.1109/38.963459]
3. Pratt P, Ives M, Lawton G, Simmons J, Radev N, Spyropoulou L, et al. Through the HoloLens™ looking glass: augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels. Eur Radiol Exp 2018;2(1):2 [FREE Full text] [doi: 10.1186/s41747-017-0033-2] [Medline: 29708204]
4. Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 2016 Oct;30(10):4174-4183 [FREE Full text] [doi: 10.1007/s00464-016-4800-6] [Medline: 26905573]
5. Chicchi Giglioli IA, Pallavicini F, Pedroli E, Serino S, Riva G. Augmented Reality: A Brand New Challenge for the Assessment and Treatment of Psychological Disorders. Comput Math Methods Med 2015;2015:862942 [FREE Full text] [doi: 10.1155/2015/862942] [Medline: 26339283]
6. U.S. Department of Health and Human Services. Smoking Cessation: A Report of the Surgeon General. The Surgeon General's Report 2020. [Medline: 32255575]
7. Creamer MR, Wang TW, Babb S, Cullen KA, Day H, Willis G, et al. Tobacco Product Use and Cessation Indicators Among Adults - United States, 2018. MMWR Morb Mortal Wkly Rep 2019 Nov 15;68(45):1013-1019 [FREE Full text] [doi: 10.15585/mmwr.mm6845a2] [Medline: 31725711]
8. Vinci C, Haslam A, Lam CY, Kumar S, Wetter DW. The use of ambulatory assessment in smoking cessation. Addictive Behaviors 2018 Aug;83:18-24. [doi: 10.1016/j.addbeh.2018.01.018]
9. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17 Suppl 1:44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]
10. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 02;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
11. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin R, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]

12. Mobile Fact Sheet. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-03-01]
13. Patel V, Johnson C. Trends in Individuals? Access and Use of Online Medical Records and Technology for Health Needs: 2017-2018. ONC Data Brief, no 48 2019.
14. Siegel S. Classical conditioning, drug tolerance, and drug dependence. In: Research advances in alcohol and drug problems. Boston, MA: Springer; 1983:207-246.
15. Stewart J, de Wit H, Eikelboom R. Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. Psychol Rev 1984 Apr;91(2):251-268. [Medline: 6571424]
16. Unrod M, Drobes DJ, Stasiewicz PR, Ditre JW, Heckman B, Miller RR, et al. Decline in cue-provoked craving during cue exposure therapy for smoking cessation. Nicotine Tob Res 2014 Mar;16(3):306-315 [FREE Full text] [doi: 10.1093/ntr/ntt145] [Medline: 24078760]
17. Monti PM, Rohsenow DJ. Coping-skills training and cue-exposure therapy in the treatment of alcoholism. Alcohol Res Health 1999;23(2):107-115 [FREE Full text] [Medline: 10890804]
18. Bouton ME. A learning theory perspective on lapse, relapse, and the maintenance of behavior change. Health Psychol 2000 Jan;19(1S):57-63. [doi: 10.1037/0278-6133.19.suppl1.57] [Medline: 10709948]
19. Conklin CA, Tiffany ST. Applying extinction research and theory to cue-exposure addiction treatments. Addiction 2002 Feb;97(2):155-167. [doi: 10.1046/j.1360-0443.2002.00014.x] [Medline: 11860387]
20. LaRowe SD, Saladin ME, Carpenter MJ, Upadhyaya HP. Reactivity to nicotine cues over repeated cue reactivity sessions. Addict Behav 2007 Dec;32(12):2888-2899 [FREE Full text] [doi: 10.1016/j.addbeh.2007.04.025] [Medline: 17537583]
21. Wray JM, Godleski SA, Tiffany ST. Cue-reactivity in the natural environment of cigarette smokers: the impact of photographic and in vivo smoking stimuli. Psychol Addict Behav 2011 Dec;25(4):733-737 [FREE Full text] [doi: 10.1037/a0023687] [Medline: 21553947]
22. Carpenter MJ, Saladin ME, DeSantis S, Gray KM, LaRowe SD, Upadhyaya HP. Laboratory-based, cue-elicited craving and cue reactivity as predictors of naturally occurring smoking behavior. Addict Behav 2009;34(6-7):536-541 [FREE Full text] [doi: 10.1016/j.addbeh.2009.03.022] [Medline: 19395178]
23. Conklin CA, Perkins KA, Robin N, McClernon FJ, Salkeld RP. Bringing the real world into the laboratory: personal smoking and nonsmoking environments. Drug Alcohol Depend 2010 Sep 01;111(1-2):58-63 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.03.017] [Medline: 20510552]
24. Conklin CA, Tiffany ST. Applying extinction research and theory to cue-exposure addiction treatments. Addiction 2002 Feb;97(2):155-167. [doi: 10.1046/j.1360-0443.2002.00014.x] [Medline: 11860387]
25. Haas JK. A history of the unity game engine. Worcester: Worcester Polytechnic Institute; 2014. URL: https://digitalcommons.wpi.edu/iqp-all/3207/ [accessed 2020-09-01]
26. Blender Community. Blender--a 3D modellingrendering package. Blender Foundation. Amsterdam; 2018. URL: http://www.blender.org [accessed 2020-06-09]
27. Biener L, Abrams DB. The Contemplation Ladder: validation of a measure of readiness to consider smoking cessation. Health Psychol 1991;10(5):360-365. [doi: 10.1037//0278-6133.10.5.360] [Medline: 1935872]
28. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
29. Baus O, Bouchard S. Moving from Virtual Reality Exposure-Based Therapy to Augmented Reality Exposure-Based Therapy: A Review. Front. Hum. Neurosci 2014 Mar 04;8. [doi: 10.3389/fnhum.2014.00112]
30. Brandon TH, Copeland AL, Saper ZL. Programmed therapeutic messages as a smoking treatment adjunct: Reducing the impact of negative affect. Health Psychology 1995;14(1):41-47. [doi: 10.1037/0278-6133.14.1.41]
31. Pericot-Valverde I, Germeroth LJ, Tiffany ST. The Use of Virtual Reality in the Production of Cue-Specific Craving for Cigarettes: A Meta-Analysis. Nicotine Tob Res 2016 May;18(5):538-546. [doi: 10.1093/ntr/ntv216] [Medline: 26453669]

Abbreviations

CO: carbon monoxide
FTND: Fagerström Test for Nicotine Dependence

Edited by L Buis; submitted 19.06.20; peer-reviewed by D Amiras; comments to author 13.07.20; revised version received 16.07.20; accepted 20.11.20; published 31.12.20.

Please cite as:
Vinci C, Brandon KO, Kleinjan M, Hernandez LM, Sawyer LE, Haneke J, Sutton SK, Brandon TH
Augmented Reality for Smoking Cessation: Development and Usability Study
JMIR Mhealth Uhealth 2020;8(12):e21643
URL: http://mhealth.jmir.org/2020/12/e21643/
doi:10.2196/21643
PMID:33382377

Corrigenda and Addenda

Correction: COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control

Li Du[1], MBBS, LLB, PhD; Vera Lúcia Raposo[1,2], BA, LLB, LLM, PhD; Meng Wang[1], LLB, LLM

[1]Faculty of Law, University of Macau, Macau, SAR, China
[2]Faculty of Law, University of Coimbra, Coimbra, Portugal

Corresponding Author:
Li Du, MBBS, LLB, PhD
Faculty of Law
University of Macau
Avenida da Universidade, Taipa
Macau, SAR, 999078
China
Phone: 853 88224733
Email: stephendu@um.edu.mo

Related Article:

Correction of: https://mhealth.jmir.org/2020/11/e23194/

(JMIR Mhealth Uhealth 2020;8(12):e23194) doi:10.2196/26239

In "COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control" (JMIR Mhealth Uhealth 2020;8(11):e23194) the authors noted one error.

This article was inadvertently published with the incorrect country code for the phone number of the Corresponding Author. The phone number was originally published as "86 88224733". This has been corrected to "853 88224733".

The correction will appear in the online version of the paper on the JMIR Publications website on December 7, 2020, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 04.08.20; this is a non–peer-reviewed article;accepted 03.11.20; published 04.12.20.

Please cite as:
Du L, Raposo VL, Wang M
Correction: COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control
JMIR Mhealth Uhealth 2020;8(12):e23194
URL: http://mhealth.jmir.org/2020/12/e26239/
doi:10.2196/26239
PMID:33275590

Corrigenda and Addenda

Correction: ColistinDose, a Mobile App for Determining Intravenous Dosage Regimens of Colistimethate in Critically Ill Adult Patients: Clinician-Centered Design and Development Study

Xueliang Hua[1*], BSc; Chen Li[2,3,4*], PhD; Jason M Pogue[5], PharmD; Varun S Sharma[4], MSc; Ilias Karaiskos[6], MD, PhD; Keith S Kaye[7], MD, MPH; Brian T Tsuji[8,9], PharmD; Phillip J Bergen[3,10], PhD; Yan Zhu[3,10], PhD; Jiangning Song[2,3,11], PhD; Jian Li[3,10], PhD

[1]Independent Researcher, Santa Clara, CA, United States
[2]Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia
[3]Infection and Immunity Program, Biomedicine Discovery Institute, Monash University, Melbourne, Australia
[4]Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, Switzerland
[5]Department of Clinical Pharmacy, University of Michigan College of Pharmacy, Ann Arbor, MI, United States
[6]1st Internal Medicine and Infectious Diseases Department, Hygeia Hospital, Marousi, Greece
[7]Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, United States
[8]Laboratory for Antimicrobial Dynamics, NYS Center of Excellence in Bioinformatics & Life Sciences, Buffalo, NY, United States
[9]School of Pharmacy and Pharmaceutical Sciences, University at Buffalo, Buffalo, NY, United States
[10]Department of Microbiology, Monash University, Melbourne, Australia
[11]Monash Centre for Data Science, Faculty of Information Technology, Monash University, Melbourne, Australia
[*]these authors contributed equally

Corresponding Author:
Jian Li, PhD
Department of Microbiology
Monash University
19 Innovation Walk, Monash University
Melbourne, VIC 3800
Australia
Phone: 61 3 9903 9702
Email: jian.li@monash.edu

Related Article:

Correction of: https://mhealth.jmir.org/2020/12/e20525/

(JMIR Mhealth Uhealth 2020;8(12):e26593) doi:10.2196/26593

In “ColistinDose, a Mobile App for Determining Intravenous Dosage Regimens of Colistimethate in Critically Ill Adult Patients: Clinician-Centered Design and Development Study” (JMIR Mhealth Uhealth 2020;8(12):e20525) the authors noted one error due to an editing issue.

In the Abbreviations section, two abbreviations were listed incorrectly as follows:

CVVD: continuous veno-venous hemofiltration

CVVH: continuous veno-venous hemofiltration

These abbreviations have been corrected to:

CVVH: continuous veno-venous hemofiltration

CVVHD: continuous veno-venous hemodialysis

The correction will appear in the online version of the paper on the JMIR Publications website on December 18, 2020, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 17.12.20; this is a non–peer-reviewed article;accepted 17.12.20; published 18.12.20.

Please cite as:
Hua X, Li C, Pogue JM, Sharma VS, Karaiskos I, Kaye KS, Tsuji BT, Bergen PJ, Zhu Y, Song J, Li J
Correction: ColistinDose, a Mobile App for Determining Intravenous Dosage Regimens of Colistimethate in Critically Ill Adult Patients: Clinician-Centered Design and Development Study
JMIR Mhealth Uhealth 2020;8(12):e26593
URL: http://mhealth.jmir.org/2020/12/e26593/
doi:10.2196/26593
PMID:33338026

Corrigenda and Addenda

Correction: Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality

Atiyeh Vaezipour[1], PhD; Jessica Campbell[2], PhD; Deborah Theodoros[1], PhD; Trevor Russell[1], PhD

[1]RECOVER Injury Research Centre, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Australia
[2]Queensland Aphasia Research Centre, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Australia

Corresponding Author:
Atiyeh Vaezipour, PhD
RECOVER Injury Research Centre
Faculty of Health and Behavioural Sciences
The University of Queensland
288 Herston Road
Brisbane, 4006
Australia
Phone: 61 7 3365 5560
Email: a.vaezipour@uq.edu.au

Related Article:

Correction of: https://mhealth.jmir.org/2020/10/e18858/

(JMIR Mhealth Uhealth 2020;8(12):e18858) doi:10.2196/26309

In “Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality” (JMIR Mhealth Uhealth 2020; 8(10):e18858) the authors noted three errors.

In the Results section, App Quality subsection, the top 5 apps were originally listed as developed by Tactus Therapy solutions as follows:

> *The top 5 apps were Naming Therapy, Speech Flipbook Standard (4.6/5), Number Therapy (4.5/5), Answering Therapy, and Constant Therapy (4.4/5). All of these top apps were made by Tactus Therapy solutions.*

This was found to be incorrect as the Constant Therapy app is made by Constant Therapy Health and not by Tactus Therapy solutions. Therefore, the sentence "All of these top apps were made by Tactus Therapy solutions" has been removed from the text to be consistent with the rest of the article which did not describe the name of the app developers.

Similarly, in Multimedia Appendix 1, the sentence “Developed by Tactus Therapy Solutions Ltd” has been deleted.

The contact number of the corresponding author was originally published as "61 7 3346 4824". This has been changed to "61 7 3365 5560".

The correction will appear in the online version of the paper on the JMIR Publications website on December 11, 2020, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1
Summary of findings.
[PNG File, 529KB - mhealth_v8i12e26309_app1.png]

Submitted 23.03.20; this is a non–peer-reviewed article;accepted 27.09.20; published 11.12.20.

Please cite as:
Vaezipour A, Campbell J, Theodoros D, Russell T
Correction: Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality
JMIR Mhealth Uhealth 2020;8(12):e18858
URL: http://mhealth.jmir.org/2020/12/e26309/
doi:10.2196/26309
PMID:33306482

Original Paper

Using Smartphone Sensor Data to Assess Inhibitory Control in the Wild: Longitudinal Study

Vincent WS Tseng[1], BSc; Jean Dos Reis Costa[2,3], PhD; Malte F Jung[3], PhD; Tanzeem Choudhury[1], PhD

[1]Department of Information Science, Cornell University, New York, NY, United States

[2]DawnLight Technologies, Palo Alto, CA, United States

[3]Department of Information Science, Cornell University, Ithaca, NY, United States

Corresponding Author:
Vincent WS Tseng, BSc
Department of Information Science
Cornell University
2 W Loop Rd
New York, NY, 10044
United States
Phone: 1 6073790290
Email: vincent@infosci.cornell.edu

Abstract

Background: Inhibitory control, or inhibition, is one of the core executive functions of humans. It contributes to our attention, performance, and physical and mental well-being. Our inhibitory control is modulated by various factors and therefore fluctuates over time. Being able to continuously and unobtrusively assess our inhibitory control and understand the mediating factors may allow us to design intelligent systems that help manage our inhibitory control and ultimately our well-being.

Objective: The aim of this study is to investigate whether we can assess individuals' inhibitory control using an unobtrusive and scalable approach to identify digital markers that are predictive of changes in inhibitory control.

Methods: We developed InhibiSense, an app that passively collects the following information: users' behaviors based on their phone use and sensor data, the ground truths of their inhibition control measured with stop-signal tasks (SSTs) and ecological momentary assessments (EMAs), and heart rate information transmitted from a wearable heart rate monitor (Polar H10). We conducted a 4-week in-the-wild study, where participants were asked to install InhibiSense on their phone and wear a Polar H10. We used generalized estimating equation (GEE) and gradient boosting tree models fitted with features extracted from participants' phone use and sensor data to predict their stop-signal reaction time (SSRT), an objective metric used to measure an individual's inhibitory control, and identify the predictive digital markers.

Results: A total of 12 participants completed the study, and 2189 EMAs and SST responses were collected. The results from the GEE models suggest that the top digital markers positively associated with an individual's SSRT include phone use burstiness (P=.005), the mean duration between 2 consecutive phone use sessions (P=.02), the change rate of battery level when the phone was not charged (P=.04), and the frequency of incoming calls (P=.03). The top digital markers negatively associated with SSRT include the standard deviation of acceleration (P<.001), the frequency of short phone use sessions (P<.001), the mean duration of incoming calls (P<.001), the mean decibel level of ambient noise (P=.007), and the percentage of time in which the phone was connected to the internet through a mobile network (P=.001). No significant correlation between the participants' objective and subjective measurement of inhibitory control was found.

Conclusions: We identified phone-based digital markers that were predictive of changes in inhibitory control and how they were positively or negatively associated with a person's inhibitory control. The results of this study corroborate the findings of previous studies, which suggest that inhibitory control can be assessed continuously and unobtrusively in the wild. We discussed some potential applications of the system and how technological interventions can be designed to help manage inhibitory control.

(JMIR Mhealth Uhealth 2020;8(12):e21703) doi:10.2196/21703

KEYWORDS

self-control; mobile phone

Introduction

Background

Inhibitory control, or inhibition, is the ability to inhibit prepotent responses to goal-irrelevant stimuli. It is one of our executive functions and is essential for sustained attention [1], working memory [2], and emotion regulation [3], which in turn contribute to our performance and well-being. Studies have shown that students who demonstrate better inhibitory control tend to be better at time management and achieve higher academic performance [4-9] and that employees who have higher inhibitory control are more likely to have higher motivation and productivity [10]. On the other hand, reduced inhibitory control can lead to attention problems and impulsive and addictive behaviors [11] (eg, alcohol and drug addiction) [12]. Furthermore, many mental disorders are associated with impaired inhibitory control, including eating disorders [13], posttraumatic stress disorder [14], bipolar disorder [15], and schizophrenia [16].

Although some individuals have lower inhibitory control than others, our inhibitory control is not a stable trait. Instead, it fluctuates over time based on internal and environmental factors, such as our moods, activities, and the context or the surrounding environment we are in. For example, it has been shown that physical activities [17] and good sleep habits [18] have positive effects on inhibitory control. Furthermore, exposure to natural landscapes can enhance one's inhibitory control. People who were exposed to natural landscapes showed a higher ability for future valuation (ie, an individual's willingness to wait for a longer period of time in exchange for a larger reward) and delayed gratification compared with those exposed to urban environments [19]. Moreover, even just showing participants images of natural environments could significantly reduce their impulsive decision making compared with showing them images of human-made environments or geometric shapes [20].

As the ebbs and flows of inhibitory control consequently influence an individual's daily behaviors, such as attentional ability [21], alcohol consumption [22,23], and antisocial and criminal behaviors [24], researchers have examined different approaches to assessing inhibitory control, including psychometric tests, physiological signals, and self-report questionnaires. For example, stop-signal task (SST) [25] is one of the most widely used psychometric tests; it measures participants' reaction time to inhibit their prepotent response after seeing a stop signal. It has been used to predict the level of smokers' craving for cigarettes [26]. Physiological signals, such as heart rate variability (HRV), also have been shown to be indicators of the level of an individual's inhibitory control. Individuals with higher HRV tend to have higher inhibitory control, as HRV is associated with one's ability to adjust in response to changes in the environment [27]. A number of self-control scales were used in studies that examined how people's self-control contributes to regular exercise [28] and web game addiction [29]. However, most of the studies were conducted in laboratory settings and only focused on the inhibitory control of individuals with cravings or addiction. To measure inhibitory control in the wild, asking users to periodically complete psychometric tests or self-report questionnaires can be burdensome, as each session of psychometric tests normally takes about 5 min. Moreover, self-reported responses are likely to be biased by responders' current inhibitory control [30,31]. Thus, assessing inhibitory control using self-reports may be subject to subjectivity and unreliability.

Some of the aforementioned factors that modulate inhibitory control, such as locations visited and physical activities, can be captured with smartphones or wearables. Behavioral patterns that are associated with changes in inhibitory control may also manifest themselves in people's phone sensor data, such as their phone use patterns. Some previous work suggested that phone sensor data can be used to predict an individual's cognitive performance [32]. However, the relationship between an individual's inhibitory control and behavioral patterns has yet to be fully studied. HRV can also be collected continuously in the wild using a chest strap or a wristband. However, the reliability of phone sensor data, particularly for the purpose of assessing inhibitory control, has not been fully investigated so far.

Objectives

To this end, the goal of this study is to investigate measuring inhibitory control using a scalable approach within a broader population in addition to people who have problems with inhibitory control and markers indicative of changes in inhibitory control. We developed an iOS app and conducted a 4-week in-the-wild study to collect users' phone sensor data, HRV, and subjective and objective measurements of their inhibitory control. We used the data to infer the changes in participants' inhibitory control and to examine the factors that were associated with the changes. The contributions of this paper are 3-fold:

1. To the best of our knowledge, this is the first in-the-wild study that collected participants' inhibitory control with objective and subjective measurements along with their phone sensor data and HRV. We also made our code for our InhibiSense publicly available on Github [33] so that other researchers can use the platform to conduct future studies.
2. We identified how different behaviors and contexts influenced people's inhibitory control in the wild based on more than 1100 measurements collected from this study. We also used these predictors as features to train machine learning models to determine whether we could predict high and low inhibitory control.
3. Finally, we discussed the implications of the study and several applications of the system that can help us manage our behaviors and well-being in real-world scenarios.

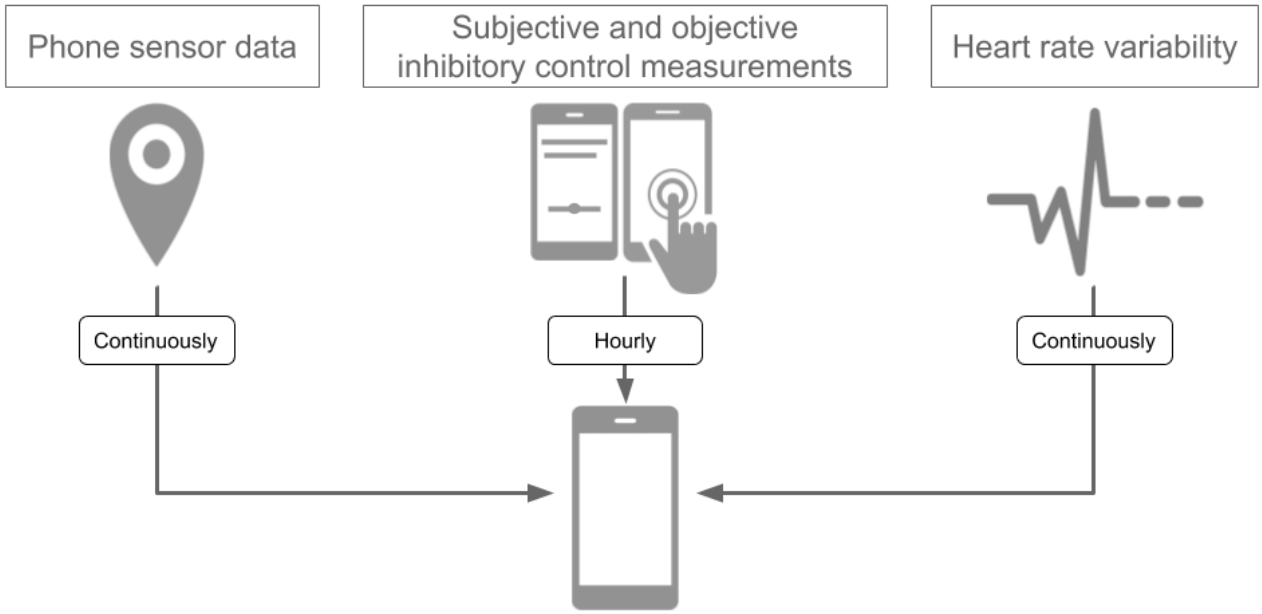
Methods

System

We developed InhibiSense, an iOS study app that collects participants' phone sensor data, HRV, ecological momentary assessment (EMA), and performance of SSTs (Figure 1). The collected data were temporarily stored on a participant's phone

and uploaded to our server when the participant's phone was charged and connected to the internet through a Wi-Fi network. The details of each data type are described below.

Figure 1. Data collected on participants' smartphones.



Phone Sensor Data

InhibiSense collected users' phone sensor data continuously using the AWARE framework [34]. The AWARE framework provides application programming interfaces (APIs) that allow researchers to integrate phone sensor data logging and storing into their study apps. Specifically, the types of data we collected include participants' GPS locations, Wi-Fi device names (service set identifier [SSID]), activities, ambient noise, networks, phone use, calls, battery status, and time.

HRV

HRV calculation is based on a person's interbeat intervals, or RR intervals. We collected participants' RR intervals using Polar H10 [35]. The RR intervals' recordings were transmitted to InhibiSense via Bluetooth.

EMA

EMA is a method to repeatedly sample participants' experiences in real time during a study [36]. We employed a valid EMA for inhibitory control that was suited for the context of our research [28]. It consisted of 6 questions regarding one's inhibitory control at the present moment, such as "I have to force myself to stay focused" (Table 1). Participants were prompted to select a score for each question on a Likert scale ranging from 0 (*Not at all)* to 6 (*Very much so*) using a slider.

Table 1. Types of data collected during the study.

Type	Description	Frequency
Demographic information	• Barratt Impulsiveness Scale • Sex • Age	During the onboarding meeting
Phone sensor	• Physical activity • Phone use • Call • Battery level • GPS location • Wi-Fi service set identifier • Ambient noise • Network • Time	Continuously
Heart rate variability	Interbeat interval	Continuously
Ecological momentary assessment	Six self-report questions asking about one's inhibitory control at the moment, including the following: • "I have to force myself to stay focused" • "I am full of willpower" • "I am having trouble pulling myself together" • "I could resist any temptation" • "I am having trouble paying attention" • "I have no trouble bringing myself to do disagreeable things"	Prompted every hour from 7 AM to 11 PM
stop-signal task	Each session consisted of 80 trials, 60 of which were Go trials and 20 of which were stop trials. The order of the trials was randomized. A tracking method was used where the SSD[a] was increased by 25 milliseconds if a stop error was made and decreased by 25 milliseconds if a stop was successful.	Prompted every hour from 7 AM to 11 PM following the completion of the ecological momentary assessment

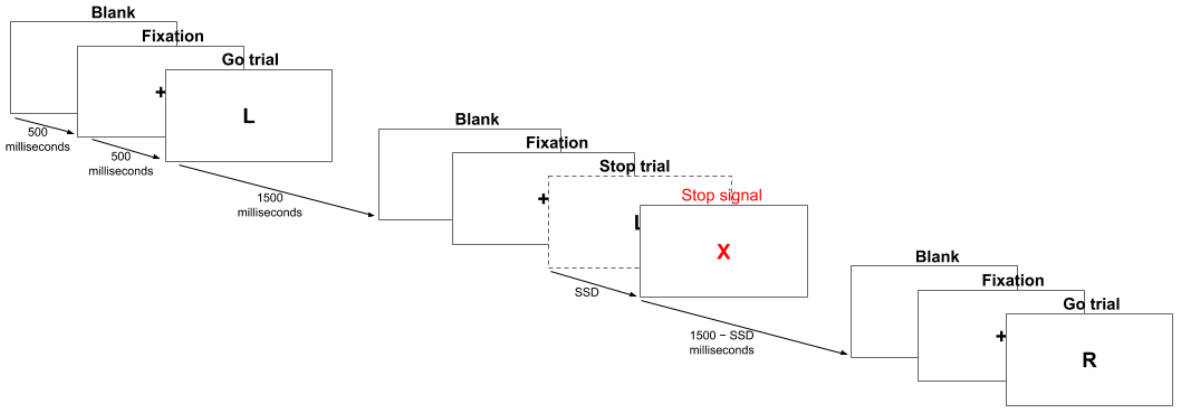
[a]SSD: stop-signal delay.

SST

SSTs are valid and widely used objective measurements for assessing inhibitory control [37]. They typically consist of 2 types of signals: go signals and stop signals. stop signals appear at some random intervals, which are often referred to as stop-signal delays (SSDs), after go signals with a predetermined probability. Participants have to respond to go signals as quickly as they can and inhibit their response when they see a stop signal.

Our implementation of SST (Figure 2) consisted of 80 trials, 75% (60/80) of which were go trials and the rest were stop trials. The order of the trials was randomized during each session. Having a higher probability of showing a go signal is a recommended approach to preventing participants from strategically delaying their response and waiting for a stop signal. The interval of each trial was 1500 milliseconds. The stimuli disappeared immediately when a participant touched the screen. In each go trial, a letter *L* or *R* was shown with equal probability, and a participant had to tap the left or the right side of the screen, respectively. In a stop trial, the go signal was replaced with a stop signal *X* after an SSD, and a participant had to inhibit their response immediately. To estimate a participant's stop-signal reaction time (SSRT) more accurately, we used the tracking method (the one-up-one-down method). The initial SSD was set to 250 milliseconds and then was increased by 25 milliseconds in the next stop trial after a successful stopping and was decreased by 25 milliseconds after an unsuccessful stopping.

Figure 2. Illustration of the stop-signal task implementation. During the task, Go trials and stop trials appeared in random orders. During a Go trial, a Go stimulus, either a letter "L" or a letter "R" was displayed, to which a participant had to respond. During a stop trial, a Go stimulus was first displayed and then followed by a stop signal "X" after a certain stop Signal Delay, in which a participant had to inhibit their response. At the beginning of each trial, the screen turned blank for 500ms followed by 500ms fixation duration with a "+" displayed at the center of the screen. SSD: stop-signal delay; SST: stop-signal task.



Study

The study lasted for 4 weeks. During the study, participants received push notifications for answering the EMA questions and completing the SST every hour, from 7 AM to 11 PM. Each task would expire in an hour. They were encouraged to wear the heart rate monitor for the entire day and were able to proceed with the task only if they were wearing it.

Recruitment

The study was approved by the Institutional Review Board at Cornell University. We used convenience sampling for our recruitment. During the onboarding meeting, participants were asked to fill out a demographic questionnaire, including their age range and sex, and a Barratt Impulsiveness Scale (BIS 11) questionnaire [38]. Then, they were informed of the flow of the study and signed a consent form. Each participant was given a Polar H10 to wear throughout the study. Finally, the participants installed InhibiSense on their phones and were instructed to complete a couple of practice sessions for SST to ensure that they fully understood the process. Participants were given compensation at the end of the study based on the number of sessions (US $1 for each session) they had completed.

Data Preprocessing and Feature Extraction

Our goal was to assess an individual's inhibitory control at a more granular level, specifically every hour. As such, we needed to use the information on the context and physiological signals during the past hour to infer a participant's inhibitory control at the end of the 1-hour window. We extracted passive sensing features from the phone sensor data and HRV features from their RR intervals during the 1-hour window before a participant started an SST. Specifically, the features from phone sensor data capture information on factors that are associated with changes in inhibitory control, including activities, phone use, surrounding environments, and sleep. We will describe our data preprocessing and feature extraction below.

Individual Inhibitory Control Baseline

The questions asked in BIS correspond to different order factors (dimensions). Specifically, 6 out of the total 30 questions correspond to order factor *self-control*, also known as trait inhibitory control. We added up the scores for the 6 questions and used it as the baseline inhibitory control.

Phone Sensor Features

Activity

Activity can be broken down into the following categories: total number of steps, distance traveled, activity type, and acceleration. We computed the percentage of time an individual was stationary, walking, running, automotive, cycling, and nonstationary (which includes all the activities except for being stationary) during each 1-hour window based on the inferred activities provided by Apple's activity recognition API [39]. For acceleration, we computed the mean and standard deviation of a user's acceleration (m/s^2).

Phone Use

We computed phone use–related features as follows:

1. Phone use sessions: On the basis of screen unlock and lock events, we computed the burstiness (the number of phone use sessions), the number of short sessions (sessions less than 30 seconds) [40], the total duration of phone use, the mean and standard deviation of phone use duration, and the intervals between consecutive phone use sessions.
2. Call: The number and the mean duration of incoming and outgoing calls.
3. Battery: The frequency and percentage of time a user's phone was being charged and the rate of battery level change when the phone was and was not being charged, respectively.

Environment

We computed environment-related features as follows:

1. Ambient noise: We computed the mean, median, and standard deviation of the frequency, loudness (dB), and power, namely root mean square, of ambient noise.
2. Location: We used the GPS coordinates and Foursquare API [41] to retrieve the semantic locations, namely, the categories of the places a participant visited. As the GPS

locations might sometimes be inaccurate, we obtained the categories of a user's nearby locations within 50 m for each GPS location and aggregated all the distinct location categories within each window. As we were interested in investigating the relationship between the consistency of environments and a person's inhibitory control, we also retrieved the locations that a participant had been to in the previous window (ie, between 1 and 2 hours before the start of an SST) and computed the similarity of the locations visited between the 2 consecutive windows using the Jaccard, Dice, Second-Kulczynski, and Ochiai distance metrics [42-45].

3. Wi-Fi: We computed the similarity between the Wi-Fi device names (SSID) in 2 consecutive windows using the Jaccard, Dice, Second_Kulczynski, and Ochiai distance metrics. The presence of the individual SSIDs during a window was encoded as a binary vector, and the similarity metrics were then applied to the vectors.
4. Network: The percentage of time a user's phone was connected to the internet and the percentage of time the phone was connected to the network through a mobile and Wi-Fi network, respectively.

Sleep

Sleep duration, sleep onset, and sleep offset were inferred based on participants' phone use [46]. Sleep duration was clipped to 2 standard deviations from the mean to replace the outliers.

Time

Information on the day of the week, weekday or weekend, and the hour of the day was obtained based on the timestamps.

The features extracted for each type of sensor data are summarized in Textbox 1.

Textbox 1. Features extracted from the sensor data.

Physical activity

- Mean and standard deviation of acceleration
- Number of steps
- Distance traveled
- Percentage of time a participant was stationary, walking, running, automotive, cycling

Phone use

- Phone use burstiness
- Frequency of short phone use sessions
- Mean and standard deviation of phone use durations (seconds) and intervals between consecutive phone use sessions

Call

- Frequency and duration (seconds) of incoming and outgoing calls

Battery

- Frequency and duration (seconds) of charging
- Mean change rate of battery level (units/second) when the phone was charged and not charged

GPS location

- Location category
- Jaccard, Dice, Second_Kulczynski, and Ochiai coefficients of location similarity

Wi-Fi

- Jaccard, Dice, Second_Kulczynski, and Ochiai coefficients for Wi-Fi service set identifier (SSID) similarity

Ambient noise

- Mean and standard deviation of frequency (Hz), loudness (dB), and sound root mean square (RMS)

Network

- Percentage of time the phone was connected to the internet through a Wi-Fi or mobile network or was not connected

Time

- Weekend or weekday
- Day of the week
- Hour of the day

Sleep

- Sleep onset and offset
- Hours of sleep

Interbeat interval

- Mean and standard deviation of heart rate
- Standard deviation of NN intervals
- Standard deviation of 5-min average NN intervals
- Root-mean-squared NN interval differences
- Standard deviation of NN interval differences
- Triangular index
- Relative power of low (0.04-0.15 Hz) and high (0.15-0.4 Hz) frequency band computed using Fast Fourier transform (FFT), Lomb-Scargle periodogram, and autoregressive method

HRV Features

To ensure the quality of the HRV features, HRV features were extracted if and only if an interval contained at least 5 min of continuous RR interval recordings. For windows that met the criteria, we first removed outlier and ectopic beats [47] and replaced them using cubic interpolation using the HRVanalysis package [48]. Then, the time-domain, frequency-domain, and nonlinear features were extracted using the pyHRV package [49] (Textbox 1).

SST Performance Metrics

SSRT is the amount of time it takes for a participant to inhibit their response, which is inversely correlated with one's inhibitory control. A longer SSRT means lower inhibitory control and vice versa. We used the integration method based on the horse race theorem [25] to compute the central SSRT, the most reliable way to estimate SSRT, in each session. For each SSD in a given session, we computed the stop unsuccessful rate p(response | signal) and the corresponding *n*th RT (the RT at the *n*th percentile) among the RT distribution in that session, where n is equal to (p(response | signal)×100). Then, we used linear interpolation to find the SSD where p(response | signal) equals 0.5. Finally, the central SSRT in that session was obtained by subtracting the SSD from the 50th RT.

Outlier Exclusion

We excluded data points in which the stop unsuccessful rate was below 0.25 or above 0.75 and the estimated SSRT was below 50 milliseconds or above 1500 milliseconds [50].

Data Analysis

The Relationship Between Subjective and Objective Measurements of Inhibitory Control

To investigate how the different constructs of inhibitory control in the EMA are related to each other and how people perceive their own inhibitory control in comparison with their inhibitory control measured using SST, we computed the correlation coefficients between the different dependent variables using repeated measures correlation (RMCORR) [51,52]. Repeated measures correlation is used to measure the strength of the relationship between 2 variables in repeated measurements across different participants, which accounts for the intraindividual associates between 2 measures.

Features Predictive of Changes in Inhibitory Control

To examine the predictive features for inhibitory control, namely, features that have significant main effects for estimating SSRT, we fitted 2 generalized estimating equation (GEE) [53] models with 2 different feature sets, one with only phone sensor features and the other with phone sensor features along with HRV features, to predict participants' SSRT. GEE is a statistical model that is used to identify feature variables that have significant effects during repeated measurements and in the meantime to account for individual differences, which is particularly useful for analyzing the relationship between the predictors and outcomes in longitudinal studies. Developing models with these 2 feature sets helped us get a better understanding of how models perform using the least obtrusive manner, namely, by using only phone sensor data, and whether the models' performance will improve when the information on HRV is accessible and incorporated.

In our analysis, the independent variables were the sensor features and their inhibitory control baseline derived from their BIS responses; the dependent variable was SSRT. We chose SSRT to be the dependent variable because we assumed that an individual's inhibitory control is a continuum. Before fitting the models, we first removed features whose collinearity was above the variance inflation factor threshold [54] and then used the remaining features as the independent variables for the models. Gamma distribution was used in the GEE models to model the distribution of SSRT.

Predicting States of Inhibitory Control

After determining which features are more predictive of changes in inhibitory control, our next research question was whether we can automatically infer the state of a person's inhibitory control using these features. To this end, we trained gradient boosting tree (GBT) [55] classifiers to classify whether an individual was in a high or low inhibitory control state with both feature sets (phone sensor features only and phone sensor features plus HRV features) after the predictive features were identified. A high or low inhibitory control state refers to whether an individual's SSRT during an SST session was lower or higher than their overall median SSRT. GBT is a type of classifier that is more robust to outliers. The features were z-standardized before fitting into the classifiers. We used leave-one-subject-out cross-validation, where 1 participant's data were held out for testing and a model was trained on the remaining participants' data during each iteration, to evaluate the model performance. The metrics for the cross-validation are the mean accuracy and the area under the receiver operating characteristic curve (AUC-ROC) [56].

Results

Descriptive Statistics

A total of 22 participants signed up for the study; 10 of them dropped out in the middle of the study due to their school workloads. The remaining participants (their demographics are summarized in Table 2) completed 2189 tasks over the course of the study. After applying data preprocessing and outlier exclusion, data from 1107 sessions (mean 92.3, SD 67.4) were used for our analysis. Table 3 shows the cumulative number of tasks completed during the different hours; the data were roughly uniformly distributed across 9 AM to 11 PM. Figure 3 shows the mean SSRT at different times of day averaged across all the participants. Overall, participants had lower SSRT around 10 AM and 4 PM and higher SSRT around 7 AM and 3 PM; however, no statistically significant difference was found.

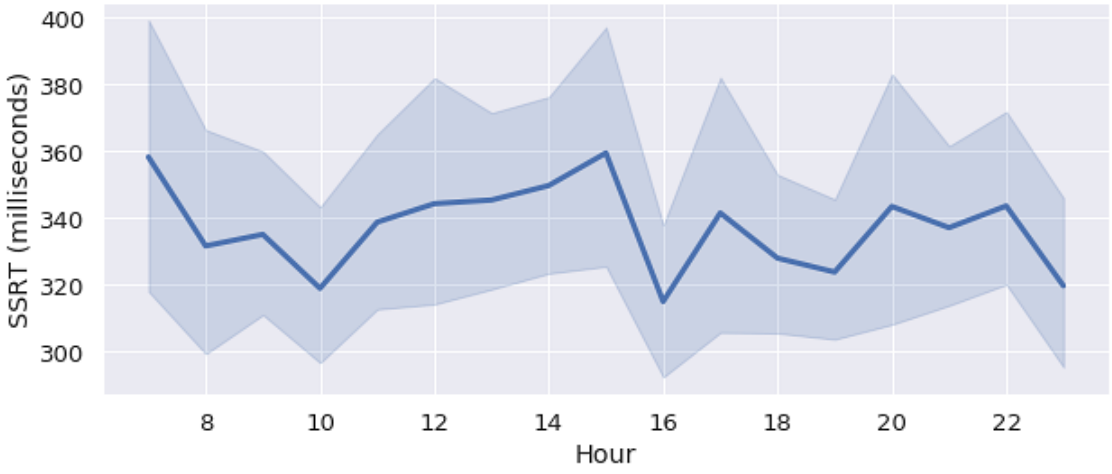
Table 2. Demographics of the participants (N=12).

Variable	Participants, n (%)
Sex	
Female	10 (83)
Male	2 (17)
Age (years)	
18-22	7 (58)
23-27	2 (17)
28-32	2 (17)
38-42	1 (8)

Table 3. The cumulative numbers of completed tasks at different times of day N=1107.

Time (hours)	Count, n (%)
7	26 (2)
8	34 (3)
9	55 (5)
10	61 (6)
11	65 (6)
12	66 (6)
13	82 (7)
14	68 (6)
15	80 (7)
16	84 (8)
17	73 (7)
18	65 (6)
19	73 (7)
20	64 (6)
21	76 (7)
22	73 (7)
23	60 (5)

Figure 3. Mean stop-signal reaction time at different times of day. The shaded area indicates 95% CIs. SSRT: stop-signal reaction time.



The distribution of participants' responses to the individual EMA questions and their SSRT are shown in Figures 4 and 5, respectively. Participant P9 had an overall shorter SSRT compared with the other participants. In the meantime, this participant also had a higher average stop unsuccessful rate. We found that all the EMA questions had strong RMCORR with each other (all the *P* values were <.001). Conversely, none of the EMA questions had significant RMCORR with SSRT, except that the responses for the EMA question "I am having trouble pulling myself together" had a marginal positive correlation with SSRT (RMCORR=0.041; 95% CI −0.01 to 0.09; *P*=.09). No statistically significant RMCORR between the aggregated EMA score (the scores for the negatively worded questions were first inverted before being added together) [28] and SSRT was found. Figure 6 shows the relationship between participants' aggregated EMA scores and their SSRTs.

Figure 4. Distribution of the responses for the individual ecological momentary assessment questions. EMA: ecological momentary assessment.

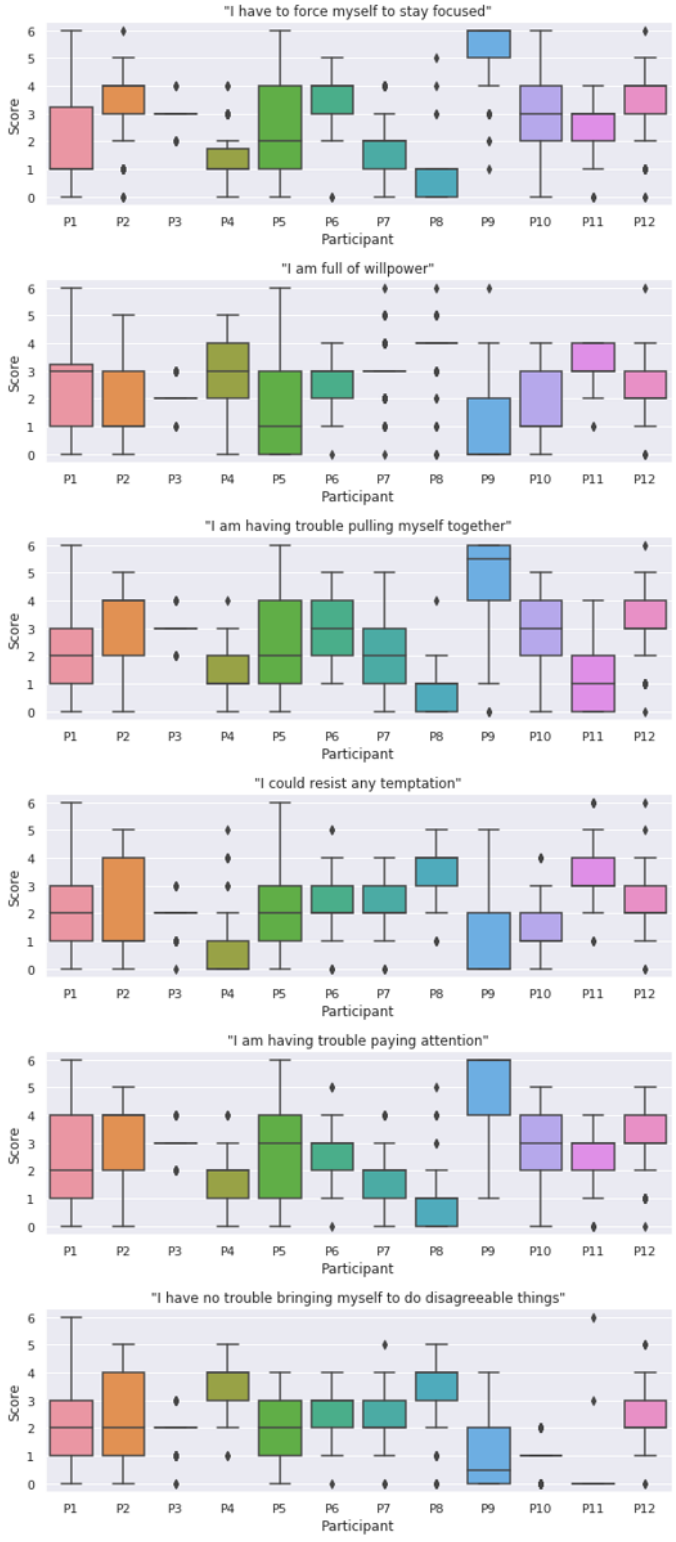


Figure 5. Distribution of participants' stop-signal reaction time. SSRT: stop-signal reaction time.

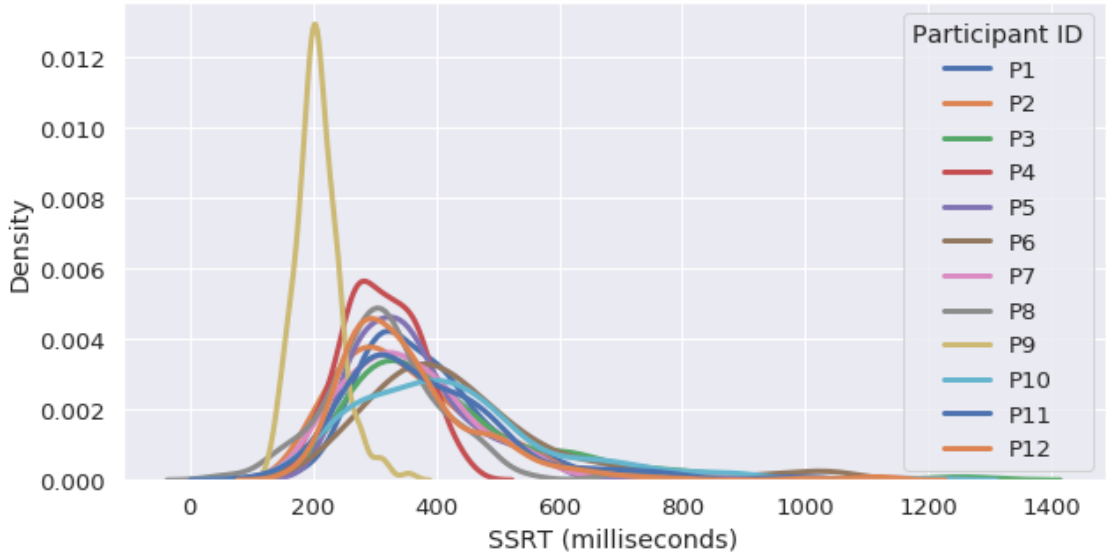
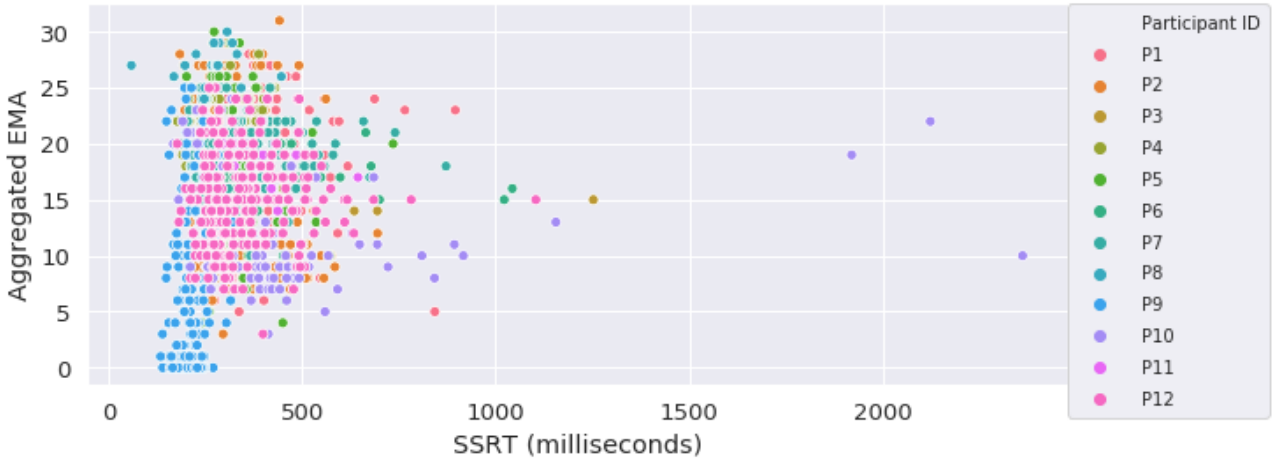


Figure 6. Relationship between participants' aggregated ecological momentary assessment scores and stop-signal reaction times. EMA: ecological momentary assessment; SSRT: stop-signal reaction times.



Features Predictive of Changes in Inhibitory Control

Table 4 shows the coefficients and their corresponding confidence intervals of the individual phone sensor features for predicting SSRT estimated by the GEE model. The features that had statistically significant effects on predicting SSRT include an individual's internal factors and factors associated with their activity, phone use, and the surrounding environment. It is worth nothing that an individual's SSRT and inhibitory control are inversely related. In other words, a higher SSRT implies a lower inhibitory control, and vice versa.

Table 4. The coefficients for all the phone sensor features for predicting stop-signal reaction time estimated by the generalized estimating equation model. A positive coefficient implies that the feature is associated with higher stop-signal reaction time (lower inhibitory control).

Feature	Estimated coefficient (95% CI)	*P* value
Individual inhibitory control baseline	1.04 (0.42 to 1.66)	.001
Phone use burstiness	0.0213 (0.006 to 0.036)	.005
Mean duration between 2 consecutive phone use sessions	4.791e-08 (8.11e-09 to 8.77e-08)	.02
Frequency of short phone use sessions	−0.0204 (−0.028 to −0.012)	<.001
Change rate of battery level when the phone was not charged	0.0479 (0.002 to 0.093)	.04
Frequency of incoming calls	0.1333 (0.011 to 0.256)	.03
Mean duration of incoming calls	−0.0006 (−0.001 to 0.000)	<.001
Standard deviation of acceleration	−0.4009 (−0.624 to −0.177)	<.001
Standard deviation of ambient noise frequency	−3.811e-05 (−7.04e-05 to −5.83e-06)	.02
Mean decibel level of ambient noise	−0.5267 (−0.912 to −0.141)	.007
Percentage of time connected to the internet through a mobile network	−0.1717 (−0.270 to −0.074)	.001
Jaccard coefficient of location similarity	−0.0415 (−0.078 to −0.005)	.03
In outdoor or recreational locations during the hour before the past hour	−0.0682 (−0.120 to −0.016)	.01
In residential buildings during the hour before the past hour	0.0719 (0.001 to 0.142)	.045
In residential buildings during the past hour	0.1185 (0.002 to 0.235)	.046
Mean change rate of battery level when the phone was charged	−0.0074 (−0.016 to 0.093)	.08
Frequency of charging	0.0116 (−0.037 to 0.060)	.64
Total duration of charging	1.084e-05 (−2.26e-05 to 4.43e-05)	.53
Hours of sleep	−0.0102 (−0.023 to 0.003)	.12
Mean phone use duration	3.971e-08 (−6.78e-09 to 8.62e-08)	.09
Standard deviation of phone use durations	3.536e-08 (−3.29e-08 to 1.04e-07)	.31
Standard deviation of durations between 2 consecutive screen unlocks	6.596e-08 (−2.11e-08 to 1.53e-07)	.14
Frequency of outgoing calls	0.0151 (−0.026 to 0.056)	.47
Total duration of outgoing calls	−6.287e-05 (−0.001 to 0)	.50
Percentage of time being stationary	−0.1276 (−0.275 to 0.020)	.09
Standard deviation of the decibel level of ambient noise	0.0282 (−0.294 to 0.351)	.86
Mean RMS[a] level of ambient noise	0.0003 (−0.001 to 0.001)	.50
Standard deviation of the RMS level of ambient noise	−0.0002 (−0.002 to 0.001)	.77
In locations for arts or entertainment during the past hour	−0.0259 (−0.157 to 0.105)	.70
In professional or other places during the past hour	0.1102 (−0.028 to 0.248)	.12

[a]RMS: root mean square.

Specifically, in terms of internal factors, individual inhibitory control baseline had a significant positive effect on predicting SSRT (95% CI 0.421-1.662; *P*=.001), that is, lower trait inhibitory control (higher self-reported score in BIS) was associated with higher SSRT (lower inhibitory control). Regarding activity, the standard deviation of acceleration had a significant negative effect (95% CI −0.624 to −0.177; *P*<.001); larger changes in one's movement were associated with higher inhibitory control. As for phone use, phone use burstiness (95% CI 0.006-0.036; *P*=.005), the mean duration between 2 consecutive phone use sessions (95% CI 8.11e-09-8.77e-08; *P*=.02), the change rate of battery level when the phone was not charged (95% CI 0.002-0.093; *P*=.04), and the frequency of incoming calls (95% CI 0.011-0.256; *P*=.03) had significant positive effects on predicting SSRT (associated with lower inhibitory control), whereas the frequency of short phone use sessions (95% CI −0.028 to −0.012; *P*<.001) and the mean duration of incoming calls (95% CI −0.001 to 0.000; *P*<.001) had significant negative effects (associated with higher inhibitory control).

With regard to surrounding environments, the standard deviation of the frequency of ambient noise (95% CI −7.04e-05 to −5.83e-06; *P*=.02), the mean decibel level of ambient noise (95% CI −0.912 to −0.141; *P*=.007), the percentage of time in which the phone was connected to the internet through a mobile

network (95% CI −0.027 to −0.074; *P*=.001), the Jaccard similarity coefficient for the locations visited during the past 2 hours (95% CI −0.078 to −0.005; *P*=.03), and whether an outdoor recreational place was visited during the past hour (95% CI −0.12 to −0.016; *P*=.01) had significant negative effects on predicting SSRT (associated with higher inhibitory control), whereas whether a participant was in a residential building in the past hour (95% CI 0.02-0.235; *P*=.046) and the hour before (95% CI 0.001-0.142; *P*=.045) had significant positive effects (associated with lower inhibitory control). The GEE model had an R^2 score of 0.22.

Predicting High and Low Inhibitory Control

The mean accuracy and ROC-AUC of leave-one-subject-out cross-validation were 55.4% and 57.4%, respectively, for models trained only with phone sensor features (Table 5). To further examine whether including partial data from an unseen participant can improve a model's performance, we trained mixed models using different amounts (from 10% to 30%) of the test participant's data. In other words, during each iteration, a certain amount of the test participant's data along with the other participants' data were used for training, and the model was tested on the remaining test participant's data (which was not seen by the model during the training). The ROC-AUC slightly increased after more data from each test participant were included for training.

Table 5. Comparison of prediction accuracy and the area under the receiver operating characteristic curve of the different machine learning models (%).

Metrics	Model				
	Baseline[a]	LOSO[b]	Mixed (10%[c])	Mixed (20%)	Mixed (30%)
ACC[d]	50.0	55.4	52.8	53.0	53.4
AUC-ROC[e]	50.0	57.4	57.9	58.3	60.0

[a]Model that gives predictions by chance.

[b]LOSO: leave-one-subject-out.

[c]The portion of data from the test participant used for training.

[d]ACC: accuracy.

[e]AUC-ROC: the area under the receiver operating characteristic curve.

Features Predictive of Inhibitory Control After Including HRV Features

To include HRV features for our analysis, we first removed data points that contained missing or incomplete HRV features (details about data preprocessing are given in the *Methods* section). After data cleaning, we had a total of 577 data points. For comparison, we applied the same set of analyses as we did for phone sensor features. After removing features that had high collinearity, the features that were included in the analysis were the standard deviation of heart rate, the standard deviation of the average NN intervals during each 5-minute segment (SDANN), the root mean square of the successive differences (RMSSD), and the relative power of high frequency (HF) bands.

The predictive features after HRV features were taken into account are similar to the predictive features given by the GEE model fitted with phone sensor features. Specifically, individual inhibitory control baseline had a significant positive effect on predicting SSRT (95% CI 0.177-1.498; *P*=.01), that is, lower trait inhibitory control was associated with lower inhibitory control. Regarding activity, the standard deviation of acceleration had a significant negative effect on predicting SSRT (95% CI −0.518 to −0.052; *P*=.02), whereas the percentage of time being stationary had a significant positive effect (95% CI 0.012-0.229; *P*=.03). In other words, larger changes in one's movements were associated with higher inhibitory control. With regard to phone use, the mean phone use duration (95% CI 1.77e-08-1.05e-07; *P*=.006) and the mean change rate of battery level when the phone was not charged (95% CI 0.005-0.164; *P*=.004) had significant positive effects on predicting SSRT (associated with lower inhibitory control), whereas the mean duration of incoming calls (95% CI −0.001 to 0.000; *P*=.007) and the number of short sessions (95% CI −0.047 to −0.025; *P*<.001) had significant negative effects (associated with higher inhibitory control). With regard to surrounding environments (also start this sentence with a new paragraph), the standard deviation of the frequency of ambient noise (95% CI −9.16e-05 to −3.01e-05; *P*<.001), the percentage of time the phone was connected to the internet through a mobile network (95% CI −0.249 to −0.039; *P*=.007), and the Jaccard coefficient of location similarity between 2 consecutive hours (95% CI −0.095 to −0.041; *P*<.001) had significant negative effects (associated with higher inhibitory control). In addition, we found that hours of sleep (95% CI −0.034 to −0.005; *P*=.008) and the standard deviation of heart rate (95% CI −7.62e-05 to 0.000; *P*=.002) also had significant effects on predicting SSRT. The estimated coefficients for the individual features are summarized in Table 6. The GEE model had an R^2 score of 0.25. Similarly, we trained GBT classifiers to examine whether adding additional HRV features could improve the performance of the models. With leave-one-subject-out cross-validation, the mean AUC-ROC slightly improved to 0.62.

Table 6. The coefficients for all the phone sensor and heart rate variability features for predicting stop-signal reaction time estimated by the generalized estimating equation model. A positive coefficient implies that the feature is associated with higher stop-signal reaction time (lower inhibitory control).

Feature	Estimated coefficient (95% CI)	*P* value
Individual inhibitory control baseline	0.8378 (0.177 to 1.498)	.01
Hours of sleep	−0.0193 (−0.034 to −0.005)	.008
Phone use burstiness	0.0357 (0.023 to 0.048)	<.001
Mean phone use duration	6.148e-08 (1.77e-08 to 1.05e-07)	.006
Mean duration between 2 consecutive phone use sessions	8.787e-08 (5.23e-08 to 1.23e-07)	<.001
Frequency of short phone use sessions	−0.0357 (−0.047 to −0.025)	<.001
Change rate of battery level when the phone was not charged	0.0843 (0.005 to 0.164)	.04
Mean duration of incoming calls	−0.0005 (−0.001 to 0)	.007
Percentage of time being stationary	0.1205 (0.012 to 0.229)	.03
Standard deviation of acceleration	−0.2852 (−0.518 to −0.052)	.02
Standard deviation of ambient noise frequency	−6.081e-05 (−9.16e-05 to −3.01e-05)	<.001
Percentage of time connected to the internet through a mobile network	−0.1436 (−0.249 to −0.039)	.007
Jaccard coefficient of location similarity	−0.0682 (−0.095 to −0.041)	<.001
Standard deviation of heart rate	−0.0002 (−7.62e-05 to 0)	.002
Frequency of incoming calls	0.13 (−0.01 to 0.28)	.08
Frequency of outgoing calls	0.0249 (−0.023 to 0.073)	.31
Total duration of outgoing calls	−6.084e-05 (−0.001 to 0.000)	.58
Standard deviation of phone use durations	7.581e-08 (−1.72e-08 to 1.69e-07)	.11
Standard deviation of durations between 2 consecutive screen unlocks	7.451e-08 (−4.62e-08 to 1.95e-07)	.23
Frequency of charging	0.0038 (−0.037 to 0.045)	.86
Total duration of charging	−7.94e-06 (−4.78e-05 to 3.2e-05)	.70
Mean change rate of battery level when the phone was charged	0.0174 (−0.034 to 0.069)	.51
Mean decibel level of ambient noise	−0.2963 (−0.663 to 0.070)	.11
Standard deviation of the decibel level of ambient noise	0.1669 (−0.305 to 0.639)	.49
Mean RMS[a] level of ambient noise	−0.0003 (−0.002 to 0.001)	.70
Standard deviation of the RMS level of ambient noise	−0.0001 (−0.002 to 0.001)	.87
In outdoor or recreational locations during the hour before the past hour	−0.0123 (−0.044 to 0.019)	.44
In residential buildings during the hour before the past hour	0.1663 (−0.030 to 0.363)	.10
In locations for arts or entertainment during the past hour	−0.0816 (−0.218 to 0.054)	.24
In professional or other places during the past hour	0.0688 (−0.134 to 0.271)	.51
In residential buildings during the past hour	−0.0608 (−0.234 to 0.113)	.49
SDANN[b]	−0.0003 (−0.001 to 0.000)	.32
RMSSD[c]	−0.0003 (−0.001 to 0.001)	.58
Relative power of FFT[d] high-frequency band	0.0011 (−0.003 to 0.005)	.59

[a]RMS: root mean square.

[b]SDANN: standard deviation of the average NN.

[c]RMSSD: root mean square of the successive differences.

[d]FFT: Fast Fourier transform.

Discussion

Principal Findings

The aim of this study is to investigate the feasibility of using smartphones and wearable sensor data to unobtrusively assess an individual's inhibitory control in the wild. We conducted a 4-week longitudinal study with 12 participants, collecting their sensor data continuously along with more than 1000 SST responses. We fitted GEE models with features extracted from their sensor data to analyze the main effects of the individual features and trained GBT classifiers to examine whether it is possible to classify different states of inhibitory control using these predictive features.

Activity

The results showed that higher levels of physical activity, which was reflected in their accelerometer data, had a positive association with an individual's inhibitory control. This corroborates findings from the literature regarding the link between inhibitory control and exercise across different individuals [17,57]. In our study, the link between exercise and inhibitory control was also observed at an individual level. With more granular information on an individual's physical activity, we can even potentially investigate how the intensity and the duration of physical activity impact an individual's inhibitory control and how long the effects last in the future.

Phone Use

We found that certain phone use patterns were associated with changes in inhibitory control. For example,higher burstiness and longer mean intervals between 2 consecutive phone use sessions within an hour were positively associated with SSRT, which suggests that more frequent and longer phone use might be linked to a decrease in inhibitory control. However, interestingly, the number of short phone use sessions (sessions less than 30 seconds) had a positive main effect on predicting an individual's inhibitory control, which suggests that short phone use sessions might be less detrimental than longer ones. We suspect that the short and longer phone uses might be associated with different types of phone interruptions: endogenous interruptions (sometimes referred to as internal interruptions) and exogenous interruptions (sometimes referred to as external interruptions), respectively [58]. An example of endogenous interruptions is self-initiated phone activities, such as checking one's social media, which usually leads to a chain of other phone activities, whereas exogenous interruptions often involve phone activities resulting from external interruptions, such as receiving phone-related alerts or notifications [59,60]. Therefore, they might impact one's inhibitory control differently.

Environment

The estimated coefficients for the different types of environments show that environments also play an essential role in predicting an individual's inhibitory control. Specifically, outdoor environments, state parks for instance, had a positive association with an individual's inhibitory control, whereas being in a residential building, such as being at home or in a dorm room, had a negative association with their inhibitory control. Moreover, the mean battery change rate when the phone was not charged and the percentage of time a user's phone was connected to a mobile network, both of which are indicators of whether a user was outdoors, also had positive associations with their inhibitory control. Taken together, the results suggest that higher inhibitory control might be associated with people spending more time outdoors, which corroborates the findings of previous studies [19,61,62]. Future studies are needed to further examine the causal relationship between environments and inhibitory control by considering other factors, such as personality types [63].

HRV

According to the literature, heart rate and HRV features, particularly the relative HF power, have been found to be related to different levels of inhibitory control. We collected participants' interbeat intervals over the course of the study. However, the standard deviation of heart rate was the only feature that was found to have a significant main effect on predicting the participants' inhibitory control. We suspected that this may be due to some confounding factors, which were difficult to control during in-the-wild studies, such as the environment, posture, and physical activity [64]. For example, although both exercise and higher HRV contribute to higher inhibitory control, an individual's HRV actually decreases during and after exercise due to sympathetic activation and parasympathetic withdrawal [65-67]. For future work, the information on the intensity of and the time since last physical activity will be helpful for assessing one's inhibitory control more accurately.

Sleep

Studies have shown that sleep plays an essential role in inhibitory control [18]. However, in our study, sleep duration was only found to have a significant effect in predicting SSRT in the second GEE model. As hours of sleep were inferred based on a participant's screen locks and unlocks, sometimes there might be some noise in the data (eg, a participant might wake up and check his or her phone in the middle of the night), which consequently affected the results of sleep inferencing. Future studies can combine more accurate sleep data collected through other sleep trackers, such as Oura Ring [68], along with mobile sensor data to better understand the relationship between sleep and inhibitory control.

Relationship Between Subjective and Objective Measurements

In this study, we employed both subjective and objective measurements to assess and compare participants' responses for both types of measurements. We did not find any significant correlation between their subjective and objective responses. The distributions of different participants' EMA responses were very distinct, whereas the distributions of their responses for objective measurements, SST, were relatively similar. When we looked at the correlations by location, we found that the responses for the EMA question "I am having trouble pulling myself together" had a mild positive correlation with SSRT (RMCORR=0.098; 95% CI 0.02 to 0.18; *P*=.02) in *College and University* locations. Also, in professional places (eg, office building), the responses for the EMA question "I am full of

willpower" also had a slight negative correlation with SSRT (RMCORR=−0.156; 95% CI −0.31 to 0.01; *P*=.06). In other words, on campus or in the workplace, people might feel that it was harder for them to pull themselves together or had less willpower when their inhibitory control was lower. On the contrary, in outdoor or recreational places, the responses for the EMA question "I have no trouble bringing myself to do disagreeable things" had a positive correlation with SSRT (RMCORR=0.138; 95% CI 0.01-0.27; *P*=.04). In nightlife spots (eg, bars), the responses to the EMA question "I am having trouble paying attention" had a marginal negative correlation with SSRT (RMCORR=−0.216; 95% CI −0.44 to 0.03; *P*=.09).

The little correlation between subjective and objective measurements was also observed in some recent studies. Some plausible explanations for the inconsistency suggested by the researchers include (1) questions in self-report scales may capture multiple constructs rather than inhibitory control alone, whereas inhibitory control tasks typically measure a narrower construct; (2) behavioral inhibition tasks, including SSTs, were designed to minimize between-subject variability to achieve high reliability (which corroborates our findings that the distributions of participants' SSRTs were fairly similar), a phenomenon called reliability paradox, which limits the tasks' ability to capture individual differences; (3) what inhibitory control tasks measure might be participants' *maximum* rather than their *actual* inhibitory control at that moment, as participants had been instructed to perform as accurately as possible; (4) there might be some publication bias in the earlier literature that overestimated the correlation between self-reports and inhibitory control tasks [69-71]. In addition to these explanations, we also suspect that there might be some discrepancy between the perceived level of inhibitory control and the actual level of their inhibitory control. In other words, people might not be aware of their diminished inhibitory control and still report high inhibitory control. Further, their self-reported inhibitory control might be biased by the level of their inhibitory control at the moment [31]. The gap between one's perceived and actual inhibitory control might depend on different contexts, such as locations. Future studies that combine both automatic and manual location logging will allow collection of more accurate information on semantic locations and help examine how different types of locations influence the way people perceive and self-report their inhibitory control.

In summary, we identified several markers that can be used to infer inhibitory control in the wild. However, these are only a subset of markers that manifest the effects of the factors modulating an individual's inhibitory control. There is information on other factors, such as diet, stimulant intake, and exposure to natural light, which is difficult to directly and accurately measure using smartphone sensors, which in turn limits the performance of our prediction models. That said, being able to track patterns that are associated with inhibitory control is still of great use in many real-world applications, such as productivity management technologies. Showing statistics about users' phone use patterns, physical activities, location patterns, and inferred levels of inhibitory control during the past few weeks can provide insights into how their inhibitory control fluctuates over time. As such, users can adjust their behaviors to change the pattern of their inhibitory control. Alternatively, they can schedule their daily tasks according to their inhibitory control patterns so that they work on tasks requiring a great level of attention when they are in states of high inhibitory control.

Potential Use Scenarios

Workplace Productivity and Well-Being

Previous studies have shown that there are contexts in the workplace where people are more susceptible to distractions. For example, people are more likely to cyberloaf after returning from a break [72]. Therefore, some intervention designs have been proposed to help mitigate the proneness to distractors [73,74]. The triggering moments of the interventions are usually rule based, for instance, the first 15 min after a physical break or pomodoro technique. However, the amount of time it takes for an individual to get to a focused state, namely, a high inhibitory control state, might vary from person to person depending on factors such as the environment and the type of work. If we can integrate the technology of unobtrusive inhibitory control assessment into those intervention technologies, it will allow such intervention technologies to more accurately assess an individual's focus state and provide more personalized interventions based on that information. For example, an intervention system can infer the level of a user's inhibitory control based on their phone use patterns, such as patterns indicating self-interrupts, and restrict their access to contents or applications that might render them vulnerable to a chain of distractions, or even delay the arrival of notifications, to keep them stay focused.

Addictions

There has been growing awareness of digital well-being and phone addiction. Android and iOS have included screen time to help users restrain the amount of time they spend on their phones and prevent excessive phone use. However, prolonged phone use time is just one of the behaviors [75] that indicates addictive phone use. Other behaviors, such as frequently checking the phone, might also be indicators of excessive phone use. As phone addiction is highly related to the level of an individual's inhibitory control [76], tracking a person's inhibitory control continuously and prompting them to take preemptive steps when they have low inhibitory control will be more effective in preventing phone addiction.

In addition to excessive phone use, there are other types of addictive behaviors, such as food addiction and substance dependence, which also result from reduced inhibitory control. Detecting the downward trends in those individuals' inhibitory control and the signatures and contexts that are associated with the decrease in their inhibitory control can prevent them from falling off the wagon by suggesting that they take some preemptive steps. For example, a system can track whether a user is in an environment that is likely to trigger alcohol craving based on the type of location and the loudness in the surrounding environment [77] and suggest that the person take an alternative route when he or she is in proximity to that type of environment.

Mental Disorders and Illnesses

The level of inhibitory control is also related to which episode an individual with mental disorders is in. For example, studies have shown that patients with bipolar disorder had impaired inhibitory control when they were in their manic and depressive episodes [78]. As such, a system that is able to continuously track the level of a bipolar patient's inhibitory control can potentially help detect the early signs of mood swings and provide timely interventions.

Limitations and Future Work

The size of our study was relatively small and homogeneous (the participants were mainly recruited from a college campus and were limited to iOS users). Follow-up studies with a large sample size and more diverse demographics will need to be conducted to examine how different demographic factors impact inhibitory control. Other information, such as light and app use, which we hypothesized would be useful for inferring inhibitory control, could not be collected on iOS devices (because the majority of the participants on campus were iOS users). The phone use patterns of users who installed productivity apps might be different from the phone use patterns of those who did not. Besides, outliers (of SST performance and inferred sleep) were removed from our analysis. Further research is needed to investigate whether such extreme values will affect the assessment of inhibitory control. For future work, in addition to phone sensor and heart rate data, we are planning to collect data using other types of wearables, such as fitness or sleep trackers, to obtain more granular information, such as physical activity, sleep quality, and chronotype, to account for the influences of these confounding factors. Being able to account for these pieces of information can potentially improve models' performance for inferring different states of inhibitory control.

Conclusions

This paper presents a preliminary study using mobile and wearable sensor data to assess individuals' inhibitory control in the wild. To the best of our knowledge, this is the first in-the-wild study aimed to develop a new approach to accessing inhibitory control unobtrusively and continually. On the basis of the data from 12 participants during a 4-week study, we found that factors including an individual's inhibitory control baseline, the number of phone use sessions, the change rate of battery level, the mean duration of incoming calls, physical activity, the mean frequency and loudness of ambient noise, whether a user was outdoor, and the types of locations a user visited had significant main effects on predicting an individual's inhibitory control. We also trained a machine learning model to predict whether an individual had high or low inhibitory control at different moments using those features, which achieved an AUC-ROC of 60%. The findings provide some insights into how we can potentially design personalized technologies that can help support and manage productivity in school or the workplace and intervention tools that can prevent individuals with mental disorders such as substance abuse from relapsing or engaging in problematic behaviors.

Acknowledgments

The study was funded by the National Science Foundation (grant number 1840025). The authors would like to thank Daniel A Adler for helpful comments on the earlier drafts of this paper.

Conflicts of Interest

None declared.

References

1. Reck SG, Hund AM. Sustained attention and age predict inhibitory control during early childhood. J Exp Child Psychol 2011 Mar;108(3):504-512. [doi: 10.1016/j.jecp.2010.07.010] [Medline: 20801457]
2. Passolunghi M, Siegel LS. Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. J Exp Child Psychol 2001 Sep;80(1):44-57. [doi: 10.1006/jecp.2000.2626] [Medline: 11511134]
3. Bartholomew ME, Heller W, Miller GA. Inhibitory control of emotional processing: theoretical and empirical considerations. Int J Psychophysiol 2019 Mar 29:- epub ahead of print. [doi: 10.1016/j.ijpsycho.2019.03.015] [Medline: 30936042]
4. MacCann C, Roberts RD. Do time management, grit, and self-control relate to academic achievement independently of conscientiousness? APA PsycInfo 2010:79-90 [FREE Full text]
5. Spinella M, Miley WM. Impulsivity and academic achievement in college students. Coll Stud J 2003;37(4):545-549 [FREE Full text]
6. Foy JG, Mann VA. Executive function and early reading skills. Read Writ 2012 Apr 25;26(3):453-472. [doi: 10.1007/s11145-012-9376-5]
7. Bock A, Gallaway K, Hund A. Specifying links between executive functioning and theory of mind during middle childhood: cognitive flexibility predicts social understanding. J Cogn Dev 2014 Apr 3;16(3):509-521. [doi: 10.1080/15248372.2014.888350]
8. Tangney JP, Baumeister RF, Boone AL. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. J Pers 2004 Apr;72(2):271-324. [doi: 10.1111/j.0022-3506.2004.00263.x] [Medline: 15016066]
9. Duckworth AL, Taxer JL, Eskreis-Winkler L, Galla BM, Gross JJ. Self-control and academic achievement. Annu Rev Psychol 2019 Jan 4;70:373-399. [doi: 10.1146/annurev-psych-010418-103230] [Medline: 30609915]

10. Cherian J, Jacob J. Impact of self efficacy on motivation and performance of employees. Int J Bus 2013 Jun 19;8(14). [doi: 10.5539/ijbm.v8n14p80]
11. Fino E, Melogno S, Iliceto P, D'Aliesio S, Pinto M, Candilera G, et al. Executive functions, impulsivity, and inhibitory control in adolescents: a structural equation model. Adv Cogn Psychol 2014;10(2):32-38 [FREE Full text] [doi: 10.5709/acp-0154-5] [Medline: 25157298]
12. Galaverna FS, Morra CA, Bueno AM. Attention in patients with chronic schizophrenia: deficit in inhibitory control and positive symptoms. Eur. J. Psychiat 2012 Sep;26(3):185-195. [doi: 10.4321/s0213-61632012000300005]
13. Claes L, Nederkoorn C, Vandereycken W, Guerrieri R, Vertommen H. Impulsiveness and lack of inhibitory control in eating disorders. Eat Behav 2006 Aug;7(3):196-203. [doi: 10.1016/j.eatbeh.2006.05.001] [Medline: 16843221]
14. Reinhard M, Allen N, Wong L, Schwartz B. Neuropsychological measurement of inhibitory control in posttraumatic stress disorder: an exploratory antisaccade paradigm. J Clin Exp Neuropsychol 2017 Dec;39(10):1002-1012. [doi: 10.1080/13803395.2017.1301389] [Medline: 28349778]
15. Larson ER, Shear P, Krikorian R, Welge J, Strakowski SM. Working memory and inhibitory control among manic and euthymic patients with bipolar disorder. J Int Neuropsychol Soc 2005 Mar;11(2):163-172. [doi: 10.1017/s1355617705050228] [Medline: 15962704]
16. Müller U, Rowe J, Rittman T, Lewis C, Robbins T, Sahakian B. Effects of modafinil on non-verbal cognition, task enjoyment and creative thinking in healthy volunteers. Neuropharmacology 2013 Jan;64:490-495 [FREE Full text] [doi: 10.1016/j.neuropharm.2012.07.009] [Medline: 22820554]
17. Boat R, Cooper SB. Self-control and exercise: a review of the bi-directional relationship. Brain Plast 2019 Dec 26;5(1):97-104 [FREE Full text] [doi: 10.3233/BPL-190082] [Medline: 31970063]
18. Rossa KR, Smith SS, Allan AC, Sullivan KA. The effects of sleep restriction on executive inhibitory control and affect in young adults. J Adolesc Health 2014 Aug;55(2):287-292. [doi: 10.1016/j.jadohealth.2013.12.034] [Medline: 24602612]
19. van der Wal AJ, Schade HM, Krabbendam L, van Vugt M. Do natural landscapes reduce future discounting in humans? Proc Biol Sci 2013 Dec 22;280(1773):20132295 [FREE Full text] [doi: 10.1098/rspb.2013.2295] [Medline: 24197412]
20. Berry M, Sweeney M, Morath J, Odum A, Jordan K. The nature of impulsivity: visual exposure to natural environments decreases impulsive decision-making in a delay discounting task. PLoS One 2014;9(5):e97915 [FREE Full text] [doi: 10.1371/journal.pone.0097915] [Medline: 24841421]
21. Riley E, Esterman M, Fortenbaugh FC, DeGutis J. Time-of-day variation in sustained attentional control. Chronobiol Int 2017;34(7):993-1001. [doi: 10.1080/07420528.2017.1308951] [Medline: 28622051]
22. Jones A, Tiplady B, Houben K, Nederkoorn C, Field M. Do daily fluctuations in inhibitory control predict alcohol consumption? An ecological momentary assessment study. Psychopharmacology (Berl) 2018 May;235(5):1487-1496 [FREE Full text] [doi: 10.1007/s00213-018-4860-5] [Medline: 29497782]
23. Jones A, Christiansen P, Nederkoorn C, Houben K, Field M. Fluctuating disinhibition: implications for the understanding and treatment of alcohol and other substance use disorders. Front Psychiatry 2013 Oct 22;4:140 [FREE Full text] [doi: 10.3389/fpsyt.2013.00140] [Medline: 24155728]
24. Baumeister RF. Self‐control, fluctuating willpower, and forensic practice. J Forensic Psychol Pract 2013 May 10;15(2):85-96. [doi: 10.1108/14636641311322278]
25. Band GP, van der Molen MW, Logan GD. Horse-race model simulations of the stop-signal procedure. Acta Psychol (Amst) 2003 Feb;112(2):105-142. [doi: 10.1016/s0001-6918(02)00079-3] [Medline: 12521663]
26. Charles-Walsh K, Furlong L, Munro DG, Hester R. Inhibitory control dysfunction in nicotine dependence and the influence of short-term abstinence. Drug Alcohol Depend 2014 Oct 1;143:81-86. [doi: 10.1016/j.drugalcdep.2014.07.008] [Medline: 25070928]
27. Krypotos A, Jahfari S, van Ast VA, Kindt M, Forstmann BU. Individual differences in heart rate variability predict the degree of slowing during response inhibition and initiation in the presence of emotional stimuli. Front Psychol 2011;2:278 [FREE Full text] [doi: 10.3389/fpsyg.2011.00278] [Medline: 22059080]
28. Schöndube A, Bertrams A, Sudeck G, Fuchs R. Self-control strength and physical exercise: an ecological momentary assessment study. Psychol Sport Exerc 2017 Mar;29:19-26. [doi: 10.1016/j.psychsport.2016.11.006]
29. Kim EJ, Namkoong K, Ku T, Kim SJ. The relationship between online game addiction and aggression, self-control and narcissistic personality traits. Eur Psychiatry 2008 Apr;23(3):212-218. [doi: 10.1016/j.eurpsy.2007.10.010] [Medline: 18166402]
30. Piquero A, Rosay A. The reliability and validity of Grasmick et al.'s self-control scale: a comment on Longshore et al. Criminology 1998 Feb;36(1):157-174. [doi: 10.1111/j.1745-9125.1998.tb01244.x]
31. Piquero AR, Macintosh R, Hickman M. Does self-control affect survey response? Applying exploratory, confirmatory, and item response theory analysis to Grasmick et al.'s self-control scale*. Criminology 2000 Aug;38(3):897-930. [doi: 10.1111/j.1745-9125.2000.tb00910.x]
32. Hamatani T, Ochiai K, Inagaki A, Yamamoto N, Fukazawa Y, Kimoto M, et al. Automated Inference of Cognitive Performance by Fusing Multimodal Information Acquired by Smartphone. In: Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International

Symposium on Wearable Computers. 2019 Presented at: UbiComp/ISWC'19; September 9-13, 2019; London, UK. [doi: 10.1145/3341162.3346275]
33. Tseng V. InhibiSense. GitHub. 2019. URL: https://github.com/vtseng/inhibition-sensing.git [accessed 2020-11-20]
34. Ferreira D, Kostakos V, Dey AK. AWARE: mobile context instrumentation framework. Front ICT 2015 Apr 20;2. [doi: 10.3389/fict.2015.00006]
35. Polar H10 Heart Rate Sensor. Polar USA: Heart Rate Monitors, Activity Trackers and Bike. URL: https://www.polar.com/us-en/products/accessories/h10_heart_rate_sensor [accessed 2020-11-20]
36. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
37. Band G, Van DM, Logan G. Horse-Race Model Simulations of the stop-Signal Procedure. Acta Psychologica. 2003. URL: www.elsevier.com/locate/actpsy [accessed 2020-11-24]
38. Stanford MS, Mathias CW, Dougherty DM, Lake SL, Anderson NE, Patton JH. Fifty years of the Barratt Impulsiveness Scale: an update and review. Pers Individ Differ 2009 Oct;47(5):385-395. [doi: 10.1016/j.paid.2009.04.008]
39. CMMotionActivity. Apple Developer Documentation. URL: https://developer.apple.com/documentation/coremotion/cmmotionactivity [accessed 2020-11-20]
40. Abdullah S, Murnane E, Matthews M, Kay M, Kientz J, Gay G, et al. Cognitive Rhythms: Unobtrusive and Continuous Sensing of Alertness Using a Mobile Phone. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Presented at: UbiComp'16; September 12-16, 2016; Heidelberg. [doi: 10.1145/2971648.2971712]
41. Documentation. Foursquare Developers. URL: https://developer.foursquare.com/docs/build-with-foursquare/categories/ [accessed 2020-11-20]
42. Tulloss RE. Assessment of Similarity Indices for Undesirable Properties and a new Tripartite Similarity Index Based on Cost Functions. Semantic Scholar. 1997. URL: https://pdfs.semanticscholar.org/a531/345cded9d0dc5d2f4f94d1f93f4a6239f4cf.pdf [accessed 2020-11-24]
43. Choi SS, Cha SH, Tappert CC. A Survey of Binary Similarity and Distance Measures. Systemics, Cybernetics and Informatics. 2010. URL: http://www.iiisci.org/journal/CV$/sci/pdfs/GS315JG.pdf [accessed 2020-11-24]
44. Chao A, Chazdon R, Colwell R, Tsung-Jen S. Abundance-based similarity indices and their estimation when there are unseen species in samples. Biometrics 2006 Jun;62(2):361-371. [doi: 10.1111/j.1541-0420.2005.00489.x] [Medline: 16918900]
45. Alroy J. A new twist on a very old binary similarity coefficient. Ecology 2015 Feb;96(2):575-586 [FREE Full text] [doi: 10.1890/14-0471.1]
46. Abdullah S, Matthews M, Murnane E, Gay G, Choudhury T. Towards Circadian Computing: 'Early to Bed and Early to Rise' Makes Some of Us Unhealthy and Sleep Deprived. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp'14; September 13-17, 2014; Seattle. [doi: 10.1145/2632048.2632100]
47. Kemper K, Hamilton C, Atkinson M. Heart rate variability: impact of differences in outlier identification and management strategies on common measures in three clinical populations. Pediatr Res 2007 Sep;62(3):337-342. [doi: 10.1203/PDR.0b013e318123fbcc] [Medline: 17597640]
48. Pichot V, Roche F, Celle S, Jean-Claude B, Chouchou F. HRVanalysis: a free software for analyzing cardiac autonomic activity. Front Physiol 2016;7:557. [doi: 10.3389/fphys.2016.00557] [Medline: 27920726]
49. Gomes P. pyHRV-OpenSource Python Toolbox for Heart Rate Variability Documentation. pyHRV Documentation. 2019. URL: https://pyhrv.readthedocs.io/en/latest/ [accessed 2020-11-20]
50. Congdon E, Mumford JA, Cohen JR, Galvan A, Canli T, Poldrack RA. Measurement and reliability of response inhibition. Front Psychol 2012;3:37. [doi: 10.3389/fpsyg.2012.00037] [Medline: 22363308]
51. Bakdash JZ, Marusich LR. Repeated measures correlation. Front Psychol 2017 Apr 7;8:-. [doi: 10.3389/fpsyg.2017.00456]
52. Vallat R. Pingouin: statistics in Python. JOSS 2018 Nov;3(31):1026. [doi: 10.21105/joss.01026]
53. Hardin J. Generalized estimating equations (GEE). In: Encyclopedia of Statistics in Behavioral Science. Chichester, UK: John Wiley & Sons, Ltd; 2005.
54. Dormann C, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2012 May 18;36(1):27-46. [doi: 10.1111/j.1600-0587.2012.07348.x]
55. Friedman J. Greedy function approximation: a gradient boosting machine internet. Ann Stat 2001:- epub ahead of print. [doi: 10.2307/2699986]
56. Davis J, Goadrich M. The Relationship Between Precision-recall and ROC Curves. In: Proceedings of the 23rd international conference on Machine learning. 2006 Presented at: CML'06; June 25-29, 2006; Pittsburgh, UK. [doi: 10.1145/1143844.1143874]
57. Sofis M, Carrillo A, Jarmolowicz D. Maintained physical activity induced changes in delay discounting. Behav Modif 2017 Jul;41(4):499-528. [doi: 10.1177/0145445516685047] [Medline: 28033718]

58. Wilmer H, Sherman L, Chein J. Smartphones and cognition: a review of research exploring the links between mobile technology habits and cognitive functioning. Front Psychol 2017;8:605 [FREE Full text] [doi: 10.3389/fpsyg.2017.00605] [Medline: 28487665]
59. Fisher CD. Effects of external and internal interruptions on boredom at work: two studies. J Organiz Behav 1998 Sep;19(5):503-522. [doi: 10.1002/(sici)1099-1379(199809)19:5<503::aid-job854>3.0.co;2-9]
60. Canale N, Vieno A, Doro M, Rosa Mineo E, Marino C, Billieux J. Emotion-related impulsivity moderates the cognitive interference effect of smartphone availability on working memory. Sci Rep 2019 Dec 6;9(1):18519 [FREE Full text] [doi: 10.1038/s41598-019-54911-7] [Medline: 31811205]
61. Beute F, de Kort Y. Natural resistance: exposure to nature and self-regulation, mood, and physiology after ego-depletion. Artic J Environ Psychol 2014 Dec;40:167-178. [doi: 10.1016/j.jenvp.2014.06.004]
62. Jenkin R, Frampton I, White MP, Pahl S. The relationship between exposure to natural and urban environments and children's self-regulation. Landscape Res 2017 Jul 11;43(3):315-328. [doi: 10.1080/01426397.2017.1316365]
63. Newman KP, Brucks M. When are natural and urban environments restorative? The impact of environmental compatibility on self-control restoration. J Consum Psychol 2016 Oct;26(4):535-541. [doi: 10.1016/j.jcps.2016.02.005]
64. Fatisson J, Oswald V, Lalonde F. Influence diagram of physiological and environmental factors affecting heart rate variability: an extended literature overview. Heart Int 2016;11(1):e32-e40 [FREE Full text] [doi: 10.5301/heartint.5000232] [Medline: 27924215]
65. Javorka M, Zila I, Balhárek T, Javorka K. Heart rate recovery after exercise: relations to heart rate variability and complexity. Braz J Med Biol Res 2002 Aug;35(8):991-1000 [FREE Full text] [doi: 10.1590/s0100-879x2002000800018] [Medline: 12185393]
66. Michael S, Graham KS, Davis GM. Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals-a review. Front Physiol 2017;8:301 [FREE Full text] [doi: 10.3389/fphys.2017.00301] [Medline: 28611675]
67. Hunt KJ, Saengsuwan J. Changes in heart rate variability with respect to exercise intensity and time during treadmill running. Biomed Eng Online 2018 Sep 24;17(1):128 [FREE Full text] [doi: 10.1186/s12938-018-0561-x] [Medline: 30249267]
68. Oura Ring. URL: https://ouraring.com/ [accessed 2020-11-20]
69. Hedge C, Powell G, Sumner P. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. Behav Res Methods 2018 Jun;50(3):1166-1186 [FREE Full text] [doi: 10.3758/s13428-017-0935-1] [Medline: 28726177]
70. Saunders B, Milyavskaya M, Etz A, Randles D, Inzlicht M. Reported self-control is not meaningfully associated with inhibition-related executive function: A Bayesian analysis. Collabra: Psychol 2018:-. [doi: 10.1525/collabra.134]
71. Wennerhold L, Friese M. Why self-report measures of self-control and inhibition tasks do not substantially correlate. Collabra Psychol 2020;6(1):28. [doi: 10.1525/collabra.276]
72. Epstein D, Avrahami D, Biehl J. Taking 5: Work-breaks, Productivity, and Opportunities for Personal Informatics for Knowledge Workers. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16; May 7–12, 2016; San Jose, USA.
73. Kim J, Cho K, Lee K. Technology Supported Behavior Restriction for Mitigating Self-interruptions in Multi-device Environments. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2017 Presented at: ACM'17; September 11-15, 2017; Maui, UK URL: http://doi.org/10.1145/3130932 [doi: 10.1145/3130932]
74. Tseng V, Lee M, Denoue L, Avrahami D. Overcoming Distractions during Transitions from Break to Work using a Conversational Website-Blocking System. In: Conference on Human Factors in Computing Systems. 2019 Presented at: HFCS'19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300697]
75. Lin Y, Chiang C, Lin P, Chang L, Ko C, Lee Y, et al. Proposed diagnostic criteria for smartphone addiction. PLoS One 2016;11(11):e0163010 [FREE Full text] [doi: 10.1371/journal.pone.0163010] [Medline: 27846211]
76. Billieux J, Maurage P, Lopez-Fernandez O, Kuss DJ, Griffiths MD. Can disordered mobile phone use be considered a behavioral addiction? An update on current evidence and a comprehensive model for future research. Curr Addict Rep 2015 Apr 9;2(2):156-162. [doi: 10.1007/s40429-015-0054-y]
77. Ghiţă A, Teixidor L, Monras M, Ortega L, Mondon S, Gual A, et al. Identifying triggers of alcohol craving to develop effective virtual environments for cue exposure therapy. Front Psychol 2019;10:74 [FREE Full text] [doi: 10.3389/fpsyg.2019.00074] [Medline: 30761042]
78. García-Blanco AC, Perea M, Salmerón L. Attention orienting and inhibitory control across the different mood states in bipolar disorder: an emotional antisaccade task. Biol Psychol 2013 Dec;94(3):556-561. [doi: 10.1016/j.biopsycho.2013.10.005] [Medline: 24161800]

Abbreviations

API: application programming interface
AUC-ROC: area under the receiver operating characteristic curve
BIS: Barratt Impulsiveness Scale

EMA: ecological momentary assessment
GBT: gradient boosting tree
GEE: generalized estimating equation
HF: high frequency
HRV: heart rate variability
RMCORR: repeated measures correlation.
SSD: stop-signal delay
SSID: service set identifier
SSRT: stop-signal reaction time
SST: stop-signal task

Edited by G Eysenbach, L Buis; submitted 22.06.20; peer-reviewed by M Ghozali, K Frie; comments to author 25.08.20; revised version received 29.09.20; accepted 26.10.20; published 04.12.20.

Please cite as:
Tseng VWS, Costa JDR, Jung MF, Choudhury T
Using Smartphone Sensor Data to Assess Inhibitory Control in the Wild: Longitudinal Study
JMIR Mhealth Uhealth 2020;8(12):e21703
URL: https://mhealth.jmir.org/2020/12/e21703
doi:10.2196/21703
PMID:33275106

Original Paper

A Real-Time Eating Detection System for Capturing Eating Moments and Triggering Ecological Momentary Assessments to Obtain Further Context: System Development and Validation Study

Mehrab Bin Morshed[1], BSc; Samruddhi Shreeram Kulkarni[1], MS; Richard Li[2], MS; Koustuv Saha[1], BTech; Leah Galante Roper[1], MS; Lama Nachman[3], MS; Hong Lu[3], PhD; Lucia Mirabella[4], PhD; Sanjeev Srivastava[4], PhD; Munmun De Choudhury[1], PhD; Kaya de Barbaro[5], PhD; Thomas Ploetz[1], PhD; Gregory D Abowd[1], DPhil

[1]Georgia Institute of Technology, Atlanta, GA, United States
[2]University of Washington, Seattle, WA, United States
[3]Intel Labs, Santa Clara, CA, United States
[4]Corporate Technology, Siemens Corporation, Princeton, NJ, United States
[5]The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Mehrab Bin Morshed, BSc
Georgia Institute of Technology
756 W Peachtree St NW
Atlanta, GA, 30308
United States
Phone: 1 4049552729
Email: mehrab.morshed@gatech.edu

Abstract

Background: Eating behavior has a high impact on the well-being of an individual. Such behavior involves not only when an individual is eating, but also various contextual factors such as with whom and where an individual is eating and what kind of food the individual is eating. Despite the relevance of such factors, most automated eating detection systems are not designed to capture contextual factors.

Objective: The aims of this study were to (1) design and build a smartwatch-based eating detection system that can detect meal episodes based on dominant hand movements, (2) design ecological momentary assessment (EMA) questions to capture meal contexts upon detection of a meal by the eating detection system, and (3) validate the meal detection system that triggers EMA questions upon passive detection of meal episodes.

Methods: The meal detection system was deployed among 28 college students at a US institution over a period of 3 weeks. The participants reported various contextual data through EMAs triggered when the eating detection system correctly detected a meal episode. The EMA questions were designed after conducting a survey study with 162 students from the same campus. Responses from EMAs were used to define exclusion criteria.

Results: Among the total consumed meals, 89.8% (264/294) of breakfast, 99.0% (406/410) of lunch, and 98.0% (589/601) of dinner episodes were detected by our novel meal detection system. The eating detection system showed a high accuracy by capturing 96.48% (1259/1305) of the meals consumed by the participants. The meal detection classifier showed a precision of 80%, recall of 96%, and F1 of 87.3%. We found that over 99% (1248/1259) of the detected meals were consumed with distractions. Such eating behavior is considered “unhealthy” and can lead to overeating and uncontrolled weight gain. A high proportion of meals was consumed alone (680/1259, 54.01%). Our participants self-reported 62.98% (793/1259) of their meals as healthy. Together, these results have implications for designing technologies to encourage healthy eating behavior.

Conclusions: The presented eating detection system is the first of its kind to leverage EMAs to capture the eating context, which has strong implications for well-being research. We reflected on the contextual data gathered by our system and discussed how these insights can be used to design individual-specific interventions.

(JMIR Mhealth Uhealth 2020;8(12):e20625) doi:10.2196/20625

KEYWORDS

eating detection; eating behavior; eating context; well-being; smartwatch; ecological momentary assessment

Introduction

Dietary habits have been studied by health researchers for many decades, and it is now well understood that eating-related habits play a critical role in overall human health [1]. Such habits consist of a variety of social, temporal, and spatial factors [1]. Despite the known relationship between dietary patterns and wellbeing, measuring dietary patterns on a daily basis is challenging [2,3]. Most assessment methodologies of dietary patterns rely on self-reports by individuals to reflect on their meals [4,5]. Self-reported food consumption quantities suffer from under-report bias and recall bias [6]. This issue poses a challenge for regular dietary assessment.

Human activity recognition using passive sensing can address some of the challenges of dietary assessment methods [7-10]. For example, identifying when individuals eat can be used to infer if individuals are consuming food at regular intervals of time. Recent ubiquitous computing research has shown promise in eating detection, primarily showing various ways to infer when an individual is eating [10-14]. However, dietary patterns of an individual are not exclusively related to their interactions with food.

Several contextual factors are directly or indirectly related to eating and, consequently, wellbeing, including with whom a person is eating [15,16], where they are eating [17], what other activities are being performed while eating [18,19], and mood around the time of eating [20]. For example, regular family meals are associated with positive well-being. Hence, it is valuable to understand in what context people eat for assessing their well-being. There are several eating detection approaches that utilize passive sensing methods to detect when an individual is eating. Such detection systems can be categorized into the following three primary categories, based on the sensing modality used to infer eating activities: (1) acoustic sensing [7,8,21]; (2) camera-based sensing [22,23]; and (3) inertial sensing [9,24]. However, using current technology, it is not feasible to passively and reliably detect relevant contextual data (eg, company, mood, kind of food, and nutrition value of food) regarding eating without being intrusive (eg, camera and microphone).

A widely adopted [25-28] way of collecting subjective contextual data is by using ecological momentary assessment (EMA). EMAs are short questionnaires that can capture contextual information from individuals [29]. EMA questions can be delivered via platforms such as text messages [30], voice calls [31,32], and smart devices [33,34]. While self-reported surveys are prone to recall bias, EMAs are most effective when asked near real time of the actual event of interest [29,35]. Owing to the above advantages, EMAs have successfully been used to facilitate a number of eating-related studies, such as examining mood and binge eating [36], environmental factors and obesity [37], night eating [38], and eating disorders [39]. As such, a real-time eating episode detector can harness EMAs to gather insights about an individual's dietary patterns and use these insights to gauge the eating habits of individuals.

Motivated by the above, our work builds on a baseline recognition system for passively recognizing eating events using a smartwatch's three-axis accelerometer to capture eating movements. Through a machine learning pipeline, we first predicted individuals' hand-to-mouth movements and then obtained aggregated meal-scale eating episodes. By leveraging such a machine learning technique, we designed an eating detection system that not only focuses on real-time detection with high predictive accuracy but also allows us to recognize people's eating contexts. In particular, the real-time eating recognizer prompts eaters with EMA questions (designed after an online study) for capturing relevant contextual information, while at the same time preserving privacy and remaining minimally intrusive, as required for real-world deployment.

This work aimed to develop and evaluate a novel approach of gathering eating context through short EMA questions that are triggered by an automated meal detection system. We deployed and validated our system in a college student population. Young adults in the age group of 18 to 25 years are likely to develop a poor diet for a variety of reasons, such as embarking on higher education or employment, beginning independent living, and starting to live with partners [40,41]. Through our research, we made the following contributions:

- We designed and deployed a real-time meal detection system using a commercial smartwatch that triggers EMAs to validate prediction, which reliably predicted major meals with an F1 score of 87.3%.
- Using the real-time meal detection system, we demonstrated how a variety of contextual data can be captured using EMAs in a college student population.

Methods

Development of a Real-Time Passive Meal Detection System

Automated detection of eating behavior would entail selecting a sensing modality that can detect an eating episode while it is in progress. Furthermore, the respective sensing modality should be feasible for regular use. Several eating detection systems place a microphone on the neck [7,13]. However, such a solution is not practical to implement in a study that focuses on capturing eating episodes of individuals on a daily basis because it might be considered too socially awkward for everyday use.

There has been relevant research from the eating detection community that involves the use of hand movements [9,42,43] as a proxy for estimating when an individual is eating. For example, Thomaz et al [9] collected and released a data set on hand movements that were related to both eating and noneating activities. The data set was collected from participants in a

laboratory setting and a semicontrolled setting. The researchers used three-axis accelerometer data from Pebble smartwatches, which were worn by the participants on their dominant hand, to collect data on eating- and noneating-related hand movements.

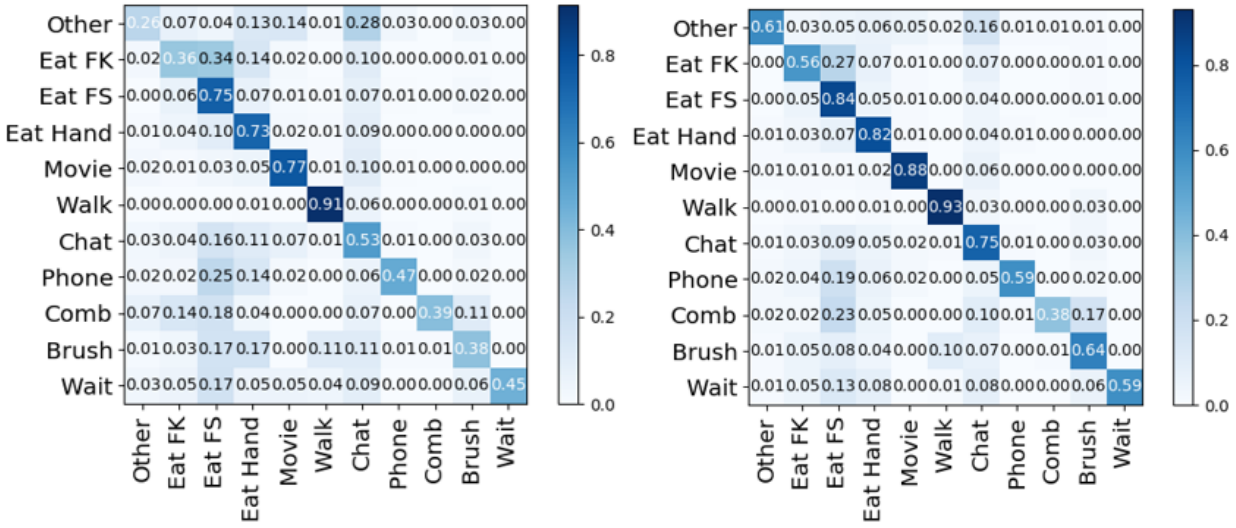
Since commercial smartwatches are becoming a part of day-to-day life, especially for college students, we chose to build a meal detection system based on the data set of Thomaz et al [9]. In this section, we provide a brief description of the baseline eating recognition system by Thomaz et al [9], clarify why we needed to extend and enhance the baseline system, and finally show the improvements in recognition performance provided by our extended approach over the baseline system.

Baseline Eating Detection System

Thomaz et al [9] built and evaluated an offline eating detection pipeline for recognizing eating moments in 60-minute intervals. For detecting an eating episode, the authors collected a data set in a laboratory setting that comprised 21 participants and contained both eating and noneating hand movements. The authors also collected another data set in a semicontrolled setting outside the laboratory with only “eating” and “noneating” labels and named the data set as Wild-7. The data from an integrated three-axis accelerometer were collected using a first-generation Pebble watch and transmitted to a companion smartphone app. After annotating the data, the authors employed an eating moment recognition pipeline, which is similar to the conventional activity recognition chain [44].

Drawing from the work of Thomaz et al, we created a baseline offline eating detection system initially to replicate the results. For creating the baseline classifier, we used a 50% overlapping 6-second sliding window to extract the following five statistical features along each axis of the accelerometer: mean, variance, skewness, kurtosis, and root mean square. Figure 1 shows the replication results for detecting eating and noneating gestures with the Lab-21 data set.

Figure 1. Eating gesture recognition performance. (A) Eating gesture recognition performance (F1 score) using the baseline system. (B) Eating gesture recognition performance using our system. For both figures, "Eat FK" represents eating gestures with a fork and knife, "Eat FS" represents eating gestures with a fork and spoon, "Eat Hand" represents eating gestures with the hands, and the rest of the classes are nontarget classes. The gesture recognition performance was observed in the Wild-7 data set released by Thomaz et al [9].



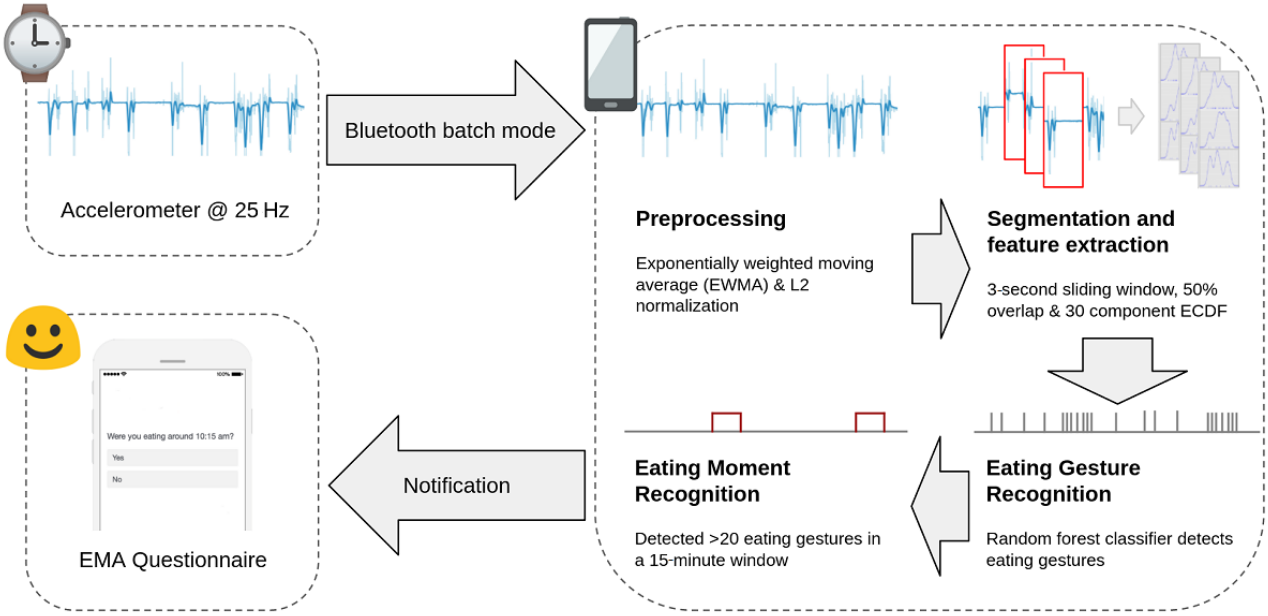
Motivation for Changes in the Eating Detection Pipeline

The recognition system of Thomaz et al took an offline approach, which can be used for passively logging eating episodes (typically at the meal level, ie, a major eating event). However, for capturing the contextual factors of eating, as they are of relevance for the assessment of well-being aspects [23], we require a real-time recognition system, which can reliably recognize eating moments and then, with minimal delay, prompt the user to answer EMA questions about their eating episode, ideally while the eating episode is still in progress. The baseline system, while serving as an excellent starting point for our work, needs to be extended such that it can be used for our purposes as outlined above. The main directions of improvement are as follows: (1) real-time recognition of eating episodes and (2) improvement of the accuracy of automated recognition. The baseline eating detection system was not robust enough to distinguish between eating and noneating gestures. Hence, we improved upon the baseline eating detection system by incorporating features that represent the temporal aspect of sensor data.

Real-Time Meal Detection System

The system architecture for real-time meal detection using a smartwatch and smartphone is presented in Figure 2. Upon detecting 20 eating gestures in a 15-minute span, the smartphone prompts the user with EMAs to capture in-situ eating-related information. After we trained a random forest classifier offline using the Python package sklearn, we ported the best classifier to run on Android using sklearn porter. This model used for making predictions on the smartphone runs every 10 minutes. When tested on a Google Pixel 2 device, the meal detection app on average consumed 30 MB of space on the phone while passively receiving data and 140 MB of RAM while the classifier was running. Using a Pebble 1 smartwatch, data were sent in batch mode to conserve the battery life of the device, which was approximately 36 hours.

Figure 2. System architecture for real-time meal detection using a smartwatch and smartphone. ECDF: empirical cumulative distribution function; EMA: ecological momentary assessment.



Changes in Feature Representation

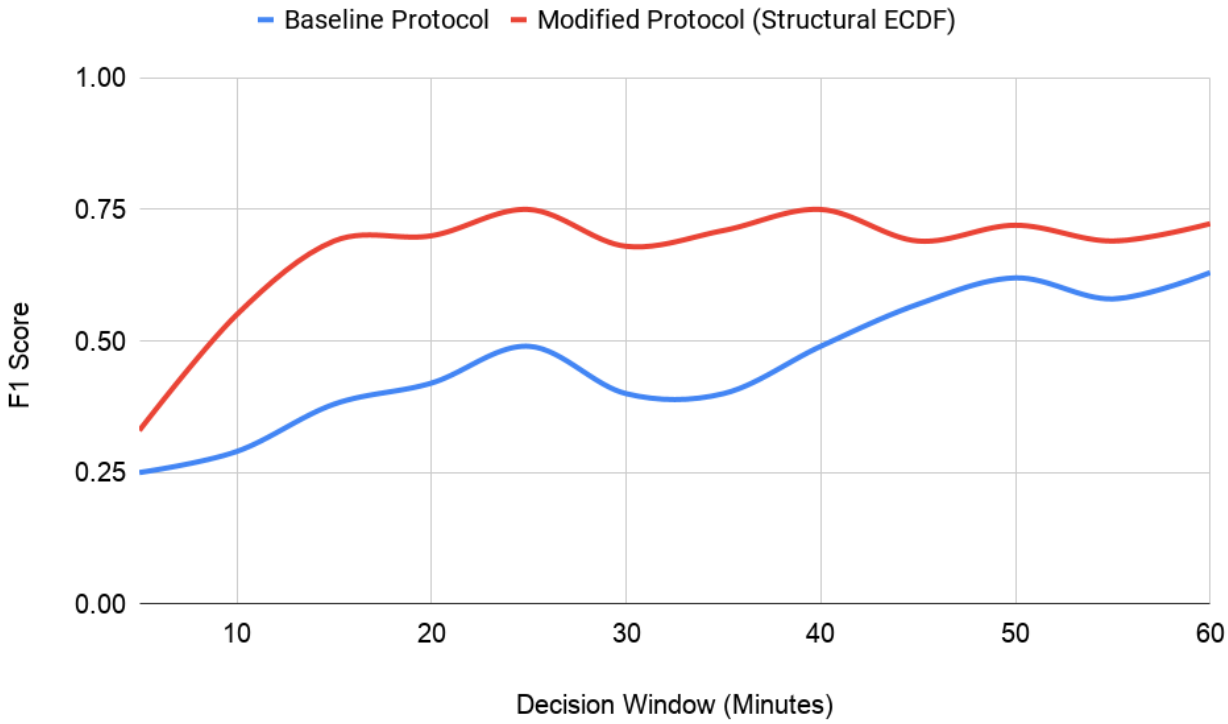
Before porting our system to run the analysis of sensor data that were recorded through the smartwatch in real time on smartphones, we extended the baseline system aiming for improved low-level gesture detection results. The baseline system misclassified some nontarget classes that appear very similar to typical eating-related hand movements, examples of which include brushing, combing, talking on the phone, etc. Upon closer inspection, we conclude that this failure was due to the fact that the feature representation used was unable to capture the temporal aspect of the signal. For example, talking on the phone would require someone to take the phone with their hand close to their head, which is similar to the hand-to-mouth movement during eating. If the feature representation does not capture the fact that the hand is not coming back down, as is the case for eating, both movements appear very similar in the feature space, which leads to confusion.

Hence, our first points of investigation were whether we can improve the feature representation and how changes in the feature space can affect the gesture recognition classification.

In response to the observation of the need to differentiate temporal dynamics, we employed the structural empirical cumulative distribution function (ECDF) feature representation [45], which specifically captures the temporal aspect of movement data at the feature level. Structural ECDF is a variation of the distribution-based feature representation ECDF [46].

Using a window size of 3 seconds with a 50% overlap generated the best results (Figure 3) on the Wild-7 data set made available by Thomaz et al [9]. The experimental results can be seen in Figure 1. It can be seen that for nontarget class detection (other, phone, chat, brush, etc), the system based on structural ECDF features performed much better. In particular, “brush” was not well recognized by the baseline system, but through the structural ECDF feature representation, we were able to classify this gesture with more than 20% higher accuracy. The recognition accuracy for “chat” also improved by more than 20%. The “chat” class contained gesticulation while the participant talked to other people. Recognition of the target classes (“eating with a fork and knife,” “eating with a fork and spoon,” and “eating with the hands”) improved overall by 38%, with “eating with a fork and knife” improving by 20%.

Figure 3. Eating moment recognition performance (F1 score) using the baseline system and our system. This analysis was performed on the same data set that Thomaz et al [9] collected for their study. ECDF: empirical cumulative distribution function.



Moving Away From the Clustering Approach

Thomaz et al indicated that they found the best performance when predicted gestures were clustered within a window of 60 minutes, that is, they needed at least 60 minutes of sensor data to infer whether an individual had an eating episode [9]. However, since the goal of our study was to capture eating behavior with respect to major meals, we needed to gather eating-related information from participants during/after each meal. Some of these insights about major meals can only be provided by participants, for example, whether the system predicted the meal correctly, since the system is not always accurate with meal predictions. If the detection was correct, one could ask a variety of questions that cannot be inferred passively.

Hence, to maximize recognition performance and mitigate the effects of noisy frame-level classification, we aggregated the results of the frame-level recognizer with a window of size *W* accumulating the frame-level results. A threshold-based approach was adopted in which *N* frames within the window must be recognized as one of the target classes for the window to be considered an eating episode, thus triggering an EMA. We used the window size mentioned by Thomaz et al of *W*=60 as a starting point and found that *N*=39 frames produced the highest F1 score (71.38%). Since our goal was to make the detection system as real time as possible, we started reducing the prediction window *W* by increments of 5 minutes at a time and optimized for the F1 score. We found that at *W*=25 minutes, our system performed the best (F1=74.63% with *N*=34 frames). However, when we considered *W*=15 minutes of sensor data, the F1 score was 69.44% (*N*=20 frames), which was not less than the F1 score at *W*=25 minutes but was closer to the actual eating episode for triggering the EMA. We finally decided upon using a window of *W*=15 minutes with *N*=20 eating gestures for detecting meal-level eating episodes.

Development of EMA Questions

Once we finalized a functional real-time meal detection system, the next step was to go beyond detecting major meal episodes and use the system for answering questions related to the mental well-being of college students. We designed a 3-week-long study to passively detect the meal consumption patterns of college students. However, given that we wanted to use EMAs to capture the context of an eating episode, it was important to understand what questions should be asked regarding an individual's meal. Hence, we first conducted an online survey study that addressed the following questions: (1) How much time do students generally spend on each meal? (2) Why do students miss certain meals? (3) What are the factors that constitute the "quality" eating experience of students?

We used the responses to this online survey to inform the design of the EMA questionnaire administered to the participants of the 3-week-long study.

Online Survey Study Design

Since we were interested in the three questions formulated above, we asked the below three open-ended and structured questions to the online participants. For the below questions 2 and 3, we provided some preset options that were informed by conducting structured interviews with 25 students (15 male and 10 female students) from the same university. We conducted qualitative coding on the interview data to derive themes and use those themes as available options for questions. In addition,

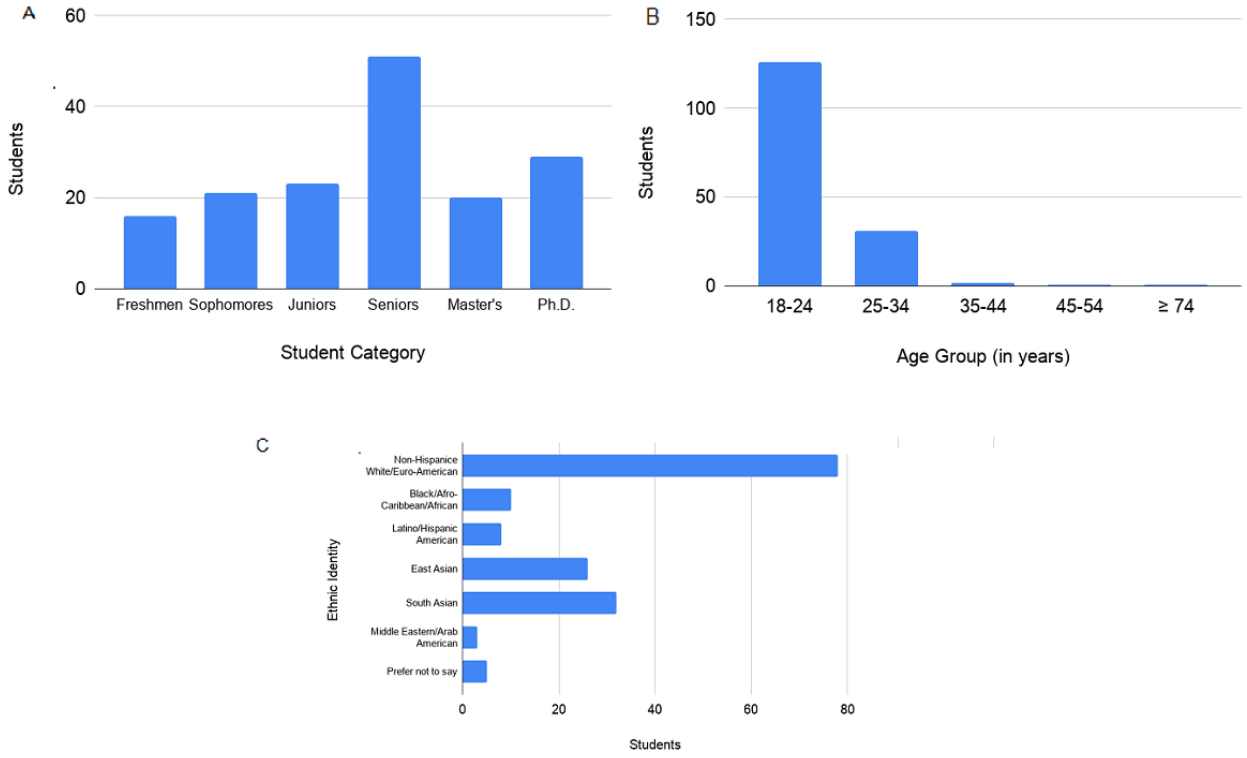
the students had the option of giving their own responses. We wanted to validate whether the themes reflected the responses of a larger subset of students. The questions were as follows: (1) How much time do you spend on major eating episodes (eg, breakfast, lunch, and dinner)? (2) If you ever miss some of your major meals (ie, breakfast, lunch, and dinner), please briefly mention why you miss these meals; (3) What does "quality" eating mean to you? We intend to learn about what you consider important as part of your eating experience. You are encouraged to come up with your own answer.

In addition to these questions, the students had to report their demographic information, which included their age, ethnicity, self-identified gender, and current academic status in the school. The demographic information was asked after the eating-related questions. The demographic information was used to ensure that our data sampling was representative of the college campus. Recruitment for the survey was conducted through various online communication channels such as email, Reddit, Facebook groups, etc. The timeline for the survey distribution was throughout summer 2018 and fall 2018.

Participant Demographics

A total of 162 participants responded to the survey. Among these respondents, 82 were female, 74 were male, one was nonbinary, and five did not disclose their gender identity. Figure 4 shows other demographic information of the student population that responded to the online survey.

Figure 4. Online survey response. (A) Student categories that responded to the online survey. (B) Age groups (in years) of students who responded to the online survey. (C) Ethnic identity of students who responded to the online survey.



Leveraging Responses From the Online Survey

Time Spent Per Meal

The average self-reported meal consumption times for breakfast, lunch, and dinner were 10 minutes, 20 minutes, and 25 minutes, respectively. Hence, we did not attempt to further improve our classifier since the minimum average meal consumption time was approximately 10 minutes for the student population of the target university. We used this information to decide upon the eating moment prediction window for capturing meals.

Factors for Missing Meals

We performed qualitative coding to extract themes from the responses to why students missed their meals. The themes found were workload, personal choice (ie, intermittent fasting), eating disorder (ie, anorexia), food insecurity, and mental health (ie, stress and mood). The responses in this section were crucial for us to derive our exclusion criteria for the meal consumption and mental well-being study. We were unaware of the fact that parts of the student population may experience food insecurity. However, we did expect some students to miss major meals due to eating disorders. Some responses included self-identified stress and mood when skipping a meal. In addition, some responses identified academic/professional workload as one of the reasons for missing meals.

Perception of "Quality" Eating Experience

We analyzed responses to this question with a similar process used for the previous questions. The emergent themes were contextual factors, perception of "healthiness" of the meal, and eating without distraction. Some of the contextual factors identified by the students were taking a meal with family, the location where the meal is being eaten, the noise around the eating location, etc. Some students mentioned that they would consider their meal as a "quality" meal if they were just taking their meal and doing nothing else while consuming the food.

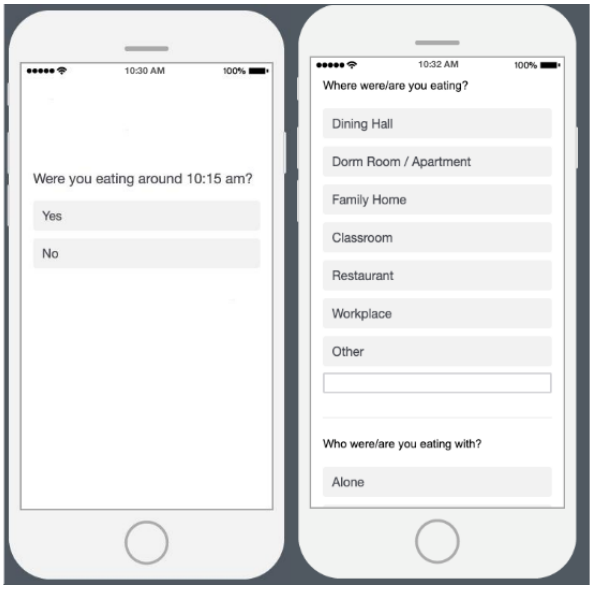
Finally, some of the students identified that if they took a healthy meal, they would consider it as a quality meal. The perceived healthiness of meals, company during meals, location of meals, and types of meals were the most common themes that came up as responses to this question. They were factored in the EMA questions, which are described below.

Study Protocol

Prompting EMA Questions

Whenever our meal detection system detected a meal-level eating episode, we prompted the user to answer questions on their smartphone (Figure 5) to validate whether they were actually having an eating episode. If the user responded with "yes" to the question, we asked them follow-up questions regarding (1) what kind of meal (eg, breakfast, lunch, and dinner) they were eating; (2) with whom and where they were eating; (3) what kind of activities they were performing while taking the meal; and (4) whether the meal was perceived as healthy. In order to obtain the ground truth total number of eating episodes, at the end of each day, participants were asked which meals they had during that day.

Figure 5. Prompt for validation. (A) Sample prompt for user validation on whether a meal was being taken. (B) Sample question users would get if they select "yes" in the validation.



Passive Sensing From the Smartwatch

For collecting and sending the raw accelerometer data from the Pebble smartwatch to a companion Android device, we wrote a native Pebble watch app (in C) that sampled the watch's accelerometer at 25 Hz and sent the data in batches to the phone approximately every 5 minutes. The battery of the Pebble watch lasted approximately 36 hours on a single charge.

Compensation

The timeline for our study was 3 weeks. If participants participated for more than 2 weeks in our study, they received an AmazeFit Bip watch valued at US $80. If they participated for more than 1 week but less than 2 weeks, they received an Amazon Gift Card valued at US $25. If participants did not participate for at least 1 week, they did not receive any compensation.

Exclusion Criteria

The results of our survey revealed that some students miss meals for a variety of reasons. Two of these reasons were the presence of an eating disorder and food insecurity. For these students, such a precondition can trigger stress. For example, participants with an eating disorder may have a relapse when they journal food since it makes them more self-conscious. Given that we were not in the position to effectively intervene if it was ethically required, we did not include students with food insecurity in our study. We used a validated eating disorder questionnaire [47] and a validated survey for identifying food insecurity [48] in our participants.

Recruitment

Our 3-week study was conducted in two semesters (summer 2019 and fall 2019). During summer, we recruited nine participants (four female and five male participants), and during fall, we recruited 21 participants (11 female and 10 male participants). In total, we obtained data from 28 participants (15 female and 13 male participants).

Results

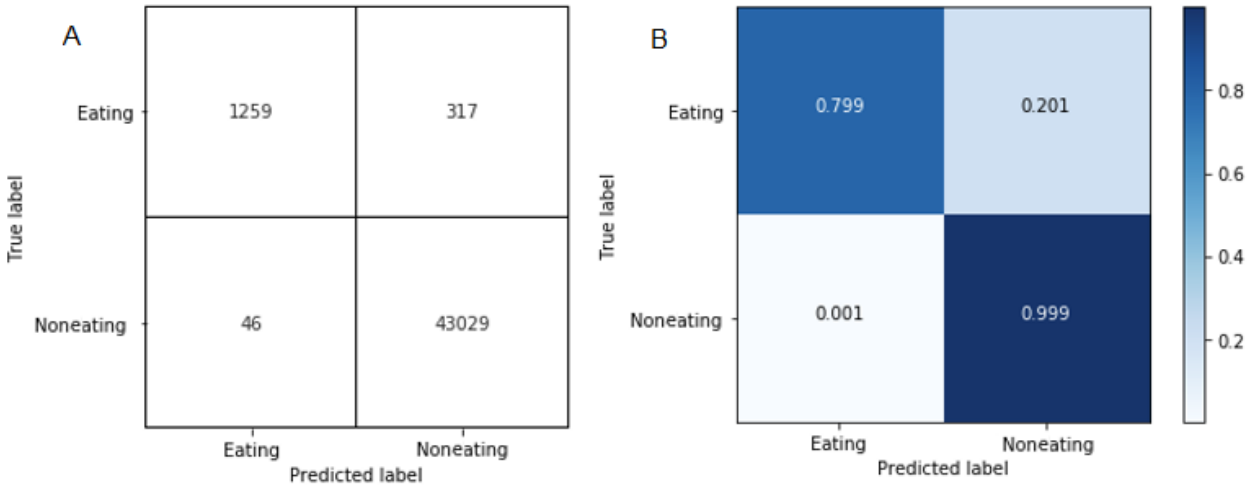
Performance of the Meal Detection System

We reflected upon the validity and reliability of the meal detection system that we deployed for approximately 3 weeks. We report the confusion matrix for the recognized eating events, explain in detail how we gathered the ground truth for eating and noneating events, and mention what kind of eating episodes were particularly challenging for our system to detect.

Recall that our real-time system (Figure 2) prompted participants with EMAs to capture eating-related information whenever it detected an eating episode. The first question in the series of EMA questions was to understand whether the participants were having a meal (Figure 5). If the participants answered "Yes,"

we considered it as a true positive, and if the participants answered "No," we considered it as a false positive. To capture false negatives, we asked participants at the end of the day which meals (eg, breakfast, lunch, and dinner) they actually had on that particular day. If our system did not detect that meal, we considered that meal as a false negative. It allowed us to understand how well or poorly our meal detector performed compared with the ground truth. Figure 6 shows the confusion matrices for eating episodes.

Figure 6. Confusion matrix. (A) Confusion matrix with number of meals. (B) Confusion matrix with percentage of meals.



The unweighted average F1 score for predicting major meals was 87.3%. The false-positive rate was 0.7%. The unweighted F1 score is particularly useful for cases where there is a class imbalance. In our study, there were only 1305 out of 44,651 instances that resembled eating episodes, which justifies F1-score analysis.

In addition, we wanted to investigate for which kinds of activities our meal detection system was making wrong predictions. Hence, during exit interviews, we asked participants whether they could recall for which activities the meal detector was erroneously prompting them (false positives) or around what kinds of eating episodes the meal detector was not detecting eating events (false negatives).

We analyzed the misclassification as follows. For false-positive predictions, we found that if participants performed hand movements similar to eating-related movements over an extended period of time (eg, brushing teeth and trimming beard), our meal detector was confusing these with eating episodes. For false-negative predictions, we found that short eating episodes (eg, eating a banana and taking a few spoons of yogurt in the morning as breakfast) were generally not detected by our meal detector. Table 1 presents the percentage of eating episodes that were detected by our meal detection system throughout the study.

Table 1. Percentage of meals detected by our meal detection system.

Meal type	Total episodes	Total detected episodes	Percentage of detected episodes
Breakfast	294	264	90
Lunch	410	406	99
Dinner	601	589	98

As can be seen, breakfast was the most frequently skipped meal by our participants throughout the study. It should be noted that seven of our participants self-identified themselves as individuals who did not have breakfast. Lunch was skipped more than dinner.

Context During Eating Episodes

We now report the contextual factors that were captured by our meal detection system. Our EMAs asked about various aspects that are challenging to be passively detected without invading an individual's privacy. These include the company of a participant during the meal, whether they were hungry when they had the meal, etc.

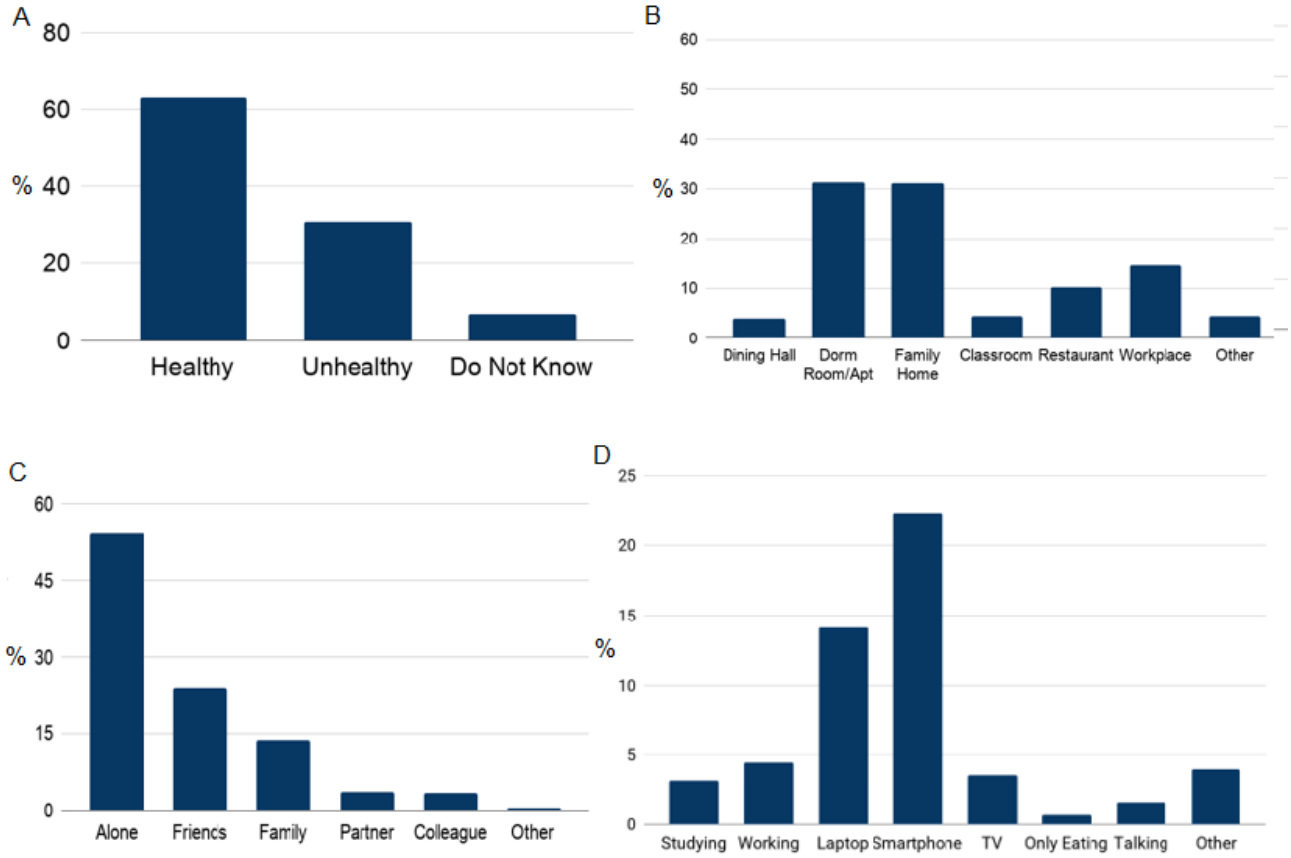
We found that 62.99% (793/1259) of meals were perceived as healthy and 31.05% (391/1259) of meals were perceived as unhealthy, and for the rest 5.95% (75/1259), the participants did not know whether the meal was healthy.

Since students generally operate on a busy and mobile schedule, we were interested to know where they were having their meals. We found that most meals were consumed either at the apartment/dorm room (393/1259, 31.22%) or family home (390/1259, 30.98%). Additionally, 14.54% (183/1259) of meals were consumed at workplaces, 10.25% (129/1259) were consumed at restaurants, and 4.13% (52/1259) were consumed in classes. Other than the predefined options, students could report places under the "other" option, and example responses included church, party, ministry, supermarket, and car.

The company during meals is strongly associated with well-being. By asking participants their company via EMAs, we found that participants had 54.17% (682/1259) of the detected meals alone, 24.17% (304/1259) with friends, 13.82% (174/1259) with family, 3.81% (48/1259) with partners, and 3.49% (44/1259) with colleagues.

Distracted eating is one of the most important factors behind many unhealthy eating behaviors, such as overeating, undereating, and binge eating. We gathered information on what noneating activities students were doing while they were having their meal (Figure 7). The two most common activities during eating were using a smartphone (281/1259, 22.32%) and laptop (178/1259, 14.14%). Only 0.87% (11/1259) of meal episodes were without any distractions.

Figure 7. Meal data. (A) Percentage of meals that were healthy and unhealthy. (B) Percentage of meals that were consumed at various locations. (C) Percentage of meals that were consumed with various companies. (D) Percentage of activities that were performed during various meals.



Discussion

Principal Results

Our work shows that major meal episodes can be detected using our meal detection system with an F1 score of 87.3%, a precision of 80%, and a recall of 96%. We demonstrated how an EMA-based design can augment a meal detection system to gather contextual information on eating behavior. This is the first-of-its-kind real-time meal detection system. When deployed for over a period of 3 weeks with 28 participants, our system showed a low false-positive rate of 0.7%, which is practical for daily usage considering that too many false positives may be bothersome to participants.

Among all consumed meals, 54.17% (682/1259) were consumed in isolation and 31.22% (393/1259) were consumed at apartment/dorm rooms. Most of the meal activities were often performed with another activity. Smartphone use and laptop use were the two most dominant activities (281/1259, 22.32% and 178/1259, 14.14%, respectively) during meals. Less than 1% (11/1259, 0.87%) of meal episodes were “only eating” episodes, which means for the rest of the cases, our participants were engaged in some other activities during a meal. These findings uncover previously unexplored and difficult to glean information, namely college students’ eating behaviors at a longitudinal scale. Our work can inform the design of well-being interventions in student populations.

Engaging in noneating activities during eating is considered as a distraction, and distraction during eating reduces the ability to assess internal sensory cues such as taste perception, which can lead to overeating [49,50]. Given the high percentage of distracted meals, we argue that college students can benefit from healthy eating behavior technologies that can build on our meal detection system.

Comparison With Prior Work

Thomaz et al built an offline meal detection system that could detect eating episodes in a period of 60 minutes with an F1 score of 71.3%. We improved upon this baseline detection system in two ways. First, we made the detection system detect an eating moment within 15 minutes, with an F1 score of 69.44%. This

improvement over the state-of-the-art wrist-worn meal detection system allowed us to prompt participants with EMA questions to capture various contexts during meals, which was missing from most meal detection systems in prior work. Previous work leveraging EMAs relied solely on nonautomated self-reports of eating behaviors, which are prone to recall bias and potentially can be a source of erroneous data. For example, we found that students are often doing other activities while having their meals, and such activities are a cause for distraction during eating. Given that we have provided a way to gauge an individual's eating behavior, relevant interventions can be designed to support healthy eating behaviors.

In a recent literature survey, Bell et al reported that 33 research studies performed an in-field assessment with a meal detection system [14]. The in-field assessment entailed participants using the sensor setup in a "free-living" condition. The authors reported the sample size of participants and how long they participated in "free-living" sessions. With respect to the number of participants, there were only two studies [14] that had more participants than our study; however, the rest of the studies (n=31) had fewer participants. For both of these studies, the timeline for the free-living condition was only 1 day, which is much less than our study timeline of 21 days. In fact, our work is the longest longitudinal study for any real-time meal detection system.

Limitations and Future Work

Though we argue that a smartwatch is more practical for detecting when an individual is eating, our study is limited in the sense that we asked our participants to wear the smartwatch on their dominant hand. Hence, we do not have insights into how robust our system is if the smartwatch is worn on the nondominant hand. However, Thomaz et al found in their study that wearing a smartwatch on the nondominant hand produced similar kinds of results compared with wearing a smartwatch on the dominant hand [51]. We did not validate this observation in our study.

In addition, our system is likely not robust enough to capture short snacking episodes. Snacking behavior has strong implications for mental and physical well-being [52]. However, solely based on wrist movements, it is difficult, if not impossible, to detect if a hand movement close to the mouth is for eating or some other activity [9]. Our future work will focus on appropriate eating detection technologies to capture snacking behavior and contexts during snacking.

Conclusions

We present the first real-time meal detection system that leverages EMA to capture context during meals, which has strong implications for well-being research. Through our paper, we reflected on how meaningful contextual data can be used for well-being research.

Acknowledgments

MBM was partly supported by a grant from Siemens FutureMaker Fellowship Task Order #7. MDC was partly supported by NIH grant #R01MH117172. The project was partly supported by a grant from Semiconductor Research Corporation in collaboration with Intel Labs.

Conflicts of Interest

None declared.

References

1. Ogden J. The Psychology of Eating: From Healthy to Disordered Behavior. UK: John Wiley & Sons; 2011.
2. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. J Multidiscip Healthc 2016;9:211-217 [FREE Full text] [doi: 10.2147/JMDH.S104807] [Medline: 27217764]
3. Shim J, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health 2014;36:e2014009 [FREE Full text] [doi: 10.4178/epih/e2014009] [Medline: 25078382]
4. Karvetti RL, Knuts LR. Validity of the 24-hour dietary recall. J Am Diet Assoc 1985 Nov;85(11):1437-1442. [Medline: 4056262]
5. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999 Mar;69(2):243-249. [doi: 10.1093/ajcn/69.2.243] [Medline: 9989687]
6. Fries E, Green P, Bowen DJ. What did I eat yesterday? Determinants of accuracy in 24-hour food memories. Appl. Cognit. Psychol 1995 Apr;9(2):143-155 [FREE Full text] [doi: 10.1002/acp.2350090204]
7. Yatani K, Truong K. BodyScope: a wearable acoustic sensor for activity recognition. In: UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012 Presented at: ACM Conference on Ubiquitous Computing; September 5-8, 2012; New York, NY p. 341-350. [doi: 10.1145/2370216.2370269]
8. van den Boer J, van der Lee A, Zhou L, Papapanagiotou V, Diou C, Delopoulos A, et al. The SPLENDID Eating Detection Sensor: Development and Feasibility Study. JMIR Mhealth Uhealth 2018 Sep 04;6(9):e170 [FREE Full text] [doi: 10.2196/mhealth.9781] [Medline: 30181111]
9. Thomaz E, Essa I, Abowd G. A Practical Approach for Recognizing Eating Moments with Wrist-Mounted Inertial Sensing. In: UbiComp '15: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.

2015 Sep Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan p. 1029-1040. [doi: 10.1145/2750858.2807545]
10. Bedri A, Li R, Haynes M, Kosaraju RP, Grover I, Prioleau T, et al. EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Sep;1(3) [FREE Full text] [doi: 10.1145/3130902] [Medline: 30135957]
11. Dong Y, Scisco J, Wilson M, Muth E, Hoover A. Detecting periods of eating during free-living by tracking wrist motion. IEEE J Biomed Health Inform 2014 Jul;18(4):1253-1260. [doi: 10.1109/JBHI.2013.2282471] [Medline: 24058042]
12. Heydarian H, Adam M, Burrows T, Collins C, Rollo ME. Assessing Eating Behaviour Using Upper Limb Mounted Motion Sensors: A Systematic Review. Nutrients 2019 May 24;11(5) [FREE Full text] [doi: 10.3390/nu11051168] [Medline: 31137677]
13. Rahman T, Czerwinski M, Gilad-Bachrach R, Johns P. Predicting "About-to-Eat" Moments for Just-in-Time Eating Intervention. In: DH '16: Proceedings of the 6th International Conference on Digital Health Conference. 2016 Presented at: 6th International Conference on Digital Health Conference; April 2016; Montréal, QC, Canada. [doi: 10.1145/2896338.2896359]
14. Bell BM, Alam R, Alshurafa N, Thomaz E, Mondol AS, de la Haye K, et al. Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review. NPJ Digit Med 2020;3:38 [FREE Full text] [doi: 10.1038/s41746-020-0246-2] [Medline: 32195373]
15. Utter J, Denny S, Robinson E, Fleming T, Ameratunga S, Grant S. Family meals and the well-being of adolescents. J Paediatr Child Health 2013 Nov 31;49(11):906-911 [FREE Full text] [doi: 10.1111/jpc.12428] [Medline: 24251656]
16. Yiengprugsawan V, Banwell C, Takeda W, Dixon J, Seubsman S, Sleigh AC. Health, happiness and eating together: what can a large Thai cohort study tell us? Glob J Health Sci 2015 Jan 14;7(4):270-277 [FREE Full text] [doi: 10.5539/gjhs.v7n4p270] [Medline: 25946941]
17. Utter J, Denny S, Lucassen M, Dyson B. Adolescent Cooking Abilities and Behaviors: Associations With Nutrition and Emotional Well-Being. J Nutr Educ Behav 2016 Jan;48(1):35-41.e1. [doi: 10.1016/j.jneb.2015.08.016] [Medline: 26411900]
18. Khan Z, Zadeh ZF. Mindful Eating and it's Relationship with Mental Well-being. Procedia - Social and Behavioral Sciences 2014 Dec;159:69-73. [doi: 10.1016/j.sbspro.2014.12.330]
19. Lyzwinski LN, Edirippulige S, Caffery L, Bambling M. Mindful Eating Mobile Health Apps: Review and Appraisal. JMIR Ment Health 2019 Aug 22;6(8):e12820 [FREE Full text] [doi: 10.2196/12820] [Medline: 31441431]
20. Carroll E, Czerwinski M, Roseway A, Kapoor A, Johns P, Rowan K, et al. Food and Mood: Just-in-Time Support for Emotional Eating. 2013 Presented at: Humaine Association Conference on Affective Computing and Intelligent Interaction; September 2-5, 2013; Geneva, Switzerland p. 252. [doi: 10.1109/acii.2013.48]
21. Cheng J, Zhou B, Kunze K, Rheinländer C, Wille S, Wehn N, et al. Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband Internet. In: UbiComp '13 Adjunct: Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication. 2013 Presented at: ACM conference on Pervasive and ubiquitous computing adjunct publication; September 2013; Zurich, Switzerland. [doi: 10.1145/2494091.2494143]
22. Liu J, Johns E, Atallah L, Pettitt C, Lo B, Frost G, et al. An Intelligent Food-Intake Monitoring System Using Wearable Sensors. 2012 Presented at: Ninth International Conference on Wearable and Implantable Body Sensor Networks; May 9-12, 2012; London, UK. [doi: 10.1109/bsn.2012.11]
23. Noronha J, Hysen E, Zhang H, Gajos K. Platemate: crowdsourcing nutritional analysis from food photographs. In: UIST '11: Proceedings of the 24th annual ACM symposium on User interface software and technology. 2011 Presented at: 24th annual ACM symposium on User interface software and technology; October 16-19, 2011; Santa Barbara, CA, USA. [doi: 10.1145/2047196.2047198]
24. Amft O, Troster G. On-Body Sensing Solutions for Automatic Dietary Monitoring. IEEE Pervasive Comput 2009 Apr;8(2):62-70. [doi: 10.1109/mprv.2009.32]
25. Lu H, Frauendorfer D, Rabbi M, Mast M, Chittaranjan G, Campbell A, et al. StressSense: detecting stress in unconstrained acoustic environments using smartphones. In: UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012 Presented at: ACM Conference on Ubiquitous Computing; September 2012; Pittsburgh, USA p. 351-360. [doi: 10.1145/2370216.2370270]
26. Morshed MB, Saha K, Li R, D'Mello SK, De Choudhury M, Abowd GD, et al. Prediction of Mood Instability with Passive Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2019 Sep 09;3(3):1-21 [FREE Full text] [doi: 10.1145/3351233]
27. Saha K, Chan L, De Barbaro K, Abowd GD, De Choudhury M. Inferring Mood Instability on Social Media by Leveraging Ecological Momentary Assessments. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2017 Sep 11;1(3):1-27. [doi: 10.1145/3130960]
28. Bin Morshed M, Saha K, De Choudhury M, Abowd GD, Ploetz T. Measuring Self-Esteem with Passive Sensing. 2020 Presented at: 14th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '20); May 18-20, 2020; Atlanta, GA, USA. [doi: 10.1145/3421937.3421952]
29. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]

30. Hébert ET, Stevens EM, Frank SG, Kendzor DE, Wetter DW, Zvolensky MJ, et al. An ecological momentary intervention for smoking cessation: The associations of just-in-time, tailored messages with lapse risk factors. Addict Behav 2018 Mar;78:30-35 [FREE Full text] [doi: 10.1016/j.addbeh.2017.10.026] [Medline: 29121530]
31. Scott LN, Stepp SD, Hallquist MN, Whalen DJ, Wright AG, Pilkonis PA. Daily shame and hostile irritability in adolescent girls with borderline personality disorder symptoms. Personal Disord 2015 Jan;6(1):53-63 [FREE Full text] [doi: 10.1037/per0000107] [Medline: 25580673]
32. Silk JS, Dahl RE, Ryan ND, Forbes EE, Axelson DA, Birmaher B, et al. Pupillary reactivity to emotional information in child and adolescent depression: links to clinical and ecological measures. Am J Psychiatry 2007 Dec;164(12):1873-1880 [FREE Full text] [doi: 10.1176/appi.ajp.2007.06111816] [Medline: 18056243]
33. van Roekel E, Goossens L, Verhagen M, Wouters S, Engels RC, Scholte RH. Loneliness, Affect, and Adolescents' Appraisals of Company: An Experience Sampling Method Study. J Res Adolesc 2013 Jun 21;24(2):350-363. [doi: 10.1111/jora.12061]
34. Russell MA, Wang L, Odgers CL. Witnessing substance use increases same-day antisocial behavior among at-risk adolescents: Gene–environment interaction in a 30-day ecological momentary assessment study. Dev Psychopathol 2015 Dec 09;28(4pt2):1441-1456 [FREE Full text] [doi: 10.1017/s0954579415001182]
35. Shiffman S, Stone A. Ecological momentary assessment: A new tool for behavioral medicine research. In: Krantz DS, Baum A, editors. Technology and methods in behavioral medicine. Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc; 1998:117-131.
36. Wegner KE, Smyth JM, Crosby RD, Wittrock D, Wonderlich SA, Mitchell JE. An evaluation of the relationship between mood and binge eating in the natural environment using ecological momentary assessment. Int J Eat Disord 2002 Nov;32(3):352-361. [doi: 10.1002/eat.10086] [Medline: 12210650]
37. Thomas JG, Doshi S, Crosby RD, Lowe MR. Ecological momentary assessment of obesogenic eating behavior: combining person-specific and environmental predictors. Obesity (Silver Spring) 2011 Aug 27;19(8):1574-1579 [FREE Full text] [doi: 10.1038/oby.2010.335] [Medline: 21273995]
38. Boseck JJ, Engel SG, Allison KC, Crosby RD, Mitchell JE, de Zwaan M. The application of ecological momentary assessment to the study of night eating. Int J Eat Disord 2007 Apr;40(3):271-276. [doi: 10.1002/eat.20359] [Medline: 17177212]
39. Stein KF, Corte CM. Ecologic momentary assessment of eating-disordered behaviors. Int J Eat Disord 2003 Nov 26;34(3):349-360. [doi: 10.1002/eat.10194] [Medline: 12949927]
40. Niemeier HM, Raynor HA, Lloyd-Richardson EE, Rogers ML, Wing RR. Fast food consumption and breakfast skipping: predictors of weight gain from adolescence to adulthood in a nationally representative sample. J Adolesc Health 2006 Dec;39(6):842-849 [FREE Full text] [doi: 10.1016/j.jadohealth.2006.07.001] [Medline: 17116514]
41. Poobalan A, Aucott L, Clarke A, Smith W. Diet behaviour among young people in transition to adulthood (18-25 year olds): a mixed method study. Health Psychol Behav Med 2014 Jan 01;2(1):909-928 [FREE Full text] [doi: 10.1080/21642850.2014.931232] [Medline: 25750826]
42. Kyritsis K, Diou C, Delopoulos A. Modeling Wrist Micromovements to Measure In-Meal Eating Behavior From Inertial Sensor Data. IEEE J Biomed Health Inform 2019 Nov;23(6):2325-2334. [doi: 10.1109/JBHI.2019.2892011] [Medline: 30629523]
43. Kyritsis K, Tatli CL, Diou C, Delopoulos A. Automated analysis of in meal eating behavior using a commercial wristband IMU sensor. Annu Int Conf IEEE Eng Med Biol Soc 2017 Jul;2017:2843-2846. [doi: 10.1109/EMBC.2017.8037449] [Medline: 29060490]
44. Bulling A, Blanke U, Schiele B. A tutorial on human activity recognition using body-worn inertial sensors. ACM Comput. Surv 2014 Jan;46(3):1-33. [doi: 10.1145/2499621]
45. Kwon H, Abowd G, Plötz T. Adding structural characteristics to distribution-based accelerometer representations for activity recognition using wearables. Proceedings of the 2018 ACM International Symposium on Wearable Computers New York, NY. In: ISWC '18: Proceedings of the 2018 ACM International Symposium on Wearable Computers. 2018 Presented at: ACM International Symposium on Wearable Computers; October 8-12, 2018; Singapore p. 72-75. [doi: 10.1145/3267242.3267258]
46. Hammerla N, Kirkham R, Andras P, Ploetz T. On preserving statistical characteristics of accelerometry data using their empirical cumulative distribution. Proceedings of the 2013 International Symposium on Wearable Computers New York, NY. In: ISWC '13: Proceedings of the 2013 International Symposium on Wearable Computers. 2013 Presented at: International Symposium on Wearable Computers; September 9-12, 2013; Zurich, Switzerland p. 65-68. [doi: 10.1145/2493988.2494353]
47. Luce K, Crowther J. The reliability of the eating disorder examination—Self-report questionnaire version (EDE-Q). Int. J. Eat. Disord 1999 Apr;25(3):349-351. [doi: 10.1002/(SICI)1098-108X(199904)25:3<349::AID-EAT15>3.0.CO;2-M]
48. Pia Chaparro M, Zaghloul SS, Holck P, Dobbs J. Food insecurity prevalence among college students at the University of Hawai'i at Mānoa. Public Health Nutr 2009 Aug 04;12(11):2097-2103. [doi: 10.1017/s1368980009990735]
49. van der Wal RC, van Dillen LF. Leaving a flat taste in your mouth: task load reduces taste perception. Psychol Sci 2013 Jul 01;24(7):1277-1284. [doi: 10.1177/0956797612471953] [Medline: 23722984]
50. Wallis DJ, Hetherington MM. Stress and eating: the effects of ego-threat and cognitive demand on food intake in restrained and emotional eaters. Appetite 2004 Aug;43(1):39-46. [doi: 10.1016/j.appet.2004.02.001] [Medline: 15262016]

51. Thomaz E, Bedri A, Prioleau T, Essa I, Abowd G. Exploring Symmetric and Asymmetric Bimanual Eating Detection with Inertial Sensors on the Wrist Internet. In: DigitalBiomarkers '17: Proceedings of the 1st Workshop on Digital Biomarkers. 2017 Presented at: 1st Workshop on Digital Biomarkers; June 2017; Niagara Falls, NY, USA. [doi: 10.1145/3089341.3089345]
52. Smith AP, Rogers R. Positive effects of a healthy snack (fruit) versus an unhealthy snack (chocolate/crisps) on subjective reports of mental and physical health: a preliminary intervention study. Front Nutr 2014;1:10 [FREE Full text] [doi: 10.3389/fnut.2014.00010] [Medline: 25988113]

Abbreviations

ECDF: empirical cumulative distribution function
EMA: ecological momentary assessment

Edited by G Eysenbach; submitted 24.05.20; peer-reviewed by J van den Boer, A Adams; comments to author 19.06.20; revised version received 14.08.20; accepted 30.10.20; published 18.12.20.

Please cite as:
Bin Morshed M, Kulkarni SS, Li R, Saha K, Roper LG, Nachman L, Lu H, Mirabella L, Srivastava S, De Choudhury M, de Barbaro K, Ploetz T, Abowd GD
A Real-Time Eating Detection System for Capturing Eating Moments and Triggering Ecological Momentary Assessments to Obtain Further Context: System Development and Validation Study
JMIR Mhealth Uhealth 2020;8(12):e20625
URL: http://mhealth.jmir.org/2020/12/e20625/
doi:10.2196/20625
PMID:33337336

Viewpoint

Evaluation of the Design and Implementation of a Peer-To-Peer COVID-19 Contact Tracing Mobile App (COCOA) in Japan

Ichiro Nakamoto[1], PhD; Ming Jiang[1], PhD; Jilin Zhang[1], PhD; Weiqing Zhuang[1], PhD; Yan Guo[1], PhD; Ming-Hui Jin[1], PhD; Yi Huang[1], MA; Kuotai Tang[1]

School of Internet Economics and Business, Fujian University of Technology, Fuzhou, China

Corresponding Author:
Ichiro Nakamoto, PhD
School of Internet Economics and Business
Fujian University of Technology
999 Dongsanhuang Road
JinAn District
Fuzhou
China
Phone: 86 059183503931
Email: ichiro2019@163.com

Abstract

We evaluate a Bluetooth-based mobile contact-confirming app, COVID-19 Contact-Confirming Application (COCOA), which is being used in Japan to contain the spread of COVID-19, the disease caused by the novel virus termed SARS-COV-2. The app prioritizes the protection of users' privacy from a variety of parties (eg, other users, potential attackers, and public authorities), enhances the capacity to balance the current load of excessive pressure on health care systems (eg, local triage of exposure risk and reduction of in-person hospital visits), increases the speed of responses to the pandemic (eg, automated recording of close contact based on proximity), and reduces operation errors and population mobility. The peer-to-peer framework of COCOA is intended to provide the public with dynamic and credible updates on the COVID-19 pandemic without sacrificing the privacy of their information. However, cautions must be exercised to address critical concerns, such as the rate of participation and delays in data sharing. The results of a simulation imply that the participation rate in Japan needs to be close 90% to effectively control the spread of COVID-19.

(JMIR Mhealth Uhealth 2020;8(12):e22098) doi:10.2196/22098

KEYWORDS

COVID-19; contact tracing; mobile app; peer-to-peer; Bluetooth-based; telehealth; privacy protection; load balancing; close contact; decentralized

Introduction

As of August 23, 2020, over 23 million cumulative cases of COVID-19 and nearly 800,000 deaths from the disease have been reported worldwide [1]. Since the first cases were reported in late 2019, the world has witnessed the rapid spread of the pathogen, and it has been declared a global public health crisis by the World Health Organization [2]. Due to the infectiousness of the disease and the dynamics of interperson interactions, the spread of COVID-19 could advance in a way that is unnoticeable to individuals, as evidenced by subclinical presymptomatic and asymptomatic cases [3], which refer respectively to cases in which infection started before the onset of symptoms and infections without the emergence of symptoms. Research has demonstrated the risk of person-to-person transmission of COVID-19 between individuals, especially for those in close contact (ie, close proximity) [4]. When infections are established where individuals are unable to readily self-triage their exposure risk, timely responses to COVID-19 will be challenging; these responses could also be weakened by delayed data sharing, impeded privacy preservation, and impaired security [3].

A variety of measures based on digital health have been adopted to control the spread of COVID-19 [5-13]. These containment measures display heterogeneities in terms of their design: decentralized (ie, mostly privacy-first) versus centralized (ie, mainly data-first) deployment frameworks have emerged alongside Bluetooth-, GPS-, and quick response (QR)–based sensor technologies [5,9]. Countermeasures such as contact tracing can play remarkable roles in the containment of the pandemic, including inference of exposure risk, identification

of infections, and quarantining or isolation of individuals being traced [5,10,11]. However, there is major divergence among nations regarding which digital health approach to employ (eg, a centralized approach that collects private data at the expense of potential illegal use vs a decentralized approach that stores data on local devices and leaves individuals in charge of their sensitive information at the cost of constrained accessibility for others) to address critical concerns such as privacy preservation, health care pressure load-balancing, speed of response, and ease of operations [9,12,13]. The centralized approach highlights a data-first methodology and involves the collection of privacy-sensitive information; this approach can enhance the capacity of unified administration, but the identities of individuals can be readily inferred [6,8,9]. GPS- and QR-based contact tracing apps facilitate evidence-based inference and improve the traceability of contact tracing [6,8,9,13]. However, multiple crucial concerns must be addressed to achieve effective containment. The first concern is that the information gathered through GPS is not strictly equivalent to close contact; hence, bias could be undoubtedly introduced. The second concern is the unlawful use or abuse of sensitive personal information obtained using GPS or QR [5-9]. Third, there is some debate that centralized approaches could cause discrimination, reduce confidence, and negatively impact the health of individuals if their private data are misused or breached [9,13]. In contrast, the typical application of the decentralized approach includes Bluetooth-based digital health, which does not theoretically identify individuals; hence, this approach is desirable for settings where concealing users' identities and preventing accessibility of their contact information are valued by the population [5,8,10,13]. Bluetooth digital health approaches rely on microwave and millimeter-wave technologies to sense the proximity between local devices, enabling the tracking of social contacts with a high degree of precision [5,10,13]. Hence, infections due to close contact (ie, geographical proximity) with pre-asymptomatic or asymptomatic patients can be easily detected and recorded by a Bluetooth-based approach. It is feasible for exposed people to evaluate and self-identify their exposure risk without disclosing either their own identity or the identities of their counterparts. This approach is more rapid and efficient in close contact diagnosis, less labor-intensive, and less susceptible to human error than extant approaches. Bluetooth contact tracing is currently being adopted by numerous countries, including Japan, India, and Singapore [5,9,13-15].

The first cases of COVID-19 in Japan were reported in late January 2020; since then, numerous cases, including asymptomatic and presymptomatic infections, have been identified, and the national medical system has been overburdened, with ever-increasing risk of collapse [14,15]. The excessive strain on the capacity of the medical system is expected to be alleviated. Unidentifiable discrete spreading events could lead to a later outbreak of infections; thus, it is important for individuals to gain an updated understanding of the pathogen [16]. When people can locally track their exposure risk, self-triage, and make differentiated responses based on digitally provided instructions, cross-transmission (eg, cluster infections at crowded locations) and unnecessary in-person visits can be reduced [15].

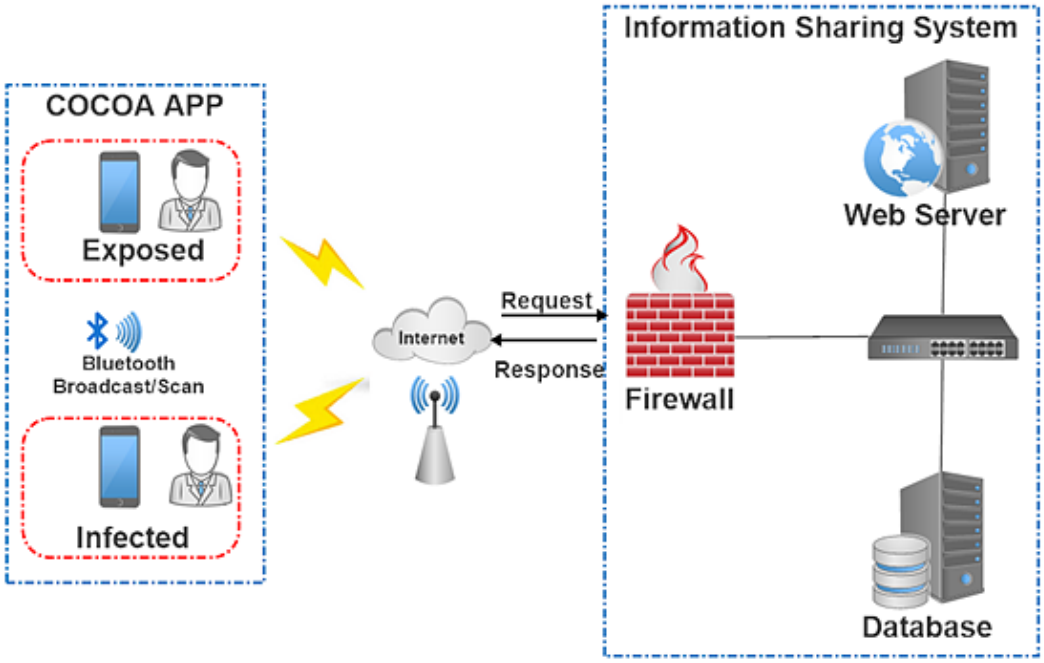
In this viewpoint, we discuss a decentralized and GPS-free Bluetooth digital health approach, COVID-19 Contact-Confirming Application (COCOA). This approach is mainly used to address the issues of privacy protection, efficacy enhancement, load balancing of pressure on the health care system, population mobility, and manual operation errors, and it principally complies with the Apple and Google contact tracing technology frameworks [5,13]. The major aim is to appraise how the approach can be used as a routine tool to contain the spread of COVID-19, with emphasis on privacy preservation and load-balancing. Prior research has revealed that other factors, such as the rate of participation, play remarkable roles in contributing to the effectiveness of containment [17]. The results of the simulation in our study are consistent with the findings in other empirical research.

Framework and Core Mechanism of COCOA

The Architecture and Prototype of COCOA

A schematic of the general architecture of the COCOA system is shown in Figure 1. The app automatically records close contact (ie, defined as within 1 meter of proximity for at least 15 minutes in COCOA) on Android and iOS devices by employing Bluetooth technology [18]. The COCOA system consists of three major sections: two mobile terminal apps for individuals (ie, infected and potentially exposed), and an infection information sharing system maintained by public authorities and health care providers. COCOA complies with the decentralized framework; this means that the COCOA app only locally tracks close contacts and performs matching inference of exposure risk, during which no personal private information is requested or collected through COCOA.

Figure 1. System architecture and prototype of the COCOA app. COCOA: COVID-19 Contact-Confirming Application.

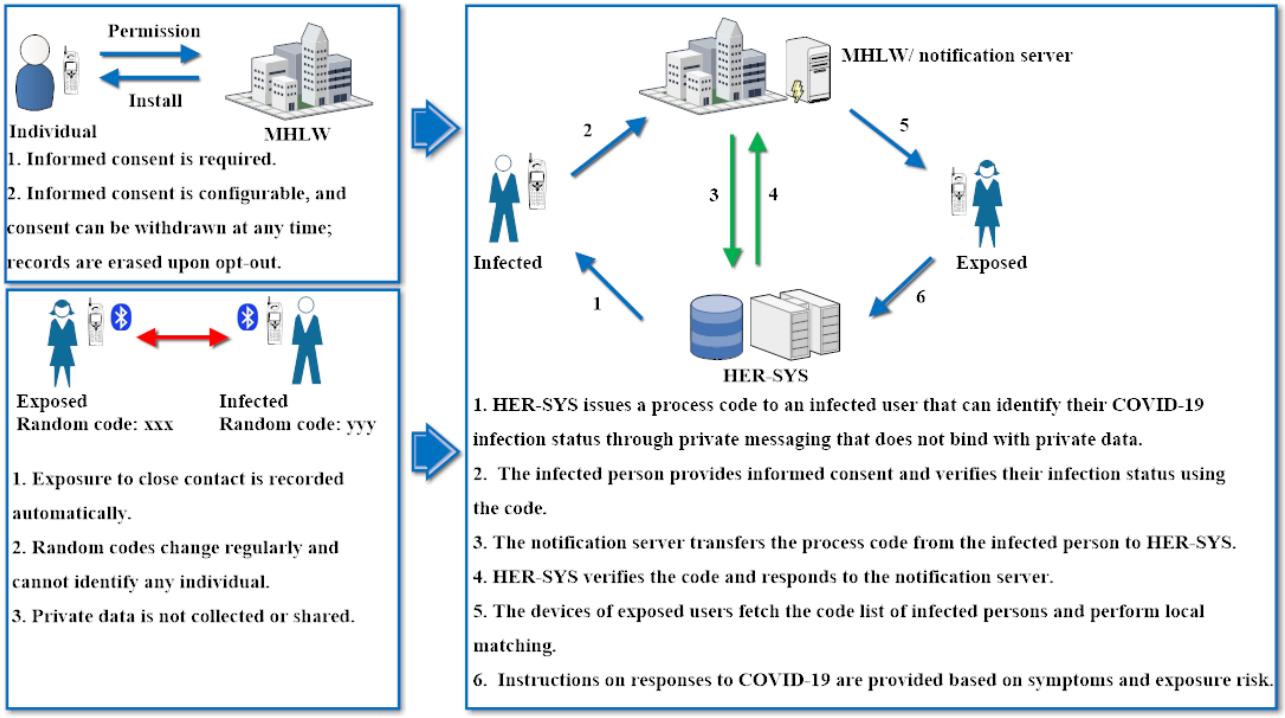


Core Mechanisms

To conceptualize how COCOA optimizes its core functionality, including privacy protection of individuals and load-balancing of health care stress, in comparison with other centralized digital health approaches, we illustrate the core mechanism and diagram in Figure 2. COCOA integrates the following principal features:

1. Individuals (ie, either infected or potentially exposed) receive informed consent to participate and authorize data sharing.
2. The informed consent feature is configurable, and consent can be withdrawn at any time.
3. Prior records are erased when the user opts out.
4. No sensitive personal information that enables the app user to be identified, such as date of birth, gender, address, telephone number, email address, or location, is requested or collected through COCOA.
5. Close-contact data are encrypted, saved only on users' local devices, and automatically deleted after 14 days, which is the period generally considered to be the average incubation interval of COVID-19.
6. If the user is infected, informed consent of the COVID-19–positive patient is required to authenticate and distribute their infection status.
7. Upon completion of verification, the process code that is used by the infected person to verify the accuracy of their infection status with the central server is eliminated from the COCOA app, the notification server, and the management system.
8. Exposure risk matching is performed on local devices.

Figure 2. Core mechanisms and diagram of the COVID-19 Contact-Confirming Application (COCOA) framework. HER-SYS: Health Center Real-time Information-sharing System on COVID-19; MHLW: Ministry of Heath, Labor and Welfare.



When individuals (whether symptomatic, presymptomatic, or asymptomatic) are in close contact, the COCOA app records this status by automatically exchanging generated random codes, which change periodically and thus cannot be exploited to identify either the infected or potentially exposed users. The codes are not shared with the information sharing system unless the individuals are COVID-19–positive. The codes will be saved only on local devices and erased after 14 days. As these codes are generated randomly and changed periodically, they cannot be exploited to uniquely identify any individuals; this guarantees the preservation of privacy of users' data from infected people, potentially exposed people, attackers, and public authorities. In this way, concerns regarding privacy-preserving issues inherent to other technologies (eg, GPS and QR) can be waved. The detection of close contacts can run automatically in the background without requiring COCOA to be active, all of which can be unnoticeable to the individuals in contact. This feature improves the efficacy of detection, increases the ease of operations, and reduces manual errors [15,18].

The Health Center Real-time Information-sharing System on COVID-19 (HER-SYS) is operated and maintained by prefecture-level or local health care providers. It issues a process code when an individual tests positive for COVID-19 by polymerase chain reaction (PCR). These process codes are distributed to the patients through private messaging (eg, emails) and are not exploited to bind private information (eg, telephone numbers that can identify individuals). Hence, privacy protection issues during the data dissemination steps are also not of concern. The notification server is administered by the Ministry of Health, Labor and Welfare in Japan, and its functionality is considerably constrained for privacy protection. The notification server does not store the patients' infection status or any other sensitive personal information. When one individual is notified that they have been infected, they are encouraged to share their status with potential close contacts. However, informed consent and authorization are requested. To prevent malicious inquiries and to guarantee the accuracy of data, the patient must input a process code and authenticate the correctness of their data with the notification server, which transfers the request to HER-SYS. HER-SYS authenticates the accuracy and returns the outcome to the notification server, which then distributes the random code of the infected person to all potentially exposed people upon request. The COCOA apps of the exposed people then perform local matching inference based on the retrieved anonymized list of random codes for individuals infected with COVID-19. Note that asymptomatic or presymptomatic infections can be traced effectively because the detection mechanism of COCOA hinges on geographical proximity, which is generally considered to be the critical factor contributing to the transmission of highly infectious disease. If a match is found and the exposure risk is identified, health care instructions on the outcome and the severity of risk will be provided. The potential exposed person can respond appropriately according to their corresponding symptoms. For severely symptomatic individuals, urgent care must be scheduled. In contrast, asymptomatic or mild cases may choose to self-isolate at local sites. As the matching calculation is performed locally and individuals can self-triage the exposure risk, risky in-person visits and cross-transmission at health care facilities can be curtailed. This can prioritize limited health care resources for more severely ill patients and enhance the load-balancing of the pressure on medical systems, thus lowering the risk of collapse of the health care system [14,15,18].

Different countries differ in their digital health participation rates [9,12,13]. According to a report by MIT Technology Review [19], we outlined the statistics of countries adopting or partially adopting Bluetooth digital health by the end of July 2020 (Table 1). On average, the rate of participation is higher

for countries using centralized Bluetooth digital health frameworks than for countries using decentralized Bluetooth frameworks; however, the latter frameworks generally outperform the former in privacy protection. Further, countries using decentralized approaches, most of which hinge on voluntary participation, mostly sustain low rates of participation. None of these countries currently have a participation rate ≥60% [19].

Table 1. Estimated participation rates of countries employing Bluetooth frameworks.

Country	Participation rate (%)	Centralized or decentralized
Japan	12.6[a]	Decentralized
Australia	25.8	Centralized
Austria	6.8	Decentralized
Bahrain	25.5	Centralized
Czech	2.6	Decentralized
Fiji	3.1	Decentralized
France	2.8	Centralized
Germany	16.9	Decentralized
Gibraltar	26.7	Decentralized
Hungary	0.10	Centralized
India	7.39	Centralized
Indonesia	7.10	Centralized
Italy	3.64	Decentralized
Malaysia	0.32	Decentralized
New Zealand	12.1	Centralized
Norway	26.6	Centralized
Philippines	1.1	Decentralized
Poland	0.1	Decentralized
Qatar	91	Centralized
Singapore	37.2	Centralized
Switzerland	5.8	Decentralized
Thailand	5.1	Decentralized
Tunisia	0.2	Centralized
Turkey	17.3	Centralized
Vietnam	0.4	Decentralized

[a]Data for Japan as of August 27, 2020.

To illustrate how countries using Bluetooth digital health can differ in the capacity of privacy preservation, we provide details of two centralized frameworks: one voluntary (ie, Aarogya Setu in India [20]) and one involuntary (ie, TraceTogether in Singapore [21]). Emphasis is placed on the core qualitative concepts employed to clarify the major differences; detailed quantitative technical specifications are not examined in this paper (Textbox 1).

Textbox 1. Core concepts of the COVID-19 Contact-Confirming Application (COCOA) framework and of Bluetooth digital health frameworks in other countries.

COCOA (Bluetooth-based digital health framework in Japan)

- Personal information, including names, telephone numbers, and GPS locations, is not requested or collected.
- Participation is voluntary. Informed consent to participate is requested.
- Close contact detection automatically runs in the background without requiring COCOA to be active, resulting in ease of use and low power consumption.
- Mobile phones generate and exchange periodically changing random codes with close contacts.
- Close contact information is saved only on local mobile phones for 14 days and is not transmitted. Individuals poll the central server (without sharing private information) to retrieve the list of infected people, not the reverse.

Aarogya Setu (Bluetooth-based digital health framework in India)

- Personal data, including name, gender, travel history, and telephone numbers, are requested and shared with the central server.
- GPS locations are collected and used to trace the paths of infected individuals.
- Participation is voluntary.
- There are risks of data inaccuracy and illegal data use.
- It is difficult to operate.
- Its power consumption is high.

TraceTogether (Bluetooth-based digital health framework in Singapore)

- Random tokens recording close contacts are shared with the central server, which maintains a database linking tokens and telephone numbers. There is a likelihood of linkage attacks and unlawful use.
- Infected individuals are required by law to share their infection status, including telephone and unique identification numbers.
- Individuals are notified of their exposure risk via identifiable information (eg, telephone numbers).
- GPS location data are not tracked; however, telephone and unique identification numbers are collected by public authorities.
- Participation is generally mandatory.
- Inference of exposure match is performed on the authority-administered central server.

COCOA differentiates from Aarogya Setu in privacy preservation in that the latter framework requests individuals' self-reported personal data, including gender, travel history, and telephone numbers; this raises concerns regarding the accuracy of data in the case where users share wrong information, as no mechanism is provided for authentication. Further, because GPS locations are collected by Aarogya Setu, it is debated that these data could be used to identify individuals without improvement of contact tracing precision (eg, individuals on different floors of the same building) [20]. In contrast, in Singapore, individuals are legally required to share their infection status; hence, the rate of participation can be guaranteed. However, personal data such as telephone numbers and unique identification numbers are also collected, which creates concerns regarding the possibility of illegal use of private information. Further, inference of exposure is performed on the central server; hence, technical pressure on the medical system must be optimized [21].

Screenshots and Diagram of the COCOA App

Figure 3 shows screenshots of the COCOA app. The app can be divided into components for infected people (Figure 3A), potentially exposed people (Figure 3B), and general settings (Figure 3C). The process can be described as follows:

1. An individual is tested and identified as COVID-19–positive using PCR. The individual receives a process code from HER-SYS, verifies the accuracy of their status through the notification server, and authorizes anonymized sharing of infection status (Figure 3A).
2. Potentially exposed individuals retrieve the up-to-date list of COVID-19 infections (Figure 3B).
3. If an exposure match is identified, exposure statistics and subsequent response guidance are promptly provided. Otherwise, a message indicating no exposure is promptly provided (Figure 3B).
4. Close-contact recording and COCOA participation are configurable. Records will be erased if the user opts out (Figure 3C).

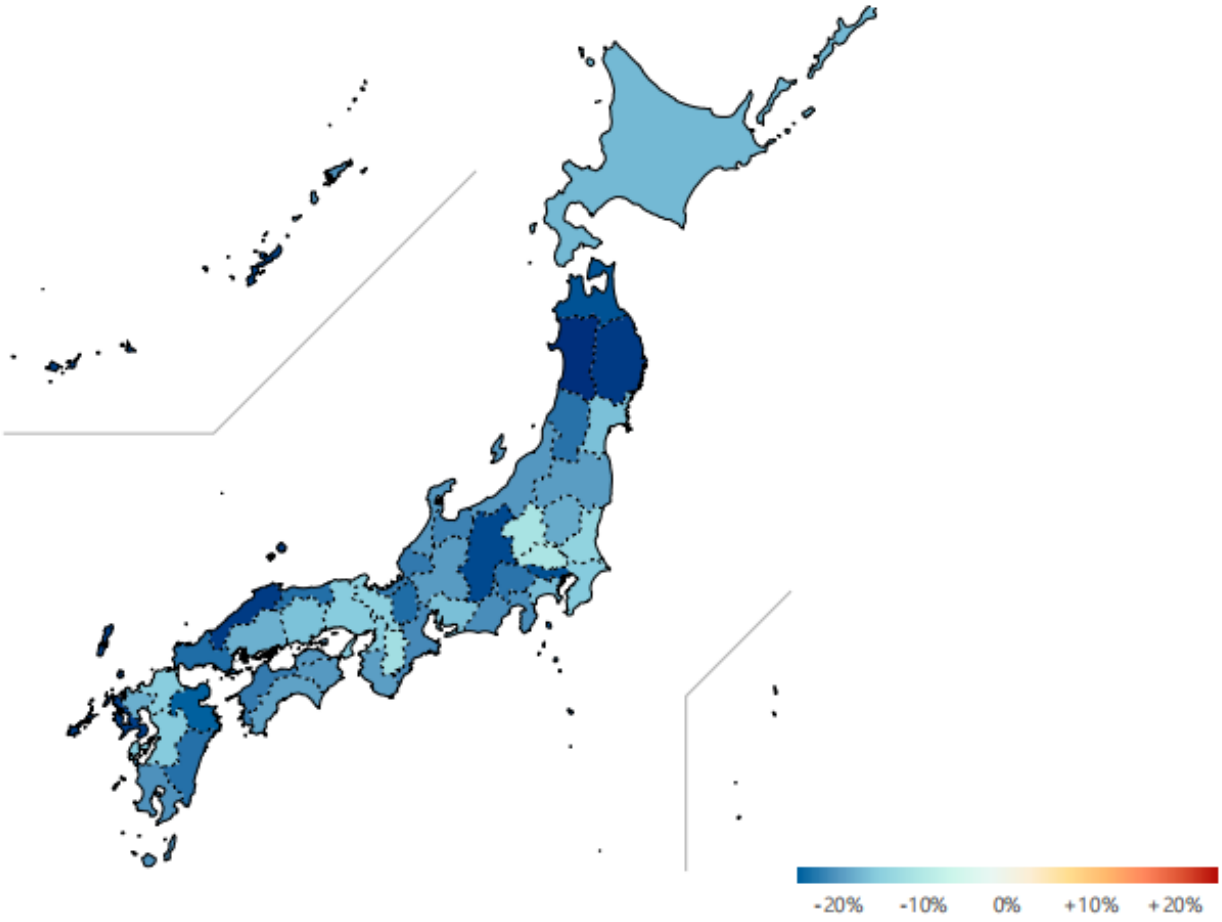
Figure 3. Screenshots and diagram of the use of the COVID-19 Contact-Confirming Application (COCOA) app. (A) Infected users verify their infection status and authorize data sharing; (B) potentially exposed users retrieve the infection list and triage their exposure risk; (C) users can adjust their close contact settings.

Statistical Analysis and Simulation Results

Digital health can enhance individuals' knowledge and risk perceptions of COVID-19, thus decreasing population mobility [22-25]. Reducing mobility is a controversial but effective measure to flatten the curve and control global pandemics; it decreases the generation of crowded spaces that aggravate cluster infections [23-25]. A simulation was conducted that implies that the spread of COVID-19 in Japan will be gradually contained by reducing the population mobility and the amount of time spent in crowded spaces [22]. We observed the dynamics of population mobility at the prefecture level in Japan by comparing data from August 2020 (ie, when COCOA was

deployed) with data from the same interval of the previous year (ie, when COCOA was not deployed) (Figure 4). The analysis suggests that the nationwide population mobility in 2020 decreased by 20% on average (ranging from the minimum of 12% in Saitama Prefecture to the maximum of 30% in Akita Prefecture). Reducing population mobility lowers the risk of exposure to COVID-19 and the risk of infection [18,25,26].

Figure 4. Comparison of the population mobility in Japan in August 2020 versus August 2019.



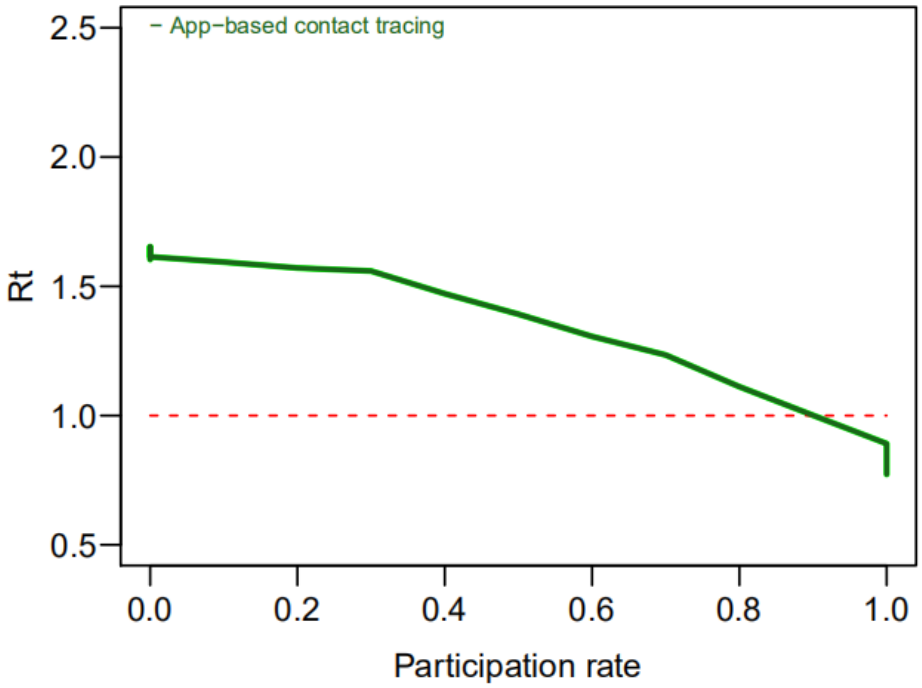
Since its official deployment on June 30, 2020, COCOA has been used by approximately 15 million individuals as of August 27, 2020 [18]. This finding denotes that around 12.6% of the population of Japan (15 million/118.6 million people aged ≥15 years) chose to participate by August 2020 [27]. The intervention measures deployed in Japan are noncompulsory, informed consent is requested prior to participation, and no legal penalty is imposed in the case of noncompliance; these features are less aggressive than those of digital strategies in some other countries [22-25].

In Japan, 69,001 confirmed cases of COVID-19 and 1307 deaths were reported from January 14 to September 2, 2020; the numbers of cases and deaths during this time were 3,769,523 and 66,333 for India and 56,852 and 27 for Singapore, respectively [26]. The effectiveness of containment in Japan is greater than that in India in terms of both infected cases and mortality. However, countries with compulsory contact tracing measures (eg, Singapore) appear to outperform those with noncompulsory measures (eg, India and Japan). It was estimated that older people (ie, age ≥65 years) in Japan would comprise 28.7% of the total population by the end of August 2020 [18,25-27]; the rates of severe cases and mortality among older people were 2.0% and 1.9%, respectively, during the same period, which are lower than the global average mortality rate of 3.3% (852,758/25,602,665) [18,26].

A simulation performed at Oxford University suggested that digital contact tracing would fail to decrease the spread of COVID-19 if the rate of participation fell to <60% (600,000/1,000,000) [28]. Similar analyses by other researchers reinforce that varying adoption rates of peer-to-peer contact tracing apps can influence the trajectory of the pandemic [17,29]. By employing the simulation model described in [30], we estimated how the rate of participation would affect the trajectory of the COVID-19 pandemic in Japan (Figure 5). The model evaluates scenarios in which the epidemic is established and countermeasures such as contact tracing are employed to control the spread of COVID-19; it can be observed how the trend of the effective reproduction number (R_t) and thus of the outbreak would dynamically change. Prior research identified that when R_t, which is defined as the average number of secondary cases generated by a single infectious case, decreases to less than one (ie, R_t<1), transmission of the disease will stop and the pandemic will ultimately be contained [22-24]. The simulation was calibrated to the demographic attributes in Japan

[14,22,25-27,29], the basic reproduction number (ie, 2.56) found for Japan [30], and the ratios of symptomatic patients in the report by the National Institute of Infectious Diseases in Japan [31]. The assumptions for the simulation are as follows: (1) the population can freely choose to opt in or opt out from COCOA; (2) there are no delays in data sharing; and (3) all the populations in households, schools, workplaces, and other scenarios can be successfully digitally traced. The simulation outcome (Figure 5) shows that when the participation rate increases starting from zero, the effective reproduction number decreases gradually from a value >1. However, the pandemic would finally be contained when a threshold was exceeded and more people chose to opt in. To meaningfully contain the spread of COVID-19 (ie, R_t<1), approximately 90% participation of the population would be required, which reinforces prior findings that controlling COVID-19 requires an estimated population uptake ranging from 56%-95% for contact-tracing apps [32].

Figure 5. Simulation of the association between the rate of participation in the COVID-19 Contact-Confirming Application (COCOA) framework and the R_t of COVID-19. R_t: effective reproduction number.



Strengths, Limitations, and Future Directions

The COVID-19 pandemic has imposed unprecedented challenges upon individuals, health care providers, and public authorities; meanwhile, different countries are using differentiated digital technologies and strategies to contain the spread of COVID-19, taking into account both technical and nontechnical factors. Digital health is not a panacea that will solve all difficulties; however, it does enhance the potential to counteract the disease compared to manual contact tracing [33,34].

In comparison with the Bluetooth-based digital health frameworks in other countries (eg, India or Singapore) or centralized approaches, COCOA more effectively protects the privacy of individuals from their counterparts, potential attackers, and public authorities without sacrificing accuracy or efficiency. First, users are not tapped to self-report personal data (eg, names or telephone numbers) through the app; this enhances the efficacy and eliminates the need for proofreading when users input incorrect data. Additionally, concerns about malicious use or illegal breach of private data can be waived. Moreover, persons are not requested to share their private information when infected or provide sensitive data, which may be subject to linkage attacks, when notified of infection. The authentication of infection precludes malicious exploitation of or attack on accurate infection information by other individuals, preventing misinformation regarding exposure. Location details, which are an unsuitable proxy for exposure, are not collected; this ensures that the movement paths of individuals will not be tracked and the identities of the individuals will hence not be disclosed.

Further, because the matching of exposure inference is performed on local sites, and subsequent provision of instructions when exposure risk is identified can curtail unnecessary in-person visits and risk of crosstransmission, the load of pressure on the health care system can be substantially balanced [9]. Bluetooth digital health uses proximity to identify close contact; hence, the speed of close contact detection is faster than that of other non-Bluetooth digital approaches. COCOA can run automatically in the background without interfering with other apps, which reduces errors from manual operation and enhances its efficiency. The deployment of the app contributes to increased risk perception and the reduction of nationwide population mobility.

Contact tracing is an essential part of transitioning back to normal economic rhythms while simultaneously managing the risk of subsequent cyclical outbreaks [9,33-35]. The benefits

can be multiple for a variety of responders. For potentially exposed or infected individuals, it is possible to know whether risky exposure has been established or disseminate knowledge to others without disclosure of identity or leakage of confidential information, which can increase individuals' confidence and trust in the health care system and self-awareness of their own behavior changes. This may facilitate appropriate and timely responses to the disease. For potential attackers, as no private information is available regarding either infected or exposed people for illegal or unauthorized exploitation, abnormal activities inherent to centralized or other digital health frameworks can be waved. For public authorities, the triage of patients and cumbersome matching inference of exposure are significantly trimmed; hence, the pressure on the medical system is expected to be alleviated. Further, as public authorities do not store personal private information, the risk from any attack on or misuse of their data is minimized. Bluetooth digital health has great potential to be used as a routine and mainstream tool in future outbreaks [36]. The rate of participation is expected to increase over time. COCOA supplies a new approach that is supplemental to extant digital health frameworks that fail, either partially or completely, in these facets. It could perform or match well in contexts where the population is highly privacy-sensitive and where limited health care resources are at risk of collapse.

Although decentralized telehealth has a variety of benefits and strengths, it has disadvantages as well [33,34,37]. Multiple critical concerns must be addressed to achieve effective containment.

First, the participation rate can essentially affect the trajectory of the outbreak. Studies have shown that societal level benefit hinges on broad and diverse user participation [19,38-43]. A low rate of participation can be associated with factors such as users' altruism, the population in rural or remote areas, wireless connectivity, availability of digital health, level of digital illiteracy, and legitimation regulations. In some countries, the use of private data is protected legally; whether this applies to other settings may need require more study and more time [5,9]. Solutions that have performed well for some communities may not work well in other communities with different cultural norms, legitimate regulations, and shared perceptions of privacy. From the legislation perceptive, public disclosure of individuals' protected data may be a violation of law in some contexts [39]. A higher level of participation may be achieved through mandatory legal regulations, enforcing adoption or substantial enhancement of shared public awareness. However, the rapid adoption of compulsory digital health measures without public consensus and discussions could provoke debates due to the fundamental heterogeneity in the attitudes regarding how digital health should function and, crucially, who should have access to the generated data [38-42,44,45]. Residents' perception of privacy and trust in public authorities can vary from culture to culture, which can impact the captured definition of individual privacy preservation [39-41]. A survey conducted in five other countries (ie, France, Germany, Italy, the United Kingdom, and the United States) found that people in settings with stronger public privacy and security concerns are relatively less supportive of app-based contact tracing, and individuals with less trust in public authorities are also less supportive [46,47].

Second, the delays in data sharing could allow the spread of COVID-19 to continue, increasing the time and effort needed to contain it [9,46,47]. If infectious individuals and their close contacts could be identified with efficacy, the effectiveness of digital health could be increased remarkably, and limited health care resources could thus be prioritized for the quarantining and treatment of the most severe cases [42]. However, this mechanism is compromised during a pandemic, in which delays of data sharing occur. Voluntary participation could cause noncompliance, generating a latency in responses [42,43]. The spread of COVID-19 hinges partially on the efficacy of data sharing and promptness of responses, given the infectiousness of the pathogen [42]. The greater the delays, the more difficult it is to contain the outbreak. Hence, timely sharing of information is critical to prevent subsequent cyclical outbreaks [43]. Finally, as data are automatically erased after a periodic interval, it is difficult to evaluate the long-term effects of a decentralized Bluetooth approach [13].

Future research could examine how privacy-enabled noncompulsory Bluetooth digital health can both quantitatively and qualitatively reduce the effectiveness of contact tracing relative to compulsory interventions. It could also examine ways to improve critical factors such as participation rate and delays of data sharing in these settings to enhance the effectiveness of containment. With the combined efforts of a variety of responders, the negative impacts of these factors are expected to be minimized. Coupled with the advancement in digital technologies and scientific understanding, telehealth can be enhanced to serve as a sustainable and mainstream solution to counter the COVID-19 pandemic, and it can be simultaneously employed as a routine tool to protect the privacy and well-being of the public [16].

Conclusions

The balance between privacy protection, public health, and other objectives is controversial [13]. COCOA contributes to prioritizing the preservation of users' privacy more effectively than the centralized Bluetooth digital health frameworks used in some other countries. The matching inference of exposure is performed locally, and individuals can self-triage their risk of exposure, which facilitates the load balancing of pressure on the medical system. It works better in load-balancing than centralized frameworks. As public authorities do not collect or manage users' sensitive personal information, concerns regarding illegal use or malicious attacks on private data can be disregarded. The detection of close contact is rapid and effective, and it reduces the likelihood of crosstransmission and in-person contacts. The background running feature enhances the efficacy of the approach and reduces errors of operation, which could be vital in the fight against highly infectious diseases such as COVID-19.

Since the deployment of COCOA, an average of 20% reduction in population mobility has been observed in Japan, which has affected the trajectory of the outbreak. With the wide spread of wireless connections and advancements in digital technologies, digital health can reduce inequality in access to health resources, promote health literacy, and improve risk perceptions. The

Tokyo area has observed faster growth in the number of infected cases than other prefectures in Japan [22,24]; hence, substantial improvements in the participation rate and speed of data sharing are of great concern in these densely populated communities or in places where the risk of close contact is high [22-24].

Countries diverge in their digital health frameworks and technologies. Decentralized privacy-first Bluetooth approaches can protect citizens' sensitive information, but possibly at the expense of compromised participation and impeded central surveillance. In contrast, a centralized data-first framework can warrant traceable data but may substantially violate individuals' privacy. Cultures differ in the perception and definition of privacy. The lack of a consensus on privacy protection in contact tracing incurs risks of noncompliance, as evidenced by recent privacy scandals [42,43]. This has hindered governments' capacity to effectively respond to the pandemic. The deployment and acceptance of telehealth in specific settings reflect both technical and nontechnical factors such as regional heterogeneity, cultural conflicts, shared altruism, and legal regulations [9,44].

Given that participation and data sharing are nonbinding, the privacy-first approach could consistently generate skepticism but ideally will enable the implementation to mitigate current and subsequent cyclical pandemics [41]. Coupled with the efforts from a variety of responders, the rate of participation and delays in data sharing are expected to improve over time. Countries using the decentralized Bluetooth approach must prioritize deliberation of how currently unresolved problems can be addressed to contain the spread of COVID-19. Digital health itself cannot overcome all these challenges; however, by combining it with other countermeasures, such as social distancing, early case isolation, and hygiene practice, it is feasible to achieve meaningful containment [45]. With these improvements, it could be feasible to achieve a balance between privacy preservation and public health by enabling individuals to have full control over sensitive data, identify local exposure risk, share their data in a timely fashion, and enact prompt responses [42].

This decentralized Bluetooth approach will undoubtedly upgrade its definition with advancement in digital health, digital technologies, and a more accurate scientific understanding of the disease. Lessons learned from this current deployment will play paramount roles in future pandemics, further aid the establishment of an effective routine surveillance approach, and provide meaningful insights for other countries and regions.

Acknowledgments

Fujian Provincial Social Science Planning Project in 2020: Research on Capacity Building of Regional Medical Collaborative Innovation Community under the Background of "Internet + Medical" (No. FJ2020B038), the Science and Technology Development Center of the Ministry of Education (No. 2019J01008), Key Projects of the National Social Science Foundation (No. 19AGL017), Fujian Natural Science Foundation (No. 2018J01648), Development Fund of Scientific Research from Fujian University of Technology (No. GY-S18109), FuJian Philosophy and Social Sciences Association (No. FJ2019B081), and the Ministry of Education's Humanities and Social Sciences Research Project (No. 18YJAZH153).

Authors' Contributions

IN designed the approach and edited the draft of the manuscript. MHJ, YH, GTT, and MJ analyzed the mechanism and framework. IN, YG, SW, and WQZ contributed to the modification of the manuscript. IN performed the simulation and provided critical analyses of the manuscript. All authors confirmed and approved all sections of the final manuscript.

Conflicts of Interest

None declared.

References

1. Coronavirus disease (COVID-19) Situation Report– 161. World Health Organization. 2020 Jun 29. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200629-covid-19-sitrep-161.pdf?sfvrsn=74fde64e_2 [accessed 2020-08-27]
2. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health 2020 May;5(5):e261-e270 [FREE Full text] [doi: 10.1016/S2468-2667(20)30073-6] [Medline: 32220655]
3. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol 2004 Sep 15;160(6):509-516 [FREE Full text] [doi: 10.1093/aje/kwh255] [Medline: 15353409]
4. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health 2020 Apr;8(4):e488-e496. [doi: 10.1016/s2214-109x(20)30074-7]
5. Li J, Guo X. Global Deployment Mappings and Challenges of Contact-tracing Apps for COVID-19. SSRN Journal Preprint posted online on May 26, 2020. [doi: 10.2139/ssrn.3609516]

6. Wang S, Ding S, Xiong L. A New System for Surveillance and Digital Contact Tracing for COVID-19: Spatiotemporal Reporting Over Network and GPS. JMIR Mhealth Uhealth 2020 Jun 10;8(6):e19457 [FREE Full text] [doi: 10.2196/19457] [Medline: 32499212]
7. Fisk M, Livingstone A, Pit SW. Telehealth in the Context of COVID-19: Changing Perspectives in Australia, the United Kingdom, and the United States. J Med Internet Res 2020 Jun 09;22(6):e19264 [FREE Full text] [doi: 10.2196/19264] [Medline: 32463377]
8. Ekong I, Chukwu E, Chukwu M. COVID-19 Mobile Positioning Data Contact Tracing and Patient Privacy Regulations: Exploratory Search of Global Response Strategies and the Use of Digital Tools in Nigeria. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e19139 [FREE Full text] [doi: 10.2196/19139] [Medline: 32310817]
9. Nakamoto I, Wang S, Guo Y, Zhuang W. A QR Code-Based Contact Tracing Framework for Sustainable Containment of COVID-19: Evaluation of an Approach to Assist the Return to Normal Activity. JMIR Mhealth Uhealth 2020 Sep 07;8(9):e22321 [FREE Full text] [doi: 10.2196/22321] [Medline: 32841151]
10. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 Contact Tracing and Data Protection Can Go Together. JMIR Mhealth Uhealth 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: 10.2196/19359] [Medline: 32294052]
11. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491) [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
12. McCall B. Shut down and reboot-preparing to minimise infection in a post-COVID-19 era. Lancet Digit Health 2020 Jun;2(6):e293-e294 [FREE Full text] [doi: 10.1016/S2589-7500(20)30103-5] [Medline: 32346671]
13. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
14. Kobayashi G, Sugasawa S, Tamae H, Ozu T. Predicting intervention effect for COVID-19 in Japan: state space modeling approach. Biosci Trends 2020 Jul 17;14(3):174-181 [FREE Full text] [doi: 10.5582/bst.2020.03133] [Medline: 32461511]
15. Hayasaki E. Covid-19: how Japan squandered its early jump on the pandemic. BMJ 2020 Apr 24;369:m1625. [doi: 10.1136/bmj.m1625] [Medline: 32332010]
16. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol 2004 Sep 15;160(6):509-516 [FREE Full text] [doi: 10.1093/aje/kwh255] [Medline: 15353409]
17. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-Peer Contact Tracing: Development of a Privacy-Preserving Smartphone App. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
18. COVID-19 pandemic update. Webpage in Japanese. Ministry of Health, Labour and Welfare in Japan. 2020 Aug. URL: https://www.mhlw.go.jp/stf/houdou/houdou_list_202008.html [accessed 2020-08-28]
19. Covid Tracing Tracker. MIT Technology Review. URL: https://www.technologyreview.com/tag/covid-tracing-tracker/ [accessed 2020-09-08]
20. Garg S, Bhatnagar N, Gangadharan N. A Case for Participatory Disease Surveillance of the COVID-19 Pandemic in India. JMIR Public Health Surveill 2020 Apr 16;6(2):e18795 [FREE Full text] [doi: 10.2196/18795] [Medline: 32287038]
21. Cho H, Ippolito D, Yu Y. Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs. ArXiv Preprint posted online on March 03, 2020. [FREE Full text]
22. Karako K, Song P, Chen Y, Tang W. Analysis of COVID-19 infection spread in Japan based on stochastic transition model. Biosci Trends 2020 May 21;14(2):134-138 [FREE Full text] [doi: 10.5582/bst.2020.01482] [Medline: 32188819]
23. Yabe T, Tsubouchi K, Fujiwara N, Wada T, Sekimoto Y, Ukkusuri SV. Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. Sci Rep 2020 Oct 22;10(1):18053 [FREE Full text] [doi: 10.1038/s41598-020-75033-5] [Medline: 33093497]
24. Parady G, Taniguchi A, Takami K. Travel behavior changes during the COVID-19 pandemic in Japan: Analyzing the effects of risk perception and social influence on going-out self-restriction. Transportation Research Interdisciplinary Perspectives 2020 Sep;7:100181. [doi: 10.1016/j.trip.2020.100181]
25. Shiina A, Niitsu T, Kobori O, Idemoto K, Hashimoto T, Sasaki T, et al. Relationship between perception and anxiety about COVID-19 infection and risk behaviors for spreading infection: A national survey in Japan. Brain Behav Immun Health 2020 Jul;6:100101 [FREE Full text] [doi: 10.1016/j.bbih.2020.100101] [Medline: 32835297]
26. WHO Coronavirus Disease (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2020-09-03]
27. Statistics Bureau of Japan. URL: https://www.stat.go.jp/english/index.html [accessed 2020-08-30]
28. Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown. University of Oxford. 2020 Apr 16. URL: https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown [accessed 2020-08-29]
29. Bilinski A, Mostashari F, Salomon JA. Modeling Contact Tracing Strategies for COVID-19 in the Context of Relaxed Physical Distancing Measures. JAMA Netw Open 2020 Aug 03;3(8):e2019217 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19217] [Medline: 32821920]

30. Kurita J, Sugawara T, Ohkusa Y. Estimated effectiveness of school closure and voluntary event cancellation as COVID-19 countermeasures in Japan. J Infect Chemother 2021 Jan;27(1):62-64 [FREE Full text] [doi: 10.1016/j.jiac.2020.08.012] [Medline: 32896479]
31. Descriptive epidemiology of 516 confirmed cases of novel coronavirus infection reported by the national epidemiological surveillance of infectious diseases (NESID) system and active epidemiological surveillance (as of March 23, 2020). National Institute of Infectious Diseases. 2020 Apr 09. URL: https://www.niid.go.jp/niid/en/2019-ncov-e/2484-idsc/9555-covid19-14-200323-e.html [accessed 2020-09-04]
32. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. Lancet Digit Health 2020 Nov;2(11):e607-e621. [doi: 10.1016/s2589-7500(20)30184-9]
33. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, CMMID COVID-19 working group. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect Dis 2020 Oct;20(10):1151-1160 [FREE Full text] [doi: 10.1016/S1473-3099(20)30457-6] [Medline: 32559451]
34. 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]
35. Liu Y, Wang Z, Ren J, Tian Y, Zhou M, Zhou T, et al. A COVID-19 Risk Assessment Decision Support System for General Practitioners: Design and Development Study. J Med Internet Res 2020 Jun 29;22(6):e19786 [FREE Full text] [doi: 10.2196/19786] [Medline: 32540845]
36. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491) [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
37. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Mar 20;26(5):309-313. [doi: 10.1177/1357633x20916567]
38. Kapa S, Halamka J, Raskar R. Contact Tracing to Manage COVID-19 Spread-Balancing Personal Privacy and Public Health. Mayo Clin Proc 2020 Jul;95(7):1320-1322 [FREE Full text] [doi: 10.1016/j.mayocp.2020.04.031] [Medline: 32622440]
39. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science 2020 May 15;368(6492):742-746 [FREE Full text] [doi: 10.1126/science.abb4557] [Medline: 32269067]
40. Maese JR, Seminara D, Shah Z, Szerszen A. Perspective: What a Difference a Disaster Makes: The Telehealth Revolution in the Age of COVID-19 Pandemic. Am J Med Qual 2020 Jun 11;35(5):429-431 [FREE Full text] [doi: 10.1177/1062860620933587] [Medline: 32525394]
41. Gao Y, Liu R, Zhou Q, Wang X, Huang L, Shi Q, COVID-19 Evidence and Recommendations Working Group. Application of telemedicine during the coronavirus disease epidemics: a rapid review and meta-analysis. Ann Transl Med 2020 May;8(10):626. [doi: 10.21037/atm-20-3315] [Medline: 32566563]
42. Legido-Quigley H, Asgari N, Teo YY, Leung GM, Oshitani H, Fukuda K, et al. Are high-performing health systems resilient against the COVID-19 epidemic? Lancet 2020 Mar;395(10227):848-850. [doi: 10.1016/S0140-6736(20)30551-1]
43. Li DKT, Zhu S. Contributions and challenges of general practitioners in China fighting against the novel coronavirus crisis. Fam Med Com Health 2020 Mar 23;8(2):e000361. [doi: 10.1136/fmch-2020-000361]
44. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733. [doi: 10.1056/nejmoa2001017]
45. Riou J, Althaus C. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. Euro Surveill 2020 Jan;25(4):25 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.4.2000058] [Medline: 32019669]
46. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. JMIR mHealth uHealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
47. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science 2020 Jul 10;369(6500):eabb9789 [FREE Full text] [doi: 10.1126/science.abb9789] [Medline: 32414780]

Abbreviations

COCOA: COVID-19 Contact-Confirming Application
HER-SYS: Health Center Real-time Information-sharing System
PCR: polymerase chain reaction
QR: quick response
R_t: effective reproduction number

Edited by G Eysenbach; submitted 04.07.20; peer-reviewed by J Li, A Dawood, M Adly, A Adly, A Adly; comments to author 24.08.20; revised version received 08.09.20; accepted 13.09.20; published 01.12.20.

Please cite as:
Nakamoto I, Jiang M, Zhang J, Zhuang W, Guo Y, Jin MH, Huang Y, Tang K
Evaluation of the Design and Implementation of a Peer-To-Peer COVID-19 Contact Tracing Mobile App (COCOA) in Japan
JMIR Mhealth Uhealth 2020;8(12):e22098
URL: https://mhealth.jmir.org/2020/12/e22098
doi:10.2196/22098
PMID:33170801

Original Paper

Use of WhatsApp for Polyclinic Consultation of Suspected Patients With COVID-19: Retrospective Case Control Study

Ramazan Sabırlı[1*], MD; Emre Karsli[1*], MD; Omer Canacik[1*], MD; Dogan Ercin[1*], MD; Handan Çiftçi[1*], MD; Levent Sahin[1*], MD; Turgut Dolanbay[1*], MD; Emin Ediz Tutuncu[2*], Prof Dr

[1]Kafkas University Faculty of Medicine, Department of Emergency Medicine, Kars, Turkey

[2]Kafkas University Faculty of Medicine, Department of Infectious Diseases, Kars, Turkey

[*]all authors contributed equally

Corresponding Author:
Ramazan Sabırlı, MD
Kafkas University Faculty of Medicine
Department of Emergency Medicine
Kafkas University Hospital
Kars, 36000
Turkey
Phone: 90 5387344298
Email: ramazan_sabirli@hotmail.com

Abstract

Background: Telephones, internet-connected devices (phablets, personal computers), chat platforms, and mobile apps (eg, Skype, Facebook Messenger, WhatsApp) can be exploited for telemedicine applications. WhatsApp and similar apps are also widely used to facilitate clinical communication between physicians. Moreover, WhatsApp is used by emergency department (ED) physicians and consulting physicians to exchange medical information during ED consultations. This platform is regarded as a useful app in the consultation of dermatological and orthopedic cases. Preventing overcrowding in the ED is key to reducing the risk of disease transmission, and teleconsulting practice is thought to be effective in the diagnosis, treatment, and reduction of transmission risk of disease, most notably during the COVID-19 pandemic. Video consultation is highly recommended in some countries on the grounds that it is likely to reduce the risk of transmission. WhatsApp-like apps are among the video consultation platforms that are assumed to reduce the risk of contamination by minimizing patient-physician contact.

Objective: The aim of this study was to investigate the effects of WhatsApp video consultation on patient admission and discharge times in comparison to bedside consultation in the evaluation of potential patients with COVID-19 visiting a COVID-19 outpatient clinic during the pandemic.

Methods: Patients who presented to the ED COVID-19 outpatient clinic between March 11 and May 31, 2020, and for whom an infectious disease specialist was consulted (via WhatsApp or at bedside) were included in the study in accordance with the inclusion and exclusion criteria. Eventually, 54 patients whose consultations were performed via WhatsApp and 90 patients whose consultations were performed at bedside were included in our study.

Results: The median length of stay in the ED of discharged patients amounted to 103 minutes (IQR 85-147.75) in the WhatsApp group and 196 minutes (IQR 141-215) in the bedside group. In this regard, the length of stay in the ED was found to be significantly shorter in the WhatsApp group than in the bedside group (P<.001). Among the consulted and discharged patients, 1 patient in each group tested positive for SARS-CoV-2 by polymerase chain reaction test and thus was readmitted and hospitalized (P=.62). The median length of stay of the inpatients in the ED was found to be 116.5 minutes (IQR 85.5-145.5) in the WhatsApp group and 132 minutes (IQR 102-168) in the bedside group. The statistical analysis of this time difference revealed that the length of stay in the ED was significantly shorter for patients in the WhatsApp group than in the bedside group (P=.04).

Conclusions: Consultation via WhatsApp reduces both contact time with patients with COVID-19 and the number of medical staff contacting the patients, which contributes greatly to reducing the risk of COVID-19 transmission. WhatsApp consultation may prove useful in clinical decision making as well as in shortening process times. Moreover, it does not result in a decreased accuracy rate. The shortened discharge and hospitalization timespans also decreased the length of stay in the ED, which can have an impact on minimizing ED crowding.

Trial Registration: ClinicalTrials.gov NCT04645563; https://clinicaltrials.gov/ct2/show/NCT04645563.

(JMIR Mhealth Uhealth 2020;8(12):e22874) doi:10.2196/22874

KEYWORDS

telemedicine; WhatsApp; ED crowding; COVID-19; emergency department; messaging; communication; consultation; clinic; infectious disease

Introduction

Telephones, internet-connected devices (phablets, personal computers), chat platforms, and mobile apps (eg, Skype, Facebook Messenger, WhatsApp) can be exploited for telemedicine applications. Telemedicine applications can be classified under four headings: communication method, timing of the delivered information, purpose of the consultation, and interaction between the individuals involved. The importance of video-based telemedicine applications lies in enabling real-time communication that is as close as possible to face-to-face meetings, facilitating the diagnosis of patients, and viewing patients and their data by health care provider professionals [1].

Preventing crowding of the emergency department (ED) is key to reducing disease transmission risk; teleconsulting practice is thought to be effective in the diagnosis, treatment, and reduction of the transmission risk of disease, most notably during the COVID-19 pandemic [2]. In addition, the British Medical Association recommends that consultations be performed over video in England, Wales, and Scotland to minimize the risk of COVID-19 infection [3].

Providing text, video, and audio transfer, the WhatsApp app was developed for smartphones and supports iOS, Android, and Windows operating systems. It enables users to send photographs, videos, voice and text messages, and documents as well as to make free calls to each other via 2G, 3G, 4G, or Wi-Fi internet connections [4]. WhatsApp, one of the most popular messaging apps, was used by around 2 billion people worldwide in March 2020 [5].

WhatsApp and similar platforms are also widely used to facilitate clinical communication between physicians [6-8]. This platform is considered to be a useful application to coordinate intradepartmental communication [9]. Moreover, WhatsApp is used by ED physicians and consulting physicians to exchange medical information during ED consultations [10,11].

Although some studies on WhatsApp consultations in the ED exist in the literature, we did not encounter any studies investigating the impact of WhatsApp-mediated consultations in polyclinics during a pandemic.

The aim of this study was thus to investigate the effects of WhatsApp video consultation on patient admission and discharge times in comparison to bedside consultation in the evaluation of potential patients with COVID-19 visiting the COVID-19 outpatient clinic during the pandemic period.

Methods

Study Design

Ethical approval (number 80576354-050-99/175) was granted by the Clinical Research Ethics Committee of Kafkas University before launching this retrospective observational study. The study was registered prior to launch (ClinicalTrials.gov NCT04645563).

Study Population

During the course of the COVID-19 pandemic, patients underwent consultations with the Infectious Disease (ID) department through bedside consultation or video transmission from WhatsApp. The consultation methods of the patients were scanned retrospectively (from the fixed ED telephone and the hospital system) and recorded in the data set. The method of consultation through video transmission was implemented at the beginning of the pandemic to increase communication between the emergency service and the ID department.

The durations of the videos that were sent by WhatsApp for consultation, the response of the ID specialist to the video, and the clinician's response to the bedside consultation were saved to the data set. Additionally, the time from the patient's application to computed tomography (CT), the time from the patient's application to the ED to receiving their laboratory results, the time spent by the patient in the ED, and the time periods of consultation within working hours (8 AM to 5 PM) or outside working hours (5 PM to 8 AM) were also saved in the data set.

Patients who presented to the COVID-19 outpatient clinic at the ED between March 11 and May 31, 2020, and for whom an ID physician was consulted (via WhatsApp or at bedside) were included in the study in accordance with the inclusion and exclusion criteria.

In this cohort, patients whose consultations were held before admission to another department and patients afflicted with problems involving multiple departments were excluded from the study.

Consultation

To isolate potential patients with COVID-19 during the pandemic, these patients were examined at the COVID-19 outpatient clinic under the supervision of the ED. Following their examination at this clinic, consultations for patients conforming to the classification of potential COVID-19 were held with the ID department. The potential patients with COVID-19 were assessed by an ED physician as specified in the Potential COVID-19 Cases Guide of the Ministry of Health [12]. During the COVID-19 pandemic, not every consultation could be performed at the patient's bedside because the schedules of both the ED and ID physicians were twice as busy

as normal due to the congestion of working conditions and the organization of pandemic areas. During the course of the pandemic, consultations via WhatsApp emerged naturally due to the overcrowding in the COVID-19 outpatient clinic.

Eligible patients who were examined by an ED physician in consultation with an ID physician at bedside or via WhatsApp were evaluated in the study.

Common Points in Both Types of Consultations

The ID physician evaluated the patients' laboratory and CT data in both types of consultations. The ID physician evaluated these data via smartphone in the WhatsApp consultation and through the hospital database in the bedside consultation.

The ID physician personally performed a physical examination of the patient in the ED during the bedside consultation; in the WhatsApp consultation, this assessment was performed after the patient was referred to the ID department from the ED.

In both types of consultations, the ID physician was called for a consultation following the finalization of the CT report and laboratory parameter results. All the consultations were evaluated by a single ID physician. Seven ED physicians sent the patient data to the ID physician via WhatsApp.

Bedside Consultation

The ED physician wrote a consultation note in the hospital information system specifying the patient's clinical status, history, and laboratory parameters. A physician was consulted for all eligible participants after their laboratory results and CT reports were complete. The ID physician examined the patients at the bedside within 30 minutes (the legal response time in Turkey) of seeing the consultation request [13]. The consultation response time was saved as the time from the entry of the consultation information in the system to the completion of the consultation note.

Although the consultant ID physicians were informed via WhatsApp, they held consultations at the bedside of patients for whom they deemed this type of consultation was appropriate.

WhatsApp Consultation

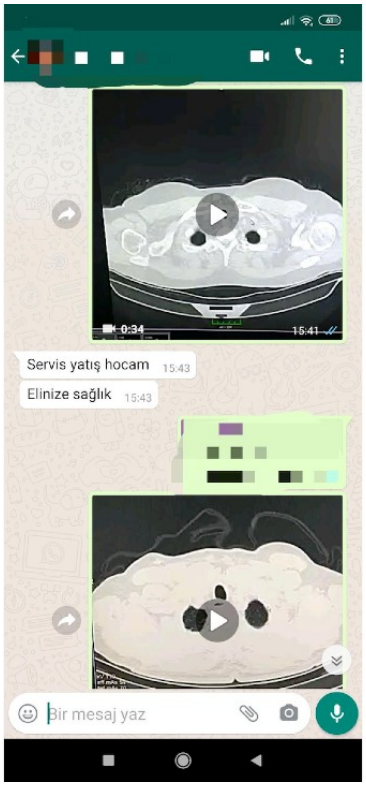
The ID physician was consulted for all participants once complete laboratory results and CT reports were obtained. All the consultations were performed with the same smartphone, and every WhatsApp consultation held since the beginning of the pandemic was evaluated.

In this type of consultation, the patient's thorax CT images were converted into a video approximately 30-35 seconds in length, and during this video recording, the patient's clinical condition and laboratory results were also transferred to the ID physician. The ID physician then stated their admission-discharge decision via WhatsApp as "hospitalization" or "discharge." Eventually, the consultation result was recorded in the patient's folder. The moment the video was sent was recorded as the beginning of the patient's consultation period, and the response time to the WhatsApp video was saved as the consultation response time.

Video Shooting and Features

The videos were captured using a Xiaomi Redmi Note 8 smartphone (Xiaomi Corporation) with a 48-megapixel rear camera in such a way to capture the entire computer screen while concealing the patient's name. The mediastinal window image of the thorax CT scan was recorded at a rate of approximately 10 sections per second from start to finish, and all the patient's clinical and laboratory information was dubbed in the video (Figure 1). Once completed, the video recording was forwarded to the ID physician via WhatsApp Messenger. The smartphone provided for the consultation was only used in the routine functioning of the ED to inform the physician in charge.

Figure 1. Screenshot of a WhatsApp consultation message.



Video and Audio Quality

The features of the videos after transmission to the other device are as follows: video screen size, 480 × 848 pixels; data transfer rate, 1353 kb/sec; bit rate, 1413 kb/sec; frames per second, 30. The audio features were bitrate 60 kb/sec and two-channel stereo, and the sound sample value was 44.100 kHz.

Exclusion Criteria

The exclusion criteria in the present study were as follows: patients whose consultations were held before admission to another department, patients afflicted with problems involving multiple departments and whose consultations were performed via WhatsApp, patients whose consultations were held by sending photographs, patients who were referred from another outpatient clinic and whose consultation procedures were completed in the ED, and patients whose consultations were not performed on the same phone.

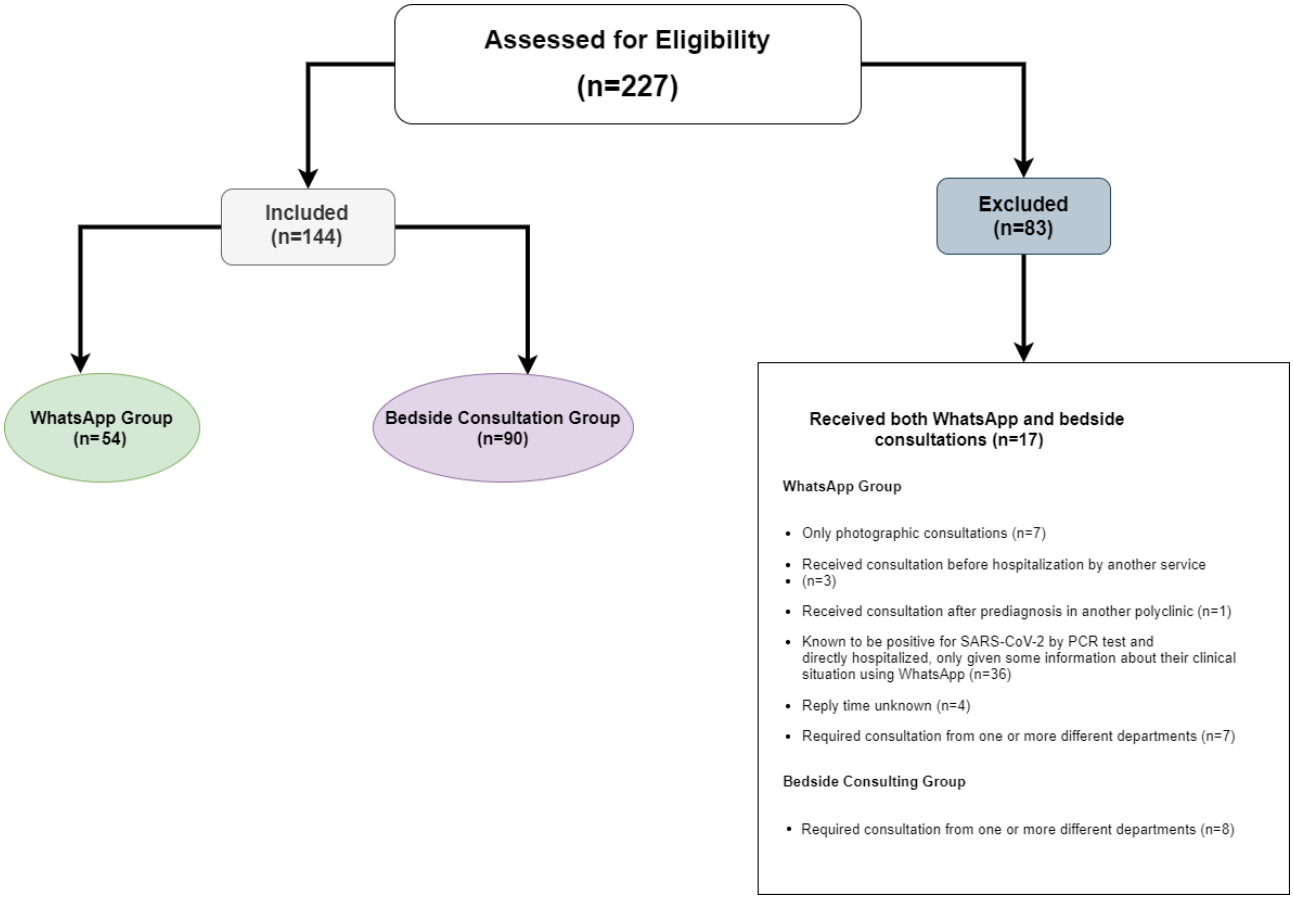
Data Analysis

Because there is no reference study framed in a similar design, at least 90 people (45 people for each group) were needed to obtain 80% power with 95% confidence based on the power analysis conducted in line with the assumptions, presuming that the effect size we expected from the study would be high (f=0.6). The data were analyzed with SPSS (IBM Corporation). Continuous variables are presented as median (IQR), while categorical variables are provided as numbers and percentages. Kolmogorov-Smirnov analysis was used in the normality distribution analysis. When the parametric assumptions were not met, the Mann-Whitney U test was used to compare differences. For all the analyses, P<.05 was considered statistically significant.

Results

Out of 227 total consultations by the ID physician between March 10 and May 31, 2020, 112 were performed only via WhatsApp, 17 both via WhatsApp and at bedside, and 98 only at bedside. Out of 112 patients who received consultations via WhatsApp, the excluded participants consisted of 7 patients who were assessed by sending photographs only, 3 patients who consulted the ID service before being hospitalized in another department, 7 patients requiring consultation by more than one department, 1 patient who received a consultation after being examined in the outpatient clinic, 36 patients known to have tested positive for COVID-19 by polymerase chain reaction (PCR) and who were directly hospitalized, and 4 patients for whom the response return times after consultation were unrecorded. Eventually, 54 patients whose consultations were performed via WhatsApp were included in our study. In the bedside group, 8 patients were excluded from the study because of problems concerning other services; therefore, 90 patients were ultimately included in this group. A flowchart of the selection of the study participants is shown in Figure 2. As a result of the post-hoc power analysis, the effect size was calculated as f=0.63, and 95.49% power was achieved within a 95% confidence interval.

Figure 2. Flowchart of the selection of the study participants.



Of the 90 patients assessed by the bedside consultation method, 15 (17%) were discharged, while 75 (83%) were hospitalized. On the other hand, 20 (37%) of the 54 patients consulted via WhatsApp were discharged, while 34 (63%) were hospitalized (Table 1).

Considering the time intervals in which the patients received consultations, in the WhatsApp group, 32/54 patients (59%) received consultations between 8 AM and 5 PM and 22/54 patients (40.7%) received consultations between 5 PM and 8 AM. In contrast, in the bedside consultation group, 65/90 patients (72%) received consultations between 8 AM and 5 PM, and 25/90 patients (28%) received consultations between 5 PM and 8 AM. However, no significant difference was noted between the groups in terms of consultation time interval (P=.10) (Table 1).

Among all the patients, 2/154 (1.3%) tested positive for COVID-19 by PCR, one in each consultation group; these patients were readmitted and hospitalized (P=.62) (Table 1).

The median ages of the patients consulted via WhatsApp and at the bedside were 59 years (IQR 48-69) and 55 years (IQR 32.5-68.25), respectively. Further, the WhatsApp group consisted of 33/54 (61%) male and 21/54 female (39%) participants, whereas the bedside group (43%) included 51 male (56.7%) and 39 female patients. There was no difference in median age and gender between the groups. (P=.23 and P=.72, respectively) (Table 1).

Table 1. Demographic data of the study participants in both groups (N=154).

Characteristic	WhatsApp group (n=54)	Bedside consultation group (n=90)	*P* value
Age (years), median (IQR)	59 (48-79.5)	55 (32.5-77.9)	.23[a]
Gender, n (%)			.72[b]
Male			
Hospitalized	14 (26)	39 (43)	
Discharged	13 (24)	12 (13)	
Female			
Hospitalized	20 (34)	36 (40)	
Discharged	7 (13)	3 (3)	
Consultation time intervals, n (%)			.10[b]
8 AM to 5 PM	32 (59)	65 (72)	
5 PM to 8 AM	22 (41)	25 (28)	
Readmitted, n (%)	1 (2)	1 (1)	.62[b]

[a]*P* value is derived from the Student *t* test.

[b]*P* value is derived from the chi-square test.

In the WhatsApp group, the median time from admittance to the CT scan was 17.5 minutes (IQR 11-52.5) for inpatients and 26 minutes (IQR 15-40.5) for discharged patients; meanwhile, the median times amounted to 32 minutes (IQR 15-51) and 27 minutes (IQR 16-47), respectively, in the bedside group. The median time period from admittance to the release of laboratory results was 54.5 minutes (IQR 48-67.75) for inpatients and 62.5 minutes (IQR 56.25-82.5) for discharged patients in the WhatsApp group, whereas the median of this period was found to be 60 minutes (IQR 49-76) for inpatients and 65 minutes (IQR 61-69) for discharged patients in the bedside group. Despite the varying numbers, no significant difference was found between the WhatsApp and bedside consultation groups (*P*=.05 and *P*=.29, respectively) in relation to the times to release of CT scan and laboratory results among the inpatients and discharged patients (*P*=.93 and *P*=.73, respectively) (Table 2).

The median duration of the videos sent during the WhatsApp consultation was 32 seconds (18.25-37.25) for the inpatients and 28.5 (16.25-36) seconds for the discharged patients. In the WhatsApp group, the median response time to consultation was calculated as 8 minutes (IQR 6-9) for the inpatients and 7.5 (IQR 4-13) for the discharged patients. In the bedside consultation group, the median response time to the consultation was calculated as 25 minutes (IQR 24-28) for the inpatients and 16 minutes (IQR 7-22) for the discharged patients. There was a significant difference between the WhatsApp and bedside consultation groups in consultation response times for both hospitalized and discharged patients (*P*=.004 and *P*<.001, respectively) (Table 2).

Table 2. Time data for video length and response consultation time in the WhatsApp and bedside consultation groups (N=144). All *P* values were derived from the Mann-Whitney U test.

Variable	WhatsApp group (n=54)			Bedside consultation group (n=90)			*P* values[c]	
	Hospitalized, median (IQR)	Discharged, median (IQR)	*P* value[a]	Hospitalized, median (IQR)	Discharged, median (IQR)	*P* value[b]	Hospitalized	Discharged
Video length (seconds)	32 (18.25-37.25)	28.5 (16.25-36)	.24	N/A[d]	N/A	N/A	N/A	N/A
Response to consultation time (minutes)	8 (6-9)	7.5 (4-13)	.89	25 (24-28)	16 (7-22)	<.001	.004	<.001

[a]Comparison of the times of hospitalized and discharged patients in the WhatsApp group.

[b]Comparison of the times of hospitalized and discharged patients in the bedside consultation group.

[c]Comparison of patients in the WhatsApp and bedside consultation groups.

[d]N/A: not applicable.

Concerning the length of stay of the discharged patients in the ED, the median time amounted to 103 minutes (IQR 85-147.75) in the WhatsApp group and 196 minutes (IQR 141-215) in the bedside group. In this regard, there is a statistically significant difference between the groups in terms of median length of stay in the ED (*P*<.001) (Table 3).

Table 3. Time data for CT results, laboratory results, and length of stay the in the WhatsApp and bedside consultation groups (N=144). All *P* values were derived from the Mann-Whitney U test.

Variable	WhatsApp group (n=54)			Bedside consultation group (n=90)			*P* values[c]	
	Hospitalized, median (IQR)	Discharged, median (IQR)	*P* value[a]	Hospitalized, median (IQR)	Discharged, median (IQR)	*P* value[b]	Hospitalized	Discharged
Time from admission to CT[d] results (minutes)	17.5 (11-52.5)	26 (15-40.5)	.32	32 (15-51)	27 (16-47)	.58	.05	.93
Time from admission to laboratory results (minutes)	54.5 (48-67.75)	62.5 (56.25-82.5)	.06	60 (49-76)	65 (61-69)	.17	.29	.73
Length of stay in ED[e] (minutes)	116.5 (85.5-145.5)	103 (85-147.75)	N/A[f]	132 (102-168)	196 (141-215)	N/A	.04	<.001

[a]Comparison of the times of hospitalized and discharged patients in the WhatsApp group.

[b]Comparison of the times of hospitalized and discharged patients in the bedside consultation group.

[c]Comparison of patients in the WhatsApp and bedside consultation groups.

[e]CT: computed tomography.

[d]ED: emergency department.

[e]N/A: not applicable.

Considering the length of stay of the inpatients in the ED, the median time was found to be 116.5 minutes (IQR 85.5-145.5) in the WhatsApp group and 132 minutes (IQR 102-168) in the bedside group. The statistical analysis of this time difference revealed that the length of stay of hospitalized patients in the ED was significantly shorter in the WhatsApp group than the bedside group (*P*=.04) (Table 3).

Discussion

Principal Findings

This study explored the impact of WhatsApp video consultations on patient admission and discharge times in comparison to bedside consultation in the evaluation of potential patients with COVID-19 visiting a COVID-19 outpatient clinic during the pandemic. Overall, our results reveal that the length of stay in the ED of patients who received consultations via WhatsApp was shorter than that of patients who received consultations at the bedside in both the discharged and inpatient cohorts. We also observed that the proportion of patients whose COVID-19 diagnosis was dismissed after being readmitted and hospitalized was similar in both groups.

Indeed, the relevant literature reports a wide range of devices or programs available for teleconsultation. Letters, telephone calls, pagers, fax machines, computer-based consultation systems, mobile phones, smartphones, and web-based medical recording programs have all been used for consultation over the course of history [14]. As technology improves, a number of computer programs and smartphone apps are being developed to prevent the emergence of congestion in the ED [15,16].

Studies on the use of WhatsApp in the ED tend to focus on its usefulness and patients' length of stay in the ED during the consultation [10,17]. A study probing into consultations performed through WhatsApp in the ED reported that WhatsApp is a useful app, particularly for communication with ED consultants located outside the hospital. Another study concluded that WhatsApp use might may be beneficial to patients requiring orthopedic consultation; however, although this approach is suggested to shorten the patients' length of stay in the ED, no specific data is available on the length by which this duration can be shortened [17]. Further, a study of patients with maxillofacial trauma suggested that the transfer of CT videos via WhatsApp enables rapid completion of consultation. Another study assessing patients in the ED through real-time videoconferencing reported that the duration of the consultations ranges between 16 and 19 minutes [18]. The legal time required for the completion of emergency consultation in Turkey is 30 minutes [13]. In our study, the response time to consultation through video transfer was found to be much shorter than the bedside consultation response time. In addition, the proportions of patients whose clinical diagnosis changed after WhatsApp and bedside consultations were similar, which overall supports the recommendations regarding the utility of WhatsApp in clinical consultation [17].

The transmission risk of COVID-19 infection tends to increase with the length of contact time. Therefore, it is crucial to reduce the overcrowding of pandemic outpatient clinics, where various precautions are introduced to avoid contact between the patient and the physician. In addition, shortening the duration of patients' stay in the ED is likely to minimize patient-physician contact. It can be assumed that WhatsApp use contributes indirectly to minimizing the infection risk of health care professionals during the COVID-19 pandemic, when one-to-one contact should be reduced to the greatest extent possible. In this regard, our study may shed light on the benefits of future video consultation systems in the ED.

One factor known to lead to ED crowding is the length of the consultation period [19]. When the consultation period is long, both the waiting period for patients and the number of patients waiting in the ED tend to increase. Waiting for laboratory and imaging results also adds to the length of stay in the ED [14]. In a prospective randomized controlled trial comparing standard telephone with WhatsApp in emergency consultations, the length of stay in the ED was reported to be shorter in the WhatsApp group than in the telephone group [20]. However, despite the interest in the use of secure messaging apps for consultations in the ED, no study has investigated WhatsApp consultation similarly to this study to the best of our knowledge. We argue that consultation via WhatsApp can shorten the length of stay in the ED for both discharged patients and inpatients. Because this type of consultation can shorten the length of hospital stays, it may be effective in minimizing ED crowding.

Limitations

Although the number of consultations we evaluated in the study is relatively low, our study reached sufficient power. Because there is only one ID physician in our hospital, we could not reach a judgment on whether the effectiveness of the WhatsApp consultation would change for people who have different smartphone use habits in the presence of more than one consultant. Furthermore, the fact that the decision of whether to perform a consultation via WhatsApp or at the bedside was made by an ID physician may have created some bias; however, given the consultation hours, conducting consultations of both groups at similar times is likely to reduce these biases to some extent. With a prospective and randomized study in a larger facility where more emergency medicine and ID specialists work, the effects of WhatsApp consultation on ED crowding and patients' time of stay in the ED can be stated more clearly, and bias may be minimalized.

Conclusion

The standard consultation method in our hospital is bedside consultation; in this study, we documented the data from our observations of consultations of patients performed by clinicians via the WhatsApp application during the natural course of the COVID-19 pandemic. Although the lack of a prospective study may decrease the power of the study, we still believe that the WhatsApp app is of particular importance in terms of highlighting this specific use.

Consultation via WhatsApp reduces both contact time with patients with COVID-19 (due to the decreased length of stay of patients in the ED) and the number of medical staff who contact the patient. In addition, the shortened discharge and hospitalization times also trimmed the patients' lengths of stay in the ED, with an impact on reducing the congestion in the ED. It can be concluded that WhatsApp consultation may prove useful in clinical decision making and increase the speed of the process, as the accuracy rates of the clinical decisions made as a result of the WhatsApp and bedside consultations were similar.

Conflicts of Interest

None declared.

References

1. Telemedicine Practice Guidelines. Board of Governors in Supersession of the Medical Council of India. 2020 Mar 25. URL: https://www.mohfw.gov.in/pdf/Telemedicine.pdf [accessed 2020-06-16]
2. Teleconsultations during a pandemic. Pan American Health Organization. URL: https://www.paho.org/ish/images/docs/covid-19-teleconsultations-en.pdf?ua= [accessed 2020-06-16]
3. COVID-19: video consultations and homeworking. British Medical Association. URL: https://www.bma.org.uk/advice-and-support/covid-19/adapting-to-covid/covid-19-video-consultations-and-homeworking [accessed 2020-06-16]
4. WhatsApp. Webpage in Turkish. Wikipedia. URL: https://tr.wikipedia.org/wiki/WhatsApp [accessed 2020-06-16]
5. Number of monthly active WhatsApp users worldwide from April 2013 to March 2020. Statista. URL: https://www.statista.com/statistics/260819/number-of-monthly-active-whatsapp-users/ [accessed 2020-06-16]
6. Giordano V, Koch HA, Mendes CH, Bergamin A, de Souza FS, do Amaral NP. WhatsApp Messenger is useful and reproducible in the assessment of tibial plateau fractures: inter- and intra-observer agreement study. Int J Med Inform 2015 Feb;84(2):141-148. [doi: 10.1016/j.ijmedinf.2014.11.002] [Medline: 25468642]
7. Astarcioglu MA, Sen T, Kilit C, Durmus HI, Gozubuyuk G, Kalcik M, et al. Time-to-reperfusion in STEMI undergoing interhospital transfer using smartphone and WhatsApp messenger. Am J Emerg Med 2015 Oct;33(10):1382-1384. [doi: 10.1016/j.ajem.2015.07.029] [Medline: 26299691]
8. Wani SA, Rabah SM, Alfadil S, Dewanjee N, Najmi Y. Efficacy of communication amongst staff members at plastic and reconstructive surgery section using smartphone and mobile WhatsApp. Indian J Plast Surg 2013 Sep;46(3):502-505 [FREE Full text] [doi: 10.4103/0970-0358.121990] [Medline: 24459338]
9. Graziano F, Maugeri R, Iacopino DG. Telemedicine versus WhatsApp: from tradition to evolution. Neuroreport 2015 Jul 08;26(10):602-603. [doi: 10.1097/WNR.0000000000000393] [Medline: 26053704]
10. Gulacti U, Lok U, Hatipoglu S, Polat H. An Analysis of WhatsApp Usage for Communication Between Consulting and Emergency Physicians. J Med Syst 2016 Jun;40(6):130. [doi: 10.1007/s10916-016-0483-8] [Medline: 27083574]
11. Vural S, Ramadan H. A short review on the comparison of consultation systems and tools in the emergency department practice: Tele-consultation. Cumhur Medical J 2019;42:239-243. [doi: 10.7197/223.vi.568582]

12. COVID-19 (New Coronavirus Disease) Algorithms. Republic of Turkey General Directorate of Public Health. URL: https://hsgm.saglik.gov.tr/depo/covid19/Ingilizce/Algoritmalar/COVID19-PLKACILHASTAYONETIMI_ENG.pdf [accessed 2020-12-09]
13. Emergency Service in Bed Healthcare Facilities: Notice About Application Procedures and Principles [In Turkish]. Republic of Turkey Official Newspaper. URL: https://www.resmigazete.gov.tr/eskiler/2018/02/20180220-4.htm [accessed 2020-12-09]
14. Institute of Medicine. The Role of Telehealth in an Evolving Health Care Environment: Workshop Summary. Washington, DC: National Academies Press; 2012.
15. Cho SJ, Jeong J, Han S, Yeom S, Park SW, Kim HH, et al. Decreased emergency department length of stay by application of a computerized consultation management system. Acad Emerg Med 2011 Apr;18(4):398-402 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01039.x] [Medline: 21496143]
16. Agarwal A, Gaieski D, Perman S, Abella B, Carr B. Emergency Department Based Telemedicine: Testing the Feasibility of Using Telemedicine as a Communication and Consultation Tool for Critically Ill Patients. Ann Emerg Med 2013 Oct;62(4):S142-S143. [doi: 10.1016/j.annemergmed.2013.07.222]
17. Gulacti U, Lok U, Çelik M. Use of WhatsApp application for orthopedic consultations in the ED. Am J Emerg Med 2016 Jul;34(7):1305-1307. [doi: 10.1016/j.ajem.2016.04.004] [Medline: 27117464]
18. Kelahmetoglu O, Firinciogullari R, Yagmur C. Efficient Utility of WhatsApp: From Computer Screen to the Surgeon's Hand to Determine Maxillofacial Traumas. J Craniofac Surg 2015 Jun;26(4):1437. [doi: 10.1097/SCS.0000000000001627] [Medline: 26080228]
19. Erenler AK, Akbulut S, Guzel M, Cetinkaya H, Karaca A, Turkoz B, et al. Reasons for Overcrowding in the Emergency Department: Experiences and Suggestions of an Education and Research Hospital. Turk J Emerg Med 2014 Jun;14(2):59-63 [FREE Full text] [doi: 10.5505/1304.7361.2014.48802] [Medline: 27331171]
20. Gulacti U, Lok U. Comparison of secure messaging application (WhatsApp) and standard telephone usage for consultations on Length of Stay in the ED. A prospective randomized controlled study. Appl Clin Inform 2017 Jul 19;8(3):742-753 [FREE Full text] [doi: 10.4338/ACI-2017-04-RA-0064] [Medline: 28880047]

Abbreviations

CT: computed tomography
ED: emergency department
ID: infectious disease
PCR: polymerase chain reaction

Edited by G Eysenbach; submitted 25.07.20; peer-reviewed by L Seuren, A Barker; comments to author 29.08.20; revised version received 13.10.20; accepted 03.12.20; published 11.12.20.

Please cite as:
Sabırlı R, Karsli E, Canacik O, Ercin D, Çiftçi H, Sahin L, Dolanbay T, Tutuncu EE
Use of WhatsApp for Polyclinic Consultation of Suspected Patients With COVID-19: Retrospective Case Control Study
JMIR Mhealth Uhealth 2020;8(12):e22874
URL: http://mhealth.jmir.org/2020/12/e22874/
doi:10.2196/22874
PMID:33276318

Original Paper

Attitudes Toward Using COVID-19 mHealth Tools Among Adults With Chronic Health Conditions: Secondary Data Analysis of the COVID-19 Impact Survey

Marlene Camacho-Rivera[1], BS, MS, MPH, SCD; Jessica Yasmine Islam[2], BS, MPH, PhD; Argelis Rivera[3], BSc, MPH, MD; Denise Christina Vidot[4], BSc, PhD

[1]Department of Community Health Sciences, State University of New York Downstate Health Sciences University, Brooklyn, NY, United States

[2]University of North Carolina, Lineberger Comprehensive Cancer Center, Chapel Hill, NC, United States

[3]Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[4]University of Miami School of Nursing and Health Studies, Coral Gables, FL, United States

Corresponding Author:
Marlene Camacho-Rivera, BS, MS, MPH, SCD
Department of Community Health Sciences
State University of New York Downstate Health Sciences University
450 Clarkson Avenue, MSC 43
Brooklyn, NY, 11203
United States
Phone: 1 7182704386
Email: marlene.camacho-rivera@downstate.edu

Abstract

Background: Adults with chronic conditions are disproportionately burdened by COVID-19 morbidity and mortality. Although COVID-19 mobile health (mHealth) apps have emerged, research on attitudes toward using COVID-19 mHealth tools among those with chronic conditions is scarce.

Objective: This study aimed to examine attitudes toward COVID-19, identify determinants of COVID-19 mHealth tool use across demographic and health-related characteristics, and evaluate associations between chronic health conditions and attitudes toward using COVID-19 mHealth tools (eg, mHealth or web-based methods for tracking COVID-19 exposures, symptoms, and recommendations).

Methods: We used nationally representative data from the COVID-19 Impact Survey collected from April to June 2020 (n=10,760). Primary exposure was a history of chronic conditions, which were defined as self-reported diagnoses of cardiometabolic, respiratory, immune-related, and mental health conditions and overweight/obesity. Primary outcomes were attitudes toward COVID-19 mHealth tools, including the likelihood of using (1) a mobile phone app to track COVID-19 symptoms and receive recommendations; (2) a website to track COVID-19 symptoms, track location, and receive recommendations; and (3) an app using location data to track potential COVID-19 exposure. Outcome response options for COVID-19 mHealth tool use were extremely/very likely, moderately likely, or not too likely/not likely at all. Multinomial logistic regression was used to compare the likelihood of COVID-19 mHealth tool use between people with different chronic health conditions, with not too likely/not likely at all responses used as the reference category for each outcome. We evaluated the determinants of each COVID-19 mHealth intervention using Poisson regression.

Results: Of the 10,760 respondents, 21.8% of respondents were extremely/very likely to use a mobile phone app or a website to track their COVID-19 symptoms and receive recommendations. Additionally, 24.1% of respondents were extremely/very likely to use a mobile phone app to track their location and receive push notifications about whether they have been exposed to COVID-19. After adjusting for age, race/ethnicity, sex, socioeconomic status, and residence, adults with mental health conditions were the most likely to report being extremely/very or moderately likely to use each mHealth intervention compared to those without such conditions. Adults with respiratory-related chronic diseases were extremely/very (conditional odds ratio 1.16, 95% CI 1.00-1.35) and moderately likely (conditional odds ratio 1.23, 95% CI 1.04-1.45) to use a mobile phone app to track their location and receive push notifications about whether they have been exposed to COVID-19.

Conclusions: Our study demonstrates that attitudes toward using COVID-19 mHealth tools vary widely across modalities (eg, web-based method vs app) and chronic health conditions. These findings may inform the adoption of long-term engagement with

COVID-19 apps, which is crucial for determining their potential in reducing disparities in COVID-19 morbidity and mortality among individuals with chronic health conditions.

(JMIR Mhealth Uhealth 2020;8(12):e24693) doi:10.2196/24693

KEYWORDS

smartphone; mHealth; COVID-19; chronic health conditions; health disparities; chronic disease; attitude; perception; data analysis; contact tracing; mobile app; disparity

Introduction

Since the declaration of the COVID-19 pandemic in the United States, over 7 million COVID-19 cases have been identified, and more than 200,000 lives have been lost [1]. Public health strategies for reducing the transmission of COVID-19 have included the enactment of policies, such as quarantine and social distancing, closures of nonessential businesses, and recommendations for preventive behaviors, such as hand washing and wearing face masks [2,3]. Emerging studies have reported that the implementation of policies and adherence to preventive best practices have varied widely in different geographic locations and demographic subgroups [4,5]. Public health and clinical researchers have reported on the disproportionate burden of COVID-19 morbidity and mortality among racial and ethnic minorities, older adults, and individuals with preexisting chronic health conditions [6-21]. Based on clinical and population-based studies, non-Hispanic Black and Latino individuals and communities are at increased risk for COVID-19 exposure [17,19,20,22], morbidity [8,9,23,24], and mortality [7,25,26]. Numerous chronic health conditions, including hypertension [8,23,26-29], diabetes [11,23,27,28,30], cancer [15,31-33], chronic obstructive pulmonary disease (COPD) [34-36] and asthma [28,37-39], and obesity [40-43] have been associated with an increased risk for poor COVID-19 outcomes.

The proliferation of COVID-19 mobile health (mHealth) tools for tracking COVID-19 statistics, monitoring potential COVID-19 symptoms, and reducing the social and mental health impacts of the COVID-19 pandemic has emerged simultaneously with the emergence of the pandemic [44-64]. Several COVID-19 apps that have been developed for public health surveillance provide up-to-date statistics, including the number of new cases, hospitalizations, and confirmed deaths [46,62,65]. Among the apps that allow for real-time symptom monitoring, trackers and telemedicine systems have helped elucidate the frequency of symptoms associated with COVID-19, thereby allowing patients and health care providers the opportunity to respond early to symptom progression [44,47,48,66]. Reviews of contact tracing apps have highlighted their effectiveness in improving the spatiotemporal reporting of new cases, management and follow-up of COVID-19 cases, and education on preventive behaviors [52,53,57,58,67-70]. To address the mental and social impacts of the COVID-19 pandemic, several apps have focused on reducing social isolation, providing positive coping strategies, and monitoring mental health symptoms [55,56,63,71-74]. Although researchers have noted the utility of COVID-19 apps in providing education and public health surveillance amid a rapidly evolving landscape, reviews of COVID-19 apps have highlighted several barriers to the availability, safety, and long-term sustainability of these technologies, including cost, the use of evidence-based guidelines, and user-centered design considerations for functionality and content [45,61,75-77]. Although differences in user preferences and perceptions of COVID-19 susceptibility based on individual risk factors exist across various social and demographic groups, few studies have examined attitudes toward COVID-19 mHealth tools among adults with and without preexisting health conditions.

The aims of this study were (1) to identify differences in attitudes toward COVID-19 mHealth tools across individuals with chronic health conditions; (2) to evaluate associations between having a preexisting chronic health condition and attitudes toward using COVID-19 mobile-based apps and websites to track potential COVID-19 exposure and symptoms; and (3) to identify determinants of mHealth intervention use. To accomplish these study objectives, we used publicly available data from the COVID-19 Impact Survey, which was designed to provide a nationally representative sample of the US adult population and offer national insights about the American population's experiences during the COVID-19 pandemic.

Methods

COVID-19 Impact Survey Dataset

Data for our analyses were obtained from the publicly available COVID-19 Household Impact Survey, which was conducted by the nonpartisan and objective research organization NORC (National Opinion Research Center) at the University of Chicago for the Data Foundation. The COVID-19 Household Impact Survey is a philanthropic effort to provide national and regional statistics about physical health, mental health, economic security, and social dynamics in the United States [78]. The survey is designed to provide weekly estimates of the US adult (ie, aged ≥18 years) household population nationwide. Currently, data from week 1 (ie, April 20-26, 2020), week 2 (ie, May 4-10, 2020), and week 3 (ie, May 30 to June 8, 2020) are available. These data were merged for this analysis. As the AmeriSpeak analytic sample of the COVID-19 Impact Survey was derived from deidentified publicly available data, institutional review board approval was not required for this study.

AmeriSpeak is a probability-based panel that is funded and operated by the NORC at the University of Chicago. It is designed to be representative of the US household population. During the initial recruitment phase of the AmeriSpeak panel for the COVID-19 Impact Survey, randomly selected US households were sampled using area probability and address-based sampling methods. These sampled households were then contacted by US mail, telephone, and field

interviewers (ie, face-to-face interview). The panel provides sample coverage for approximately 97% of the US household population. Those excluded from the sample include people with post office box–only addresses, addresses not listed in the US Postal Service Delivery Sequence File, and newly constructed dwellings. Although most AmeriSpeak households can participate in surveys via the internet, households without web access are able to participate in AmeriSpeak surveys by telephone. Interviews were conducted in English and Spanish. Interviews were conducted with adults who were aged ≥18 years and represented the 50 states and the District of Columbia. Panelists were offered a US $5 monetary incentive for completing the survey. With regard to the number of participants invited and percentage of interviews completed by week, 11,133 panelists were invited and 19.7% of interviews were completed in week 1, 8570 panelists were invited and 26.1% of interviews were completed in week 2; and 10,373 panelists were invited and 19.7% of interviews were completed in week 3 [78]. The analytic sample includes 10,760 nationwide adults and is weighted to reflect the US population of adults aged ≥18 years. The demographic weighting variables were obtained from the 2020 Current Population Survey [79].

Attitudes Toward Using COVID-19 mHealth Tools

Our primary outcomes for this analysis were participants' attitudes and willingness toward using COVID-19 mHealth tools to track potential COVID-19 exposure and symptoms. To characterize attitudes toward the use of COVID-19 mHealth tools for tracking potential COVID-19 exposure, we used participant's responses to the following question: "If this option was available to you, how likely would you install an app on your phone that tracks your location and sends push notifications if you might have been exposed to COVID-19"?

To characterize attitudes toward the use of COVID-19 mHealth tools for tracking potential COVID-19 symptoms and recommendations, we used participant's responses to the following 2 questions: (1) "If this option was available to you, how likely would you use a website to log your symptoms and location and get recommendations about COVID-19"; and (2) "If this option was available to you, how likely would you install an app on your phone that asks you questions about your own symptoms and provides recommendations about COVID-19"?

For each of the 3 outcomes, the provided response options were extremely likely, very likely, moderately likely, not too likely, and not likely at all. Due to the small sample sizes, the extremely likely and very likely response options were combined into 1 category, and the not too likely and not likely at all response options were also combined into 1 category. This was done for each outcome.

History of Chronic Health Conditions

The primary predictor for this analysis was participants' self-reports on a chronic health condition. Within the survey, participants were asked to reply "yes, no, or not sure" to the following question: "Has a doctor or other health care provider ever told you that you have any of the following: Diabetes; High blood pressure or hypertension; Heart disease, heart attack or stroke; Asthma; Chronic lung disease or COPD; Bronchitis or emphysema; Allergies; a Mental health condition; Cystic fibrosis; Liver disease or end-stage liver disease; Cancer; a Compromised immune system; or Overweight or obesity." Given the small sample sizes reported among certain categories of health conditions, we further aggregated health conditions into the following 5 categories: cardiometabolic (ie, diabetes, high blood pressure, heart disease, heart attack or stroke, and liver disease or end-stage liver disease), overweight/obesity, respiratory (ie, allergies, asthma, chronic lung disease or COPD, and bronchitis or emphysema), immune-related (ie, cystic fibrosis, cancer, and a compromised immune system), and mental health conditions.

Covariates

The following covariates were included in the multivariable analyses: age categories (ie, 18-29, 30-44, 45-59, ≥60 years), sex (ie, male or female), race/ethnicity categories (ie, non-Hispanic White, non-Hispanic Black, Hispanic or Latino, Asian, and non-Hispanic other), education categories (ie, no high school diploma, high school graduate, some college, and baccalaureate or above), residence (ie, rural or urban), and household income (ie, <US $50,000, US $50,000-US $100,000, ≥US $100,000). These covariates were chosen based on the review of mHealth disparities literature and COVID-19 disparities literature [5,80-83].

Data Analysis

Descriptive statistics are presented as percentages for all respondents unless otherwise labeled, and include a margin of error of 3.0 percentage points for 95% confidence intervals among all adults. Chi-square tests were used for the univariate comparison of categorical variables, including age, sex, race/ethnicity, education, income, residence, and chronic health conditions. For COVID-19 mHealth outcomes, we used multinomial logistic regression to compare the likelihood of COVID-19 mHealth tool use across people with different chronic health conditions after adjusting for age, sex, race/ethnicity, education, income, and residence. For each COVID-19 mHealth outcome, we compared participants' likelihood of using COVID-19 mHealth tools to those in the not likely to use group (ie, the reference category). To estimate determinants of being extremely/very likely to use each COVID-19 mHealth intervention, we computed prevalence ratios via Poisson regression using the robust estimation of standard errors [84-86]. Potential variables for inclusion in the model were assessed using available sociodemographic variables and bivariate Poisson regression analysis. Due to the exploratory nature of this analysis, which used a predictive framework, an arbitrary P value of <.10 was used as criteria to include variables in the multivariable Poisson regression model. For multivariable Poisson regression models, adjusted prevalence ratios and 95% confidence intervals for each independent variable were calculated.

The Type I error was maintained at 5%. Due to the exploratory nature of this analysis, we did not include an adjustment for multiple comparisons [87,88]. All statistical analyses were conducted using Stata IC 15 (StataCorp LLC). Sampling weights were applied to provide results that were nationally

representative of the US adult population. As such, absolute n values could not be reported.

Results

Descriptive COVID-19 Impact Survey Results

Tables 1 and 2 display the descriptive characteristics of the analytic sample. Of the 10,760 respondents, 61.6% of respondents were non-Hispanic White, 11.9% were non-Hispanic Black, 16.5% were Hispanic, and 8.6% were of another non-Hispanic race or ethnicity. With regard to education, 9.8% of respondents had less than a high school diploma, 28.2% had a high school diploma, 27.7% had some college education, and 34.3% had a baccalaureate degree or higher. With regard to age, 20.5% of participants were aged 18-29 years, 25.4% were aged 30-44 years, 24.3% were aged 45-59 years, and 29.8% were aged ≥60 years. With regard to sex, 48.3% of respondents were male and 51.7% were female. With regard to residence, 72.2% of participants lived in urban areas and 27.9% lived in rural or suburban areas.

Histories of chronic health conditions ranged from 15.5% of respondents reporting a history of mental health conditions to 37.8% reporting a history of cardiometabolic diseases. The most common chronic conditions reported were cardiometabolic diseases (37.8%) and overweight/obesity (33.1%). Furthermore, 21.8% of adults reported that they were extremely or very likely to install a mobile phone app to record symptoms and obtain recommendations about COVID-19, and 21.1% reported that they were extremely or very likely to use a website to log their symptoms and location and receive recommendations about COVID-19. Additionally, 24.1% of adults reported that they were extremely or very likely to install an app on their phone to track their location and receive push notifications about whether they may have been exposed to COVID-19.

Table 1. Demographic characteristics of the COVID-19 Impact Survey respondents (N=10,760) from April to June 2020. This survey is a nationally representative survey of the United States.

Characteristic	Total, %[a] (95% CI)
Age (years)	
18-29	20.5 (19.3-21.8)
30-44	25.4 (24.4-26.5)
45-59	24.3 (23.2-25.4)
≥60	29.8 (28.6-30.9)
Race/ethnicity	
Non-Hispanic White	61.6 (60.3-62.9)
Non-Hispanic Black	11.9 (11.0-12.7)
Hispanic	16.5 (15.5-17.7)
Non-Hispanic Asian	5.1 (4.4-5.8)
Other non-Hispanic race/ethnicity	3.5 (3.1-3.9)
Sex	
Male	48.3 (47.0-49.6)
Female	51.7 (50.4-53.0)
Employed in the past 7 days	49.7 (48.4-51.1)
Education	
No high school diploma	9.8 (8.8-10.8)
High school graduate	28.2 (27.0-29.6)
Some college	27.7 (26.7-28.7)
Baccalaureate or above	34.3 (33.1-35.5)
Household income (US $)	
<50,000	45.8 (44.5-47.1)
50,000-100,000	32.1 (30.9-33.3)
≥100,000	22.1 (21.1-23.2)
Residence	
Rural	9.1 (8.4-9.8)
Suburban	18.8 (17.8-19.7)
Urban	72.2 (71.0-73.3)
Insurance type or health coverage plans	
Purchased plan	17.4 (16.4-18.5)
Employer sponsored	51.7 (50.3-53.0)
TRICARE	4.9 (4.4-5.4)
Medicaid	23.5 (22.4-24.7)
Medicare	25.3 (24.2-26.4)
Dually eligible (ie, Medicare and Medicaid)	9.7 (9.0-10.4)
VA	4.5 (4.0-5.0)
Indian Health Service	1.2 (0.9-1.6)
No insurance	8.8 (8.1-9.6)
Cardiometabolic-related chronic diseases[b]	37.8 (36.5-39.0)
Respiratory-related chronic diseases[c]	23.6 (22.5-24.7)

Characteristic	Total, %[a] (95% CI)
Immune-related chronic diseases[d]	13.1 (12.3-13.9)
Overweight/obesity	33.1 (31.9-34.3)
Mental health–related conditions[e]	15.5 (14.6-16.4)

[a]Absolute n values are not displayed due to the inclusion of survey weights to provide nationally representative estimates.

[b]Cardiometabolic-related chronic diseases include diabetes, high blood pressure, heart disease, and liver disease/end-stage liver disease.

[c]Respiratory-related diseases include asthma, chronic lung disease/chronic obstructive pulmonary disease, and bronchitis/emphysema.

[d]Immune-related chronic diseases include cystic fibrosis, cancer, and a compromised immune system.

[e]Mental health–related conditions include at least 1 mental health condition.

Table 2. Survey characteristics of the COVID-19 Impact Survey respondents (N=10,760) from April to June 2020. This survey is a nationally representative survey of the United States.

Questions and responses	Total, %[a] (95% CI)
How likely would be to participate in installing an app on your phone that asks you questions about your own symptoms and provides recommendations about COVID-19?	
Extremely/very likely	21.8 (20.7-22.9)
Moderately likely	20.9 (19.8-22.1)
Not too likely/not likely at all	57.3 (55.9-58.6)
How likely would you be to participate in using a website to log your symptoms and location and get recommendations about COVID-19?	
Extremely/very likely	21.1 (20.1-22.2)
Moderately likely	23.5 (22.3-24.7)
Not likely	55.4 (54.1-56.7)
How likely would you be to participate in installing an app on your phone that tracks your location and sends push notifications if you might have been exposed to COVID-19?	
Extremely/very likely	24.1 (22.9-25.2)
Moderately likely	19.9 (18.8, 20.9)
Not likely	56.0 (54.7-57.4)

[a]Absolute n values are not displayed due to the inclusion of survey weights to provide nationally representative estimates.

Attitudes Toward Using COVID-19 mHealth Tools Across People With Chronic Health Conditions

As shown in Tables 3 and 4, differences in attitudes toward the use of COVID-19 mHealth tools emerged across various chronic health conditions. Compared to adults without mental health conditions, adults with a history of mental health conditions were significantly more likely to potentially download and use an app to track COVID-19 symptoms and recommendations (P=.004), potentially download and use an app to track their location and potential COVID-19 exposure (P<.001), and use a website to log COVID-19 symptoms and receive recommendations (P=.005). Compared to nonobese participants, respondents who reported being obese were significantly more likely to potentially download and use an app to track their location and potential COVID-19 exposure (P=.039). In the univariate analysis, no other significant differences in attitudes toward using COVID-19 mHealth tools were observed in respondents with a history of respiratory conditions or immune-related conditions.

Table 3. Attitudes toward mHealth interventions for COVID-19 testing and tracking to help slow the spread of the virus, stratified by cardiometabolic-related[a], respiratory-related[b], and immune-related[c] chronic diseases.

Questions and responses	Cardiometabolic-related chronic disease		*P* value[d]	Respiratory-related chronic disease		*P* value[d]	Immune-related chronic disease		*P* value[d]
	No	Yes		No	Yes		No	Yes	
	Total, %[e] (95% CI)	Total, % (95% CI)		Total, % (95% CI)	Total, % (95% CI)		Total, % (95% CI)	Total, % (95% CI)	
How likely would be to participate in installing an app on your phone that asks you questions about your own symptoms and provides recommendations about COVID-19?			.02			>.99			.74
Extremely/very	20.7 (19.40-22.2)	23.5 (21.7-25.3)		21.8 (20.6-23.1)	21.7 (19.5-24.1)		21.7 (20.5-22.9)	22.7 (19.8-25.8)	
Moderately	21.8 (20.3-23.3)	19.5 (17.9-21.2)		20.9 (19.6-22.2)	21.0 (18.7-23.4)		20.9 (19.7-22.1)	21.3 (18.7-24.2)	
Not likely	57.5 (55.7-59.2)	57.0 (54.9-59.1)		57.3 (55.8-58.8)	57.3 (54.5-60.0)		57.5 (56.0-58.9)	56.0 (52.6-59.4)	
How likely would you be to participate in installing an app on your phone that tracks your location and sends push notifications if you might have been exposed to COVID-19?			.65			.08			.55
Extremely/very	23.9 (22.5-25.4)	24.4 (22.7-26.3)		23.8 (22.5-25.1)	25.1 (22.8-27.5)		23.9 (22.7-25.2)	25.2 (22.3-28.3)	
Moderately	20.2 (18.9-21.7)	19.3 (17.7-20.9)		19.4 (18.2-20.6)	21.5 (19.3-24.0)		20.1 (18.9-21.2)	18.6 (16.1-21.5)	
Not likely	55.9 (54.2-57.6)	56.3 (54.2-58.3)		56.9 (55.4-58.3)	53.4 (50.6-56.1)		56.0 (54.6-57.4)	56.1 (52.7-59.5)	
How likely would you be to participate in using a website to log your symptoms and location and get recommendations about COVID-19?			.45			.32			.91
Extremely/very	20.6 (19.3-22.0)	22.0 (20.3-23.80		21.6 (20.4-22.9)	19.6 (17.6-21.8)		21.2 (20.0-22.4)	20.9 (18.2-23.80	
Moderately	23.5 (22.0-25.0)	23.5 (21.8-25.3)		23.3 (22.0-24.7)	24.1 (21.8-26.6)		23.4 (22.2-24.7)	24.1 (21.2-27.2)	
Not likely	55.9 (54.2-57.6)	54.5 (52.4-56.6)		55.1 (53.6-56.6)	56.2 (53.5-58.9)		55.4 (54.0-56.9)	55.0 (51.6-58.4)	

[a]Cardiometabolic-related chronic diseases include diabetes, high blood pressure, heart disease, and liver disease/end-stage liver disease.

[b]Respiratory-related chronic diseases include asthma, chronic lung disease/chronic obstructive pulmonary disease, and bronchitis/emphysema.

[c]Immune-related chronic diseases include cystic fibrosis, cancer, and a compromised immune system.

[d]Refers to Chi-square *P* values.

[e]Absolute n values are not displayed due to the inclusion of survey weights to provide nationally representative estimates.

Table 4. Attitudes toward mHealth interventions for COVID-19 testing and tracking to help slow the spread of the virus, stratified by mental health–related[a] chronic diseases and overweight/obese.

Questions and responses	Mental health-related chronic disease		*P* value[b]	Overweight/obese		*P* value[b]
	No	Yes		No	Yes	
	Total, %[c] (95% CI)	Total, % (95% CI)		Total, % (95% CI)	Total, % (95% CI)	
How likely would be to participate in installing an app on your phone that asks you questions about your own symptoms and provides recommendations about COVID-19?			<.001			.54
Extremely/very	21.6 (20.4-22.8)	22.8 (20.2-25.7)		21.6 (20.2-23.0)	22.2 (20.4-24.1)	
Moderately	20.2 (19.0-21.5)	24.7 (21.8-27.7)		21.3 (19.9-22.8)	20.1 (18.3-21.9)	
Not likely	58.6 (56.7-59.6)	52.5 (49.2-55.8)		57.1 (55.4-58.7)	57.7 (55.5-59.9)	
How likely would you be to participate in installing an app on your phone that tracks your location and sends push notifications if you might have been exposed to COVID-19?			<.001			.04
Extremely/very	23.6 (22.4-24.9)	26.7 (23.9-29.7)		23.1 (21.7-24.5)	26.1 (24.2-28.1)	
Moderately	19.2 (18.1-20.4)	23.4 (20.6-26.5)		19.9 (18.7-21.3)	19.7 (18.0-21.6)	
Not likely	57.2 (55.7-58.6)	49.9 (46.6-53.2)		57.0 (55.3-58.6)	54.2 (52.0-56.3)	
How likely would you be to participate in using a website to log your symptoms and location and get recommendations about COVID-19?			.01			.26
Extremely/very	20.9 (19.7-22.1)	22.5 (19.9-25.4)		21.2 (19.9-22.6)	21.0 (19.3-22.8)	
Moderately	22.9 (21.6-24.1)	26.9 (24.1-30.0)		22.8 (21.4-24.3)	24.8 (22.9-26.8)	
Not likely	56.3 (54.8-57.7)	50.5 (47.2-53.8)		55.9 (54.3-57.6)	54.2 (52.0-56.4)	

[a]Mental health–related chronic disease include at least 1 mental health condition.

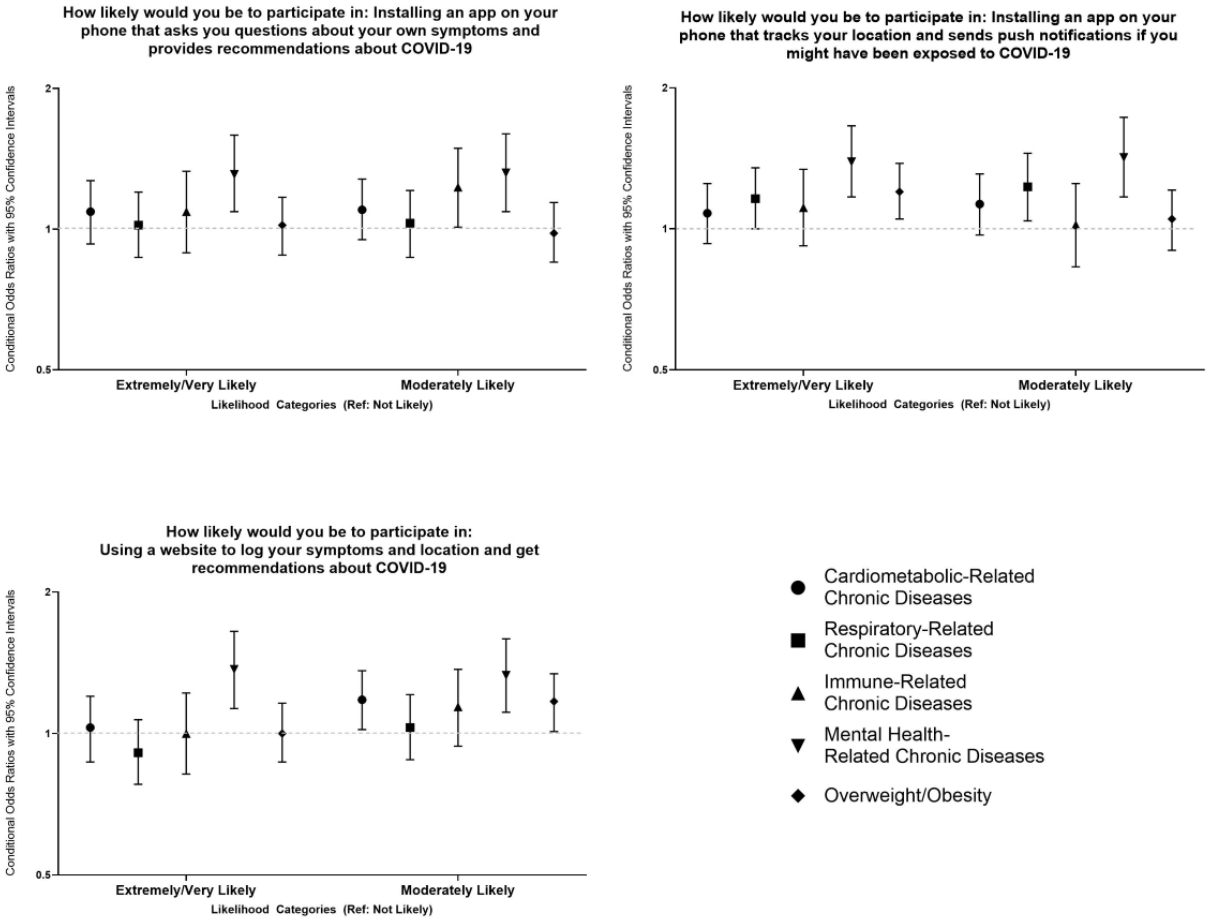
[b]Refers to Chi-square *P* values.

[c]Absolute n values are not displayed due to the inclusion of survey weights to provide nationally representative estimates.

Multinomial Logistic Regression Results for Attitudes Toward Using COVID-19 mHealth Tools Across People With Chronic Health Conditions

Figure 1 displays the multivariable results for the multinomial logistic regression models for each COVID-19 mHealth outcome and chronic health condition category. These models were adjusted for age, sex, race/ethnicity, education, income, and residence. The point estimates are available in Table S1 in Multimedia Appendix 1. Compared to adults without a history of cardiometabolic diseases, adults with cardiometabolic conditions were moderately likely to use a website to log their COVID-19 symptoms (conditional odds ratio [cOR] 1.18, 95% CI 1.02-1.36). Compared to adults without a history of respiratory conditions, adults with a history of respiratory diseases were moderately likely to download and use an app to track their location and potential COVID-19 exposure (cOR 1.23, 95% CI 1.04-1.45). Compared to adults without immune-related conditions, adults with a history of immune-related diseases were moderately likely to potentially download and use an app to track COVID-19 symptoms and receive health recommendations about COVID-19 (cOR 1.23, 95% CI 1.01-1.49).

Figure 1. Associations between COVID-19 mHealth intervention acceptability and different chronic disease groups.



Compared to adults without a history of mental health conditions, adults a history of mental health conditions were extremely likely and moderately likely to potentially download and use an app to track COVID-19 symptoms and receive health recommendations about COVID-19 (cOR 1.31, 95% CI 1.09-1.59; cOR 1.32, 95% CI 1.09-1.60, respectively). Adults with a history of mental health conditions were extremely likely and moderately likely to download and use an app to track location and potential COVID-19 exposure compared to adults without a history of mental health conditions (cOR 1.39, 95% CI 1.17-1.66; cOR 1.42, 95% CI 1.17-1.73, respectively). Compared to adults without a history of mental health conditions, adults with a history of mental health conditions were extremely likely and moderately likely to use of a website to log their COVID-19 symptoms (cOR 1.37, 95% CI 1.13-1.65; cOR 1.33, 95% CI 1.11-1.59, respectively).

Compared to nonobese participants, adults who reported being obese were extremely likely to download and use an app to track location and potential COVID-19 exposure (cOR 1.20, 95% CI 1.05-1.38). Participants who were obese were also moderately likely to use a website to log their COVID-19 symptoms compared to nonobese participants (cOR 1.17, 95% CI 1.01-1.34).

Determinants of Being Extremely/Very Likely to Use COVID-19 mHealth Interventions Among US Adults

Table 5 summarizes the results of the analysis for identifying determinants of using each mHealth intervention among US adults. Across each mHealth intervention, there were several common determinants. Women had a higher prevalence of being extremely/very likely to use each mHealth intervention than men. Compared to those with a college degree, adults with a high school degree and some college education had a lower prevalence of being extremely/very likely to use each mHealth intervention. Racial/ethnic minorities, including non-Hispanic Black, Hispanic, and non-Hispanic Asian adults, had a higher prevalence of being extremely/very likely to use each mHealth intervention compared to non-Hispanic White adults. Additionally, adults with at least 1 COVID-19 related symptom had a higher prevalence of being extremely/very likely to install and use mHealth apps.

Table 5. Determinants of being extremely/very[a] likely to use COVID-19 mHealth interventions based on data from the COVID-19 Impact Survey collected from April to June 2020. The survey is a nationally representative survey of the United States.

Variable	Installing an app on your phone that asks you questions about your own symptoms and provides recommendations about COVID-19, adjusted PR[b] (95% CI)	Installing an app on your phone that tracks your location and sends push notifications if you might have been exposed to COVID-19, adjusted PR (95% CI)	Using a website to log your symptoms and location and get recommendations about COVID-19, adjusted PR (95% CI)
Age (years)			
18-29	0.71 (0.59-0.86)	—[c,d]	0.74 (0.61-0.91)
30-44	0.80 (0.70-0.92)	—	0.90 (0.78-1.03)
45-59	1.03 (0.90-1.17)	—	1.02 (0.89-1.16)
≥60	Ref[e]	—	Ref
Sex			
Male	Ref	Ref	Ref
Female	1.12 (1.01-1.24)	1.12 (1.02-1.23)	1.11 (1.00-1.23)
Education			
No high school diploma	1.07 (0.87-1.33)	0.99 (0.80-1.21)	0.92 (0.73-1.16)
High school graduate	0.75 (0.65-0.87)	0.74 (0.64-0.85)	0.68 (0.59-0.80)
Some college	0.80 (0.71-0.90)	0.79 (0.71-0.88)	0.78 (0.70-0.88)
Baccalaureate or above	Ref	Ref	Ref
Race/ethnicity			
Non-Hispanic White	Ref	Ref	Ref
Non-Hispanic Black	1.58 (1.36-1.83)	1.32 (1.14-1.53)	1.67 (1.44-1.93)
Hispanic	1.49 (1.29-1.73)	1.29 (1.13-1.47)	1.47 (1.27-1.70)
Non-Hispanic Asian	1.84 (1.47-2.30)	1.73 (1.43-2.10)	1.74 (1.39-2.17)
Other non-Hispanic race/ethnicity	1.18 (0.90-1.55)	1.07 (0.83-1.38)	1.23 (0.94-1.61)
At least 1 COVID-19–related symptom[f]	1.13 (1.02-1.25)	1.19 (1.08-1.31)	—[d]
At least 1 chronic disease[g]	—[d]	—[d]	—[d]
Region			
Northeast	Ref	—[d]	Ref
Midwest	0.84 (0.71-0.99)	—	0.76 (0.64-0.90)
South	0.97 (0.83-1.13)	—	0.85 (0.73-0.99)
West	0.83 (0.71-0.98)	—	0.79 (0.67-0.92)
Employed	0.85 (0.76-0.95)	0.85 (0.77-0.94)	0.78 (0.70-0.88)
Uninsured	—[d]	—[d]	0.90 (0.74-1.10)
Household income (US $)			
<50,000	0.96 (0.84-1.10)	0.82 (0.73-0.93)	0.86 (0.74-1.00)
50,000-100,000	0.91 (0.80-1.04)	0.86 (0.76-0.97)	0.85 (0.74-0.98)
≥100,000	Ref	Ref	Ref
Residence			
Rural	0.95 (0.78-1.17)	1.04 (0.87-1.25)	0.98 (0.81-1.19)
Suburban	0.98 (0.85-1.12)	0.95 (0.83-1.08)	0.93 (0.80-1.07)
Urban	Ref	Ref	Ref

[a]Respondents were asked if they were extremely likely, very likely, moderately likely, not too likely, or not at all likely to participate in each mHealth

intervention. We categorized those who responded with extremely/very likely to participate as exposed (ie, exposed=1) and those who gave other responses as unexposed (ie, unexposed=0).

[b]PR: prevalence ratio.

[c]Not available.

[d]The corresponding *P* value was >.10 for all categories of the variable in unadjusted analyses

[e]Ref: the referent group.

[f]COVID-19–related symptoms include fever, chills, runny or stuffy nose, chest congestion, skin rash, cough, sore throat, sneezing, muscle or body aches, headaches, fatigue or tiredness, shortness of breath, abdominal discomfort, nausea or vomiting, diarrhea, changed or lost sense of taste or smell, and loss of appetite.

[g]Chronic disease include diabetes, high blood pressure, heart disease/heart attack/stroke, asthma, chronic obstructive pulmonary disease, bronchitis or emphysema, cystic fibrosis, liver disease, cancer, a compromised immune system, and overweight/obesity.

Discussion

Principal Results

This study analyzed differences and associations between attitudes toward using COVID-19 mHealth tools and various chronic health conditions. Compared to adults without a history of chronic health conditions, adults with chronic health conditions were more likely to potentially use COVID-19 apps or websites to monitor potential COVID-19 exposure and symptoms. Yet, attitudes toward COVID-19 mHealth tools varied significantly across people with different chronic disease conditions and modalities (ie, app vs website). Our study findings highlight the potential for mHealth tools to improve disease self-management and reduce health disparities among individuals with chronic health conditions.

Our study highlighted disparities in attitudes toward COVID-19 mHealth tools across age, sex, race/ethnicity, education, and region. We observed that women in our sample were more likely to report positive attitudes toward using an app or website to track potential COVID-19 exposures or symptoms compared to men. This is consistent with prior studies that focused on disparities in mHealth use among the general population [80,89]. Previous studies have also documented lower rates of using mHealth tools among individuals with lower socioeconomic backgrounds compared to individuals with higher socioeconomic backgrounds [81,90]. In our study, irrespective of self-reported chronic health conditions, participants with lower levels of education were less likely to use an app or website for tracking COVID-19 symptoms or exposure compared to respondents with a college degree or higher. Similar patterns were observed in the different income groups [91].

Our study also highlighted several novel findings. The first was that respondents with a racial/ethnic minority background had a greater likelihood of using COVID-19 mHealth tools than non-Hispanic White respondents in our entire sample. Hispanic (adjusted prevalence ratio [aPR] 1.49, 95% CI 1.29-1.73), Asian (aPR 1.84, 95% CI 1.47-2.30), and non-Hispanic Black (aPR 1.58, 95% CI 1.36-1.83) respondents were significantly more likely to potentially use mHealth tools for tracking COVID-19 exposure and symptoms. These attitudes may indicate greater awareness among racial and ethnic minority communities with regard to increased susceptibility to COVID-19 morbidity and mortality [22,25,92]. Additionally, the use of mHealth tools has been increasing among all racial/ethnic groups over the last decade. This is indicative of all racial/ethnic groups having greater access to mHealth tools. This is also indicative of the potential that mHealth tools have for reducing racial/ethnic disparities in chronic disease management, morbidity, and mortality.

The second novel finding was observed when examining age-related disparities in COVID-19 mHealth attitudes. Contrary to many prior studies that have documented increased mHealth tool use among younger adults, respondents aged 18-29 years reported being less likely to use apps or websites for tracking COVID-19 symptoms and exposure compared to participants aged ≥60 years. These differences may indicate that older adults have a greater awareness of the risks of COVID-19 than younger adults. These differences may also indicate potential misperceptions of COVID-19 risk among younger populations [93,94].

We also observed geographic differences in attitudes toward COVID-19 mHealth tools with participants from the Midwest, South, and West, who reported less interest in using apps or websites to track COVID-19 symptoms and exposure compared to respondents from the Northeast. These disparities are consistent with prior studies that focused on regional differences in the digital divide, as well as studies of COVID-19 preventive behaviors across the United States [95]. Our findings may also reflect geographic differences in perceived susceptibility to COVID-19 at the time the survey was administered, given the disproportionate burden of COVID-19 cases in the Northeast during the onset of the pandemic [96,97].

Our findings are important, given the accumulating evidence for cardiovascular diseases, obesity, and diabetes being risk factors for COVID-19 incidence and mortality. In our previous study, which analyzed COVID-19 preventive behaviors among US adults, we observed that adults with chronic health conditions were significantly more likely to engage in many COVID-19 preventive behaviors, such as hand washing, using a face mask, and maintaining appropriate social distancing in public [5,33]. The use of mHealth tools for COVID-19 education, risk reduction, and symptom monitoring may complement existing prevention strategies [98]. Our findings also highlight potential opportunities for mHealth tools to reduce racial/ethnic and age disparities in COVID-19 exposure and outcomes [99,100].

Lastly, our findings highlight the potential that mHealth tools have for reducing chronic disease disparities in disease management and outcomes beyond the COVID-19 pandemic. Although there has been a proliferation of mHealth tools for people with chronic health conditions, the acceptability and long-term use of mHealth tools have varied considerably across

people with different chronic conditions [101,102]. Variability in mHealth tool use across people with chronic health conditions may be related to the quality of both the content and functionality of mHealth tools, as well as user-related preferences [103-107].

Study Strengths and Limitations

Our results should be interpreted within the context of several limitations. First, a history of chronic health conditions was based on self-reported data rather than medical records. Therefore, there may be potential for recall or misclassification bias for the chronic health conditions reported in this analysis. Additionally, while we acknowledge the increasing importance of considering multiple chronic conditions, given the complexity of examining multiple chronic conditions by both number and the combination of different chronic condition types, the examination of attitudes toward using COVID-19 mHealth tools among individuals with multiple chronic conditions was beyond the scope of this analysis. Second, attitudes toward potential COVID-19 mHealth tool use were based on self-reports rather than downloads or monitoring patterns of actual mHealth tool usage. As such, there may be potential for social desirability bias in terms of the reported COVID-19 mHealth attitudes, which may not correlate with actual COVID-19 mHealth tool use.

Despite these limitations, our study has notable strengths. First, the sampling methods used to obtain a nationally representative sample of the US adult population provided the opportunity for increased generalizability. Our findings may be helpful to clinicians and public health officials when guiding mHealth messaging for COVID-19 prevention, symptom monitoring, and health recommendations. Additionally, our study results are drawn from a population that is diverse in terms of social and demographic characteristics and health statuses. Since the COVID-19 pandemic has exacerbated health inequities across different social groups and groups with chronic health conditions, public health strategies that consider the utility of mHealth approaches across diverse contexts will have a vital role in reducing COVID-19 disparities.

Conclusions

Our study extends previous research in this field by analyzing attitudes toward various COVID-19 mHealth tools across a nationally representative sample of US adults with and without chronic health conditions. Future directions for research may include the examination of COVID-19 mHealth tool use among individuals with multiple chronic conditions and associations between COVID-19 mHealth tool use and COVID-19 knowledge and preventive behaviors. Our study results have several implications concerning the development of COVID-19 mHealth tools. As the use of mHealth apps has been found to impact the improvement of health outcomes across a range of chronic health conditions, the development of COVID-19 apps should focus on user preferences and consider differences in COVID-19 susceptibility across people with different chronic disease conditions. Combining big data analytic approaches with qualitative data from individuals with chronic health conditions may yield additional insights in increasing the acceptability of and long-term engagement with COVID-19 apps. Lastly, incorporating opportunities for the real-time tracking of potential COVID-19 exposure and symptoms among existing health apps for people with chronic conditions (eg, mental health conditions, cardiometabolic conditions, and respiratory diseases) may provide additional opportunities to strengthen prevention and early detection efforts among populations that are vulnerable to COVID-19 morbidity and mortality.

Acknowledgments

MCR is supported by the Association of American Medical Colleges Herbert W. Nickens Faculty Fellowship and Translational Program of Health Disparities Research Training (5S21MD012474-02). JYI is supported by the University of North Carolina's Cancer Care Quality Training Program (2T32CA116339-11).

Authors' Contributions

MCR, JYI, and AR contributed to the drafting of the manuscript. JYI and DCV contributed to the analysis of the data. MCR, JYI, AR, DCV contributed to the interpretation and discussion of the data and review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1. Conditional odds ratios for evaluating associations between attitudes toward using COVID-19 mHealth tools and chronic disease status; estimates for Figure 1.
[DOCX File , 18 KB - mhealth_v8i12e24693_app1.docx]

References

1. CDC COVID Data Tracker. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html [accessed 2020-12-09]

2. How to Protect Yourself & Others. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2020-12-09]
3. Coronavirus disease (COVID-19) advice for the public. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public [accessed 2020-12-09]
4. Rothgerber H, Wilson T, Whaley D, Rosenfeld D, Humphrey M, Moore A, et al. Politicizing the COVID-19 Pandemic: Ideological Differences in Adherence to Social Distancing. PsyArXive Preprints Preprint posted online on April 22, 2020. [doi: 10.31234/osf.io/k23cv]
5. Camacho-Rivera M, Islam JY, Vidot DC. Associations Between Chronic Health Conditions and COVID-19 Preventive Behaviors Among a Nationally Representative Sample of U.S. Adults: An Analysis of the COVID Impact Survey. Health Equity 2020;4(1):336-344 [FREE Full text] [doi: 10.1089/heq.2020.0031] [Medline: 32783017]
6. Laurencin CT, McClinton A. The COVID-19 Pandemic: a Call to Action to Identify and Address Racial and Ethnic Disparities. J Racial Ethn Health Disparities 2020 Jun;7(3):398-402 [FREE Full text] [doi: 10.1007/s40615-020-00756-0] [Medline: 32306369]
7. CDC COVID-19 Response Team. Geographic Differences in COVID-19 Cases, Deaths, and Incidence - United States, February 12-April 7, 2020. MMWR Morb Mortal Wkly Rep 2020 Apr 17;69(15):465-471 [FREE Full text] [doi: 10.15585/mmwr.mm6915e4] [Medline: 32298250]
8. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, the Northwell COVID-19 Research Consortium, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. JAMA 2020 May 26;323(20):2052-2059 [FREE Full text] [doi: 10.1001/jama.2020.6775] [Medline: 32320003]
9. Shekerdemian LS, Mahmood NR, Wolfe KK, Riggs BJ, Ross CE, McKiernan CA, International COVID-19 PICU Collaborative. Characteristics and Outcomes of Children With Coronavirus Disease 2019 (COVID-19) Infection Admitted to US and Canadian Pediatric Intensive Care Units. JAMA Pediatr 2020 Sep 01;174(9):868-873. [doi: 10.1001/jamapediatrics.2020.1948] [Medline: 32392288]
10. Gosain R, Abdou Y, Singh A, Rana N, Puzanov I, Ernstoff MS. COVID-19 and Cancer: a Comprehensive Review. Curr Oncol Rep 2020 May 08;22(5):53 [FREE Full text] [doi: 10.1007/s11912-020-00934-7] [Medline: 32385672]
11. Hussain A, Bhowmik B, do Vale Moreira NC. COVID-19 and diabetes: Knowledge in progress. Diabetes Res Clin Pract 2020 Apr;162:108142 [FREE Full text] [doi: 10.1016/j.diabres.2020.108142] [Medline: 32278764]
12. McClure ES, Vasudevan P, Bailey Z, Patel S, Robinson WR. Racial Capitalism Within Public Health-How Occupational Settings Drive COVID-19 Disparities. Am J Epidemiol 2020 Nov 02;189(11):1244-1253 [FREE Full text] [doi: 10.1093/aje/kwaa126] [Medline: 32619007]
13. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and Racial/Ethnic Disparities. JAMA 2020 Jun 23;323(24):2466-2467. [doi: 10.1001/jama.2020.8598] [Medline: 32391864]
14. Shah M, Sachdeva M, Dodiuk-Gad RP. COVID-19 and racial disparities. J Am Acad Dermatol 2020 Jul;83(1):e35 [FREE Full text] [doi: 10.1016/j.jaad.2020.04.046] [Medline: 32305444]
15. Balogun OD, Bea VJ, Phillips E. Disparities in Cancer Outcomes Due to COVID-19-A Tale of 2 Cities. JAMA Oncol 2020 Oct 01;6(10):1531-1532. [doi: 10.1001/jamaoncol.2020.3327] [Medline: 32789508]
16. Núñez A, Madison M, Schiavo R, Elk R, Prigerson HG. Responding to Healthcare Disparities and Challenges With Access to Care During COVID-19. Health Equity 2020;4(1):117-128 [FREE Full text] [doi: 10.1089/heq.2020.29000.rtl] [Medline: 32368710]
17. Jones J, Sullivan PS, Sanchez TH, Guest JL, Hall EW, Luisi N, et al. Similarities and Differences in COVID-19 Awareness, Concern, and Symptoms by Race and Ethnicity in the United States: Cross-Sectional Survey. J Med Internet Res 2020 Jul 10;22(7):e20001 [FREE Full text] [doi: 10.2196/20001] [Medline: 32614778]
18. Waltenburg MA, Victoroff T, Rose CE, Butterfield M, Jervis RH, Fedak KM, COVID-19 Response Team. Update: COVID-19 Among Workers in Meat and Poultry Processing Facilities - United States, April-May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 10;69(27):887-892 [FREE Full text] [doi: 10.15585/mmwr.mm6927e2] [Medline: 32644986]
19. Poteat T, Millett GA, Nelson LE, Beyrer C. Understanding COVID-19 risks and vulnerabilities among black communities in America: the lethal force of syndemics. Ann Epidemiol 2020 Jul;47:1-3 [FREE Full text] [doi: 10.1016/j.annepidem.2020.05.004] [Medline: 32419765]
20. Hawkins D. Differential occupational risk for COVID-19 and other infection exposure according to race and ethnicity. Am J Ind Med 2020 Sep;63(9):817-820 [FREE Full text] [doi: 10.1002/ajim.23145] [Medline: 32539166]
21. Vidot DC, Islam JY, Camacho-Rivera M, Harrell MB, Rao DR, Chavez JV, et al. The COVID-19 cannabis health study: Results from an epidemiologic assessment of adults who use cannabis for medicinal reasons in the United States. J Addict Dis 2020 Sep 15:1-11. [doi: 10.1080/10550887.2020.1811455] [Medline: 32933383]
22. Holmes L, Enwere M, Williams J, Ogundele B, Chavan P, Piccoli T, et al. Black-White Risk Differentials in COVID-19 (SARS-COV2) Transmission, Mortality and Case Fatality in the United States: Translational Epidemiologic Perspective and Challenges. Int J Environ Res Public Health 2020 Jun 17;17(12):4322 [FREE Full text] [doi: 10.3390/ijerph17124322] [Medline: 32560363]

23. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Apr 17;69(15):458-464 [FREE Full text] [doi: 10.15585/mmwr.mm6915e3] [Medline: 32298251]
24. Wadhera RK, Wadhera P, Gaba P, Figueroa JF, Joynt Maddox KE, Yeh RW, et al. Variation in COVID-19 Hospitalizations and Deaths Across New York City Boroughs. JAMA 2020 Jun 02;323(21):2192-2195 [FREE Full text] [doi: 10.1001/jama.2020.7197] [Medline: 32347898]
25. Rodriguez-Diaz CE, Guilamo-Ramos V, Mena L, Hall E, Honermann B, Crowley JS, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. Ann Epidemiol 2020 Dec;52:46-53.e2 [FREE Full text] [doi: 10.1016/j.annepidem.2020.07.007] [Medline: 32711053]
26. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, COVID-19 Lombardy ICU Network. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. JAMA 2020 Apr 28;323(16):1574-1581 [FREE Full text] [doi: 10.1001/jama.2020.5394] [Medline: 32250385]
27. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? Lancet Respir Med 2020 Apr;8(4):e21 [FREE Full text] [doi: 10.1016/S2213-2600(20)30116-8] [Medline: 32171062]
28. Emami A, Javanmardi F, Pirbonyeh N, Akbari A. Prevalence of Underlying Diseases in Hospitalized Patients with COVID-19: a Systematic Review and Meta-Analysis. Arch Acad Emerg Med 2020;8(1):e35 [FREE Full text] [Medline: 32232218]
29. Madjid M, Safavi-Naeini P, Solomon SD, Vardeny O. Potential Effects of Coronaviruses on the Cardiovascular System: A Review. JAMA Cardiol 2020 Jul 01;5(7):831-840. [doi: 10.1001/jamacardio.2020.1286] [Medline: 32219363]
30. COVID-19: underlying metabolic health in the spotlight. Lancet Diabetes Endocrinol 2020 May 06;8(6):457. [doi: 10.1016/s2213-8587(20)30164-9]
31. Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 Transmission in Patients With Cancer at a Tertiary Care Hospital in Wuhan, China. JAMA Oncol 2020 Jul 01;6(7):1108-1110 [FREE Full text] [doi: 10.1001/jamaoncol.2020.0980] [Medline: 32211820]
32. Kutikov A, Weinberg DS, Edelman MJ, Horwitz EM, Uzzo RG, Fisher RI. A War on Two Fronts: Cancer Care in the Time of COVID-19. Ann Intern Med 2020 Jun 02;172(11):756-758 [FREE Full text] [doi: 10.7326/M20-1133] [Medline: 32219410]
33. Islam JY, Camacho-Rivera M, Vidot DC. Examining COVID-19 Preventive Behaviors among Cancer Survivors in the United States: An Analysis of the COVID-19 Impact Survey. Cancer Epidemiol Biomarkers Prev 2020 Dec;29(12):2583-2590 [FREE Full text] [doi: 10.1158/1055-9965.EPI-20-0801] [Medline: 32978173]
34. Higham A, Mathioudakis A, Vestbo J, Singh D. COVID-19 and COPD: a narrative review of the basic science and clinical outcomes. Eur Respir Rev 2020 Dec 31;29(158):200199 [FREE Full text] [doi: 10.1183/16000617.0199-2020] [Medline: 33153991]
35. Edis EC. Chronic Pulmonary Diseases and COVID-19. Turk Thorac J 2020 Sep;21(5):345-349 [FREE Full text] [doi: 10.5152/TurkThoracJ.2020.20091] [Medline: 33031727]
36. Pezzuto A, Tammaro A, Tonini G, Ciccozzi M. COPD influences survival in patients affected by COVID-19, comparison between subjects admitted to an internal medicine unit, and subjects admitted to an intensive care unit: An Italian experience. J Med Virol. Epub ahead of print 2020 Oct 07 [FREE Full text] [doi: 10.1002/jmv.26585] [Medline: 33026657]
37. Hughes-Visentin A, Paul ABM. Asthma and COVID-19: What do we know now. Clin Med Insights Circ Respir Pulm Med 2020;14:1179548420966242 [FREE Full text] [doi: 10.1177/1179548420966242] [Medline: 33173369]
38. Pignatti P, Visca D, Cherubino F, Zampogna E, Spanevello A. Impact of COVID-19 on patients with asthma. Int J Tuberc Lung Dis 2020 Nov 01;24(11):1217-1219. [doi: 10.5588/ijtld.20.0608] [Medline: 33172533]
39. Robinson LB, Fu X, Bassett IV, Triant VA, Foulkes AS, Zhang Y, et al. COVID-19 severity in hospitalized patients with asthma: A matched cohort study. J Allergy Clin Immunol Pract 2020 Oct 22;S2213-2198(20):31132-31136 [FREE Full text] [doi: 10.1016/j.jaip.2020.10.021] [Medline: 33164794]
40. Parikh R, Garcia MA, Rajendran I, Johnson S, Mesfin N, Weinberg J, et al. ICU outcomes in Covid-19 patients with obesity. Ther Adv Respir Dis 2020;14:1753466620971146 [FREE Full text] [doi: 10.1177/1753466620971146] [Medline: 33176612]
41. Akbas F, Usta Atmaca H. Obesity and COVID-19: Time to Take Action. Obes Facts 2020 Nov 09:1-3 [FREE Full text] [doi: 10.1159/000511446] [Medline: 33166962]
42. Prado-Galbarro FJ, Sanchez-Piedra C, Gamiño-Arroyo AE, Cruz-Cruz C. Determinants of survival after severe acute respiratory syndrome coronavirus 2 infection in Mexican outpatients and hospitalised patients. Public Health 2020 Sep 30;189:66-72 [FREE Full text] [doi: 10.1016/j.puhe.2020.09.014] [Medline: 33166857]
43. Mesas AE, Cavero-Redondo I, Álvarez-Bueno C, Cabrera MAS, de Andrade SM, Sequí-Dominguez I, et al. Predictors of in-hospital COVID-19 mortality: A comprehensive systematic review and meta-analysis exploring differences by age, sex and health conditions. PLoS One 2020;15(11):e0241742 [FREE Full text] [doi: 10.1371/journal.pone.0241742] [Medline: 33141836]
44. 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;26(7):1037-1040. [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]

45. Inkster B, O'Brien R, Selby E, Joshi S, Subramanian V, Kadaba M, et al. Digital Health Management During and Beyond the COVID-19 Pandemic: Opportunities, Barriers, and Recommendations. JMIR Ment Health 2020 Jul 06;7(7):e19246 [FREE Full text] [doi: 10.2196/19246] [Medline: 32484783]
46. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
47. Timmers T, Janssen L, Stohr J, Murk JL, Berrevoets MAH. Using eHealth to Support COVID-19 Education, Self-Assessment, and Symptom Monitoring in the Netherlands: Observational Study. JMIR Mhealth Uhealth 2020 Jun 23;8(6):e19822 [FREE Full text] [doi: 10.2196/19822] [Medline: 32516750]
48. Yap KYL, Xie Q. Personalizing symptom monitoring and contact tracing efforts through a COVID-19 web-app. Infect Dis Poverty 2020 Jul 13;9(1):93 [FREE Full text] [doi: 10.1186/s40249-020-00711-5] [Medline: 32660568]
49. Wang S, Ding S, Xiong L. A New System for Surveillance and Digital Contact Tracing for COVID-19: Spatiotemporal Reporting Over Network and GPS. JMIR Mhealth Uhealth 2020 Jun 10;8(6):e19457 [FREE Full text] [doi: 10.2196/19457] [Medline: 32499212]
50. García-Iglesias JJ, Martín-Pereira J, Fagundo-Rivera J, Gómez-Salgado J. [Digital surveillance tools for contact tracking of infected persons by SARS-CoV-2.]. Rev Esp Salud Publica 2020 Jun 23;94:e202006067 [FREE Full text] [Medline: 32572019]
51. Yamamoto K, Takahashi T, Urasaki M, Nagayasu Y, Shimamoto T, Tateyama Y, et al. Health Observation App for COVID-19 Symptom Tracking Integrated With Personal Health Records: Proof of Concept and Practical Use Study. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e19902 [FREE Full text] [doi: 10.2196/19902] [Medline: 32568728]
52. Echeverría P, Mas Bergas MA, Puig J, Isnard M, Massot M, Vedia C, et al. COVIDApp as an Innovative Strategy for the Management and Follow-Up of COVID-19 Cases in Long-Term Care Facilities in Catalonia: Implementation Study. JMIR Public Health Surveill 2020 Jul 17;6(3):e21163 [FREE Full text] [doi: 10.2196/21163] [Medline: 32629425]
53. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-Peer Contact Tracing: Development of a Privacy-Preserving Smartphone App. JMIR Mhealth Uhealth 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: 10.2196/18936] [Medline: 32240973]
54. Alexopoulos AR, Hudson JG, Otenigbagbe O. The Use of Digital Applications and COVID-19. Community Ment Health J 2020 Oct;56(7):1202-1203 [FREE Full text] [doi: 10.1007/s10597-020-00689-2] [Medline: 32734311]
55. Banskota S, Healy M, Goldberg EM. 15 Smartphone Apps for Older Adults to Use While in Isolation During the COVID-19 Pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525 [FREE Full text] [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
56. Conway C, Fowler E, Howes B, Blythe A. If you're happy and you know it click the app - the Covid-19 pandemic and its effect on students' use of the "Happy App". Educ Prim Care 2020 Aug 12:1. [doi: 10.1080/14739879.2020.1808859] [Medline: 32783785]
57. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
58. Walrave M, Waeterloos C, Ponnet K. Adoption of a Contact Tracing App for Containing COVID-19: A Health Belief Model Approach. JMIR Public Health Surveill 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: 10.2196/20572] [Medline: 32755882]
59. Berglund J. Tracking COVID-19: There's an App for That. IEEE Pulse 2020;11(4):14-17. [doi: 10.1109/MPULS.2020.3008356] [Medline: 32841114]
60. Zens M, Brammertz A, Herpich J, Südkamp N, Hinterseer M. App-Based Tracking of Self-Reported COVID-19 Symptoms: Analysis of Questionnaire Data. J Med Internet Res 2020 Sep 09;22(9):e21956 [FREE Full text] [doi: 10.2196/21956] [Medline: 32791493]
61. Ming LC, Untong N, Aliudin NA, Osili N, Kifli N, Tan CS, et al. Mobile Health Apps on COVID-19 Launched in the Early Days of the Pandemic: Content Analysis and Review. JMIR Mhealth Uhealth 2020 Sep 16;8(9):e19796 [FREE Full text] [doi: 10.2196/19796] [Medline: 32609622]
62. Soriano JB, Fernández E, de Astorza A, de Llano LAP, Fernández-Villar A, Carnicer-Pont D, et al. Hospital Epidemics Tracker (HEpiTracker): Description and pilot study of a mobile app to track COVID-19 in hospital workers. JMIR Public Health Surveill 2020 Sep 21;6(3):e21653 [FREE Full text] [doi: 10.2196/21653] [Medline: 32845852]
63. Mira JJ, Vicente MA, Lopez-Pineda A, Carrillo I, Guilabert M, Fernández C, et al. Preventing and Addressing the Stress Reactions of Health Care Workers Caring for Patients With COVID-19: Development of a Digital Platform (Be + Against COVID). JMIR Mhealth Uhealth 2020 Oct 05;8(10):e21692 [FREE Full text] [doi: 10.2196/21692] [Medline: 32936769]
64. O'Dowd A. Covid-19: App to track close contacts is launched in England and Wales. BMJ 2020 Sep 25;370:m3751. [doi: 10.1136/bmj.m3751] [Medline: 32978120]
65. Chan AT, Drew DA, Nguyen LH, Joshi AD, Ma W, Guo CG, COPE Consortium. The COronavirus Pandemic Epidemiology (COPE) Consortium: A Call to Action. Cancer Epidemiol Biomarkers Prev 2020 Jul;29(7):1283-1289 [FREE Full text] [doi: 10.1158/1055-9965.EPI-20-0606] [Medline: 32371551]

66. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and Management of Home-Quarantined Patients With COVID-19 Using a WeChat-Based Telemedicine System: Retrospective Cohort Study. J Med Internet Res 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: 10.2196/19514] [Medline: 32568727]
67. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Mölken MRV. COVID-19 Contact Tracing Apps: Predicted Uptake in the Netherlands Based on a Discrete Choice Experiment. JMIR Mhealth Uhealth 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: 10.2196/20741] [Medline: 32795998]
68. Kaspar K. Motivations for Social Distancing and App Use as Complementary Measures to Combat the COVID-19 Pandemic: Quantitative Survey Study. J Med Internet Res 2020 Aug 27;22(8):e21613 [FREE Full text] [doi: 10.2196/21613] [Medline: 32759100]
69. Cioffi A, Lugi C, Cecannecchia C. Apps for COVID-19 contact-tracing: Too many questions and few answers. Ethics Med Public Health 2020;15:100575 [FREE Full text] [doi: 10.1016/j.jemep.2020.100575] [Medline: 32838002]
70. Sun S, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, RADAR-CNS Consortium. Using Smartphones and Wearable Devices to Monitor Behavioral Changes During COVID-19. J Med Internet Res 2020 Sep 25;22(9):e19992 [FREE Full text] [doi: 10.2196/19992] [Medline: 32877352]
71. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental Health and Behavior of College Students During the Early Phases of the COVID-19 Pandemic: Longitudinal Smartphone and Ecological Momentary Assessment Study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
72. Drissi N, Ouhbi S, Idrissi MAJ, Ghogho M. An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps. Int J Med Inform 2020 Sep;141:104243 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104243] [Medline: 32768994]
73. Reyes AT. A Mindfulness Mobile App for Traumatized COVID-19 Healthcare Workers and Recovered Patients: A Response to "The Use of Digital Applications and COVID-19". Community Ment Health J 2020 Oct;56(7):1204-1205 [FREE Full text] [doi: 10.1007/s10597-020-00690-9] [Medline: 32772205]
74. 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]
75. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A systematic review of the smartphone applications available for coronavirus disease 2019 (COVID19) and their assessment using the mobile app rating scale (MARS). medRxiv Preprint posted online on July 4, 2020. [FREE Full text] [doi: 10.1101/2020.07.02.20144964] [Medline: 32637969]
76. Collado-Borrell R, Escudero-Vilaplana V, Villanueva-Bueno C, Herranz-Alonso A, Sanjurjo-Saez M. Features and Functionalities of Smartphone Apps Related to COVID-19: Systematic Search in App Stores and Content Analysis. J Med Internet Res 2020 Aug 25;22(8):e20334 [FREE Full text] [doi: 10.2196/20334] [Medline: 32614777]
77. Davalbhakta S, Advani S, Kumar S, Agarwal V, Bhoyar S, Fedirko E, et al. A Systematic Review of Smartphone Applications Available for Corona Virus Disease 2019 (COVID19) and the Assessment of their Quality Using the Mobile Application Rating Scale (MARS). J Med Syst 2020 Aug 10;44(9):164 [FREE Full text] [doi: 10.1007/s10916-020-01633-3] [Medline: 32779002]
78. Homepage. COVID Impact Survey. URL: https://www.covid-impact.org/ [accessed 2020-05-20]
79. Current Population Survey, September 2020. U.S. Census Bureau. URL: https://www.census.gov/programs-surveys/cps/data.html [accessed 2020-11-04]
80. Calixte R, Rivera A, Oridota O, Beauchamp W, Camacho-Rivera M. Social and Demographic Patterns of Health-Related Internet Use Among Adults in the United States: A Secondary Data Analysis of the Health Information National Trends Survey. Int J Environ Res Public Health 2020 Sep 19;17(18):6856 [FREE Full text] [doi: 10.3390/ijerph17186856] [Medline: 32961766]
81. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
82. Greenberg-Worisek AJ, Kurani S, Finney Rutten LJ, Blake KD, Moser RP, Hesse BW. Tracking Healthy People 2020 Internet, Broadband, and Mobile Device Access Goals: An Update Using Data From the Health Information National Trends Survey. J Med Internet Res 2019 Jun 24;21(6):e13300 [FREE Full text] [doi: 10.2196/13300] [Medline: 31237238]
83. Raharja A, Tamara A, Kok LT. Association Between Ethnicity and Severe COVID-19 Disease: a Systematic Review and Meta-analysis. J Racial Ethn Health Disparities 2020 Nov 12:1-10 [FREE Full text] [doi: 10.1007/s40615-020-00921-5] [Medline: 33180278]
84. Barros AJD, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003 Oct 20;3:21 [FREE Full text] [doi: 10.1186/1471-2288-3-21] [Medline: 14567763]
85. Behrens T, Taeger D, Wellmann J, Keil U. Different methods to calculate effect estimates in cross-sectional studies. A comparison between prevalence odds ratio and prevalence ratio. Methods Inf Med 2004;43(5):505-509. [Medline: 15702210]
86. Coutinho LMS, Scazufca M, Menezes PR. Methods for estimating prevalence ratios in cross-sectional studies. Rev Saude Publica 2008 Dec;42(6):992-998 [FREE Full text] [Medline: 19009156]

87. Althouse AD. Adjust for Multiple Comparisons? It's Not That Simple. Ann Thorac Surg 2016 May;101(5):1644-1645. [doi: 10.1016/j.athoracsur.2015.11.024] [Medline: 27106412]
88. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990 Jan;1(1):43-46. [Medline: 2081237]
89. Kamenidou IE, Stavrianea A, Mamalis S, Mylona I. Knowledge Assessment of COVID-19 Symptoms: Gender Differences and Communication Routes for the Generation Z Cohort. Int J Environ Res Public Health 2020 Sep 23;17(19):6964 [FREE Full text] [doi: 10.3390/ijerph17196964] [Medline: 32977632]
90. Wolf MS, Serper M, Opsasnick L, O'Conor RM, Curtis LM, Benavente JY, et al. Awareness, Attitudes, and Actions Related to COVID-19 Among Adults With Chronic Conditions at the Onset of the U.S. Outbreak: A Cross-sectional Survey. Ann Intern Med 2020 Jul 21;173(2):100-109 [FREE Full text] [doi: 10.7326/M20-1239] [Medline: 32271861]
91. Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. Proc Natl Acad Sci U S A 2020 Aug 18;117(33):19658-19660 [FREE Full text] [doi: 10.1073/pnas.2009412117] [Medline: 32727905]
92. Geldsetzer P. Knowledge and Perceptions of COVID-19 Among the General Public in the United States and the United Kingdom: A Cross-sectional Online Survey. Ann Intern Med 2020 Jul 21;173(2):157-160 [FREE Full text] [doi: 10.7326/M20-0912] [Medline: 32196071]
93. Parlapani E, Holeva V, Voitsidis P, Blekas A, Gliatas I, Porfyri GN, et al. Psychological and Behavioral Responses to the COVID-19 Pandemic in Greece. Front Psychiatry 2020;11:821 [FREE Full text] [doi: 10.3389/fpsyt.2020.00821] [Medline: 32973575]
94. Alsan M, Stantcheva S, Yang D, Cutler D. Disparities in Coronavirus 2019 Reported Incidence, Knowledge, and Behavior Among US Adults. JAMA Netw Open 2020 Jun 01;3(6):e2012403 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12403] [Medline: 32556260]
95. Clements JM. Knowledge and Behaviors Toward COVID-19 Among US Residents During the Early Days of the Pandemic: Cross-Sectional Online Questionnaire. JMIR Public Health Surveill 2020 May 08;6(2):e19161 [FREE Full text] [doi: 10.2196/19161] [Medline: 32369759]
96. Karaye IM, Horney JA. The Impact of Social Vulnerability on COVID-19 in the U.S.: An Analysis of Spatially Varying Relationships. Am J Prev Med 2020 Sep;59(3):317-325 [FREE Full text] [doi: 10.1016/j.amepre.2020.06.006] [Medline: 32703701]
97. Hamidi S, Sabouri S, Ewing R. Does Density Aggravate the COVID-19 Pandemic? J Am Plann Assoc 2020 Jun 18;86(4):495-509 [FREE Full text] [doi: 10.1080/01944363.2020.1777891]
98. O'Conor R, Opsasnick L, Benavente JY, Russell AM, Wismer G, Eifler M, et al. Knowledge and Behaviors of Adults with Underlying Health Conditions During the Onset of the COVID-19 U.S. Outbreak: The Chicago COVID-19 Comorbidities Survey. J Community Health 2020 Dec;45(6):1149-1157 [FREE Full text] [doi: 10.1007/s10900-020-00906-9] [Medline: 32780294]
99. Campos-Castillo C, Laestadius LI. Racial and Ethnic Digital Divides in Posting COVID-19 Content on Social Media Among US Adults: Secondary Survey Analysis. J Med Internet Res 2020 Jul 03;22(7):e20472 [FREE Full text] [doi: 10.2196/20472] [Medline: 32568726]
100. Gibson A, Bardach SH, Pope ND. COVID-19 and the Digital Divide: Will Social Workers Help Bridge the Gap? J Gerontol Soc Work 2020 Jun 05:1-3. [doi: 10.1080/01634372.2020.1772438] [Medline: 32500841]
101. Lorca-Cabrera J, Martí-Arques R, Albacar-Riobóo N, Raigal-Aran L, Roldan-Merino J, Ferré-Grau C. Mobile Applications for Caregivers of Individuals with Chronic Conditions and/or Diseases: Quantitative Content Analysis. Int J Med Inform 2020 Nov 02;145:104310. [doi: 10.1016/j.ijmedinf.2020.104310] [Medline: 33161319]
102. Birkhoff SD, Smeltzer SC. Perceptions of Smartphone User-Centered Mobile Health Tracking Apps Across Various Chronic Illness Populations: An Integrative Review. J Nurs Scholarsh 2017 Jul;49(4):371-378. [doi: 10.1111/jnu.12298] [Medline: 28605151]
103. Camacho-Rivera M, Vo H, Huang X, Lau J, Lawal A, Kawaguchi A. Evaluating Asthma Mobile Apps to Improve Asthma Self-Management: User Ratings and Sentiment Analysis of Publicly Available Apps. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e15076 [FREE Full text] [doi: 10.2196/15076] [Medline: 33118944]
104. Villarreal V, Berbey-Alvarez A. Evaluation of mHealth Applications Related to Cardiovascular Diseases: a Systematic Review. Acta Inform Med 2020 Jun;28(2):130-137 [FREE Full text] [doi: 10.5455/aim.2020.28.130-137] [Medline: 32742066]
105. Buss VH, Leesong S, Barr M, Varnfield M, Harris M. Primary Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus Using Mobile Health Technology: Systematic Review of the Literature. J Med Internet Res 2020 Oct 29;22(10):e21159 [FREE Full text] [doi: 10.2196/21159] [Medline: 33118936]
106. Larbi D, Randine P, Årsand E, Antypas K, Bradway M, Gabarron E. Methods and Evaluation Criteria for Apps and Digital Interventions for Diabetes Self-Management: Systematic Review. J Med Internet Res 2020 Jul 06;22(7):e18480 [FREE Full text] [doi: 10.2196/18480] [Medline: 32628125]
107. Widnall E, Grant CE, Wang T, Cross L, Velupillai S, Roberts A, et al. User Perspectives of Mood-Monitoring Apps Available to Young People: Qualitative Content Analysis. JMIR Mhealth Uhealth 2020 Oct 10;8(10):e18140 [FREE Full text] [doi: 10.2196/18140] [Medline: 33037875]

Abbreviations

aPR: adjusted prevalence ratio
COPD: chronic obstructive pulmonary disease
cOR: conditional odds ratio
mHealth: mobile health
NORC: National Opinion Research Center

Edited by L Buis; submitted 30.09.20; peer-reviewed by R Calixte, N Drissi; comments to author 13.11.20; revised version received 16.11.20; accepted 20.11.20; published 17.12.20.

Please cite as:
Camacho-Rivera M, Islam JY, Rivera A, Vidot DC
Attitudes Toward Using COVID-19 mHealth Tools Among Adults With Chronic Health Conditions: Secondary Data Analysis of the COVID-19 Impact Survey
JMIR Mhealth Uhealth 2020;8(12):e24693
URL: http://mhealth.jmir.org/2020/12/e24693/
doi:10.2196/24693
PMID:33301415

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org