

---

# JMIR mHealth and uHealth

---

Impact Factor (2022): 5.0  
Volume 7 (2019), Issue 11 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

---

## Contents

### Viewpoint

- Mind the App: Considerations for the Future of Mobile Health in Canada ([e15301](#))  
Ma'n Zawati, Michael Lang. .... 4

### Reviews

- Examining Mobile Technologies to Support Older Adults With Dementia Through the Lens of Personhood and Human Needs: Scoping Review ([e15122](#))  
Bon Koo, Lisa Vizer. .... 7
- Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review ([e14730](#))  
Aurélie Najm, Laure Gossec, Catherine Weill, David Benoist, Francis Berenbaum, Elena Nikiphorou. .... 139

### Original Papers

- Iterative Adaptation of a Mobile Nutrition Video-Based Intervention Across Countries Using Human-Centered Design: Qualitative Study ([e13604](#))  
Jasmin Isler, N Sawadogo, Guy Harling, Till Bärnighausen, Maya Adam, Moubassira Kagoné, Ali Sié, Merlin Greuel, Shannon McMahon. . .  
2 2
- Exploring the Patterns of Use and Acceptability of Mobile Phones Among People Living With HIV to Improve Care and Treatment: Cross-Sectional Study in Three Francophone West African Countries ([e13741](#))  
Phillipe Lepère, Yélamikan Touré, Alexandra Bitty-Anderson, Simon Boni, Gildas Anago, Boris Tchounga, Pendadiago Touré, Albert Minga, Eugène Messou, Guillaume Kanga, Serge Koule, Armel Poda, Alexandra Calmy, Didier Ekouevi, Patrick Coffie. .... 40
- A Model for Assessing Necessary Conditions for Rural Health Care's Mobile Health Readiness: Qualitative Assessment of Clinician-Perceived Barriers ([e11915](#))  
Bryan Weichelt, Casper Bendixsen, Timothy Patrick. .... 51
- Making Sense of Negative Findings from Mobile Attention Bias Modification Interventions for Individuals with Addictive Disorders: Quantitative Feasibility Study ([e16325](#))  
Melvyn Zhang, Jiangbo Ying, Syidda Amron, Zaakira Mahreen, Guo Song, Daniel Fung, Helen Smith. .... 65

Development of a Deep Learning Model for Dynamic Forecasting of Blood Glucose Level for Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial ( <a href="#">e14452</a> ) Syed Faruqui, Yan Du, Rajitha Meka, Adel Alaeddini, Chengdong Li, Sara Shirinkam, Jing Wang. ....	75
Mental Health Apps in China: Analysis and Quality Assessment ( <a href="#">e13236</a> ) Jie Shang, Shaoming Wei, Jianbo Jin, Puhong Zhang. ....	89
Mediators of Intervention Effects on Depressive Symptoms Among People Living With HIV: Secondary Analysis of a Mobile Health Randomized Controlled Trial Using Latent Growth Curve Modeling ( <a href="#">e15489</a> ) Mengting Zhu, Weiping Cai, Linghua Li, Yan Guo, Aliza Monroe-Wise, Yiran Li, Chengbo Zeng, Jiaying Qiao, Zhimeng Xu, Hanxi Zhang, Yu Zeng, Cong Liu. ....	99
Assessment of a Mobile App by Adolescents and Young Adults With Cystic Fibrosis: Pilot Evaluation ( <a href="#">e12442</a> ) Isa Rudolf, Katharina Pieper, Helga Nolte, Sibylle Junge, Christian Dopfer, Annette Sauer-Heilborn, Felix Ringshausen, Burkhard Tümmeler, Ute von Jan, Urs-Vito Albrecht, Jan Fuge, Gesine Hansen, Anna-Maria Dittrich. ....	113
Mood Monitoring Over One Year for People With Chronic Obstructive Pulmonary Disease Using a Mobile Health System: Retrospective Analysis of a Randomized Controlled Trial ( <a href="#">e14946</a> ) Maxine Whelan, Carmelo Velardo, Heather Rutter, Lionel Tarassenko, Andrew Farmer. ....	127
Effects of a Smartphone-Based Approach-Avoidance Intervention on Chocolate Craving and Consumption: Randomized Controlled Trial ( <a href="#">e12298</a> ) Adrian Meule, Anna Richard, Radomir Dinic, Jens Blechert. ....	149
A Mobile-Based Comprehensive Weight Reduction Program for the Workplace (Health-On): Development and Pilot Study ( <a href="#">e11158</a> ) Min Han, Belong Cho, Hyuktae Kwon, Ki Son, Hyejin Lee, Joo Lee, Jinho Park. ....	161
Development of a Mobile Phone App for Measuring Striking Response Time in Combat Sports: Cross-Sectional Validation Study ( <a href="#">e14641</a> ) Victor Coswig, Jader Sant' Ana, Maicon Coelho, Antonio Pereira Moro, Fernando Diefenthaeler. ....	174
Impact of a Mobile App-Based Health Coaching and Behavior Change Program on Participant Engagement and Weight Status of Overweight and Obese Children: Retrospective Cohort Study ( <a href="#">e14458</a> ) Victor Cueto, C Wang, Lee Sanders. ....	182
A Baby Formula Designed for Chinese Babies: Content Analysis of Milk Formula Advertisements on Chinese Parenting Apps ( <a href="#">e14219</a> ) Jing Zhao, Mu Li, Becky Freeman. ....	195
An Interactive Parent-Targeted Text Messaging Intervention to Improve Oral Health in Children Attending Urban Pediatric Clinics: Feasibility Randomized Controlled Trial ( <a href="#">e14247</a> ) Belinda Borrelli, Michelle Henshaw, Romano Endrighi, William Adams, Timothy Heeren, Rochelle Rosen, Beth Bock, Scott Wertz. ....	207
Impact of Social Determinants of Health and Demographics on Refill Requests by Medicare Patients Using a Conversational Artificial Intelligence Text Messaging Solution: Cross-Sectional Study ( <a href="#">e15771</a> ) Rena Brar Prayaga, Ridhika Agrawal, Benjamin Nguyen, Erwin Jeong, Harmony Noble, Andrew Paster, Ram Prayaga. ....	220
Development of MyTeen Text Messaging Program to Support Parents of Adolescents: Qualitative Study ( <a href="#">e15664</a> ) Joanna Chu, Angela Wadham, Yannan Jiang, Robyn Whittaker, Karolina Stasiak, Matthew Shepherd, Christopher Bullen. ....	235

<b>An Interactive Mobile Phone App (SMART 5-A-DAY) for Increasing Knowledge of and Adherence to Fruit and Vegetable Recommendations: Development and Pilot Randomized Controlled Trial (e14380)</b>	
Katherine Appleton, David Passmore, Isobel Burn, Hanna Pidgeon, Philippa Nation, Charlotte Boobyer, Nan Jiang. ....	244
<b>Evaluation of Heart Failure Apps to Promote Self-Care: Systematic App Search (e13173)</b>	
Sahr Wali, Catherine Demers, Hiba Shah, Huda Wali, Delphine Lim, Nirav Naik, Ahmad Ghany, Ayushi Vispute, Maya Wali, Karim Keshavjee. 2	6
<b>Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study (e15195)</b>	
Miguel Mayer, Octavi Rodríguez Blanco, Antonio Torrejon. ....	275
<b>Characteristics and Usage Patterns Among 12,151 Paid Subscribers of the Calm Meditation App: Cross-Sectional Survey (e15648)</b>	
Jennifer Huberty, Ana-Maria Vranceanu, Colleen Carney, Michael Breus, Michael Gordon, Megan Puzia. ....	287
<b>Improving Medication Information Presentation Through Interactive Visualization in Mobile Apps: Human Factors Design (e15940)</b>	
Don Roosan, Yan Li, Anandi Law, Huy Truong, Mazharul Karim, Jay Chok, Moom Roosan. ....	302
<b>The Use of Smartphone-Based Triage to Reduce the Rate of Outpatient Error Registration: Cross-Sectional Study (e15313)</b>	
Wanhua Xie, Xiaojun Cao, Hongwei Dong, Yu Liu. ....	312
<b>Quality Awareness and Its Influence on the Evaluation of App Meta-Information by Physicians: Validation Study (e16442)</b>	
Urs-Vito Albrecht, Theodor Framke, Ute von Jan. ....	321
<b>Continuous Digital Monitoring of Walking Speed in Frail Elderly Patients: Noninterventional Validation Study and Longitudinal Clinical Trial (e15191)</b>	
Arne Mueller, Holger Hoefling, Amir Muaremi, Jens Praestgaard, Lorcan Walsh, Ola Bunte, Roland Huber, Julian Fürmetz, Alexander Keppler, Matthias Schieker, Wolfgang Böcker, Ronenn Roubenoff, Sophie Brachat, Daniel Rooks, Ieuan Clay. ....	340
<b>Parental Perspectives of a Wearable Activity Tracker for Children Younger Than 13 Years: Acceptability and Usability Study (e13858)</b>	
Kelly Mackintosh, Stephanie Chappel, Jo Salmon, Anna Timperio, Kylie Ball, Helen Brown, Susie Macfarlane, Nicola Ridgers. ....	352
<b>Opportunities and Pitfalls in Applying Emotion Recognition Software for Persons With a Visual Impairment: Simulated Real Life Conversations (e13722)</b>	
Hendrik Buimer, Renske Schellens, Tjerk Kostelijk, Abdellatif Nemri, Yan Zhao, Thea Van der Geest, Richard Van Wezel. ....	368
<b>Real-Time Mobile Monitoring of Drinking Episodes in Young Adult Heavy Drinkers: Development and Comparative Survey Study (e13765)</b>	
Daniel Fridberg, James Faria, Dingcai Cao, Andrea King. ....	380

Viewpoint

# Mind the App: Considerations for the Future of Mobile Health in Canada

Ma'n H Zawati<sup>1</sup>, LLB, LLM, DCL; Michael Lang<sup>1</sup>, BA, BCL, LLB

Centre of Genomics and Policy, McGill University, Montreal, QC, Canada

**Corresponding Author:**

Ma'n H Zawati, LLB, LLM, DCL

Centre of Genomics and Policy

McGill University

740, Avenue Dr Penfield

Suite 5203

Montreal, QC, H3A 0G1

Canada

Phone: 1 5146686599

Email: [man.zawati@mcgill.ca](mailto:man.zawati@mcgill.ca)

## Abstract

Over the past decade, smartphone technology has become increasingly sophisticated and ubiquitous. Modern smartphones, now owned by more than three quarters of Canadians and 94% of millennials, perform an array of functions that are potentially useful in the health care context, such as tracking fitness data, enabling health record sharing, and providing user-friendly platforms for disease management. Approximately half of smartphone users have downloaded at least one health app, and clinicians are increasingly using them in their practice. However, despite widespread use, there is little evidence that supports their safety and efficacy. Few apps have been independently evaluated and many lack basic patient protections such as privacy policies. In this context, the demand for the regulation of mobile health apps has increased. Against this backdrop, regulators, including Health Canada, have begun to propose regulating the use of smartphones in health care. In this viewpoint, we respond to Health Canada's recent proposal to regulate smartphone use in Canada according to a risk-based model. We argue that although Health Canada's recent proposed approach is promising, it may require complementary regulation and oversight.

(*JMIR Mhealth Uhealth* 2019;7(11):e15301) doi:[10.2196/15301](https://doi.org/10.2196/15301)

**KEYWORDS**

smartphone; mobile phone; regulation; patients; physicians

## Introduction

As smartphone technology has become increasingly sophisticated and ubiquitous, their use and prominence in health care has grown spectacularly. Modern smartphones, now owned by more than three quarters of Canadians and 94% of millennials [1], perform an array of functions that are potentially useful in the health care context. They are capable of tracking fitness data, enabling health record sharing, and providing user-friendly platforms for disease management. Approximately half of smartphone users have downloaded at least one health app, and clinicians are increasingly using them in their practice [2]. However, despite widespread use, there is an incomplete patchwork of evidence supporting their safety and efficacy [3]. Few apps have been independently evaluated and many lack basic patient protections such as privacy policies [3]. In this context, the demand for the regulation of mobile health apps has increased [4].

## Health Canada's Draft Guidance

In January 2019, Health Canada issued draft guidance on the regulation of software as a medical device (SaMD). This important development follows early attempts to regulate the proliferation of mobile health apps in the United Kingdom and the United States. The United Kingdom's National Health Service, for example, has developed an Apps Library, which publishes lists of health apps that have been reviewed on a range of factors, including security and clinical safety [4]. The United States Food and Drug Administration has pursued an approach based on risk assessments, in which apps that function as medical devices and potentially pose risks to patient safety are subject to regulatory oversight [5]. Health Canada's draft guidance closely follows this risk-based model.

Health Canada's issuance of the draft guidance on this issue confirms the increasingly prominent role that software plays in

health care delivery in Canada and provides clarity on the regulatory status of novel technologies and devices. There are, however, numerous open, persistent questions on the issue of mobile apps that have medical functions. Indeed, for the purposes of Health Canada's regulatory oversight of mobile health apps under the *Medical Devices Regulations*, the draft guidance specifies that certain mobile apps will be considered SaMD [6]. However, mobile apps meaningfully differ from other kinds of SaMD in at least two respects. First, mobile health apps are incredibly numerous: over 100,000 of them are available on Apple and Android app stores [7]. Second, mobile health apps are ubiquitous: members of the public use them widely, often outside of clinical settings [3]. Each of these factors presents a unique regulatory challenge. Although Health Canada's recent proposed approach is promising, it will not likely be adequate on its own.

In the draft guidance, software that meets the definition of a medical device according to the *Food and Drugs Act* will be regulated according to a risk-based classification system outlined in the *Medical Devices Regulations*. There are, however, certain kinds of software that Health Canada proposes to exclude from the regulation. Software that matches patient symptoms with treatment guidelines for common illnesses, for example, will not meet the definition of a medical device. Although generally innocuous, such apps may still carry a degree of risk for patients. Consumer apps that draw associations between symptoms and disease might prompt patients to make uninformed health decisions, whereas the collection of personal health information might raise general privacy and data security concerns.

However, potentially deeper challenges exist. Health Canada's draft guidance proposes to regulate SaMD according to a manufacturer's *labeled intended use for the product* [6]. In the event of a disagreement on device classification, Health Canada retains final decision-making authority. However, apps will often not have an explicit intended use or may have ancillary functions that approximate a medical purpose even with express indications to the contrary [8]. Furthermore, apps may be made available to the public in the absence of any interaction with Health Canada. The Apple app Store, for example, does not appear to require regulatory clearance for the distribution of medical apps. App Store Review Guidelines specify only that medical apps may be subject to increased scrutiny and that developers should provide documentary evidence of regulatory clearance if such clearance has been obtained [9]. Where an app in public use falls within Health Canada's regulatory ambit, or where its classification as a medical device is contentious, it is not clear how the agency's oversight can be managed. Given the number and diversity of available apps, it is almost certain that many of them will fall through the cracks.

Health Canada's focus on apps as medical devices also potentially obscures that they are increasingly used as tools for conducting health research. Given their unprecedented data collection capacities and ease of distribution, research projects have begun using mobile apps to facilitate investigative work. As they do, the traditionally firm ethical distinction between research and clinical care is becoming significantly more difficult to manage.

## Role of Medical Colleges and Faculties

Taken together, each of these factors underscores the unique complexity of regulating mobile health apps. In fact, mobile health apps may simply be too numerous and too diverse in function to be effectively regulated through 1 government department or agency. Rather, a flexible and multifaceted approach to the regulation of mobile health apps should be pursued. The medical colleges, for one, ought to play an active role in facilitating the use of mobile apps in the clinic. Guidelines for using mobile apps in clinical practice and for recommending them to patients should be developed and periodically revised. Specific limitations on what an app can or cannot do in relation to reserved medical practice should be clarified. Although the Canadian Medical Association has released a guidance document for the prescription of mobile apps by physicians, [10] provincial medical colleges are more directly empowered to regulate the clinical use of health apps among their members, and as mobile health becomes ever more prominent, clinical training that accounts for these trends will become critically important. In parallel, medical education and training programs will play a critical role in preparing physicians to confront the mobile future. Present trends suggest that health apps will be increasingly prevalent features of the Canadian health care system, and physicians must be prepared to understand their powers and limitations. This means that Canadian medical faculties should train future physicians on how mobile apps are likely to affect their practice and their duties to inform, treat, and protect confidentiality and to follow-up. In the future, we may even see the need for a medical computing specialization.

## Conclusions

Whatever approach is ultimately pursued, the medical profession must play an active role in confronting how health apps and other technologies will change the face of health care. This will require a concerted effort by medical colleges to provide both guidance and clinical training to their members. It will also require prospective planning through medical education and, perhaps, specialization. Although an apple a day might keep the doctor away, the same cannot be said, at the moment, of smartphone apps.

---

## Conflicts of Interest

None declared.

---

## References

1. Statistics Canada. Life in the fast lane: How are Canadians managing? 2016 URL: <https://www.statcan.gc.ca/daily-quotidien/171114/dq171114a-eng.pdf> [accessed 2019-09-18]
2. Powell AC, Landman AB, Bates DW. In search of a few good apps. *J Am Med Assoc* 2014 May 14;311(18):1851-1852. [doi: [10.1001/jama.2014.2564](https://doi.org/10.1001/jama.2014.2564)] [Medline: [24664278](https://pubmed.ncbi.nlm.nih.gov/24664278/)]
3. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. *J Med Internet Res* 2016 Oct 31;18(11):e287 [FREE Full text] [doi: [10.2196/jmir.5692](https://doi.org/10.2196/jmir.5692)] [Medline: [27806926](https://pubmed.ncbi.nlm.nih.gov/27806926/)]
4. National Health Services. How We Assess Health Apps and Digital Tools URL: <https://digital.nhs.uk/services/nhs-apps-library/guidance-for-health-app-developers-commissioners-and-assessors/how-we-assess-health-apps-and-digital-tools#step-4-get-published-on-the-library> [accessed 2019-09-18]
5. Food & Drug Administration. 2015. Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff URL: <https://www.fda.gov/media/80958/download> [accessed 2019-09-18]
6. Freudenberg N. Assessing the public health impact of the mHealth app business. *Am J Public Health* 2017 Nov;107(11):1694-1696. [doi: [10.2105/AJPH.2017.304083](https://doi.org/10.2105/AJPH.2017.304083)] [Medline: [29019765](https://pubmed.ncbi.nlm.nih.gov/29019765/)]
7. Government of Canada. 2019. Draft Guidance Document: Software as a Medical Device URL: <https://www.canada.ca/en/health-canada/services/drugs-health-products/public-involvement-consultations/medical-devices/software-medical-device-draft-guidance/document.html> [accessed 2019-09-20]
8. Lang M, Zawati MH. The app will see you now: mobile health, diagnosis, and the practice of medicine in Quebec and Ontario. *J Law Biosci* 2018 May;5(1):142-173 [FREE Full text] [doi: [10.1093/jlb/lsy004](https://doi.org/10.1093/jlb/lsy004)] [Medline: [29707219](https://pubmed.ncbi.nlm.nih.gov/29707219/)]
9. Apple Developer. 2019. App Store Review Guidelines URL: <https://developer.apple.com/app-store/review/guidelines/> [accessed 2019-09-18]
10. Canadian Medical Association. 2015. Guiding Principles for Physicians Recommending Mobile Health Applications to Patients URL: <https://www.cma.ca/guiding-principles-physicians-recommending-mobile-health-applications-patients> [accessed 2019-09-18]

## Abbreviations

**SaMD:** software as a medical device

*Edited by G Eysenbach; submitted 28.06.19; peer-reviewed by B Price, M Doerr; comments to author 07.08.19; revised version received 21.08.19; accepted 21.08.19; published 04.11.19.*

*Please cite as:*

Zawati MH, Lang M

*Mind the App: Considerations for the Future of Mobile Health in Canada*

*JMIR Mhealth Uhealth* 2019;7(11):e15301

URL: <https://mhealth.jmir.org/2019/11/e15301>

doi: [10.2196/15301](https://doi.org/10.2196/15301)

PMID: [31682580](https://pubmed.ncbi.nlm.nih.gov/31682580/)

©Ma'n H Zawati, Michael Lang. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 04.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Review

# Examining Mobile Technologies to Support Older Adults With Dementia Through the Lens of Personhood and Human Needs: Scoping Review

Bon Mi Koo<sup>1</sup>, PhD; Lisa M Vizer<sup>2</sup>, PhD

<sup>1</sup>SSK Research Center for Mental Health and Communal Society, Kwangwoon University, Seoul, Republic of Korea

<sup>2</sup>Division of General Medicine and Clinical Epidemiology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

**Corresponding Author:**

Bon Mi Koo, PhD

SSK Research Center for Mental Health and Communal Society

Kwangwoon University

Chambit 604

20 Gwangun-ro, Nowon-gu

Seoul, 01897

Republic of Korea

Phone: 82 2 940 5430

Email: [bmkmmon@gmail.com](mailto:bmkmmon@gmail.com)

## Abstract

**Background:** With the world's rapidly growing older adult population, there is an increase in the number of people living with dementia. This growth leads to a strain on their caregivers and our health care system and to an increased attention on mitigating strain by using mobile technology to sustain the independence of people with dementia. However, less attention is given to whether these technologies meet the stated and unstated needs of people with dementia.

**Objective:** The aim of this study was to provide an overview of the current research on mobile technologies for people with dementia, considering the current research through the lens of personhood and human needs, and to identify any gaps that represent research opportunities.

**Methods:** We performed a systematic search in Medical Literature Analysis and Retrieval System Online (MEDLINE), Web of Science, PsycINFO, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Excerpta Medica dataBASE (EMBASE), and the Cochrane Central Register of Controlled Trials (CENTRAL) in October 2018. We screened 5560 articles and identified 24 that met our inclusion and exclusion criteria. We then performed thematic analysis to organize the articles by the types of support mobile technologies provide and mapped those types of support to human needs to identify the gaps in support.

**Results:** Articles described research on mobile technologies that support people with dementia to (1) perform daily activities, (2) maintain social interaction, (3) aid memory, (4) engage in leisure activities, (5) track location, and (6) monitor health. At least one type of support mapped to each human need, with most supporting lower-level needs such as physiological and safety needs. Little attention seems to be paid to personhood.

**Conclusions:** Mobile technologies that support daily activities, relationships, memory, leisure activities, health, and safety can partially compensate for decreased function owing to dementia, but the human needs of people with dementia are often not adequately considered. Most technologies support basic physiological and safety needs, whereas many pay little attention to higher-level needs such as self-esteem and agency. Important research opportunities include using person-centered methods to develop technology to meet higher-level needs and to preserve personhood by incorporating human and psychological needs of people with dementia along with ethical considerations.

(*JMIR Mhealth Uhealth* 2019;7(11):e15122) doi:[10.2196/15122](https://doi.org/10.2196/15122)

**KEYWORDS**

dementia; Alzheimer disease; mobile health; consumer health informatics; personhood; systematic review; smartphone; mobile phone; tablet computers

## Introduction

As the world's population rapidly ages, the number of older adults with cognitive impairment will grow as well. Nearly 47 million people worldwide are now experiencing dementia, and this number is projected to triple by 2050 [1]. Of the diseases that can occur as people age, dementia is of particular concern to patients, their family and friends, and society as people with dementia progressively lose independence and autonomy, causing a sharp increase in the burden of care. In the early stages of dementia, people can benefit from support for complex tasks of daily living, but as the disease progresses, people become fully dependent on others to complete even basic activities of daily living [2,3].

Innovative technologies may assist people with dementia and informal carers (usually family or close friends, as opposed to health professionals such as nurses) in several ways. Assistive technology can decrease the burden of care [4-6], increase the independence of people with health conditions [7], and improve the well-being of people and their carers [8]. Advanced mobile technology designed for people with dementia, their carers, and even the surrounding environment has further expanded the scope of assistive technology available for people with dementia [9]. Examples [4-9] include partial compensation for functional deficits of people with dementia and support for carer's care routines. Furthermore, ambient intelligence technologies embedded in the environment can increase safety and security by monitoring people with dementia in their homes to detect emergencies [10]. To date, however, little research has evaluated how mobile technologies specifically support daily activities of people with dementia.

Previous work [11-13] examining the role of human needs in the quality of life of people with dementia emphasizes the importance of considering the technologies designed for people with dementia in light of human needs as described by Maslow [14] and Kitwood [15]. Dewing's [16] work also contributes context with his definition of personhood as "the attributes possessed by human beings that make them a person." This definition is particularly applicable to people with dementia as it is inclusive and, unlike some other definitions of personhood, does not require that a person possess certain capabilities. Rather, it values any person's possession of any attributes of humanity [16,17]. However, we found no research examining the role of these technologies fulfilling human needs and preserving personhood.

To address this gap, we reviewed the literature on mobile technologies intended for use by people with dementia in their daily lives and considered the current research through the lens of human needs and personhood. Our review defined 6 types of support that mobile technologies provide to people with dementia and caregivers, mapped the types of technology support to the human needs defined by Maslow [14] and Kitwood [15], and discussed considerations relating to technology adoption and ethics involved in developing technology for people with dementia.

## Methods

### Overview

We conducted a scoping review based on the methodological framework suggested by Arksey and O'Malley [18] and Peters et al [19], with the goal of organizing the literature on mobile technologies to assist people with dementia with daily living. We aimed to rapidly review the extent and range of research activity and map the key concepts underpinning the research area of mobile technology support for people with dementia [18]. The methodological framework of our scoping review includes identifying data sources and search strategies, selecting relevant articles, and extracting and charting the results.

### Data Source and Search Strategies

We performed a systematic search in October 2018 for articles published in Medical Literature Analysis and Retrieval System Online (MEDLINE), Web of Science, PsycINFO, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Excerpta Medica dataBASE (EMBASE), and the Cochrane Central Register of Controlled Trials (CENTRAL). Systematic search terms included mobile devices, mobile app, smartphone, tablet, mild cognitive impairment, Alzheimer, dementia, and older adults. [Multimedia Appendix 1](#) lists search strings per database.

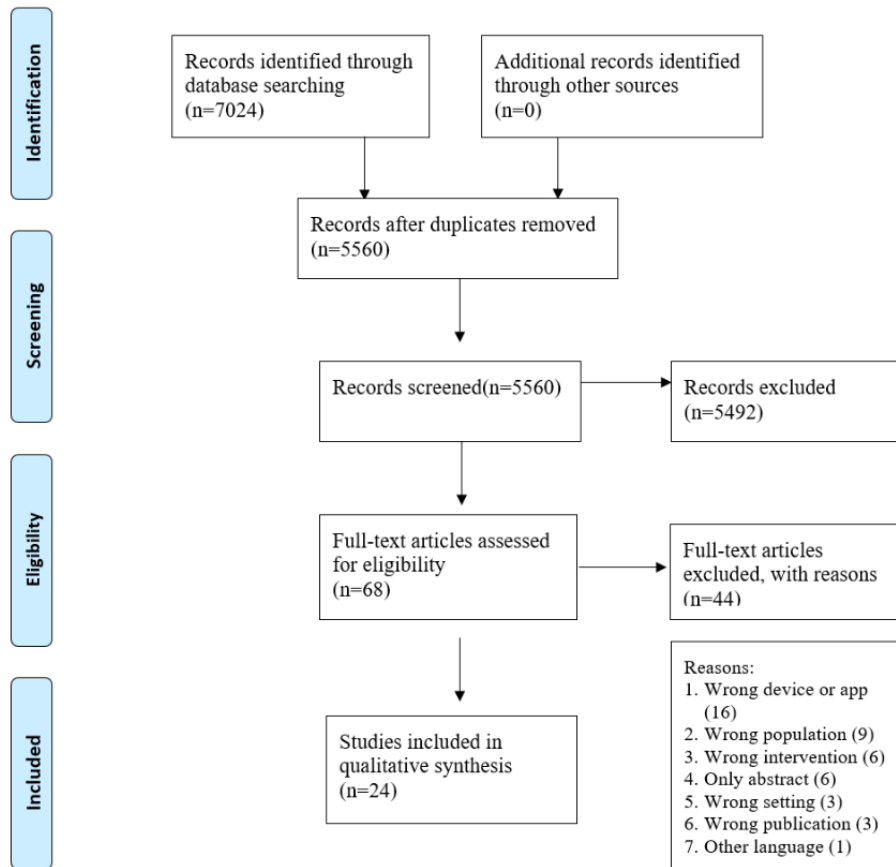
### Study Selection

We selected articles based on the following inclusion criteria: (1) the article is published in a peer-reviewed journal in English, with full text available, (2) the article describes the study participants or target users as people with dementia aged 50 years or above and their caregivers, (3) the article describes mobile technology or mobile apps used by people with dementia and their caregivers to support daily activities of people with dementia, and (4) the article describes technology or apps intended for use outside of the clinic. We included study protocol or system design articles meeting our user criteria so that the review reflects the newest research trends. We focused on systems that used only smartphones and tablets as they are widely used, and these systems do not require purchase of specialized equipment. We excluded articles describing mobile technology to support people with illnesses other than dementia.

### Procedure

We obtained 7024 articles from our initial search: 1132 from MEDLINE, 1760 from the Web of Science, 646 from PsycINFO, 211 from CINAHL, 3255 from EMBASE, and 20 from CENTRAL. We removed 1464 duplicate papers, then both authors independently performed title and abstract screening. We reconciled discrepancies in these screening results through consensus and then independently conducted full-text screening on 68 papers. Differences in the full-text screening results were also resolved through consensus. Finally, 24 papers met our inclusion criteria. The flow diagram of the search procedure is presented in [Figure 1](#).



**Figure 1.** Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram.

### Extraction and Charting of Results

After selecting the relevant articles, we conducted thematic analysis and summarized key information [19]. One researcher (BMK) undertook initial reading and used an inductive thematic analysis procedure to identify emergent themes related to the types of support the technologies provide and coded articles by the strategies used to support people with dementia. As coding progressed, new themes were incorporated, and previously coded studies were revisited. A second researcher (LMV) validated the themes after they were finalized. Discrepancies were discussed and resolved. To characterize the needs these systems can fulfill and the dignity they might afford people with dementia, we then matched the types of technology support identified in the thematic analysis to human needs using a mapping adapted from Barker and Board [20], and applied Dewing's definition of personhood [14-16,21-23].

### Results

First, we show the year-to-year trends in publication volume and the types of technologies used for research on mobile

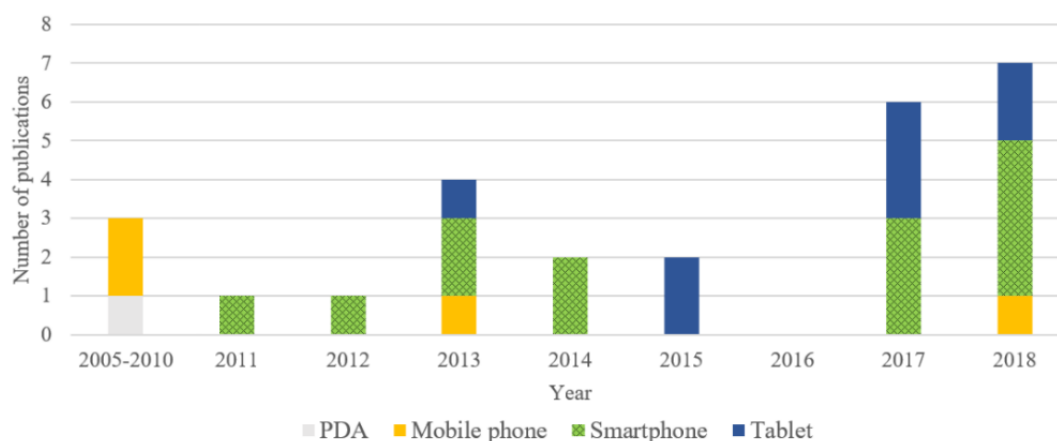
technologies to support daily living for people with dementia. We then describe the types of support offered by the technologies in each article. Finally, we map those types of support to human needs to identify research opportunities.

### Publications per Year and Types of Technology

Figure 2 presents the number of publications per year, showing the increasing interest in this topic and the evolution of the types of devices studied over time. Our search produced only 3 articles published before 2010 that used mobile technology to support daily activities of people with dementia. The number of publications grew marginally between 2011 and 2015 but increased substantially in 2017 and 2018.

Recent articles show a preference for using smartphones and tablets to assist people with dementia with daily activities, reflecting the relative popularity of these mobile devices. Of the 24 studies, 17 used smartphone (n=13) and mobile phone (n=4) technologies, about twice the number describing tablets (n=8). About 71% (n=17) of the articles focused on the use of small devices.

**Figure 2.** The number of publications per year (N=26). Publications by Lancioni et al [28] and Megges et al [38] are depicted twice as both a tablet and a smartphone were used in the study. We designated “mobile phone” when the article mentioned the use of a mobile phone without the use of apps or functions of a smartphone. PDA: personal digital assistant.



### Mobile Technology to Support Independence and Well-Being of People With Dementia

Our review produced 24 articles meeting the inclusion criteria. We categorized those articles into 6 types of support that the technologies provided to people with dementia and caregivers. A total of 9 articles described technologies that support the performance of daily activities; 2 articles discussed technologies that facilitate social interactions; 1 article defined technology to aid retrospective memory; 2 articles examined technologies

for supporting leisure activities; and 7 articles studied mobile technology for tracking location. Finally, 3 articles described technologies for health monitoring.

The next sections detail the technologies outlined in each article per category. Table 1 organizes the articles in the review by the type of support provided and whether they described developed systems or proposed systems. Table 2 provides a summary of each article organized by type of support, listing the system technologies and functions.

**Table 1.** Categorization of the articles reviewed.

Type of support	Tested systems		Proposed systems <sup>a</sup>	
	Tablet/PDA <sup>b</sup>	Smartphone/mobile phone	Tablet	Smartphone
<b>Performing daily activities (n=9)</b>				
Providing reminders	Imbeault et al [24]	El Haj et al [25], Imbeault et al [26], and Imbeault et al [27]	— <sup>c</sup>	—
Storing important information	—	Imbeault et al [26] and Imbeault et al [27]	—	—
Providing sequential instructions for daily activities	Lancioni et al [28] <sup>d</sup>	Lancioni et al [28] <sup>d</sup> and Lancioni et al [29]	Mahoney et al [30]	—
Providing sequential instruction for way finding	Chang et al [31]	Kwan et al [32]	—	—
Maintaining social interaction (n=2)	Ekström et al [33]	—	—	Fardoun et al [34]
Aiding autobiographical memory (n=1)	—	De Leo et al [35]	—	—
Engaging in leisure activities (n=2)	Tyack et al [36] and Lim et al [37]	—	—	—
Tracking location (n=7)	Megges et al [38] <sup>c</sup>	Megges et al [38] <sup>c</sup> , Faucounau et al [39], Miskelly [40], and Olsson et al [41]	—	Ko et al [42], Solanas et al [43], and Xiao et al [44]
Monitoring health (n=3)	—	Zylstra et al [45] and Kamil et al [46]	—	Lin et al [47]

<sup>a</sup>Proposed systems: Article only described a prototype or a feasibility test with healthy users.

<sup>b</sup>PDA: personal digital assistant.

<sup>c</sup>Not applicable.

<sup>d</sup>These studies tested both smartphones and tablets, so they are listed twice.

**Table 2.** Summary of the articles reviewed.

Type of support and article	Technology	System functions	Study outcome
<b>Performing daily activities</b>			
Chang et al [31]	PDA <sup>a</sup> , RFID <sup>b</sup> tag, readers, and a routing engine. the server	Provides navigation guidance when the PDA is close to RFID tags on the wall at decision points inside a building	Using the navigation prompt system increased the correct performance rates of PwD <sup>c</sup> on indoor wayfinding tasks
El Haj et al [25]	Smartphone and Google Calendar	Sends 5 alerts for each of the targeted prospective events	Providing reminders improved task performance for individual PwD
Imbeault et al [26] and [27]	Smartphone and AP@LZ app	AP@LZ supports daily activities such as recording a memo or setting reminders for scheduled activities	PwD successfully learned AP@LZ functions and used them for daily activities
Imbeault et al [24]	Tablet and calendar app	A calendar app schedules appointments and activities	PwD successfully learned to use the calendar app to add notes, appointments, or activities to calendar; set alarms; attend appointments; and log past events
Kwan et al [32]	Smartphone and a Map app	Sends verbal instructions from the Map to guide users to a destination	Study of using the smartphone app for navigation assistance demonstrated similar rates of acceptability and feasibility for groups with and without mild dementia
Lancioni et al [29]	Tablet, the Talking Alarm Clock app, and a wireless Bluetooth earpiece	Sends reminders at predetermined schedules, verbal instructions, and brief encouragement	PwD using app showed higher rates of independently starting scheduled activities and performance accuracy
Lancioni et al [28]	Smartphone or tablet, the Talking Alarm Clock app, and a wireless Bluetooth earpiece	Sends reminders at predetermined schedules, verbal instructions, and brief encouragement	PwD using app showed higher rates of independently starting and completing activities
Mahoney et al [30] <sup>d</sup>	Tablet, Microsoft Kinect, fiducial tracking system, and RFID tags	Designed to identify dressing actions and clothing items, then provide guidance using sensors	N/A <sup>e</sup>
<b>Maintaining social interaction</b>			
Ekström et al [33]	Tablet and the GoTalk NOW app	Creates a personalized communication book using GoTalk NOW and supports communication with caregiver using this app	Helped PwD and caregivers find conversation topics and helped PwD initiate communication
Fardoun et al [34] <sup>d</sup>	Smartphone, smartwatch, and cloud architecture (server and database)	Designed to perform face recognition on a picture taken with a smartwatch and to return information about the person in the picture	N/A
<b>Aiding a utobiographical memory</b>			
De Leo et al [35]	Smartphone, app, and server	Automatically takes pictures at 5-min intervals from 8 am to 8 pm, uploads pictures to a server overnight, and prepares a DVD with picture slides for delivery to the PwD	PwD demonstrated a higher correct recall rate for recent events
<b>Engaging in leisure activities</b>			
Tyack et al [36]	Tablet and the art-viewing app	Art app using tablets loaded with photos of art objects, photographs, and paintings from 3 London museums	PwD showed improved well-being score at the end of art-viewing session
Lim et al [37]	Tablet and 11 preselected apps	Uses 11 commercial leisure activity apps for creativity, simple games, and relaxation	Usable by those with little technology experience after training; half of the participants independently used apps for an average of 24 min per day

Type of support and article	Technology	System functions	Study outcome
<b>Tracking location</b>			
Faucounau et al [39]	Mobile device (GPS)	Sends regular monitoring messages to a caregiver's phone, as well as alarms when it detects activity out of a preset safety zone, long periods of inactivity, or a fall	PwD and spouse reported issues with large device size, malfunctions, and usage difficulties after testing
Ko et al [42] <sup>d</sup>	Smartphone and wandering path tracking and fall detection system	Designed to automatically take pictures and send them to a cloud system labeled with location and time	N/A
Miskelly et al [40]	Mobile phone (GPS)	Regularly sends geographical position of the mobile phone to a central server for tracking	Tracking was 90% accurate but study had a high rate of noncompliance owing to comfort and usability issues
Olsson et al [41]	Mobile phone (GPS) and the passive positioning alarm package	Sends a short message service text message with a map to the caregiver when the PwD is out of the safe zone	PwD and spouses developed trust in the alarm system over time, contributing to perception of value
Megges et al [38]	Smartphone (GPS), tablet, and the app	Sends alarm and location information to the caregiver when PwD goes out of safe zone	Good initial usability and function ratings, but usability rating decreased after 4 weeks; however, most caregivers were willing to purchase the system
Solanas et al [43] <sup>d</sup>	Smartphone (GPS) and m-Carer app	m-Carer app will send GPS data to a location server linked to a preference server with personal information	N/A
Xiao et al [44] <sup>d</sup>	Smartphone with GPS, compass, and camera with fish-eye lens	Designed to send real-time snapshots, maps, and street views to the caregiver	N/A
<b>Monitoring health</b>			
Zylstra et al [45]	Smartwatch and smartphone	Measures daily step count and maximum distance traveled from home	Trajectory of activity and GPS data provided a good estimation of functional status
Kamil et al [46]	Cell phone (text messaging)	Collects fall reports via text message	Data from text messages were more accurate than data from calendars
Lin et al [47] <sup>d</sup>	Smartphone and wandering behavior detection system: Outdoor Aider for Elders with Dementia	Designed to detect spatially disoriented behaviors by sending ongoing GPS trace to a server to analyze for pacing and lapping patterns	N/A

<sup>a</sup>PDA: personal digital assistant.

<sup>b</sup>RFID: radio-frequency identification.

<sup>c</sup>PwD: people with dementia.

<sup>d</sup>Proposed systems: article only described a prototype or a feasibility test with healthy users.

<sup>e</sup>N/A: not applicable.

### **Support for Performing Daily Activities**

People with dementia often have difficulty performing daily activities owing to loss of memory and executive function [48]. Mobile technologies have been used as an external memory aid to help people with dementia remember future activities [25,26] and to provide visual and/or audio instructions for complex sequential tasks [29]. For step-by-step instructions, a context-aware system was used to detect situations or behavior through sensors in mobile devices [30,31] or clothing [30]. The portability of mobile devices and the use of a wireless Bluetooth

earphone allowed people with dementia to be more aware of reminders and instructions in any location [28,29].

### **Provide Reminders for Upcoming Events**

Reminder systems send alarms to help people with dementia to remember future events. In a case study [25], a person with mild Alzheimer disease used Google Calendar to remind her to perform scheduled activities without any help from her caregiver. In the intervention stage, the person with dementia received 5 smartphone alerts for each of the 3 targeted events: medical appointments, community club activities, and weekly church services. The study participant successfully performed

the tasks supplemented with reminders more often than tasks without any reminders.

Imbeault et al [24] also studied the use of a commercially available calendar app on a tablet with people with dementia. The study showed that through repeated step-by-step instructions and training, a person with dementia could add notes, appointments, or activities to the calendar; set an alarm for upcoming events; and then successfully attend scheduled appointments. The participant also used the calendar app as a logbook for remembering past events, thus demonstrating the possibility of compensating for the loss of retrospective memory as well as prospective memory.

AP@LZ [26,27] is a multifunction smartphone app that provides a unique user interface specifically designed for people with dementia. This app includes reminders and other functions such as setting alarms, scheduling appointments, placing calls, and writing notes. A case study showed that participants with Alzheimer Disease could learn AP@LZ and perform daily activities with the help of reminders from AP@LZ. For example, participants used AP@LZ to set alarms to wake up, take medication, carry out activities, attend appointments, and lock the door at night.

### ***Store Important Information for Future Use***

People with dementia often experience difficulty remembering lists and important personal information, and mobile devices provide an easily accessible means of storing that information. For example, a person with dementia can keep a grocery list, maintain a medication history, or record personal information, such as name, age, address, phone number, and family contacts, for use in an emergency [26,27].

### ***Provide Sequential Instructions for Daily Activities***

The Talking Alarm Clock developed by Lancioni et al [28,29] supports people with dementia performing daily activities using reminders, step-by-step verbal instructions, encouragement, and praise. The system comprises a tablet device, the Talking Alarm Clock app, and a wireless Bluetooth earpiece linked to the tablet. Participants use the earpiece to hear alarms or instructions when they are not close to the tablet. The Talking Alarm Clock saves activity schedules with related instructions and praise statements. After an alarm and verbal reminder, specific instructions are provided to guide the person with dementia through each step. Some activities saved on the tablet include preparing coffee and breakfast, setting the table, watering plants, taking paper and soap to bathrooms, and making photocopies. Studies showed that this intervention helped people with dementia to start these activities by themselves and improved independent and accurate performance of these activities. However, the system could only provide prepared instructions and therefore could not alter instructions based on users' behaviors.

Mahoney et al [30] developed a prototype of a complex system that detects users' behaviors in context and provides accurate feedback to help people with dementia dress by themselves. The Development of Responsive Emotive Sensing System (DRESS) provided dressing guidance and track the sequence of dressing. The DRESS prototype comprised 4 parts: (1) a tablet (iPad) on top of the dresser to provides written, audio, or

visual dressing guidance, (2) a Microsoft Kinect to track movements for dressing, (3) a fiducial tracking system to ensure that clothing is worn correctly, and (4) radio-frequency identification (RFID) tags on clothes and drawers to identify the clothing items that are removed. These elements allowed DRESS to detect correct and incorrect dressing actions and provide appropriate visual and audio guidance to people with dementia.

### ***Provide Sequential Instructions for Way Finding***

People with dementia often have difficulty navigating to a destination, either indoors or outdoors. The Ubiquitous Service for Direction Guide is an indoor way finding system [31] comprising a Personal Digital Assistant (PDA), Radio-Frequency Identification (RFID) tags, RFID readers, and a routing engine. For the study, RFID tags were installed at every decision point, such as doorways, corners, elevators, exits, or intersections of hallways, in a building. When the PDA is placed near a tag on the wall at a decision point, the reader interacts with the tag and presents a directional sign on the screen to aid navigation.

Kwan et al [32] developed an outdoor way finding system to help people with dementia with navigation. The system used GPS and a map app (Maps) installed on a smartphone to provide visual and verbal instructions to direct the person with dementia to a destination. To use the system, participants pressed the home button and initiated a voice command app to activate the Maps. They interacted with the Maps using earphones and a microphone. The study demonstrated that people with dementia could successfully navigate with assistance from a smartphone.

### ***Support for Maintaining Social Interaction***

People with dementia can have difficulty remembering or recognizing others' names and faces, which can lead to social isolation. Mobile technology can facilitate social interaction by suggesting interesting conversational topics [33] or providing the persons with dementia with information about familiar people that they cannot remember [34].

The GoTalk NOW app [33] is a personalized communication book created by people with dementia and a trained specialist based on the person's personal interests and activities, communication difficulties, and needs. In the case study, the specialist took pictures of objects and places that are meaningful to the subject and video-recorded activities that are familiar to her. Main topics are displayed on a home page, and link to other pages with relevant materials. Compared with traditional reminiscence materials such as scrapbooks and picture albums, the communication app can provide easy, organized access to more information and multimedia functions, making it more interactive and enjoyable. This study showed that the personalized communication app helped the person with dementia and her caregivers find conversation topics and helped them initiate communication.

Fardoun et al [34] described a system that uses a smartwatch, a smartphone, and cloud support to implement a face-detection system intended to help people with dementia recognize familiar people. To use the system, a person would take a picture of a person with a smartwatch; the watch sends the picture to a

smartphone, and the phone uploads it to the cloud infrastructure. Face recognition is performed using a database and relevant information is delivered to the smartphone, which sends it to the smartwatch. The authors note their concern that people with dementia would have difficulty using the smartwatch's small form factor.

### ***Support for Aiding Autobiographical Memory***

For people with dementia who have trouble remembering recent events, researchers are using mobile technology to aid autobiographical memory by providing a summary of daily events. De Leo et al [35] tested a system that used pictures to help people with dementia remember everyday events. Participants wore a smartphone around their necks on a lanyard and the smartphone took pictures every 5 min during the day. The images were transferred to a secure server and saved as slideshows labeled with date and time, and DVDs of the slideshows were mailed to participants once a week. Despite the relatively infrequent mailing, the study showed that watching slideshows helped people with dementia remember recent events. The authors demonstrated the feasibility of using a smartphone to augment retrospective memory for people with dementia by capturing, storing, and viewing pictures.

### ***Support for Engaging in Leisure Activities***

Mobile devices can deliver leisure activities for people with dementia with the advantages of intuitive interfaces, multimedia functions, and a wide variety of apps. People with dementia may be able to engage with these technologies independently, which can also benefit caregivers.

An art-viewing app developed by Tyack et al [36] enables people to enjoy art at home. The app shows pictures of objects, paintings, and photographs from several museums, collectors, and artists. Researchers tested the app with people with dementia and adjusted the interface, color, and font size to ensure easy manipulation of this app. The art-based intervention via a tablet increased interaction between people with dementia and other people, or between people with dementia and technologies, and encouraged people with dementia and their caregivers to create new activities that they could enjoy together.

Lim et al [37] studied several commercial apps to investigate whether people with dementia were able to use a tablet independently during leisure time. The apps encouraged people with dementia to play musical instruments, draw pictures, play games, and listen to music on tablets. The study found that even participants with little computer and technology experience could use the tablet after training sessions, and half of the participants independently used the apps during leisure time. Notably, 48% of the participants with dementia spent an average of 24 min per day using the tablet without supervision, and their caregivers expressed an intention to keep using the tablets on a regular basis.

### ***Support for Tracking Location***

Wandering is one of the more worrisome behaviors that people with dementia exhibit [39]. Wandering behavior increases the burden on family caregivers to keep the person with dementia safe, which can compel them to consider institutional care

[39,41-43]. Mobile technology can help address this issue by providing detailed location information using GPS or photo and by providing alerts when movement is sensed outside a preset *digital fence*. Both of these approaches provide peace of mind and help caregivers find a lost family member as soon as possible. Among the 7 studies using mobile technologies for tracking location [38-44], 3 articles described new tracking systems that had not been tested with people with dementia [42-44].

For location tracking, systems usually require people with dementia to carry a GPS-integrated mobile device in a pocket [41,43], shoulder or waist bag [40], or waist belt [38,39] for continuous monitoring. Several studies described using a registered mobile phone to send geographical information to a server at predetermined time intervals for tracking and alerting caregivers based on predefined parameters [38-41]. GPS-integrated mobile phones can also be used to define a safety zone. If a person moves outside of that zone, an SMS message with location information and a map would be sent to the caregiver's phone [38,39,41]. Considering privacy issues with continuous monitoring of people with dementia, Solanas et al [43] described the m-Carer app. It is designed to enable private monitoring by encrypting the information it sends to servers. The m-Carer app would regularly send encrypted data to the location server if the person with dementia stayed within a predefined area. In emergencies, however, it would send unencrypted location information to the server, which was forwarded to caregivers.

Some approaches use mobile devices to take pictures when the person with dementia is outside, requiring the smartphone to be worn on the chest, carried in a front pocket, or attached at the waist [42,44]. Ko et al [42] proposed the wandering path tracking and fall detection system (PTFaD), which takes pictures and sends real-time images, GPS location, and time to a server. This information is saved in the cloud and can be downloaded when necessary. In addition, the proposed PTFaD includes a fall detection system that measures the current direction of the smartphone using triaxial accelerometers. Xiao et al [44] described the Canderoid system to monitor the movement of people with dementia using multiple built-in smartphone sensors such as a GPS, compass, and camera. The proposed system uses a fish-eye lens for a wide-angle view, a digital compass for orientation, and GPS for location. This system would also give caregivers remote access to the smartphone, so they could use lasers to guide the subject back home.

### ***Support for Monitoring Health***

People with dementia can have difficulty expressing discomfort, symptoms, and unmet needs, so others may find it difficult to detect functional decline and provide assistance. However, mobile devices can be used to monitor functional status, predict functional decline, and detect neuropsychological behavior symptoms.

Zylstra et al [45] employed smartwatches and smartphones to monitor the functional status of people with dementia using daily step counts and the maximum distance that the study participants move from home. The smartwatch transfers activity and GPS data to the smartphone daily, which then uploads that

information to a server. The study found that the trajectory of these data provides a good estimation of the subject's functional status.

To track falls, a person with dementia or caregiver can record falls on a calendar, but this method suffers from recall bias and missing data. Kamil et al [46] addressed these issues by collecting fall data via text messaging. A secure platform sends a daily text message asking if a fall incident occurred the day before and records a *Yes* or *No* response from the person with dementia or caregiver. The authors showed that this method was more accurate than calendar data.

Lin et al [47] developed the Outdoor Aider for Elders with Dementia to detect disorientation, falls, inactivity, and wandering. A smartphone with GPS sends mobility data to a server, which then analyzes that data for pacing and lapping movement patterns to identify wandering behaviors often exhibited by people with dementia. Testing the algorithm on prerecorded sample GPS data showed that it could detect spatially disoriented behaviors.

## Mobile Technology for People With Dementia Through the Lens of Personhood and Human Needs

Table 3 shows a mapping between the needs defined by Maslow [14] and Kitwood [15], adapted from Barker and Board [20], along with the relevant technology support the reviewed systems provide for each need.

We found that at least one system included in the review provided support for each need, and technologies tend to support more basic needs such as physiological and safety needs. Reminder systems and instructions satisfy physiological needs by assisting people with dementia with basic activities [28-30] and satisfy safety needs by helping people with dementia with tasks such as navigating to a destination [31,32] or taking medications [26,27]. These systems also support agency and, therefore, satisfy esteem and identity needs by helping carry out activities unassisted, and thus maintain independence [24,25].

Health monitoring meets physiological and safety needs by observing functional status [45,46] and alerting others to falls [46]. Location functions meet safety needs by allowing caregivers to remotely monitor a person's movements [38-44].

**Table 3.** Mapping between needs defined by Maslow (M) and Kitwood (K) and types of technology support provided by systems in the review.

Human needs	Definition	Articles	Type of support
Physiological (M)	Basic bodily and health needs	Lancioni et al [28], Imbeault et al [24], El Haj et al [25], Imbeault et al [26], Imbeault et al [27], Lancioni et al [29], Mahoney et al [30], Zylstra et al [45], Kamil et al [46], and Lin et al [47]	Send reminders, provide instructions, and monitor health
Safety (M)	Need for security	Lancioni et al [28], Megges et al [38], Imbeault et al [24], El Haj et al [25], Imbeault et al [26], Imbeault et al [27], Lancioni et al [29], Mahoney et al [30], Chang et al [31], Kwan et al [32], Faucounau et al [39], Miskelly [40], Olsson et al [41], Ko et al [42], Solanas et al [43], Xiao et al [44], Zylstra et al [45], Kamil et al [46], and Lin et al [47]	Send reminders, store personal information, provide instructions, track location, and monitor health
Comfort (K)	Need for a feeling of well-being	Ekström et al [33], Fardoun et al [34], Tyack et al [36], and Lim et al [37]	Aid interpersonal communication, provide information about familiar people, and provide entertainment
Attachment (K)	Need for reliable relationships	Ekström et al [33] and Fardoun et al [34]	Aid interpersonal communication and provide information about familiar people
Love and belonging (M); Inclusion (K)	Need for feeling accepted and included	Ekström et al [33] and Fardoun et al [34]	Aid interpersonal communication and provide information about familiar people
Occupation (K)	Need for rest and activity	Ekström et al [33], Fardoun et al [34], Tyack et al [36], and Lim et al [37]	Provide instructions, aid interpersonal communication, provide information about familiar people, and provide entertainment
Esteem (M); Identity (K)	Need for self-worth and autonomy	Lancioni et al [28], Imbeault et al [24], El Haj et al [25], Lancioni et al [29], Ekström et al [33], Fardoun et al [34], De Leo et al [35], Tyack et al [36], and Lim et al [37]	Send reminders, provide instructions, provide information about recent events and activities, aid interpersonal communication, and provide entertainment
Self-actualization (M); Agency (K)	Need for personal growth and freedom	Ekström et al [33], Fardoun et al [34], Tyack et al [36], and Lim et al [37]	Aid interpersonal communication, provide information about familiar people, and provide entertainment

Technologies for maintaining social relationships and aiding autobiographical memory directly support all needs except physiological and safety, by facilitating communication and providing information about people or events [33-35]. Spending time alone to pursue enjoyable and stimulating activities can meet the needs for comfort, occupation, esteem, identity, self-actualization, and agency [36,37]. Applying the concept of personhood [16], our review also showed that mobile technologies can facilitate positive person work [15,22] such as play [36,37], celebration [28,29], collaboration [33], and facilitation [28-32] to satisfy the psychological needs of people with dementia.

## Discussion

This scoping review surveyed the literature on using mobile devices to increase the independence and well-being of people with dementia, and then mapped those technologies to human needs [14,15]. This section will discuss trends that emerged across articles, and then discuss the implications raised by considering the research through the lens of personhood and human needs.

### Emerging Trends in Mobile Technology to Support Independence and Well-Being of People With Dementia

We organized the literature by how the technologies compensate for the loss of function that people with dementia experience in daily living. The 6 categories that emerged are (1) performing daily activities, (2) maintaining social interaction, (3) supporting autobiographical memory, (4) engaging in leisure activities, (5) tracking location, and (6) monitoring health status. Most mobile technology support was implemented for smartphones or mobile phones, implying that researchers believe that these devices are more suitable for supporting daily routines than PDAs or tablets.

The technologies that aid people with dementia with performing daily activities by providing instructions or reminders are the most mature among those in this review. Overall, 8 of the 9 articles in this category described studies in which people with dementia successfully used the technology [24,25,27-29,31,32]. Although these systems are not particularly complicated, they could fill an important role needed to promote independence.

We reviewed 2 articles concerning technologies that facilitate social interaction. The first article [33] discussed a case study that demonstrated the technology's ability to aid conversation between a person with dementia and caregiver dyad. The other article [34] described an untested prototype intended to prompt people with dementia with information about people they are having trouble recognizing. In the analysis of human needs, we saw that social interactions are involved in meeting most needs; thus, supporting communication has the potential to aid many aspects of a person's life. The lack of research in this area and value of social connection suggest an important opportunity for research.

A single article [35] described an aid for autobiographical memory and used technology that is now outdated. However, the study demonstrated that the idea of using pictures to augment memory is feasible. As autobiographical memory is vital for higher needs such as esteem and identity, and its loss is a

hallmark of dementia, further research in this area could be valuable.

Research shows that enjoyable activities enhance quality of life, can reduce some behavioral symptoms of dementia, and ease caregiver burden [49]. The 2 articles [36,37] that involved using mobile technologies for leisure activity also reported positive outcomes; one showed an increase in well-being scores, and the other showed an increase in independent activity, along with a reduction in caregiver burden. Additional research in this area has the potential to enhance quality of life for both people with dementia and their caregivers.

A total of 7 articles [38-44] involved tracking devices intended to improve safety and address wandering behaviors common in dementia. However, 3 articles [42-44] only discussed prototypes, and 3 [38-40] of the 4 studies suffered from significant usability issues. This type of technology also raises important ethical concerns. Although these technologies could afford meaningful benefits, researchers must address usability and ethical issues before they can achieve more widespread use.

Overall, 3 articles discussed technology for monitoring the health of people with dementia. One [47] described an untested prototype for detecting wandering behavior, and 2 [45,46] studied simple measures that nonetheless provided good estimates of functional health. The modest research in this area suggests an opportunity for improved measures and interactivity.

Finally, a valuable finding demonstrated in several articles is that, despite impairments, people with dementia can successfully and meaningfully use mobile devices on a daily basis [24-29,31-33,35-37]. This result confirms the value of continued research in this area.

We noted that researchers are studying passively and unobtrusively collected data for monitoring the health status of people with dementia. Daily activity level and life space show trends in physical status [45], and patterns of outdoor movement observed from GPS traces can suggest wandering behavior [47]. Although in early stages, these studies mirror the trend in digital phenotyping research that is demonstrating that passive data from mobile devices can provide indicators of function and enable early detection of problematic behavioral symptoms [50]. Goals of that work include lessening reliance on self-report and reducing burden on the health system, patients, and caregivers.

Our review included articles describing technology specifically for people with dementia, but these technologies could possibly generalize to support people experiencing similar functional limitations owing to other illnesses or disabilities such as stroke, mental illness, brain injury, and physical or sensory disability. Widely popular mobile devices offer portability, connectivity, apps, and sensors that can provide support similar to dedicated assistive devices but eliminate the burden of carrying a separate device at all times and reduce stigma associated with more visible assistive devices [51]. However, people with limited function have specific accessibility needs that must be considered during design, and further research is needed to test efficacy and usability in these other user populations.



## Mobile Technology Support Through the Lens of Human Needs

By examining the types of support provided by the technologies in this review in the context of the needs of people with dementia, we found that many systems supported basic safety and physiological needs but much fewer facilitated higher-level needs. This is probably a consequence of the relative ease of using technology for providing objective information, such as locations, reminders, and instructions, versus providing individual support for creativity and personal growth. Technology for supporting higher-level needs must integrate insights from medicine and psychology to allow personalization, esteem-oriented feedback, and creativity. Other reviews also observed this gap, as a review of intelligent assistive technologies noted too little focus on the needs of people with dementia [52] and a review of how technology can address unmet needs discussed how behavioral and psychological needs were largely ignored [53].

## Technology Adoption Through the Lens of Human Needs

Much literature discusses low uptake of health-related technology among older adults [21,52]. To understand human needs as a contributing factor, Thielke et al [21] applied Maslow's hierarchy of needs to show how health-related technologies may inadvertently undermine the user's independence. This can occur when a person believes that her lower-level needs are met, either as she is cognitively healthy or is unaware of her actual need, and therefore sees using a technology aimed at meeting those needs as threatening her autonomy. For example, a person who thinks he needs no help remembering directions is unlikely to adopt a way-finding technology. Some articles have labeled a dementia patient's refusal to adopt a technology as *curious* [54] or report other describing people with dementia as *lazy* or *stubborn* [55], but a person with dementia may simply be unable to perceive his actual needs. Considering this, people with dementia are probably more likely to use systems that address needs they consider most important and may reject support for needs they do not believe they have, especially if they lack insight into their impairments. However, they may adopt systems that provide that support in the course of meeting other priorities. We might imagine that the same person who rejects a way-finding app might eagerly accept a technology that satisfies his desire to engage in community arts activities, which happens to provide directions to those events.

Thielke et al [21] also emphasized the importance of considering *who* uses a technology and *whose* needs that technology meets. Technology developed for use by people with dementia may actually help caregivers most by providing them peace of mind or free time to engage in leisure activities. For example, technology that detects wandering behavior is used by people with dementia and intends to increase their safety. However, the solution alerts *others* to problematic behavior, serving to give caregivers peace of mind but doing nothing in the moment for the person with dementia's peace of mind. A person with dementia may reject this technology as she does not perceive that it provides for her own needs, though the solution does

ultimately increase safety. Although people with dementia may be persuaded to use technology to appease caregivers [55], technology meant for use by people with dementia may be more successful if they can clearly see that it provides them with direct benefits in addition to caregivers.

For examples of systems that meet lower-level needs while also facilitating higher-level needs, we can look to accessibility-related technology developed for people with other impairments. One example is Aira [56], a system that supports people who are visually impaired by providing assistance with visual tasks such as outdoor navigation and clothing selection. Such a system might meet the security needs of a person with dementia in the course of facilitating esteem and self-actualization through autonomy. For instance, a technology such as Aira could help a person with dementia independently explore a new neighborhood by supporting safety through basic navigation assistance while also supporting agency and personal growth by highlighting personalized points of interest and making customized restaurant recommendations.

## Ethics Through the Lens of Personhood and Human Needs

Although not often discussed in the literature, we wish to highlight the importance of ethical considerations when working with and developing technology for people with dementia. Personhood [16] engenders a fundamental respect for human life, and people with dementia are particularly vulnerable to experiencing loss of that respect owing to cognitive decline, communication difficulties, and reliance on others for basic activities. Technologies that help compensate for functional deficits can partially protect against loss of personhood and help meet fundamental human needs, thus aiding in preserving the vital senses of self and dignity. However, the ubiquity of these technologies means that they become intertwined with all dimensions of a person's life, thus introducing ethical concerns. Ethics should be considered early in the design process, with technology carefully developed within an ethical framework that serves the needs and protects the personhood of people with dementia. Technologies that undermine the personhood of people with dementia are those that work against their needs, value the needs of others over their needs, or value the personhood of others over their personhood. Considerations for mitigating the loss of personhood and meeting the needs of people with dementia include employing user-centered or value-sensitive design, ensuring informed consent, protecting privacy, and safeguarding data [54].

Using the concept of personhood [16], Maslow's hierarchy of human needs [14], and Kitwood's psychological needs of people with dementia [15] as organizing principles reveals that the significance of the technologies in this review goes beyond the surface functions of simply providing reminders or instructions. They can support personal dignity, fulfill vital needs, and help a person with dementia maintain independence and autonomy. This insight reinforces that mobile technologies for people with dementia should satisfy essential needs in a manner that retains and enhances personhood. To meet this goal, designers must understand the users of their technology and develop functionality that matches their aspirations and abilities [21].

A deep appreciation of user characteristics involves employing user-centered design methods [57] with a diverse sample of people with dementia and caregivers. Using these methods can improve adoption of mobile technology by people with dementia and promote positive outcomes. Above all, we must remember that people with dementia are not passive recipients of support but are people who actively desire to enrich their lives [58]. With continued research, we can design mobile technologies that more effectively meet their needs.

### Limitations

Limitations of this study pertain to the method and the technologies reviewed. First, we chose the scoping review approach as the research on mobile technologies to support people with dementia is still emergent and much current research is not yet published. As such, we included all articles that met our criteria and did not judge quality or exclude research still in progress, as advocated in Khalil et al [59]. We also chose to include only full-text English-language articles published in peer-reviewed journals or conference proceedings and may not have captured some pertinent research. These 2 factors combined account somewhat for the low number of included articles and limit our ability to assess the impact of these technologies and to make recommendations for practice. However, this method did allow us to illustrate the potential of and increased interest in using mobile technologies to support daily living and improve the independence and well-being of people with dementia and to identify opportunities for future research. We expect that the body of research will continue to evolve rapidly so that we can more systematically evaluate the efficacy and impact of these technologies and distill best practices.

Next, we focused only on standard mobile devices as they are widely available and increasingly already in use by people who develop dementia. However, this choice left out research on specialized technology, such as wearable technology or sensors installed in the environment [60], and systems that could be adapted for mobile technology in the future, such as reminiscence therapy using a laptop computer [61].

### Conclusions

This review summarizes the current research on mobile technologies that support daily activities for people with dementia. Our thematic analysis found that the current research describes the use of mobile devices to provide support for (1) performing daily activities, (2) maintaining social interaction, (3) aiding memory, (4) engaging in leisure activities, (5) tracking location, and (6) monitoring health. Further characterization of the literature using personhood and human needs for orientation finds that most work focuses on supporting safety and physiological needs and exposes an opportunity for work to support higher-order needs involving belonging, self-esteem, identity, and self-actualization. As research in this area matures, we expect continued innovations in technologies that meet higher-level needs, incorporate ethical considerations, employ user-centered methods, and are tested with large, diverse participant samples. Most importantly, we find that, beyond their mere functionality, the potential value of mobile technologies that support people with dementia lies in their ability to provide for vital human and psychological needs and reinforce personhood.

### Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A3A2074955) and the National Center for Advancing Translational Sciences Grant at the University of North Carolina at Chapel Hill (grant number UL1TR002489).

### Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strings per database.

[[DOCX File, 13 KB - mhealth\\_v7i11e15122\\_app1.docx](#)]

### References

1. World Health Organization. 2017. 10 Facts on Dementia URL: <http://www.who.int/features/factfiles/dementia/en/> [accessed 2018-12-16]
2. Nowrangi MA, Lyketsos CG, Rosenberg PB. Principles and management of neuropsychiatric symptoms in Alzheimer's dementia. *Alzheimers Res Ther* 2015 Jan 29;7(1):12 [FREE Full text] [doi: [10.1186/s13195-015-0096-3](https://doi.org/10.1186/s13195-015-0096-3)] [Medline: [27391771](https://pubmed.ncbi.nlm.nih.gov/27391771/)]
3. D'Onofrio G, Sancarlo D, Ricciardi F, Panza F, Seripa D, Cavallo F, et al. Information and communication technologies for the activities of daily living in older patients with dementia: a systematic review. *J Alzheimers Dis* 2017;57(3):927-935. [doi: [10.3233/JAD-161145](https://doi.org/10.3233/JAD-161145)] [Medline: [28304297](https://pubmed.ncbi.nlm.nih.gov/28304297/)]
4. Fänge AM, Schmidt SM, Nilsson MH, Carlsson G, Liwander A, Bergström CD, TECH@HOME Research Group. The TECH@HOME study, a technological intervention to reduce caregiver burden for informal caregivers of people with

- dementia: study protocol for a randomized controlled trial. *Trials* 2017 Feb 9;18(1):63 [FREE Full text] [doi: [10.1186/s13063-017-1796-8](https://doi.org/10.1186/s13063-017-1796-8)] [Medline: [28183323](https://pubmed.ncbi.nlm.nih.gov/28183323/)]
5. Kales HC, Gitlin LN, Stanislawski B, Marx K, Turnwald M, Watkins DC, et al. WeCareAdvisor™: The development of a caregiver-focused, web-based program to assess and manage behavioral and psychological symptoms of dementia. *Alzheimer Dis Assoc Disord* 2017;31(3):263-270 [FREE Full text] [doi: [10.1097/WAD.000000000000177](https://doi.org/10.1097/WAD.000000000000177)] [Medline: [27849639](https://pubmed.ncbi.nlm.nih.gov/27849639/)]
  6. Ploeg J, McAiney C, Duggleby W, Chambers T, Lam A, Peacock S, et al. A web-based intervention to help caregivers of older adults with dementia and multiple chronic conditions: qualitative study. *JMIR Aging* 2018 Apr 23;1(1):e2 [FREE Full text] [doi: [10.2196/aging.8475](https://doi.org/10.2196/aging.8475)] [Medline: [31518230](https://pubmed.ncbi.nlm.nih.gov/31518230/)]
  7. Cahill S, Begley E, Faulkner J, Hagen I. 'It gives me a sense of independence' – Findings from Ireland on the use and usefulness of assistive technology for people with dementia. *Technol Disabil* 2007;19(2-3):133-142 [FREE Full text] [doi: [10.3233/TAD-2007-192-310](https://doi.org/10.3233/TAD-2007-192-310)]
  8. Carrillo MC, Dishman E, Plowman T. Everyday technologies for Alzheimer's disease care: Research findings, directions, and challenges. *Alzheimers Dement* 2009;5(6):479-488. [doi: [10.1016/j.jalz.2009.09.003](https://doi.org/10.1016/j.jalz.2009.09.003)] [Medline: [19896587](https://pubmed.ncbi.nlm.nih.gov/19896587/)]
  9. Hakobyan L, Lumsden J, O'Sullivan D, Bartlett H. Mobile assistive technologies for the visually impaired. *Surv Ophthalmol* 2013;58(6):513-528. [doi: [10.1016/j.survophthal.2012.10.004](https://doi.org/10.1016/j.survophthal.2012.10.004)] [Medline: [24054999](https://pubmed.ncbi.nlm.nih.gov/24054999/)]
  10. Van Hoof J, Kort H, Rutten P, Duijnste M. Ageing-in-place with the use of ambient intelligence technology: perspectives of older users. *Int J Med Inform* 2011 May;80(5):310-331 [FREE Full text] [doi: [10.1016/j.ijmedinf.2011.02.010](https://doi.org/10.1016/j.ijmedinf.2011.02.010)] [Medline: [21439898](https://pubmed.ncbi.nlm.nih.gov/21439898/)]
  11. Brittain K, Corner L, Robinson L, Bond J. Ageing in place and technologies of place: the lived experience of people with dementia in changing social, physical and technological environments. *Sociol Health Illn* 2010 Feb 1;32(2):272-287 [FREE Full text] [doi: [10.1111/j.1467-9566.2009.01203.x](https://doi.org/10.1111/j.1467-9566.2009.01203.x)] [Medline: [20003041](https://pubmed.ncbi.nlm.nih.gov/20003041/)]
  12. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. *JMIR Mhealth Uhealth* 2016 Jun 9;4(2):e29 [FREE Full text] [doi: [10.2196/mhealth.5127](https://doi.org/10.2196/mhealth.5127)] [Medline: [27282195](https://pubmed.ncbi.nlm.nih.gov/27282195/)]
  13. Armstrong N, Nugent C, Moore G, Finlay D. Using smartphones to address the needs of persons with Alzheimer's disease. *Ann Telecommun* 2010;65(9-10):485-495. [doi: [10.1007/s12243-010-0165-3](https://doi.org/10.1007/s12243-010-0165-3)]
  14. Maslow AH. *Motivation And Personality*. Third Edition. London: Pearson Education; 1970.
  15. Kitwood T. The experience of dementia. *Aging & Mental Health* 1997;1(1):13-22. [doi: [10.1080/13607869757344](https://doi.org/10.1080/13607869757344)]
  16. Dewing J. Personhood and dementia: revisiting Tom Kitwood's ideas. *Int J Older People Nurs* 2008 Mar;3(1):3-13 [FREE Full text] [doi: [10.1111/j.1748-3743.2007.00103.x](https://doi.org/10.1111/j.1748-3743.2007.00103.x)] [Medline: [20925884](https://pubmed.ncbi.nlm.nih.gov/20925884/)]
  17. Post SG. Respectare: moral respect for the lives of the deeply forgetful. In: Hughes JC, Louw SJ, Sabat SR, editors. *Dementia: Mind, Meaning, and the Person (International Perspectives in Philosophy and Psychiatry)*. Oxford, England: Oxford University Press; 2006:223.
  18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Method* 2005;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
  19. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc* 2015 Sep;13(3):141-146. [doi: [10.1097/XEB.000000000000050](https://doi.org/10.1097/XEB.000000000000050)] [Medline: [26134548](https://pubmed.ncbi.nlm.nih.gov/26134548/)]
  20. Barker S, Board M. *Dementia Care in Nursing*. London: SAGE Publications; 2012.
  21. Thielke S, Harniss M, Thompson H, Patel S, Demiris G, Johnson K. Maslow's hierarchy of human needs and the adoption of health-related technologies for older adults. *Ageing Int* 2012;37(4):470-488. [doi: [10.1007/s12126-011-9121-4](https://doi.org/10.1007/s12126-011-9121-4)]
  22. Mitchell G, Agnelli J. Person-centred care for people with dementia: Kitwood reconsidered. *Nurs Stand* 2015 Oct 14;30(7):46-50. [doi: [10.7748/ns.30.7.46.s47](https://doi.org/10.7748/ns.30.7.46.s47)] [Medline: [26463810](https://pubmed.ncbi.nlm.nih.gov/26463810/)]
  23. Kaufmann EG, Engel SA. Dementia and well-being: A conceptual framework based on Tom Kitwood's model of needs. *Dementia (London)* 2016 Jul;15(4):774-788. [doi: [10.1177/1471301214539690](https://doi.org/10.1177/1471301214539690)] [Medline: [24948470](https://pubmed.ncbi.nlm.nih.gov/24948470/)]
  24. Imbeault H, Langlois F, Bocti C, Gagnon L, Bier N. Can people with Alzheimer's disease improve their day-to-day functioning with a tablet computer? *Neuropsychol Rehabil* 2018 Jul;28(5):779-796. [doi: [10.1080/09602011.2015.1133431](https://doi.org/10.1080/09602011.2015.1133431)] [Medline: [28278593](https://pubmed.ncbi.nlm.nih.gov/28278593/)]
  25. El Haj M, Gallouj K, Antoine P. Google calendar enhances prospective memory in Alzheimer's disease: a case report. *J Alzheimers Dis* 2017;57(1):285-291. [doi: [10.3233/JAD-161283](https://doi.org/10.3233/JAD-161283)] [Medline: [28222535](https://pubmed.ncbi.nlm.nih.gov/28222535/)]
  26. Imbeault H, Bier N, Pigot H, Gagnon L, Marcotte N, Fulop T, et al. Electronic organiser and Alzheimer's disease: fact or fiction? *Neuropsychol Rehabil* 2014;24(1):71-100. [doi: [10.1080/09602011.2013.858641](https://doi.org/10.1080/09602011.2013.858641)] [Medline: [24359438](https://pubmed.ncbi.nlm.nih.gov/24359438/)]
  27. Imbeault H, Gagnon L, Pigot H, Giroux S, Marcotte N, Cribier-Delande P, et al. Impact of AP@LZ in the daily life of three persons with Alzheimer's disease: long-term use and further exploration of its effectiveness. *Neuropsychol Rehabil* 2018 Jul;28(5):755-778. [doi: [10.1080/09602011.2016.1172491](https://doi.org/10.1080/09602011.2016.1172491)] [Medline: [27126266](https://pubmed.ncbi.nlm.nih.gov/27126266/)]
  28. Lancioni GE, Singh NN, O'Reilly MF, Sigafos J, D'Amico F, Laporta D, et al. Technology-based behavioral interventions for daily activities and supported ambulation in people with Alzheimer's disease. *Am J Alzheimers Dis Other Dement* 2018 Aug;33(5):318-326. [doi: [10.1177/1533317518775038](https://doi.org/10.1177/1533317518775038)] [Medline: [29742903](https://pubmed.ncbi.nlm.nih.gov/29742903/)]

29. Lancioni G, Singh N, O'Reilly M, Sigafos J, D'Amico F, Pinto K, et al. A technology-aided program for helping persons with Alzheimer's disease perform daily activities. *J Enabling Technol* 2017;11(3):85-91. [doi: [10.1108/jet-03-2017-0011](https://doi.org/10.1108/jet-03-2017-0011)]
30. Mahoney DF, Burseson W, Lozano C, Ravishankar V, Mahoney EL. Prototype Development of a Responsive Emotive Sensing System (DRESS) to aid older persons with dementia to dress independently. *Gerontechnology* 2015;13(3):345-358 [FREE Full text] [doi: [10.4017/gt.2015.13.3.005.00](https://doi.org/10.4017/gt.2015.13.3.005.00)] [Medline: [26321895](https://pubmed.ncbi.nlm.nih.gov/26321895/)]
31. Chang Y, Peng S, Wang T, Chen S, Chen Y, Chen H. Autonomous indoor wayfinding for individuals with cognitive impairments. *J Neuroeng Rehabil* 2010 Sep 14;7:45 [FREE Full text] [doi: [10.1186/1743-0003-7-45](https://doi.org/10.1186/1743-0003-7-45)] [Medline: [20840786](https://pubmed.ncbi.nlm.nih.gov/20840786/)]
32. Kwan RY, Cheung DS, Kor PP. The use of smartphones for wayfinding by people with mild dementia. *Dementia (London)* 2018 Jan 1;1471301218785461. [doi: [10.1177/1471301218785461](https://doi.org/10.1177/1471301218785461)] [Medline: [29973063](https://pubmed.ncbi.nlm.nih.gov/29973063/)]
33. Ekström A, Ferm U, Samuelsson C. Digital communication support and Alzheimer's disease. *Dementia (London)* 2017 Aug;16(6):711-731. [doi: [10.1177/1471301215615456](https://doi.org/10.1177/1471301215615456)] [Medline: [26643684](https://pubmed.ncbi.nlm.nih.gov/26643684/)]
34. Fardoun HM, Mashat AA, Castillo JR. Recognition of familiar people with a mobile cloud architecture for Alzheimer patients. *Disabil Rehabil* 2017 Feb;39(4):398-402. [doi: [10.3109/09638288.2015.1025992](https://doi.org/10.3109/09638288.2015.1025992)] [Medline: [25784202](https://pubmed.ncbi.nlm.nih.gov/25784202/)]
35. De Leo G, Brivio E, Sautter SW. Supporting autobiographical memory in patients with Alzheimer's disease using smart phones. *Appl Neuropsychol* 2011 Jan;18(1):69-76. [doi: [10.1080/09084282.2011.545730](https://doi.org/10.1080/09084282.2011.545730)] [Medline: [21390903](https://pubmed.ncbi.nlm.nih.gov/21390903/)]
36. Tyack C, Camic PM, Heron MJ, Hulbert S. Viewing art on a tablet computer: a well-being intervention for people with dementia and their caregivers. *J Appl Gerontol* 2017 Jul;36(7):864-894. [doi: [10.1177/0733464815617287](https://doi.org/10.1177/0733464815617287)] [Medline: [26675353](https://pubmed.ncbi.nlm.nih.gov/26675353/)]
37. Lim FS, Wallace T, Luszcz MA, Reynolds KJ. Usability of tablet computers by people with early-stage dementia. *Gerontology* 2013;59(2):174-182. [doi: [10.1159/000343986](https://doi.org/10.1159/000343986)] [Medline: [23257664](https://pubmed.ncbi.nlm.nih.gov/23257664/)]
38. Megges H, Freiesleben SD, Jankowski N, Haas B, Peters O. Technology for home dementia care: a prototype locating system put to the test. *Alzheimers Dement (N Y)* 2017 Sep;3(3):332-338. [doi: [10.1016/j.trci.2017.04.004](https://doi.org/10.1016/j.trci.2017.04.004)] [Medline: [29067340](https://pubmed.ncbi.nlm.nih.gov/29067340/)]
39. Faucounau V, Riguet M, Orvoen G, Lacombe A, Rialle V, Extra J, et al. Electronic tracking system and wandering in Alzheimer's disease: a case study. *Ann Phys Rehabil Med* 2009;52(7-8):579-587 [FREE Full text] [doi: [10.1016/j.rehab.2009.07.034](https://doi.org/10.1016/j.rehab.2009.07.034)] [Medline: [19744906](https://pubmed.ncbi.nlm.nih.gov/19744906/)]
40. Miskelly F. Electronic tracking of patients with dementia and wandering using mobile phone technology. *Age Ageing* 2005 Sep;34(5):497-499. [doi: [10.1093/ageing/afi145](https://doi.org/10.1093/ageing/afi145)] [Medline: [16107453](https://pubmed.ncbi.nlm.nih.gov/16107453/)]
41. Olsson A, Engström M, Lampic C, Skovdahl K. A passive positioning alarm used by persons with dementia and their spouses--a qualitative intervention study. *BMC Geriatr* 2013 Feb 5;13:11 [FREE Full text] [doi: [10.1186/1471-2318-13-11](https://doi.org/10.1186/1471-2318-13-11)] [Medline: [23384329](https://pubmed.ncbi.nlm.nih.gov/23384329/)]
42. Ko CY, Leu FY, Lin IT. A Wandering Path Tracking and Fall Detection System for People with Dementia. In: Proceedings of the 2014 Ninth International Conference on Broadband and Wireless Computing, Communication and Applications. 2014 Presented at: BWCCA'14; November 8-10, 2014; Guangdong, China p. 306-311. [doi: [10.1109/BWCCA.2014.127](https://doi.org/10.1109/BWCCA.2014.127)]
43. Solanas A, Martinez-Balleste A, Perez-Martinez PA, Pena AF, Ramos J. m-Carer: Privacy-aware monitoring for people with mild cognitive impairment and dementia. *IEEE J Select Areas Commun* 2013;31(9):19-27. [doi: [10.1109/jsac.2013.sup.0513002](https://doi.org/10.1109/jsac.2013.sup.0513002)]
44. Xiao B, Asghar M, Jamsa T, Pulli P. 'Canderoid': A Mobile System to Remotely Monitor Travelling Status of the Elderly With Dementia. In: Proceedings of the 2013 International Joint Conference on Awareness Science and Technology & Ubi-Media Computing. 2013 Presented at: iCAST'13 & UMEDIA'13; November 2-4, 2013; Aizu-Wakamatsu, Japan. [doi: [10.1109/icawst.2013.6765519](https://doi.org/10.1109/icawst.2013.6765519)]
45. Zylstra B, Netscher G, Jacquemot J, Schaffer M, Shen G, Bowhay AD, et al. Extended, continuous measures of functional status in community dwelling persons with Alzheimer's and related dementia: Infrastructure, performance, tradeoffs, preliminary data, and promise. *J Neurosci Methods* 2018 Apr 15;300:59-67 [FREE Full text] [doi: [10.1016/j.jneumeth.2017.08.034](https://doi.org/10.1016/j.jneumeth.2017.08.034)] [Medline: [28865985](https://pubmed.ncbi.nlm.nih.gov/28865985/)]
46. Kamil RJ, Bakar D, Ehrenburg M, Frankenthaler S, Wei EX, Anson E, et al. 'Text It' program to track falls in patients with Alzheimer's disease and dementia. *Alzheimers Dement (N Y)* 2018;4:137-140 [FREE Full text] [doi: [10.1016/j.trci.2018.03.001](https://doi.org/10.1016/j.trci.2018.03.001)] [Medline: [29955657](https://pubmed.ncbi.nlm.nih.gov/29955657/)]
47. Lin Q, Zhang D, Huang X, Ni H, Zhou X. Detecting Wandering Behavior Based on GPS Traces for Elders With Dementia. In: Proceedings of the 2012 12th International Conference on Control Automation Robotics & Vision. 2012 Presented at: ICARCV'12; December 5-7, 2012; Guangzhou, China. [doi: [10.1109/icarcv.2012.6485238](https://doi.org/10.1109/icarcv.2012.6485238)]
48. Guarino A, Favieri F, Boncompagni I, Agostini F, Cantone M, Casagrande M. Executive functions in Alzheimer disease: a systematic review. *Front Aging Neurosci* 2018;10:437 [FREE Full text] [doi: [10.3389/fnagi.2018.00437](https://doi.org/10.3389/fnagi.2018.00437)] [Medline: [30697157](https://pubmed.ncbi.nlm.nih.gov/30697157/)]
49. Gitlin LN, Winter L, Earland T, Herge E, Chernet N, Piersol CV, et al. The Tailored Activity Program to reduce behavioral symptoms in individuals with dementia: feasibility, acceptability, and replication potential. *Gerontologist* 2009 Jun;49(3):428-439 [FREE Full text] [doi: [10.1093/geront/gnp087](https://doi.org/10.1093/geront/gnp087)] [Medline: [19420314](https://pubmed.ncbi.nlm.nih.gov/19420314/)]
50. Insel TR. Digital phenotyping: technology for a new science of behavior. *J Am Med Assoc* 2017 Oct 3;318(13):1215-1216. [doi: [10.1001/jama.2017.11295](https://doi.org/10.1001/jama.2017.11295)] [Medline: [28973224](https://pubmed.ncbi.nlm.nih.gov/28973224/)]

51. Shinohara K, Wobbrock J. In the Shadow of Misperception: Assistive Technology Use and Social Interactions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI'11; May 7-12, 2011; Vancouver, BC, Canada. [doi: [10.1145/1978942.1979044](https://doi.org/10.1145/1978942.1979044)]
52. Ienca M, Fabrice J, Elger B, Caon M, Pappagallo A, Kressig RW, et al. Intelligent assistive technology for Alzheimer's disease and other dementias: a systematic review. *J Alzheimers Dis* 2017;56(4):1301-1340. [doi: [10.3233/JAD-161037](https://doi.org/10.3233/JAD-161037)] [Medline: [28222516](https://pubmed.ncbi.nlm.nih.gov/28222516/)]
53. Lauriks S, Reinersmann A, van der Roest HG, Meiland FJ, Davies RJ, Moelaert F, et al. Review of ICT-based services for identified unmet needs in people with dementia. *Ageing Res Rev* 2007 Oct;6(3):223-246. [doi: [10.1016/j.arr.2007.07.002](https://doi.org/10.1016/j.arr.2007.07.002)] [Medline: [17869590](https://pubmed.ncbi.nlm.nih.gov/17869590/)]
54. Bharucha AJ, Anand V, Forlizzi J, Dew MA, Reynolds CF, Stevens S, et al. Intelligent assistive technology applications to dementia care: current capabilities, limitations, and future challenges. *Am J Geriatr Psychiatry* 2009 Feb;17(2):88-104 [FREE Full text] [doi: [10.1097/JGP.0b013e318187dde5](https://doi.org/10.1097/JGP.0b013e318187dde5)] [Medline: [18849532](https://pubmed.ncbi.nlm.nih.gov/18849532/)]
55. Gibson G, Dickinson C, Brittain K, Robinson L. The everyday use of assistive technology by people with dementia and their family carers: a qualitative study. *BMC Geriatr* 2015 Jul 24;15:89 [FREE Full text] [doi: [10.1186/s12877-015-0091-3](https://doi.org/10.1186/s12877-015-0091-3)] [Medline: [26205957](https://pubmed.ncbi.nlm.nih.gov/26205957/)]
56. AIRA. 2019. How it works. URL: <https://aira.io/how-it-works>
57. Holtzblatt K, Beyer H. Contextual design: Design for life. San Francisco: Morgan Kaufmann; 2016.
58. Von Kutzleben M, Schmid W, Halek M, Holle B, Bartholomeyczik S. Community-dwelling persons with dementia: what do they need? What do they demand? What do they do? A systematic review on the subjective experiences of persons with dementia. *Aging Ment Health* 2012;16(3):378-390. [doi: [10.1080/13607863.2011.614594](https://doi.org/10.1080/13607863.2011.614594)] [Medline: [22250961](https://pubmed.ncbi.nlm.nih.gov/22250961/)]
59. Khalil H, Peters M, Godfrey CM, McInerney P, Soares CB, Parker D. An evidence-based approach to scoping reviews. *Worldviews Evid Based Nurs* 2016 Apr;13(2):118-123. [doi: [10.1111/wvn.12144](https://doi.org/10.1111/wvn.12144)] [Medline: [26821833](https://pubmed.ncbi.nlm.nih.gov/26821833/)]
60. Maresova P, Tomsone S, Lameski P, Madureira J, Mendes A, Zdravevski E, et al. Technological solutions for older people with Alzheimer's disease: review. *Curr Alzheimer Res* 2018;15(10):975-983 [FREE Full text] [doi: [10.2174/1567205015666180427124547](https://doi.org/10.2174/1567205015666180427124547)] [Medline: [29701154](https://pubmed.ncbi.nlm.nih.gov/29701154/)]
61. Lazar A, Thompson H, Demiris G. A systematic review of the use of technology for reminiscence therapy. *Health Educ Behav* 2014 Oct;41(1 Suppl):51S-61S [FREE Full text] [doi: [10.1177/1090198114537067](https://doi.org/10.1177/1090198114537067)] [Medline: [25274711](https://pubmed.ncbi.nlm.nih.gov/25274711/)]

## Abbreviations

**CENTRAL:** Cochrane Central Register of Controlled Trials  
**CINAHL:** Cumulative Index of Nursing and Allied Health Literature  
**DRESS:** Development of Responsive Emotive Sensing System  
**EMBASE:** Excerpta Medica dataBASE  
**MEDLINE:** Medical Literature Analysis and Retrieval System Online  
**PDA:** personal digital assistant  
**PTFaD:** wandering path tracking and fall detection system  
**PwD:** people with dementia  
**RFID:** radio-frequency identification

*Edited by G Eysenbach; submitted 21.06.19; peer-reviewed by W Zhang, A Baumel; comments to author 16.07.19; revised version received 25.07.19; accepted 24.09.19; published 11.11.19.*

*Please cite as:*

*Koo BM, Vizer LM*

*Examining Mobile Technologies to Support Older Adults With Dementia Through the Lens of Personhood and Human Needs: Scoping Review*

*JMIR Mhealth Uhealth* 2019;7(11):e15122

URL: <https://mhealth.jmir.org/2019/11/e15122>

doi: [10.2196/15122](https://doi.org/10.2196/15122)

PMID: [31710305](https://pubmed.ncbi.nlm.nih.gov/31710305/)

©Bon Mi Koo, Lisa M Vizer. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Iterative Adaptation of a Mobile Nutrition Video-Based Intervention Across Countries Using Human-Centered Design: Qualitative Study

Jasmin Isler<sup>1</sup>; N H el ene Sawadogo<sup>2</sup>, BSc; Guy Harling<sup>3,4,5,6,7</sup>, SCD; Till B arnighausen<sup>1,4,6</sup>, MD, SCD; Maya Adam<sup>8</sup>, MD; Moubassira Kagon e<sup>2</sup>, PhD; Ali Si e<sup>2</sup>, MD, PhD; Merlin Greuel<sup>1</sup>, MPH; Shannon A McMahon<sup>1,9</sup>, MHS, PhD

<sup>1</sup>Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany

<sup>2</sup>Nouna Health Research Center, Nouna, Nouna, Burkina Faso

<sup>3</sup>Institute for Global Health, University College London, London, United Kingdom

<sup>4</sup>Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States

<sup>5</sup>Harvard Center for Population & Development Studies, Harvard University, Cambridge, MA, United States

<sup>6</sup>Africa Health Research Institute, KwaZulu-Natal, South Africa

<sup>7</sup>MRC/Wits Rural Public Health & Health Transitions Research Unit (Agincourt), University of the Witwatersrand, Johannesburg, South Africa

<sup>8</sup>Stanford Center for Health Education, Stanford School of Medicine, Stanford University, Stanford, CA, United States

<sup>9</sup>Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

**Corresponding Author:**

Shannon A McMahon, MHS, PhD

Heidelberg Institute of Global Health

Heidelberg University

Im Neuenheimer Feld 130/3

Heidelberg, 69120

Germany

Phone: 49 06221 565344

Email: [mcmahon@uni-heidelberg.de](mailto:mcmahon@uni-heidelberg.de)

**Related Article:**

This is a corrected version. See correction statement: <https://mhealth.jmir.org/2020/1/e17666/>

## Abstract

**Background:** Mobile health (mHealth) video interventions are often transferred across settings. Although the outcomes of these transferred interventions are frequently published, the process of adapting such videos is less described, particularly within and across lower-income contexts. This study fills a gap in the literature by outlining experiences and priorities adapting a suite of South African maternal nutrition videos to the context of rural Burkina Faso.

**Objective:** The objective of this study was to determine the key components in adapting a suite of maternal nutrition mHealth videos across settings.

**Methods:** Guided by the principles of human-centered design, this qualitative study included 10 focus group discussions, 30 in-depth interviews, and 30 observations. We first used focus group discussions to capture insights on local nutrition and impressions of the original (South African) videos. After making rapid adjustments based on these focus group discussions, we used additional methods (focus group discussions, in-depth interviews, and observations) to identify challenges, essential video refinements, and preferences in terms of content delivery. All data were collected in French or Dioula, recorded, transcribed, and translated as necessary into French before being thematically coded by two authors.

**Results:** We propose a 3-pronged Video Adaptation Framework that places the aim of video adaptation at the center of a triangle framed by end recipients, health workers, and the environment. End recipients (here, pregnant or lactating mothers) directed us to (1) align the appearance, priorities, and practices of the video's protagonist to those of Burkinabe women; (2) be mindful of local realities whether economic, health-related, or educational; and (3) identify and routinely reiterate key points throughout videos and via reminder cards. Health workers (here, Community Health Workers and Mentor Mothers delivering the videos)

guided us to (1) improve technology training, (2) simplify language and images, and (3) increase the frequency of their engagements with end recipients. In terms of the environment, respondents guided us to localize climate, vegetation, diction, and how foods are depicted.

**Conclusions:** Design research provided valuable insights in terms of developing a framework for video adaptation across settings, which other interventionists and scholars can use to guide adaptations of similar interventions.

(*JMIR Mhealth Uhealth* 2019;7(11):e13604) doi:[10.2196/13604](https://doi.org/10.2196/13604)

## KEYWORDS

mHealth; Burkina Faso; mothers; Community Health Workers; pregnancy; diet; digital health

## Introduction

### Background

Cultural sensitivity is essential in public health programming and is particularly relevant when programs are transferred across settings [1]. As target groups of health interventions change, adaptation is required to create a program that respects cultural differences [2]. Culturally tailored health interventions have generally been more effective than uniform care [3], and tailored nutrition interventions specifically have proven more likely than nontailored interventions to result in healthier nutrition practices [4].

Within public health programs, mobile health (mHealth) is a promising, relatively recent approach to reaching hard-to-reach populations [5-7]. Technology usage has been shown to achieve positive health outcomes among groups that are historically underserved, and tailoring the technology to the context of the target population is a means to bolster personal relevance [8]. Variations in terms of the meaning that individuals across cultures derive from a given message can however occur, even in seemingly straightforward formats [9].

Culturally grounded, mHealth video interventions have shown positive effects in diverse settings, particularly among ethnic minorities in high-income settings with examples such as a drug prevention program among Hawaiian youth [10,11] and the *keepin'it REAL* substance prevention program among Mexican-American youth [12]. Although cultural adaptation of mHealth interventions has been studied among ethnic minorities in high-income countries [6,13,14], research on adaptations of mHealth interventions across countries—especially within or across low- and middle-income countries (LMIC)—is rare [15]. An exception is a short message service intervention for hypertension prevention conducted in 3 Latin American countries [16]. Another exception is the *Diagnosing hypertension—Engaging Action and Management in Getting Lower Bp in Aboriginal and LMIC* program that evaluated the implementation of an mHealth intervention in Tanzanian and Canadian communities [9]. We are unaware of literature describing the adaptation process of mHealth interventions across African countries.

In Burkina Faso, the use of mHealth video interventions is scarce, with limited literature on their format or development. The *Strengthening Partnerships, Results, and Innovations in Nutrition Globally* program was an exception that involved a video-based intervention that targeted maternal behaviors [17]. The intervention was transplanted from neighboring Niger to

Burkina Faso and contained locally produced videos shown monthly during women's group meetings [17]. The intervention significantly improved hygiene behavior in the Burkinabe study population, although not maternal or infant nutrition outcomes. In contrast to the original intervention development process, which was outlined in the Niger case [18], the tailoring process undertaken in Burkina Faso has not been published.

### Objectives

The objective of this formative research was to gather insights on the adaptation process when transferring a South African suite of mHealth videos [19] to the context of Burkina Faso. The South African videos, which are delivered via tablets, focus on improving nutrition during pregnancy and breastfeeding. The enclosed research involved first gathering Burkinabe perspectives on the South African videos, then rapidly adapting the videos and eliciting feedback on the adaptations. Although a primary aim was to adapt videos in preparation for a larger, randomized controlled trial that would examine intervention effects, we noted a dearth of literature outlining how mHealth interventions could be adapted for new contexts. The subaims of this formative research were thus 2-fold and entailed gathering applied insights regarding how to change videos and teasing out surface and deep structure changes that proved essential while creating a culturally grounded product. The process ultimately led to the development of a Video Adaptation Framework.

## Methods

### Study Setting and Population

Burkina Faso is among the world's least developed countries [20]. The total fertility rate is high, at 6 children per woman [21], and maternal and infant mortality are among the highest in the world [20,21]. Anemia is common among Burkinabe women of childbearing age, and although iron supplementation rates in pregnancy are over 90% [21], an estimated 58% of pregnant women are anemic [21]. Female nutritional status, the most important contributor to anemia globally [22], is particularly poor in rural areas [21].

We conducted this research in the market town of Nouna (population approximately 30,000) and surrounding villages in northwest Burkina Faso. Nouna lies within the Boucle du Mouhoun region, where 13% of the women exhibit a low body mass index, which is slightly lower than the national average, and 69 per 1000 children die within the first year of life, which is slightly higher than the national average [21]. The study area

is a health and demographic surveillance system (HDSS) site that includes a total of approximately 100,000 residents [23]. Although French is the official language in Burkina Faso, Dioula is the most commonly spoken language in the study area; however, many households primarily speak Bwaba, Dafi, or Mooré [24].

### The Intervention

The initial suite of videos was developed by researchers working with the Philani Maternal Child Health and Nutrition Trust in South Africa [25] and entailed messages related to child development, family planning, and healthy eating. The suite, which is available on the Web [26], was designed to be presented by Community Health Workers (CHWs) during

in-home health counseling sessions with mothers. As our formative study in Burkina Faso focused on maternal nutrition, we adapted a relevant subset of the South African videos that covered food groups, special nutrients, and tips for eating well on a budget (Table 1). We worked with CHWs and Mentor Mothers (MM) who visited mothers at home and informed them about maternal nutrition during pregnancy and breastfeeding by showing them the videos via tablets. CHWs are the official linkage point between the community and the health sector and engage in prevention efforts on hygiene and nutrition. MMs are elderly women who accompany pregnant women to their antenatal care appointments and sometimes assist during delivery. At present, although home visits are standard of care for CHWs and MMs, the presentation of videos is not.

**Table 1.** Concept and content of South African videos related to maternal health.

Video	Subject	Content
1	Introduction	Introduction to the protagonist (Farida) who expects her first child; importance of maternal nutrition for the child
2	Building foods	Introduction to food groups: building foods (proteins), energy foods (carbohydrates), and protection foods (fruit and vegetables); tasks of and examples for building foods (proteins)
3	Energy foods	Tasks of and examples for energy foods (carbohydrates); distinction between refined and unrefined carbohydrates
4	Protection foods	Tasks of and examples for protection foods (fruit and vegetables); encouragement to eat a variety of foods
5	Special nutrients	Tasks of and examples for iron, calcium, and vitamin A; obesity
6	Eating well on a budget	Planning ahead, using plants as protein sources, buying seasonal products, and not wasting food

### Theoretical Underpinnings of This Study

We conducted a qualitative study informed by the principles of human-centered design (HCD). These principles include empathy with the customer, product iteration, rapid refinement, and an openness to failing fast [27]. HCD directs researchers to capture details of a target audience's needs and context and to maintain a focus on those needs throughout product development. In HCD evaluations, group discussions and individual interviews are often used to explore the desires and needs of people in the target audience [28], and observations of product use allow for deeper insights into barriers and solutions that people are unaware of, hesitant to reveal, or consider implicit and thus not worthy of mention. In the field of nutrition, HCD has previously been used to design multimedia interventions to address weight gain and obesity in adolescence [29] and to design a mobile phone app to monitor vegetable consumption [30].

### Study Design and Sampling

We collected the data in 2 phases (Table 2). In the first phase, we conducted separate focus group discussions (FGD) with mothers, MMs, and CHWs. The aim of these FGDs was to learn how respondents perceived the original videos and to gather feedback on how to adapt them. In the second phase, we conducted in-depth interviews (IDIs), observations, and FGDs with several respondent groups. The aim of the IDIs was to examine how respondents perceived the adapted videos and what they thought about different delivery options. The goal of

the observations was to determine whether and what practical or logistical factors merit consideration. The aim of the second FGDs with MMs and CHWs was to learn about their experiences using the videos.

Sampling reflected the rural to urban population distribution within the Nouna HDSS, that is, 2:1. We thus conducted 4 FGDs and 20 IDIs with pregnant or lactating women in rural areas and 2 FGDs and 10 IDIs in urban areas. In line with qualitative research [31], sampling was purposive and focused on those with intimate understanding of maternal health issues, whether as pregnant or lactating mothers or health workers (male and female) working with mothers. Inclusion criteria included the following: being aged 18 years or older, being able to discuss complex topics in Dioula or French and being willing and able to give informed consent (Table 3). We embarked on data collection with an estimated sample size guided by Morse [32]. Data collection concluded near the estimated sample size as we reached saturation. During the first FGD with CHWs, 1 CHW joined belatedly. A mother ended an IDI prematurely, citing a need to attend to household duties.

The study team selected participants by drawing from catchment populations in 2 urban health centers in Nouna town and 4 rural health centers in Nouna region (Table 3). For IDIs and FGDs, participants were selected with assistance from health center staff who invited eligible participants beforehand. Data were collected in a private place within health facilities. For observations, data collectors joined MMs and CHWs as they entered villages and presented videos in women's households.



**Table 2.** Respondent groups by data collection method.

Data collection method and respondents	Number of data collection activities	Number of respondents
<b>Phase 1: FGD<sup>a</sup></b>		
Mothers	6	48
MMs <sup>b</sup>	1	8
CHWs <sup>c</sup>	1	8
<b>Phase 2</b>		
<b>FGD</b>		
MMs	1	8
CHWs	1	8
<b>In-depth interviews</b>		
Mothers	30	30
<b>Observations</b>		
Encounters mothers—MM	15	15
Encounters mothers—CHW	15	15
Total <sup>d</sup>	70	140

<sup>a</sup>FGD: focus group discussion.

<sup>b</sup>MM: Mentor Mother.

<sup>c</sup>CHW: Community Health Worker.

<sup>d</sup>Not applicable.

**Table 3.** Inclusion criteria (exclusion criteria are the inverse of the inclusion criteria).

Inclusion criteria	Mothers	CHWs <sup>a</sup>	MMs <sup>b</sup>
Age (years)	≥18	≥18	≥18
Language skills	Able to discuss complex topics in Dioula or French	Able to discuss complex topics in Dioula or French	Able to discuss complex topics in Dioula or French
Informed consent	Willing and able to give informed consent	Willing and able to give informed consent	Willing and able to give informed consent
Gender	Female	Male or female	Female
Specific criteria	Pregnant or breastfeeding	Currently employed as a CHW in the Nouna HDSS <sup>c</sup>	Currently working as a MM in the Nouna HDSS
Place of residency	Primarily a resident within the Nouna HDSS	No criteria	No criteria

<sup>a</sup>CHW: Community Health Worker.

<sup>b</sup>MM: Mentor Mother.

<sup>c</sup>HDSS: health and demographic surveillance system.

## Training and Data Collection

We trained 2 teams of 2 data collectors for 3 days on maternal nutrition, nutrition during pregnancy, ethics in research, qualitative interviewing, using tablets to display videos, and recording techniques. Data collectors were multilingual (French and Dioula), from Nouna region, and with at least a high school education. All data collectors had previously engaged in health research. We piloted and refined the interview guides for FGDs and IDIs before formal data collection, which took place between April and June 2018. Written consent preceded the collection of any data. We showed mothers foods and asked

about local availability of foods and seasonal variation, followed by open questions on maternal nutrition, the videos, and constraints to behavior change (Table 4). During FGDs, we took observational notes to capture body language. FGDs and IDIs lasted 60 to 90 min. For observations, 2 members of the study team followed a CHW or MM to the household of a pregnant or breastfeeding mother and observed video delivery, focusing on the mother's reaction, her interaction with the CHW or MM, and the latter's technical proficiency (Table 4). The video delivery was an additional task besides CHWs' and MMs' regular work, and the study team trained CHWs and MMs on the videos in terms of content (maternal nutrition) and delivery

(how to use tablets). First, we conducted a group training session that included turning the tablet on, starting the videos, and swiping across videos. Second, we had a one-on-one refresher immediately before observations to repeat technical processes and clarify questions. We did not tell CHWs and MMs whether

they should pause the videos, elicit questions, or summarize the message. Instead, we sought to gather information on how they would approach this task in the natural environment. In the interest of time, all videos were shown in 1 session.

**Table 4.** Data collection method and main subjects of interview and observation guides.

Data collection method and respondents	Main subjects of interview and observation guides <sup>a</sup>
<b>Phase 1: FGDs<sup>b</sup></b>	
Mothers	The original videos and local food choices
MMs <sup>c</sup>	An ideal intervention and the original videos
CHWs <sup>d</sup>	An ideal intervention and the original videos
<b>Phase 2</b>	
<b>FGDs</b>	
MMs and CHWs	Experiences using the videos during the observations
<b>In-depth interviews</b>	
Mothers	The adapted videos and distribution of the videos
<b>Observations</b>	
Encounters mothers—MM or CHW	Interaction between mother and MM or CHW, mother's reaction, technical proficiency of MM or CHW showing the videos on a tablet, overall approach of the MM or CHW, and family buy-in

<sup>a</sup>Interview and observation guides covered a wider range of subjects. We present here the main subjects relevant for intervention adaptation.

<sup>b</sup>FGD: focus group discussion.

<sup>c</sup>MM: Mentor Mother.

<sup>d</sup>CHW: Community Health Worker.

## Analysis

Analysis began via routine debriefings [33] between research leads and the data collection team. All data were audio recorded, transcribed and directly translated into French, and checked for completeness and quality by a bilingual research assistant. Debriefing sessions formed the basis for a codebook that included principal categories and subcategories. We discussed and agreed on a final codebook after applying initial codes to a sample of transcripts. Finalized codes were then applied to all transcripts using thematic analysis and supported by NVivo (developed by QSR International) [34,35]. We chose a semantic approach, looking for themes at an explicit level. Strategies to identify themes included searching for typologies, repetition, differences, similarities, and categories.

We applied data triangulation by comparing across data sources, namely, FGDs, IDIs, and observations. We incorporated analyst triangulation by engaging 2 researchers in the analysis process. When discrepancies occurred, we discussed them with senior researchers in the study team. Initial analysis fell in line with a Social Ecological Model [36], which emphasizes various levels that influence health behavior (from individual to superstructural) [37]. Later analysis led us to a triad that reflects the context-content-process triad with actors at the center, as outlined in the seminal framework by Walt and Gilson [38] on health policy analysis. Although our intervention is not linked to health policy, the understanding that key spheres of influence (content, context, process, and actors) inform the adaptation of

an existing model guided and grounded our framework development [38].

We conducted this study with approval of the ethical committee of the Burkinabe health ministry in Nouna (N°2018-07-/CIE/CRSN) and the ethical committee of the medical faculty Heidelberg (S-140/2018).

## Results

### Overview

For this article, we focused on the main codes emerging during analysis of the qualitative data that informed the video adaptation process specifically (see [Multimedia Appendix 1](#)). The complete codebook is provided in [Multimedia Appendix 2](#). For an overview of sociodemographic factors across participant groups, see [Table 5](#).

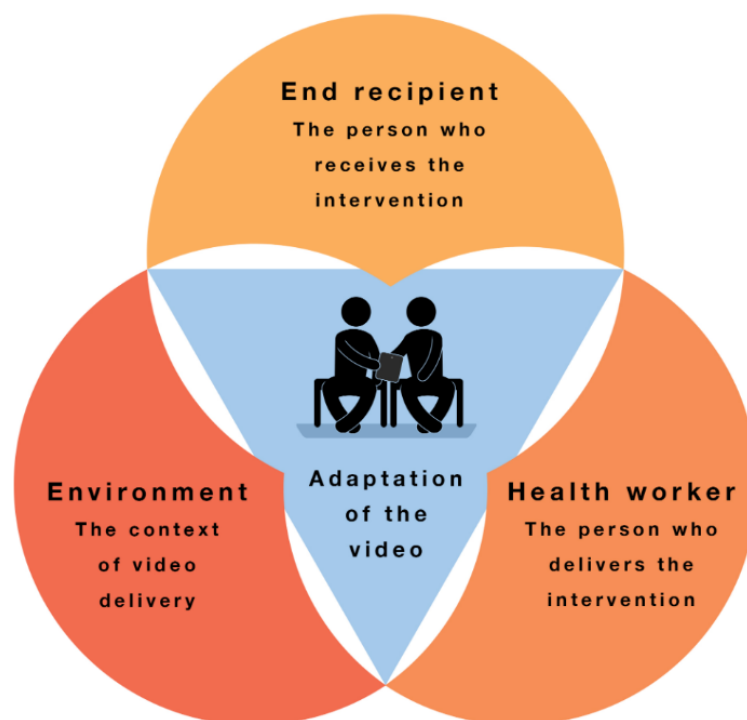
We structure our results along the main categories of our Video Adaptation Framework, which emphasizes 3 spheres of influence to mHealth video adaptation: (1) end recipients (in our case, pregnant or breastfeeding women who watch the mHealth intervention through their lens of life circumstances and experiences), (2) health workers (in our case, CHWs and MMs whose background knowledge and technological know-how informs how a video is distributed and perceived), and (3) the local environment (here referring to rural Burkina Faso where climate, local food options, and language issues proved especially pertinent; [Figure 1](#)).

**Table 5.** Sociodemographic characteristics of participant groups.

Participant group	Mothers <sup>a</sup>	Community Health Workers	Mentor mothers
Gender	Female	5 males and 3 females	Female
<b>Age (years)</b>			
Range	18 <sup>b</sup> -41	22-38	50-61
Mean	27.4	31.3	55.3
<b>Years of education</b>			
Range	0-11	5-9	0
Mean	2.5	6.3	0
<b>Number of children</b>			
Range	0 (pregnant)-9	1-5	5-9
Mean	3.2	3.4	6.9

<sup>a</sup>Sociodemographic information is only available for mothers participating in focus group discussions.

<sup>b</sup>A 15-year-old mother participated but was excluded during analysis because of her age.

**Figure 1.** The Video Adaptation Framework.

### End Recipient–Specific Considerations

Breastfeeding or pregnant mothers represent the end recipients of the mHealth video intervention. A total of 5 major themes emerged in relation to end recipient considerations amid an mHealth adaptation: appearance, priorities, and practices of the protagonist; economic status; health profile; educational background; and knowledge retention. For a comprehensive list of end recipient–related factors and how they were incorporated into this process, see [Table 6](#).

Several respondents emphasized the need to ensure that the videos' protagonist (here, a young woman named Farida) reflected the look and behavior of women in the intervention's

target group. CHWs were concerned that the skin of Farida was too light to be recognized as a woman who could live in the surrounding villages. A CHW (aged 34 years) said:

*You should take out Farida...Put in an African woman! [Laughter] [Phase 1, FGD CHWs, rural]*

We thus adapted Farida and her family to the local context of Nouna to allow for better identification with Farida and the ideal nutrition practices that she represents ([Figures 2 and 3](#)). The video designer used pictures and Google images of *Burkinabe* women and darkened Farida's skin, changed her facial features, and dressed her in a Burkinabe-patterned shirt and skirt. CHWs agreed that the adapted Farida was more adequate and could be a member of a nearby community.

**Table 6.** Video adaptations: end recipient specific.

End recipient-specific adaptations	Changes as enacted in a mobile health video in Nouna, Burkina Faso
<b>Appearance, priorities, and practices of the protagonist</b>	
Skin, facial features, and clothing	Ensure skin tone, facial features, and clothing reflect local identification
Similar struggles	Speak about financial constraints using examples that echo Burkinabe women's struggles
Cooking habits	Incorporate cooking techniques that Burkinabe women use (eg, cooking corn porridge)
Food presentation	Show vegetables and fruit as they are presented in local context (whole fruits, bowls of sauce, and minimal-to-no packaging)
<b>Economic status</b>	
Food cost	Reduce but retain reference to expensive foods (meat and poultry)
Financial resources	Delete references to consistent and robust income as many viewers are living in subsistence
<b>Health profile</b>	
Nutritional status	Include issues that are pertinent in the population (anemia and malnutrition) and remove messages that are not relevant (obesity and wasteful food habits)
Hygiene	Include a reminder about hand washing
<b>Educational background</b>	
Literacy	Restrict written text as many viewers cannot read and remove references that build upon an ability to read (lists and calendars)
<b>Knowledge retention</b>	
Information density	Reduce information density and focus on 1 special nutrient per video
Repetition	Repeat important points (the iron message)
Reminder	Provide cues to action (hard copies of key pictures) to end recipients as a reminder

**Figure 2.** Farida before adaptation.

**Figure 3.** Farida after adaptation.



Farida's manner of caring for her family was admired by women who expressed a desire to be able to look after their children the way Farida did. However, women also constantly expressed concerns that they would not be able to behave like Farida because they lacked financial means. Upon watching Farida prepare a dish, a mother (aged 22 years) said:

*You don't have the money to buy all the ingredients, so you manage as you can.* [Phase 1, FGD women, urban]

We thus slightly modified the storyline, emphasizing that Farida also struggles to provide her family a varied diet and struggles to prepare foods that are affordable, but that healthy food preparation nevertheless remains a priority for her. Women found this storyline more relatable and aspirational. In addition, we reduced references to meat and poultry as many families can only afford meat and poultry on celebration days. We retained them, however, when describing excellent sources of protein, which was acceptable to women because meat has a high local value. Women who watched the adapted videos appeared more engaged during video viewing and emphasized that they recognized small nutritional changes that they could make. For example, during one observation, a woman excitedly

told the MM and the observer that she will now focus on growing beans and groundnuts because she understands that they hold a higher nutritional value than she had previously realized.

In addition, we sought to respect the household structure and hierarchy when presenting the videos. We had initially envisioned presenting the videos in a one-on-one format, but via observations, we learned that other family members (cowives and children) were also interested in watching the videos, and some partners wanted to take a look. This preference was accommodated during observations: first, by not repressing this natural change of delivery format and later, by explicitly allowing a mother to invite cowives and husbands to join the intervention.

As a means to further reflect the health profile of families, we adapted images of children. Burkinabe research colleagues said that the initial version of a malnourished child resembled a child of average health status in rural villages. We thus adapted the image, creating a gaunter image, to better reflect the audio text regarding malnutrition (Figures 4 and 5). In addition, we stripped excess details and complex sentence structures as a means to align with the language used among end recipients.

**Figure 4.** Malnourished child before adaptation.



**Figure 5.** Malnourished child after adaptation.



### Health Worker–Specific Considerations

Health workers—CHWs and MMs—informed the video adaptation through their background knowledge and technical

proficiency. For a comprehensive list of health worker–related factors and how they were incorporated into this process, see [Table 7](#).

**Table 7.** Video adaptations: health workers–specific.

Health worker-specific adaptations	Changes as enacted in a mobile health video in Nouna, Burkina Faso
Age and educational background	Recognize that older health workers or health workers with less or no educational background will need more intense training
Technology know-how	Make technology use as simple as possible and be prepared to retrain on technology several times throughout program implementation (eg, turning the device on and off and swiping across videos)
Background knowledge	Make videos understandable without explanation

We noticed many differences between CHWs and MMs. MMs were older, aged between 50 and 61 years, whereas CHWs were aged between 22 and 38 years. None of the MMs had attended school, whereas all CHWs had attended school and could read and write. Most CHWs rapidly grasped how to use a tablet, and they handled the video delivery confidently. MMs, on the other hand, were generally hesitant to use tablets and largely relied on outside support (from mothers or the research team) to manage video delivery. As 1 MM, aged 50 years, said:

*It was the pregnant woman herself who helped me!*  
[Phase 2, FGD MMs, rural]

CHWs were eager to improve the intervention, peppering the research team with suggestions for changes. MMs were more subdued; they accepted the intervention and made no suggestions on how to modify it. While the CHWs repeated the video's key messages to mothers, most MMs contented themselves with watching alongside.

In terms of video delivery, a majority of CHWs and MMs had never used devices with touch screen or tablets. MMs especially required training and retraining on how to turn tablets on or off, swipe across screens, or open programs. MMs and CHWs alike expressed technical concerns about using tablets. A CHW told us that his fingers were shaking during his first video delivery because he was afraid to fail.

**Table 8.** Video adaptations: environment-specific.

Environment-specific adaptations	Changes as enacted in a mobile health video in Nouna, Burkina Faso
<b>Local food options</b>	
Food availability	Include only images of foods that are locally available and portray them as they are commonly seen
Source of food supply	Depict images of marketplaces rather than grocery stores
<b>Language</b>	
Language complexity	Simplify text and syntax to the extent possible and ensure illustrations convey what is being said
Speaker's accent	Ensure the voice and accent of the speaker reflects the cadence of local speakers
Cultural references	Avoid pictograms, idiomatic expressions, and abstract images (batteries or brains) that have no cultural reference
<b>Climate</b>	
Seasons	Incorporate images of seasons that reflect the local setting (Burkina Faso has 2 seasons rather than 4)
Temperature	Remove any references to climate that induce confusion (delete pictures of warm bedding or hot water bottles)

We incorporated new, healthy food suggestions such as bean leaves, okra, and hibiscus based on feedback from mothers and documents available from the Burkinabe health department. We excluded unavailable foods including peas, lentils, and pineapples and made sure foods reflected the packaging and presentation known locally. For example, in South Africa, soya

Some health workers underlined that they had no background knowledge on nutrition. They had no experience classifying foods into categories and no previous understanding that certain foods comprise a healthy diet. Consequently, health workers urged us to include all essential background nutrition knowledge in the video, rather than relying on them to amplify the video with their own insights as originally planned. We also simplified language and images, as some CHWs said that certain parts would be too difficult to explain (such as the concept of refined vs unrefined sugar).

The reactions of health workers showed that we should extend the number of engagements between end recipients and health workers, rather than condensing videos into fewer sessions. With less messages to convey, CHWs especially could have more time to pause between videos to repeat important points, explain messages in better detail, and ask mothers to summarize what they understood.

### Environment-Specific Considerations

In this case, the environment of rural Burkina Faso and its climate, language, and food options informed the video adaptation. For a comprehensive list of environment-specific factors and how they were incorporated into this process, see [Table 8](#).

packages are commonly used, whereas in Burkina Faso, this product is unknown, and soya refers to the widely used soya beans ([Figures 6 and 7](#)). In contrast to the mothers who saw the original videos, mothers who saw the adapted version were excited to recognize local foods: they repeated the food names and confirmatively clicked their tongues during video viewing.

Respondents said that seeing depictions of well-known and locally available foods in videos made it seem like the video's message was within reach. Especially MMs and mothers showed reactions of surprise and excitement upon learning that many foods that they were used to eating and that could be obtained locally had nutritional value (such as baobab leaves and beans).

Rural women highlighted that they could eat everything that they planted, whereas foods that they must buy were difficult to obtain. Urban women were more accustomed to buying food. However, in both cases, the local market is far more frequented than the rare grocery stores. Markets were, therefore, featured in the videos (Figures 8 and 9). Our illustration of the market

received positive attention because women recognized local foods and the market (designed based on photos of actual Burkinabe markets).

We worked with a local speaker to respect syntax and cadence of the local language. As Dioula is a second language for many people in our target population, some respondents had problems understanding certain foods in Dioula. Therefore, we used photos taken at the local market as a basis to make recognizable food illustrations. We also included zoom-ins to allow for quick identification of several foods.

Finally, we adapted the videos to reflect the climate of Burkina Faso (Figures 10 and 11).

**Figure 6.** Soya before adaptation.



**Figure 7.** Soya after adaptation.





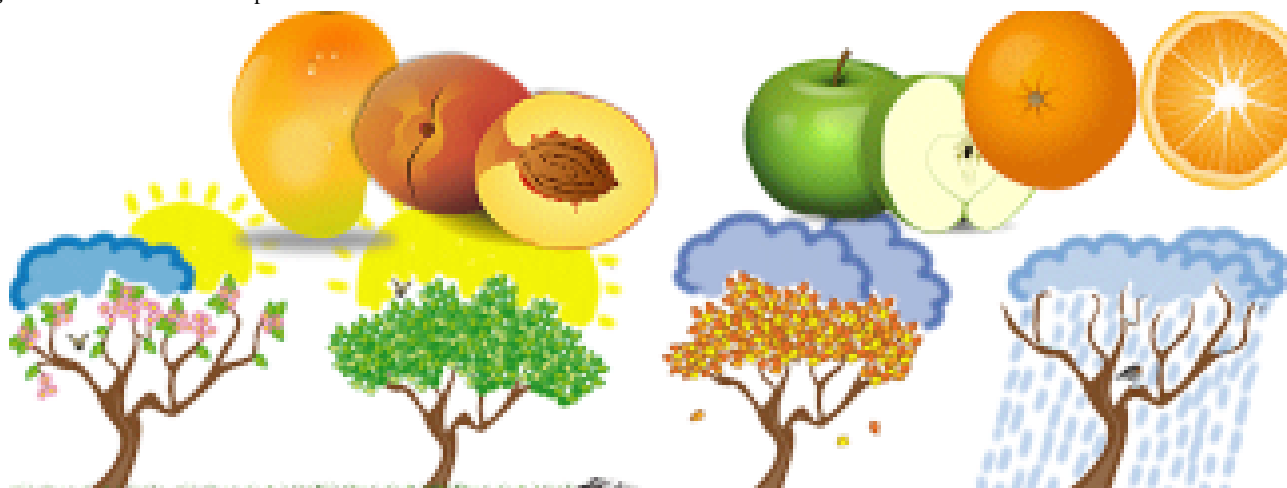
Figure 8. Market before adaptation.

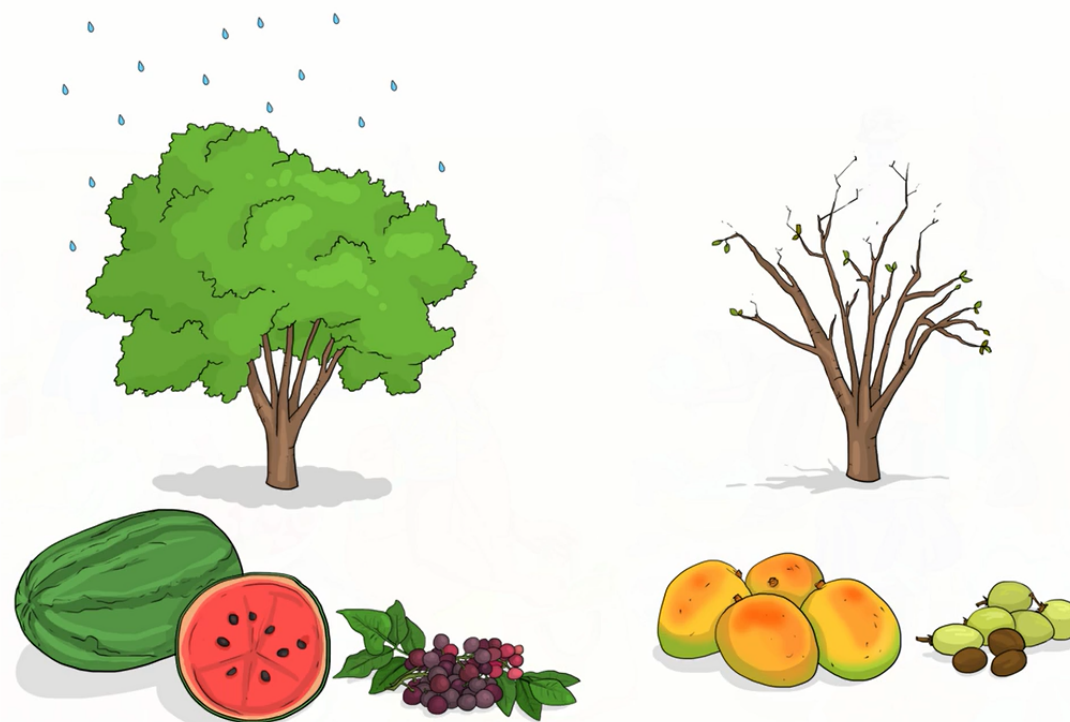


Figure 9. Market after adaptation.



Figure 10. Seasons before adaptation.



**Figure 11.** Seasons after adaptation.

## Discussion

### Principal Findings

This study used the principles of HCD to adapt maternal nutrition videos designed in urban South Africa to the context of Nouna, Burkina Faso. The HCD process led to essential modifications to address 3 priority spheres: end recipients, health workers, and environment. Respondents guided us to match the protagonist's (1) appearance, (2) actions, and (3) priorities to those of the end recipient and to adapt to their economic situation, health profile, educational background, and knowledge retention. Health workers (who differed in age, educational background, technical proficiency, and engagement) directed us to retrain them on tablets to simplify the nutrition messages and to increase their number of engagements with end recipients. In terms of the environment, locally available foods, how they are grown, packaged and sold, and the climate informed adaptation. We changed recordings and illustrations to reflect how communities, vegetation, and households look in Burkina Faso and to accommodate for cadence and syntax of the local language.

### Relation to Other Frameworks in Health and Mobile Health

Our adaptations largely reflect the priorities and insights of those who are intended to enact the videos, whether as health providers or mothers. This approach reflects both a principle of HCD and the actor-centered perspective emphasized by Walt and Gilson [38] who, in seminal research on health policy reforms, argued that an overemphasis on content masked important considerations regarding who was involved in a given reform, the process of how reforms were developed and implemented, and the context in which a reform took place. The

desires and needs of end recipients and health workers, coupled with recognition of the environment within which these actors operate (rural Burkina Faso), proved instrumental in the adaptation process and in the development of our Video Adaptation Framework.

More recent work by Maar et al [39] in 2017 reflects our findings in relation to mHealth adaptations specifically. Maar et al's [39] work on hypertension management in Canada and Tanzania emphasizes 4 human organizational levels of influence on a health outcome: patients, providers, community and organization actors, and health systems and settings. Maar posits that those 4 levels influence the outcome of mHealth interventions. Our data led us to modify the understanding of community and organization actors and narrow the definition of health system and settings, but similar to the framework of Maar et al [39], our findings highlight that patients and providers critically influence mHealth interventions, and their perspectives thus merit consideration.

### The Concept of Surface and Deep Structure of Cultural Sensitivity

Many of the changes that were feasible for us entailed surface changes such as portraying local food, climate, and communities, but ideally, deep changes would take place that adapt the videos to the family structures, economic realities, and locally informed gender roles in a given setting. Through studies with African Americans and Hispanics in the United States, Resnicow et al [40] categorized cultural sensitivity in 2 layers: the surface structure and the deep structure. Surface structure changes include changes to observable differences between cultures such as brands, foods, or locations [40]. Surface changes are needed to increase the acceptability of an intervention [40]. The adaptation of the maternal nutrition videos falls almost

exclusively into the category of surface structure adaptation. The positive reactions of respondents toward our adapted maternal nutrition videos align with Resnicow et al's [40] observation that surface structure changes increase the acceptability of an intervention.

In contrast, deep structure changes include all influences on the health behavior of the target population such as family structures, religion, politics, and economics. Deep structure changes are necessary for program impact [40]. For example, parental attitudes have a stronger influence on substance use initiation among African American youth compared with European American youth [41]. In this study, we found that the tense economic situation of our target population needed to be respected in maternal nutrition videos. The pregnant and breastfeeding women in our study underlined similarly that they were part of a family structure and their lives were embedded in social norms and gender rules that could inhibit their range of choices (despite personal preferences).

In 2010, Mier et al [42] reviewed 18 nutrition and exercise interventions tailored to Hispanics and examined the surface and deep structure features that interventions had in common. The most salient surface structure components were the usage of bilingual (and bicultural) delivery agents and materials (78%), inclusion of ethnically matched foods (33%), intervention delivery in a group setting (24%), and working with CHWs (24%) [42]. In our adaptation process, we incorporated nearly all of these surface structure components, except for the approach of delivering the intervention in a group setting, which in our case often occurred naturally. In addition, we found it crucial to provide a role model who appeared ethnically similar and to adapt the video to align with differing climates and population health profiles. Mier et al [42] do not mention or emphasize these components, possibly because the interventions in Mier et al's review took place among an ethnic minority within the same country rather than across countries.

The most salient deep structure components described by Mier et al included the following: family integration into the intervention (such as partner support; 47%); adjusting for participants' literacy levels (39%); integrating (Hispanic) cultural values such as familism, fatalism, simpatia/agreeableness, and confianza/trust (29%); and using networks and social support systems (29%) [42]. Our video was adapted along several deep structure lines. We invited cowives and men to also view the videos, although this was not our initial plan. We also adjusted videos to reflect participants' literacy levels. We did not explicitly integrate cultural values but sought to ensure that the video resonated with mothers and drew on core values, such as a sense of family and maternal pride. Moreover, we did not address networks or social support systems outside of the family in our study. Looking ahead, we aim to further refine the intervention to account for gender and to more deeply examine how men's roles in fostering child nutrition could be bolstered in our videos.

### Adaptation of Video Illustration Across Countries

In this study, we found that it was not feasible to use South African maternal nutrition videos without adapting both the voice-over and the illustrations. At least two examples from the

gray literature, however, maintain illustrations while changing language. The HealthPhone project, which provides a library of videos mainly on mother and child health, maintains the same video while changing languages globally [43]. The Scientific Animations Without Borders (SAWBO) project, which develops videos about health, agriculture, and socioeconomics [44], has an animated video about the armyworm that uses the same video but adapts the audio for several Asian and African languages [45]. SAWBO argues that the same illustrations can be used across cultures because people are willing to learn from not-local-looking characters as long as the audio's speaker has a local dialect [46]. Our differing conclusion may be because nutrition is more culturally sensitive, and more variable, than pest management. We emphasize, however, that we do not find essential guidance regarding which topics necessitate video modification and which do not.

### Study Strengths and Limitations

One of the strengths of this study is that we invested considerable time in empathizing with end recipients and health workers of the maternal nutrition videos; we considered their ideas and concerns not only on the original videos but also on the adapted versions, which gives us the assurance that the videos are attractive and relevant in this specific context. Another strength is that we observed the video viewing multiple times under real-life circumstances; we observed the abilities and struggles of health workers and could incorporate those insights into the planning of the intervention. In addition, the Video Adaptation Framework that we propose may be generalizable to other interventions that aim to adapt health communication across settings. The results of this study should, however, be seen within its limitations. First, the mothers who saw the original videos did not see the adapted videos in a pre-post sense, instead, we showed the adapted videos to different mothers. Second, we only did 2 rounds of edits instead of being able to constantly prototype as would have been ideal. Third, a disproportionate amount of insights regarding adaptations stem from FGDs with CHWs and observations of health workers and end recipients. End recipients themselves were less forthcoming about how they would like to see the videos changed. Finally, the transcripts were translated from Dioula into French and analyzed in French; thus, text could be lost or compromised in the shift from Dioula to French to English.

### Conclusions

Past research examining surface and deep structure changes in nutrition interventions has focused on outreach among ethnic minorities in high-income countries [42,47-51]. In this study, we found that there are also valuable points to consider when transferring interventions across LMIC contexts. For adaptations generally, we found that it is essential to consider the priorities and perspectives of end recipients, health workers, and the environment. End recipients guided the adaptation in terms of appearance and practices of the protagonist. Health workers' background knowledge and technical know-how further guided adaptation, particularly of video delivery. Finally, the environment guided adaptation in terms of portraying the look and feel of Burkina Faso in terms of food, climate, and language.

We hope our findings in terms of surface and deep structure changes as well as the development of a Video Adaptation Framework can serve as a guide to other interventions adapting health communication material across settings.

## Acknowledgments

The authors would like to thank Paré Marceline, Drabo Awa, and Kanazoé Ami for their diligent work during data collection. The authors would also like to thank Digital MEDIC South Africa and Shân Fischer for illustration support. The design of the Video Adaptation Framework in [Figure 1](#) was inspired by Mokitimi et al's 2018 paper on mental health policy in South Africa. The authors would like to thank Tamsyn Seimon for sharing her expertise with us during the interpretation of our findings. This study was funded by the Alexander von Humboldt Foundation. The Olympia Morata Program supports coauthor SAM.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Main codes that spoke to the adaptation process.

[\[PDF File \(Adobe PDF File\), 61 KB - mhealth\\_v7i11e13604\\_app1.pdf\]](#)

### Multimedia Appendix 2

Complete codebook.

[\[PDF File \(Adobe PDF File\), 127 KB - mhealth\\_v7i11e13604\\_app2.pdf\]](#)

## References

1. Resnicow K, Soler R, Braithwaite RL, Ahluwalia JS, Butler J. Cultural sensitivity in substance use prevention. *J Community Psychol* 2000 May;28(3):271-290. [doi: [10.1002/\(SICI\)1520-6629\(200005\)28:3<271::AID-JCOP4>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1520-6629(200005)28:3<271::AID-JCOP4>3.0.CO;2-I)]
2. Castro FG, Barrera M, Martinez CR. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prev Sci* 2004 Mar;5(1):41-45. [doi: [10.1023/b:prev.0000013980.12412.cd](https://doi.org/10.1023/b:prev.0000013980.12412.cd)] [Medline: [15058911](https://pubmed.ncbi.nlm.nih.gov/15058911/)]
3. Barrera M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: a progress report. *J Consult Clin Psychol* 2013 Apr;81(2):196-205 [FREE Full text] [doi: [10.1037/a0027085](https://doi.org/10.1037/a0027085)] [Medline: [22289132](https://pubmed.ncbi.nlm.nih.gov/22289132/)]
4. Eyles H, Mhurchu C. Does tailoring make a difference? A systematic review of the long-term effectiveness of tailored nutrition education for adults. *Nutr Rev* 2009 Aug;67(8):464-480. [doi: [10.1111/j.1753-4887.2009.00219.x](https://doi.org/10.1111/j.1753-4887.2009.00219.x)] [Medline: [19674343](https://pubmed.ncbi.nlm.nih.gov/19674343/)]
5. Mackert M, Kahlor L, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: addressing the obesity epidemic. *Telemed J E Health* 2009 Sep;15(7):672-677 [FREE Full text] [doi: [10.1089/tmj.2009.0012](https://doi.org/10.1089/tmj.2009.0012)] [Medline: [19694596](https://pubmed.ncbi.nlm.nih.gov/19694596/)]
6. Migneault JP, Dedier JJ, Wright JA, Heeren T, Campbell MK, Morisky DE, et al. A culturally adapted telecommunication system to improve physical activity, diet quality, and medication adherence among hypertensive African-Americans: a randomized controlled trial. *Ann Behav Med* 2012 Mar;43(1):62-73. [doi: [10.1007/s12160-011-9319-4](https://doi.org/10.1007/s12160-011-9319-4)] [Medline: [22246660](https://pubmed.ncbi.nlm.nih.gov/22246660/)]
7. Gustafson DH, McTavish FM, Stengle W, Ballard D, Hawkins R, Shaw BR, et al. Use and impact of eHealth system by low-income women with breast cancer. *J Health Commun* 2005;10(Suppl 1):195-218. [doi: [10.1080/10810730500263257](https://doi.org/10.1080/10810730500263257)] [Medline: [16377608](https://pubmed.ncbi.nlm.nih.gov/16377608/)]
8. Montague E, Perchonok J. Health and wellness technology use by historically underserved health consumers: systematic review. *J Med Internet Res* 2012 May 31;14(3):e78 [FREE Full text] [doi: [10.2196/jmir.2095](https://doi.org/10.2196/jmir.2095)] [Medline: [22652979](https://pubmed.ncbi.nlm.nih.gov/22652979/)]
9. Maar MA, Yeates K, Toth Z, Barron M, Boesch L, Hua-Stewart D, et al. Unpacking the black box: a formative research approach to the development of theory-driven, evidence-based, and culturally safe text messages in mobile health interventions. *JMIR Mhealth Uhealth* 2016 Jan 22;4(1):e10 [FREE Full text] [doi: [10.2196/mhealth.4994](https://doi.org/10.2196/mhealth.4994)] [Medline: [26800712](https://pubmed.ncbi.nlm.nih.gov/26800712/)]
10. Okamoto S, Helm S, McClain L, Dinson AL. The development of videos in culturally grounded drug prevention for rural native Hawaiian youth. *J Prim Prev* 2012 Dec;33(5-6):259-269 [FREE Full text] [doi: [10.1007/s10935-012-0281-0](https://doi.org/10.1007/s10935-012-0281-0)] [Medline: [23143071](https://pubmed.ncbi.nlm.nih.gov/23143071/)]
11. Okamoto S, Kulis S, Helm S, Lauricella M, Valdez J. An evaluation of the Ho'ouana Pono curriculum: a pilot study of culturally grounded substance abuse prevention for rural Hawaiian youth. *J Health Care Poor Underserved* 2016;27(2):815-833 [FREE Full text] [doi: [10.1353/hpu.2016.0061](https://doi.org/10.1353/hpu.2016.0061)] [Medline: [27180710](https://pubmed.ncbi.nlm.nih.gov/27180710/)]
12. Hecht M, Marsiglia F, Elek E, Wagstaff D, Kulis S, Dustman P, et al. Culturally grounded substance use prevention: an evaluation of the keepin' it R.E.A.L. curriculum. *Prev Sci* 2003 Dec;4(4):233-248. [doi: [10.1023/A:1026016131401](https://doi.org/10.1023/A:1026016131401)] [Medline: [14598996](https://pubmed.ncbi.nlm.nih.gov/14598996/)]

13. Shaw J, Butow P, Sze M, Young J, Goldstein D. Reducing disparity in outcomes for immigrants with cancer: a qualitative assessment of the feasibility and acceptability of a culturally targeted telephone-based supportive care intervention. *Support Care Cancer* 2013 Aug;21(8):2297-2301. [doi: [10.1007/s00520-013-1786-7](https://doi.org/10.1007/s00520-013-1786-7)] [Medline: [23519567](https://pubmed.ncbi.nlm.nih.gov/23519567/)]
14. Bender M, 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 Jul;27(4):420-428. [doi: [10.1177/1043659614565249](https://doi.org/10.1177/1043659614565249)] [Medline: [25556208](https://pubmed.ncbi.nlm.nih.gov/25556208/)]
15. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Global Problems-Global Solutions. 2010 May. Barriers and Gaps Affecting mHealth in Low and Middle Income Countries: Policy White Paper. URL: [http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth\\_Barriers\\_White\\_Paper.pdf](http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf) [accessed 2019-09-30]
16. Diez-Canseco F, Zavala-Loayza J, Beratarrechea A, Kanter R, Ramirez-Zea M, Rubinstein A, et al. Design and multi-country validation of text messages for an mHealth intervention for primary prevention of progression to hypertension in Latin America. *JMIR Mhealth Uhealth* 2015 Mar 18;3(1):e19 [FREE Full text] [doi: [10.2196/mhealth.3874](https://doi.org/10.2196/mhealth.3874)] [Medline: [25693595](https://pubmed.ncbi.nlm.nih.gov/25693595/)]
17. Spring-Nutrition. 2017 Nov. The SPRING Community Video Program in East Region, Burkina Faso: Effects on Women and Children's Nutrition and Hygiene Behaviors Analysis Report. URL: [https://www.spring-nutrition.org/sites/default/files/publications/reports/burkina\\_cv\\_report\\_eng.pdf](https://www.spring-nutrition.org/sites/default/files/publications/reports/burkina_cv_report_eng.pdf) [accessed 2019-02-01]
18. Spring-Nutrition. 2016. Seeing Is Believing The SPRING/Digital Green Experience in Niger. URL: [https://www.spring-nutrition.org/sites/default/files/publications/reports/spring\\_niger\\_seeing\\_is\\_believing.pdf](https://www.spring-nutrition.org/sites/default/files/publications/reports/spring_niger_seeing_is_believing.pdf) [accessed 2019-02-01]
19. Adam M, McMahon S, Prober C, Bärnighausen T. Human-centered design of video-based health education: an iterative, collaborative, community-based approach. *J Med Internet Res* 2019 Jan 30;21(1):e12128 [FREE Full text] [doi: [10.2196/12128](https://doi.org/10.2196/12128)] [Medline: [30698531](https://pubmed.ncbi.nlm.nih.gov/30698531/)]
20. Human Development Reports. 2018. Human Development Indicators. URL: <http://hdr.undp.org/en/countries/profiles/BFA> [accessed 2019-02-01]
21. The DHS Program. 2010. Enquête Démographique et de Santé et à Indicateurs Multiples (EDSBF-MICS IV). URL: <https://dhsprogram.com/pubs/pdf/FR256/FR256.pdf> [accessed 2019-02-01]
22. Stoltzfus RJ, Dreyfuss ML. Guidelines for the Use of Iron Supplements to Prevent and Treat Iron Deficiency Anemia. Washington, D.C: ILSI press; 1998.
23. Sié A, Louis V, Gbangou A, Müller O, Niamba L, Stieglbauer G, et al. The Health and Demographic Surveillance System (HDSS) in Nouna, Burkina Faso, 1993-2007. *Glob Health Action* 2010 Sep 14;3 [FREE Full text] [doi: [10.3402/gha.v3i0.5284](https://doi.org/10.3402/gha.v3i0.5284)] [Medline: [20847837](https://pubmed.ncbi.nlm.nih.gov/20847837/)]
24. Becher H, Kouyaté B. Health Research in Developing Countries : A Collaboration Between Germany and Bukina Faso. Berlin: Springer; 2005.
25. Rotheram-Borus M, le Roux IM, Tomlinson M, Mbewu N, Comulada W, le Roux K, et al. Philani Plus (+): a Mentor Mother community health worker home visiting program to improve maternal and infants' outcomes. *Prev Sci* 2011 Dec;12(4):372-388 [FREE Full text] [doi: [10.1007/s11121-011-0238-1](https://doi.org/10.1007/s11121-011-0238-1)] [Medline: [21850488](https://pubmed.ncbi.nlm.nih.gov/21850488/)]
26. YouTube. Health Outreach Videos Playlist. URL: <https://www.youtube.com/watch?v=g4UENBiNfSw&list=PLB0AYRZmsManRX6gXAHjWU1E6Q4B9wpe2> [accessed 2019-02-01]
27. Brown T, Wyatt J. Design thinking for social innovation. *Dev Outreach* 2010;12(1):29-43 [FREE Full text] [doi: [10.1596/1020-797X\\_12\\_1\\_29](https://doi.org/10.1596/1020-797X_12_1_29)]
28. Ideo. The Field Guide To Human-Centered Design: Design Kit. San Francisco: IDEO; 2015.
29. Knoblock-Hahn AL, LeRouge CM. A qualitative, exploratory study of predominantly female parental perceptions of consumer health technology use by their overweight and/or obese female adolescent participating in a fee-based 4-week weight-management intervention. *J Acad Nutr Diet* 2014 Apr;114(4):570-577. [doi: [10.1016/j.jand.2013.11.021](https://doi.org/10.1016/j.jand.2013.11.021)] [Medline: [24507626](https://pubmed.ncbi.nlm.nih.gov/24507626/)]
30. Mummah S, Mathur M, King A, Gardner C, Sutton S. Mobile technology for vegetable consumption: a randomized controlled pilot study in overweight adults. *JMIR Mhealth Uhealth* 2016 May 18;4(2):e51 [FREE Full text] [doi: [10.2196/mhealth.5146](https://doi.org/10.2196/mhealth.5146)] [Medline: [27193036](https://pubmed.ncbi.nlm.nih.gov/27193036/)]
31. Patton MQ. *Qualitative Evaluation And Research Methods*. Thousand Oaks, CA, US: Sage Publications Inc; 1990.
32. Morse JM. Determining sample size. *Qual Health Res* 2000;10(1):3-5. [doi: [10.1177/104973200129118183](https://doi.org/10.1177/104973200129118183)]
33. McMahon S, Winch P. Systematic debriefing after qualitative encounters: an essential analysis step in applied qualitative research. *BMJ Glob Health* 2018;3(5):e000837 [FREE Full text] [doi: [10.1136/bmjgh-2018-000837](https://doi.org/10.1136/bmjgh-2018-000837)] [Medline: [30233833](https://pubmed.ncbi.nlm.nih.gov/30233833/)]
34. Bazeley P, Jackson K. *Qualitative data analysis with NVivo*. Thousand Oaks, CA, US: Sage Publications Inc; 2013.
35. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101 [FREE Full text]
36. Bronfenbrenner U. *The Ecology Of Human Development: Experiments By Nature And Design*. Cambridge, MA: Harvard University Press; 1979.
37. Stokols D. Translating social ecological theory into guidelines for community health promotion. *Am J Health Promot* 1996;10(4):282-298. [doi: [10.4278/0890-1171-10.4.282](https://doi.org/10.4278/0890-1171-10.4.282)] [Medline: [10159709](https://pubmed.ncbi.nlm.nih.gov/10159709/)]
38. Walt G, Gilson L. Reforming the health sector in developing countries: the central role of policy analysis. *Health Policy Plan* 1994 Dec;9(4):353-370. [doi: [10.1093/heapol/9.4.353](https://doi.org/10.1093/heapol/9.4.353)] [Medline: [10139469](https://pubmed.ncbi.nlm.nih.gov/10139469/)]

39. Maar MA, Yeates K, Perkins N, Boesch L, Hua-Stewart D, Liu P, et al. A framework for the study of complex mHealth interventions in diverse cultural settings. *JMIR Mhealth Uhealth* 2017 Apr 20;5(4):e47 [FREE Full text] [doi: [10.2196/mhealth.7044](https://doi.org/10.2196/mhealth.7044)] [Medline: [28428165](https://pubmed.ncbi.nlm.nih.gov/28428165/)]
40. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. *Ethn Dis* 1999;9(1):10-21. [Medline: [10355471](https://pubmed.ncbi.nlm.nih.gov/10355471/)]
41. Vega WA, Zimmerman RS, Warheit GJ, Apospori E, Gil AG. Risk factors for early adolescent drug use in four ethnic and racial groups. *Am J Public Health* 1993 Mar;83(2):185-189. [doi: [10.2105/ajph.83.2.185](https://doi.org/10.2105/ajph.83.2.185)] [Medline: [8427320](https://pubmed.ncbi.nlm.nih.gov/8427320/)]
42. Mier N, Ory M, Medina A. Anatomy of culturally sensitive interventions promoting nutrition and exercise in hispanics: a critical examination of existing literature. *Health Promot Pract* 2010 Jul;11(4):541-554 [FREE Full text] [doi: [10.1177/1524839908328991](https://doi.org/10.1177/1524839908328991)] [Medline: [19193933](https://pubmed.ncbi.nlm.nih.gov/19193933/)]
43. HealthPhone. 2018. URL: <http://www.healthphone.org/> [accessed 2019-09-30]
44. Rodriguez-Domenech M, Bello-Bravo J, Pittendrigh B. Scientific animations without borders (SAWBO): an innovative strategy for promoting education for sustainable development. *Sustain Sci* 2018;14(4):1105-1116. [doi: [10.1007/s11625-018-0624-8](https://doi.org/10.1007/s11625-018-0624-8)]
45. Bello-Bravo J, Huesing J, Boddupalli P, Goergen G, Eddy R, Tamò M, et al. IPM-based animation for fall Armyworm: a multi-institutional and virtual international collaboration using the scientific animations without borders (SAWBO) platform. *Outlook Pest Manag* 2018;29(5):225-230. [doi: [10.1564/v29\\_oct\\_10](https://doi.org/10.1564/v29_oct_10)]
46. Bello-Bravo J, Pittendrigh B. Scientific animations without borders (SAWBO): animating IPM information and education everywhere. *Outlook Pest Manag* 2018;29(2):58-61. [doi: [10.1564/v29\\_apr\\_02](https://doi.org/10.1564/v29_apr_02)]
47. Resnicow K, Jackson A, Braithwaite R, DiIorio C, Blisset D, Rahotep S, et al. Healthy Body/Healthy Spirit: a church-based nutrition and physical activity intervention. *Health Educ Res* 2002 Oct;17(5):562-573. [doi: [10.1093/her/17.5.562](https://doi.org/10.1093/her/17.5.562)] [Medline: [12408201](https://pubmed.ncbi.nlm.nih.gov/12408201/)]
48. Nierkens V, Hartman M, Nicolaou M, Vissenberg C, Beune E, Hosper K, et al. Effectiveness of cultural adaptations of interventions aimed at smoking cessation, diet, and/or physical activity in ethnic minorities: a systematic review. *PLoS One* 2013;8(10):e73373 [FREE Full text] [doi: [10.1371/journal.pone.0073373](https://doi.org/10.1371/journal.pone.0073373)] [Medline: [24116000](https://pubmed.ncbi.nlm.nih.gov/24116000/)]
49. Lancaster K, Carter-Edwards L, Grilo S, Shen C, Schoenthaler A. Obesity interventions in African American faith-based organizations: a systematic review. *Obes Rev* 2014;15(Suppl 4):159-176. [doi: [10.1111/obr.12207](https://doi.org/10.1111/obr.12207)] [Medline: [25196412](https://pubmed.ncbi.nlm.nih.gov/25196412/)]
50. Sumlin L, Garcia A. Effects of food-related interventions for African American women with type 2 diabetes. *Diabetes Educ* 2012;38(2):236-249. [doi: [10.1177/0145721711422412](https://doi.org/10.1177/0145721711422412)] [Medline: [22454408](https://pubmed.ncbi.nlm.nih.gov/22454408/)]
51. Bender M, Clark M. Cultural adaptation for ethnic diversity: a review of obesity interventions for preschool children. *Calif J Health Promot* 2011 Sep;9(2):40 [FREE Full text] [Medline: [24159268](https://pubmed.ncbi.nlm.nih.gov/24159268/)]

## Abbreviations

**CHW:** Community Health Worker  
**FGD:** focus group discussion  
**HCD:** human-centered design  
**HDSS:** health and demographic surveillance system  
**IDI:** in-depth interview  
**LMIC:** low- and middle-income countries  
**mHealth:** mobile health  
**MM:** Mentor Mother  
**SAWBO:** Scientific Animations Without Borders

*Edited by G Eysenbach; submitted 02.02.19; peer-reviewed by M Maar, L Brandt, E Da Silva; comments to author 17.05.19; revised version received 29.06.19; accepted 27.07.19; published 11.11.19.*

### *Please cite as:*

Isler J, Sawadogo NH, Harling G, Bärnighausen T, Adam M, Kagoné M, Sié A, Greuel M, McMahan SA  
*Iterative Adaptation of a Mobile Nutrition Video-Based Intervention Across Countries Using Human-Centered Design: Qualitative Study*  
*JMIR Mhealth Uhealth* 2019;7(11):e13604  
URL: <http://mhealth.jmir.org/2019/11/e13604/>  
doi: [10.2196/13604](https://doi.org/10.2196/13604)  
PMID: [31710302](https://pubmed.ncbi.nlm.nih.gov/31710302/)

©Jasmin Isler, N H el ene Sawadogo, Guy Harling, Till B arnighausen, Maya Adam, Moubassira Kagon e, Ali Si e, Merlin Greuel, Shannon A McMahon. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Exploring the Patterns of Use and Acceptability of Mobile Phones Among People Living With HIV to Improve Care and Treatment: Cross-Sectional Study in Three Francophone West African Countries

Phillipe Lepère<sup>1</sup>, MPH; Yélamikan Touré<sup>2</sup>, MD; Alexandra M Bitty-Anderson<sup>2</sup>, MPH; Simon P Boni<sup>2</sup>, MD; Gildas Anago<sup>2</sup>, MSc; Boris Tchounga<sup>2</sup>, MPH, PhD, MD; Pendadiago Touré<sup>3</sup>, BA; Albert Minga<sup>4</sup>, MD, PhD; Eugène Messou<sup>5</sup>, MD, PhD; Guillaume Kanga<sup>6</sup>, MD; Serge Koule<sup>7</sup>, MD; Armel Poda<sup>8</sup>, PhD; Alexandra Calmy<sup>9</sup>, MD, PhD; Didier K Ekouevi<sup>2,10,11,12</sup>, MPH, MD, PhD; Patrick A Coffie<sup>2,13,14</sup>, MPH, MD, PhD

<sup>1</sup>Institut de Santé Globale, Université de Genève, Suisse, Genève, Switzerland

<sup>2</sup>Programme PACCI, Abidjan, Cote D'Ivoire

<sup>3</sup>Centre Solidarité Actions Sociales, Bouaké, Cote D'Ivoire

<sup>4</sup>Centre Médical de Suivi des donneurs de sang, Abidjan, Cote D'Ivoire

<sup>5</sup>Entre de Prise en charge et de Formation à Yopougon-Attié, Abidjan, Cote D'Ivoire

<sup>6</sup>Centre Intégré de Recherches Biocliniques d'Abidjan, Abidjan, Cote D'Ivoire

<sup>7</sup>Unité de soins ambulatoires et de conseils, Abidjan, Cote D'Ivoire

<sup>8</sup>Hôpital de Jour, Service des Maladies Infectieuses et Tropicales, CHU Souro Sanou, Bobo Dioulasso, Burkina Faso

<sup>9</sup>Unité Virus de L'immunodéficience Humaine, Service des maladies infectieuses, Hôpital Universitaire de Genève, Genève, Switzerland

<sup>10</sup>Centre Africain de Recherche en Epidémiologie et en Santé Publique, Lomé, Togo

<sup>11</sup>Département Santé Publique, Université de Lomé, Lomé, Togo

<sup>12</sup>Centre Inserm - 1219, Université de Bordeaux, Bordeaux, France

<sup>13</sup>Département de Dermatologie et d'Infectiologie, Unités de Formation et de Recherche des Sciences Médicales, Université Félix Houphouët Boigny, Abidjan, Cote D'Ivoire

<sup>14</sup>Service des Maladies Infectieuses et Tropicales, Centre Hospitalier Universitaire de Treichville, Abidjan, Cote D'Ivoire

**Corresponding Author:**

Phillipe Lepère, MPH  
Institut de Santé Globale  
Université de Genève, Suisse  
24 rue du Général-Dufour  
1205 Genève  
Genève  
Switzerland  
Phone: 41 +33673777280  
Email: [lepere.p@wanadoo.fr](mailto:lepere.p@wanadoo.fr)

## Abstract

**Background:** The use of mobile technology in health care (mobile health [mHealth]) could be an innovative way to improve health care, especially for increasing retention in HIV care and adherence to treatment. However, there is a scarcity of studies on mHealth among people living with HIV (PLHIV) in West and Central Africa.

**Objective:** The aim of this study was to assess the acceptability of an mHealth intervention among PLHIV in three countries of West Africa.

**Methods:** A cross-sectional study among PLHIV was conducted in 2017 in three francophone West African countries: Côte d'Ivoire, Burkina Faso, and Togo. PLHIV followed in the six preselected HIV treatment and care centers, completed a standardized questionnaire on mobile phone possession, acceptability of mobile phone for HIV care and treatment, preference of mobile phone services, and phone sharing. Descriptive statistics and logistic regression were used to describe variables and assess factors associated with mHealth acceptability.



**Results:** A total of 1131 PLHIV—643 from Côte d’Ivoire, 239 from Togo, and 249 from Burkina Faso—participated in the study. Median age was 44 years, and 76.1% were women (n=861). Almost all participants owned a mobile phone (n=1107, 97.9%), and 12.6% (n=140) shared phones with a third party. Acceptability of mHealth was 98.8%, with the majority indicating their preference for both phone calls and text messages. Factors associated with mHealth acceptability were having a primary school education or no education (adjusted odds ratio=7.15, 95% CI 5.05-10.12;  $P<.001$ ) and waiting over one hour before meeting a medical doctor on appointment day (adjusted odds ratio=1.84, 95% CI 1.30-2.62;  $P=.01$ ).

**Conclusions:** The use of mHealth in HIV treatment and care is highly acceptable among PLHIV and should be considered a viable tool to allow West and Central African countries to achieve the Joint United Nations Programme on HIV/AIDS 90-90-90 goals.

(*JMIR Mhealth Uhealth* 2019;7(11):e13741) doi:[10.2196/13741](https://doi.org/10.2196/13741)

## KEYWORDS

acceptability; mHealth; PLHIV; West Africa

## Introduction

In 2017, the number of people living with HIV (PLHIV) was estimated to be approximately 36.9 million people worldwide, and the number of new HIV infections was estimated to be 1.8 million [1]. In the fight against the HIV epidemic, considerable progress has been made, including an increase in the number of patients on treatment, a decrease in the number of new infections, and a reduction in morbidity and mortality [1]. The Joint United Nations Programme on HIV/AIDS (UNAIDS) Fast-Track proposes a framework for achieving a set of ambitious targets for HIV prevention (reduce the number of new infections to <500,000), treatment (reduce the number of AIDS-related deaths to <500,000), and zero discrimination by 2020. As part of this strategy, the world has embraced the 90-90-90 treatment target for the post-2015 era, setting the goal that by 2020, 90% of all PLHIV know their HIV status, 90% of all people with diagnosed HIV infection receive sustained antiretroviral therapy (ART), and 90% of all people receiving ART have durable viral suppression [2].

Today, if treatment is initiated early, adherence is maintained, and quality of care is high, the life expectancy of a person living with HIV is nearly the same as that of a person without HIV [3]. This is true in high-income countries and low-/middle-income countries where some disparities still exist between men and women [4]. As a result, HIV infection is increasingly becoming a chronic disease. Within this paradigm of HIV as a chronic disease, mobile technology can help health systems improve the quality of the HIV treatment cascade. In addition, the World Health Organization (WHO) has endorsed text messaging or short messages service (SMS) interventions for supporting individual-level adherence to ART and for improving linkage of people diagnosed with HIV to HIV-care services [5]. A WHO survey reported that 83% of the Member States used a mobile health (mHealth) initiative for promoting HIV issues [6].

Several pilot projects using mHealth for infectious diseases such as tuberculosis, malaria, and Ebola and for chronic diseases such as hypertension and diabetes have been proven to be feasible and acceptable in the African context, demonstrating the potential of mHealth in contributing to chronic patients’ treatment and care [7-12]. These technologies are also used in HIV treatment and care to improve retention in care and

adherence to treatment, which are major challenges to care programs in sub-Saharan Africa, where retention is estimated to be less than 75% at 12 months and 60% at 24 months [13,14]. A systematic review of SMS interventions found that 73% of the studies used “SMS only” interventions, and in only 27% of studies, the use of SMS was associated with other interventions [15]. In sub-Saharan Africa, most interventions using mHealth have mainly used SMS as a way to improve retention of PLHIV in care and adherence to ART. In Cameroon, two interventions among PLHIV used text messages to improve adherence: one used once-per-week motivational messages and the other used four-times-per-week educational messages [16]. Another study in South Africa used an electronic adherence monitoring device and text reminders if patients were late in taking their medications [17]. Text messages reminders for appointments were used to assess retention among PLHIV in Mozambique and were found to improve retention among urban patients and those who newly initiated ART [18].

As part of the standard of care, some interventions were used to improve retention and adherence such as phone calls to patients missing their appointments and home visits in some countries [19]. However, home visits and phone calls, despite their relative effectiveness, constitute a significant financial burden and are not sustainable in the long term [20]. mHealth interventions could decrease this financial burden, especially since mobile phone penetration in sub-Saharan Africa is high and represents an opportunity for the continent [21]. Therefore, there is a need to identify simple, acceptable, and sustainable interventions to improve retention in care and adherence to ART for PLHIV. As a systematic review of mHealth interventions in Africa indicated, successful mHealth interventions would yield positive outcomes if they are accessible, accepted, and low cost while adaptable to the local context [20].

In sub-Saharan Africa, the scientific literature on mHealth has been concentrated in East and Southern Africa, while there is a scarcity of studies in West Africa, especially francophone West Africa countries [22]. A systematic review of mHealth interventions in low-and middle-income countries identified 34 publications from Africa, only two of which were from West Africa including one from a francophone country [22]. Before suggesting mHealth interventions, it is essential to assess the acceptability and feasibility of an mHealth intervention in HIV care and treatment among PLHIV in francophone West Africa.

The aim of this study was to explore the social and financial acceptability, which has not been reported yet in Francophone Africa, and the feasibility of an mHealth intervention for the improvement of care among PLHIV in West Africa. We hypothesize that the rates of acceptability would be similar to that found in other regions of Africa.

## Methods

### Study Design and Setting

A cross-sectional study was conducted from September to December 2017 in three francophone West African countries: Côte d'Ivoire, Burkina Faso, and Togo. Six HIV clinics in Côte d'Ivoire, one HIV clinic in Burkina Faso, and one HIV clinic in Togo participated in the study. These centers are reference centers for the three countries, with a significant number of PLHIV in care.

### Participants

Inclusion criteria were age  $\geq 18$  years, receiving ART for at least 1 year, and receiving follow-up at one of the HIV clinics participating in the study. Participants were selected in each clinic by using a systematic random sampling procedure. Each day of the survey, eligibility criteria for the study were checked by a social worker. Sequential numbers were issued to eligible participants in the order in which they arrived at the HIV clinic. The first eligible participant of the day was included, and a sampling interval of five was applied to select subsequent participants. This procedure was repeated every day in each participating HIV clinic. To avoid possible duplicates, each selected participant received a marked sticker on his medical record.

### Study Size

We hypothesized that mHealth coverage would be around 80% according to the results from previous studies conducted in sub-Saharan Africa [23,24-25]. Considering a relative risk of  $\alpha=5\%$  and a precision of 5% of the value of the expected acceptability, and taking into account an assumption of 10% of missing data, based on the formula  $n=z^2 p(1-p)/\alpha^2$  (with  $z$ =confidence level;  $p$ =estimated proportion of mHealth coverage in sub-Saharan Africa;  $\alpha$ =precision level) for estimation of the sample size, the minimum sample size was estimated at 246 PLHIV on ART in each of the three participating countries for a total of 738 PLHIV [26]. This allowed comparison of the acceptability of mHealth per country.

### Data Collection

A questionnaire was elaborated by a multidisciplinary team of epidemiologists, clinicians, sociologists, social workers, and members of networks of PLHIV. After obtaining informed consent, this standardized questionnaire was administered face-to-face to each participant by trained social workers

involved in HIV care and treatment using classical pen and paper material. The questionnaire explored five aspects: sociodemographic characteristics, HIV characteristics, possession of mobile phone, phone sharing, and acceptability (of mobile health). The main outcomes for this study were possession of a phone and acceptability of mHealth (phone calls/text messages).

### Statistical Analysis

The data collected were entered into a Microsoft Access 2013 database (Microsoft Corporation, Redmond, WA) developed for this purpose. Data analysis was performed using STATA (version 11.0; Stata Corp, College Station, Texas). Quantitative variables were described using means and medians, while categorical variables were describing using numbers and proportions. Kruskal-Wallis and Wilcoxon nonparametric tests were used to compare medians for quantitative variables. Chi-square or Fisher exact tests were used to compare categorical variables. Univariate and multivariate regression analyses were carried out to assess factors associated with acceptability of text messages or phone calls to improve health conditions. Regression analyses assessed the factors associated with the acceptability of mHealth on the basis of the towns (Abidjan, Bouaké, Bobo-Dioulasso, and Lomé).

### Ethical Considerations

The protocol for this study was approved by the National Ethics Committee of the Ministry of Health in Côte d'Ivoire, Togo, and Burkina-Faso.

## Results

### Sociodemographic Characteristics

A total of 1131 PLHIV from Côte d'Ivoire ( $n=643$ , 56.9%), Togo ( $n=239$ , 21.1%), and Burkina Faso ( $n=249$ , 22.0%), with a median age of 44 years (interquartile range [IQR]=38-51; Côte d'Ivoire: median=43, IQR=38-50; Burkina-Faso: median=43, IQR: 38-50; Togo: median=44, IQR=37-49;  $P=.04$ ) participated in the study. Among them, 861 (76.1%) were women and about half ( $n=578$ , 51.1%) were single. A total of 640 (56.6%) participants had no formal education or a primary educational level and nearly one-third ( $n=393$ , 34.7%) could not read or write. Participants from Burkina Faso had the lowest rate of literacy, with 65.1% ( $n=162$ ) of PLHIV being unable to read and write compared to 25.5% ( $n=61$ ) from Togo and 26.4% ( $n=170$ ) from Côte d'Ivoire ( $P<.001$ ). The monthly income of three-quarter of the participants ( $n=837$ , 74.0%) was less than 60,000 Franc Communauté Financière Africaine (FCFA) (approximately US \$120), which corresponds to the minimum wage in most countries of francophone West Africa. With regard to the HIV treatment status, 785 (69.4%) had been on ART for over 5 years. Other patient characteristics are summarized in Table 1.

**Table 1.** The sociodemographic characteristics of patients in the three participating countries (N=1131).

Characteristic	Côte d'Ivoire (n=643), n (%)	Burkina-Faso (n=249), n (%)	Togo (n=239), n (%)	Total (N=1131), n (%)	P value
<b>Age (years)</b>					<b>.02</b>
18-34	80 (12.4)	23 (9.2)	38 (15.9)	141 (12.5)	
35-49	363 (56.5)	160 (64.3)	148 (61.9)	671 (59.3)	
≥50	200 (31.1)	66 (26.5)	53 (22.2)	319 (28.2)	
<b>Sex</b>					<b>&lt;.001</b>
Male	179 (27.8)	40 (16.1)	51 (21.3)	270 (23.9)	
Female	464 (72.2)	209 (83.9)	188 (78.7)	861 (76.1)	
<b>Marital status</b>					<b>.07</b>
Single	335 (52.1)	112 (45.0)	131 (54.8)	578 (51.1)	
In a relationship	308 (47.9)	137 (55.0)	108 (45.2)	553 (48.9)	
<b>Level of education</b>					<b>&lt;.001</b>
No education and primary school	312 (48.5)	194 (77.9)	134 (56.1)	640 (56.6)	
Secondary school and university	331 (51.5)	55 (22.1)	105 (43.9)	491 (43.4)	
<b>Literacy</b>					<b>&lt;.001</b>
Can read or write	473 (73.6)	87 (34.9)	178 (74.5)	738 (65.3)	
Cannot read or write	170 (26.4)	162 (65.1)	61 (25.5)	393 (34.7)	
<b>Monthly income (in FCFA<sup>a,b</sup>)</b>					<b>&lt;.001</b>
<60,000	400 (62.2)	222 (89.2)	215 (90.0)	837 (74.0)	
60,000-199,999	176 (27.4)	24 (9.6)	20 (8.3)	220 (19.5)	
>200,000	67 (10.4)	3 (1.2)	4 (1.7)	74 (6.5)	
<b>Duration of antiretroviral therapy (years)</b>					<b>&lt;.001</b>
<5	202 (31.4)	59 (23.7)	85 (35.6)	346 (30.6)	
5-10	207 (32.2)	135 (54.2)	101 (42.3)	443 (39.2)	
>10	234 (36.4)	55 (22.1)	53 (22.1)	342 (30.2)	

<sup>a</sup>US \$1=576.90 FCFA; 60,000 FCFA is the minimum wage in most West African countries.

<sup>b</sup>FCFA: Franc Communauté Financière Africaine.

### Mobile Phone Possession and Access to the Internet

Table 2 summarizes mobile phone possession, mobile internet access, and phone sharing. Overall, the rate of mobile phone possession was 97.9% (n=1107), with a small difference between countries (98.8%, 95.6%, and 97.9% for Côte d'Ivoire, Burkina Faso, and Togo, respectively;  $P=.01$ ). More men owned a mobile phone than women (99.6% vs 97.3%;  $P=.03$ ). With regard to the level of education status, more PLHIV who could read or write possessed phones than PLHIV who could not read or write (99.4% vs 96.7%;  $P=.002$ ). A total of 140 (12.6%) participants shared phones with somebody: 15.4% (n=99) in Côte d'Ivoire, 8.0% (n=20) in Burkina Faso, 8.8% (n=21) in Togo ( $P=.002$ ). Mobile phone sharing was also higher among women and among those with high monthly incomes.

Among PLHIV, 282 (25.5%) had mobile internet access: 30.7% (n=195) from Côte d'Ivoire, 29.9% (n=70) from Togo, and 7.1% (n=17) from Burkina Faso ( $P<.001$ ). Most of those who had a secondary school or university education had better mobile internet access, including young people, women, and those with a higher monthly income (n=208, 42.6%;  $P<.001$ ).

Among mobile phone owners, 33.4% (n=370) owned a smartphone, with a significant difference between countries (73.5%, 3.3%, and 23.2% for Côte d'Ivoire, Burkina Faso, and Togo, respectively;  $P<.001$ ). Among those who owned a smartphone, 83.7% (n=236) had access to mobile internet through mobile phones, with a difference between countries (71.2%, 2.5%, and 26.3% for Côte d'Ivoire, Burkina Faso and Togo respectively;  $P<.001$ ).

**Table 2.** Description of mobile phone possession among patients living with HIV in the three participating countries (N=1131).

Description	Mobile phone possession (n=1107)		Internet access on mobile phone (n=282)		Mobile phone sharing with someone else (n=140)	
	n (%)	P value	n (%)	P value	n (%)	P value
<b>Country</b>	N/A <sup>a</sup>	0.013	N/A	<.001	N/A	0.002
Côte d'Ivoire	635 (98.8)	N/A	195 (30.7)	N/A	99 (15.4)	N/A
Burkina Faso	238 (95.6)	N/A	17 (7.1)	N/A	20 (8.0)	N/A
Togo	234 (97.9)	N/A	70 (29.9)	N/A	21 (8.8)	N/A
<b>Age (years)</b>	N/A	0.31	N/A	<.001	N/A	0.59
18-34	136 (96.5)	N/A	61 (44.9)	N/A	20 (14.2)	N/A
35-49	660 (98.4)	N/A	175 (26.5)	N/A	85 (12.7)	N/A
≥50	311 (97.5)	N/A	46 (14.8)	N/A	35 (11.0)	N/A
<b>Sex</b>	N/A	0.03	N/A	<.001	N/A	0.03
Male	269 (99.6)	N/A	106 (39.4)	N/A	23 (16.4)	N/A
Female	838 (97.3)	N/A	176 (21.0)	N/A	117 (83.6)	N/A
<b>Marital status</b>	N/A	0.26	N/A	0.38	N/A	0.24
Single	563 (97.4)	N/A	137 (24.3)	N/A	65 (11.3)	N/A
In a relationship	544 (98.4)	N/A	145 (26.7)	N/A	75 (13.6)	N/A
<b>Level of education</b>	N/A	0.002	N/A	<.001	N/A	0.89
No education and primary school	619 (96.7)	N/A	74 (12.0)	N/A	80 (12.5)	N/A
Secondary school and university	488 (99.4)	N/A	208 (42.6)	N/A	60 (12.2)	N/A
<b>Literacy</b>	N/A	<.001	N/A	<.001	N/A	0.51
Can read or write	733 (99.3)	N/A	256 (34.9)	N/A	95 (12.9)	N/A
Cannot read or write	374 (95.2)	N/A	26 (7.0)	N/A	45 (11.5)	N/A
<b>Monthly income (in FCFA<sup>b,c</sup>)</b>	N/A	0.18	N/A	<.001	N/A	.01
<60,000	815 (97.4)	N/A	147 (18.0)	N/A	91 (10.9)	N/A
60,000-200,000	218 (99.1)	N/A	92 (42.2)	N/A	33 (15.0)	N/A
>200,000	74 (100.0)	N/A	43 (58.1)	N/A	16 (21.6)	N/A
<b>Duration of antiretroviral therapy (years)</b>	N/A	0.25	N/A	0.04	N/A	0.34
<5	335 (96.8)	N/A	97 (29.0)	N/A	48 (13.9)	N/A
5-10	435 (98.2)	N/A	93 (21.4)	N/A	47 (10.6)	N/A
>10	337 (98.5)	N/A	92 (27.3)	N/A	45 (13.2)	N/A

<sup>a</sup>N/A: not applicable.

<sup>b</sup>US \$1=576.90 FCFA; 60,000 FCFA is the minimum wage in most West African countries.

<sup>c</sup>FCFA: Franc Communauté Financière Africaine.

### Mobile Health Acceptability and Preferred Way of Communication

The overall rate of mHealth acceptability (patients' acceptability of receiving text messages or phone calls from their physicians) was 98.8% with no variation by country (98.3% in Côte d'Ivoire, 99.1% in Togo, and 100.0% in Burkina Faso;  $P=.08$ ). A total of 386 (34.9%), 708 (63.9%), and 11 (1.0%) patients were willing to receive phone calls only, text messages and phone

calls, and text messages only, respectively. Thus, there was a total of 1083 (97.8%) PLHIV who wished to receive phone calls and 708 (63.9%) who wished to receive text messages and phone calls.

There was no statistically significant difference between countries for preference for phone calls ( $n=618$  [57.1%],  $n=234$  [21.6%], and  $n=231$  [21.3%] in Côte d'Ivoire, Burkina Faso, and Togo, respectively;  $P=.47$ ) and for sex ( $n=260$  [24.0%] and  $n=823$  [for men and women, respectively;  $P=.12$ ]).

Among participants who preferred text messages, 60.7% (n=430) had a secondary or higher education level, while 39.3% (n=278) had a lower educational level ( $P<.001$ ). Conversely, 56.1% (n=607) of those who preferred phone calls had primary school level or no educational background ( $P=.55$ ).

### Financial Acceptability and Willingness to Pay

Financial acceptability (answers to the question, “How much would you be willing to spend per month, if you could contact your health care worker for your care?”) was also considered, with 50.1% (n=322), 88.7% (n=212), and 92.8% (n=231) of PLHIV in Côte d’Ivoire, Togo, and Burkina Faso, respectively, likely to financially support out-of-pocket expenditures generated by the adoption of a mHealth solution, if any. The options given for the level of monthly contribution each patient was willing to pay for a mHealth intervention varied from 0 to 3000 FCFA (nearly US \$0 to \$6). A total of 20.1% (n=227) of PLHIV were willing to pay less than 1000 FCFA (nearly US \$2), 35.7% (n=404) were willing to pay between 1000 and 3000 FCFA (nearly US \$2 to \$6), and 11.8% (n=134) were willing

to pay more than 3000 FCFA (>US \$6). However, there was no significant relationship between financial acceptability and monthly income: 67.4% (n=362) of those who had less than the minimum wage were willing to financially contribute to mHealth costs, and 68.4% (n=201) of those who had more than the minimum wage were also willing to financially contribute to mHealth costs ( $P=.47$ ).

### Factors Associated With Mobile Health Acceptability

In multivariate analysis, factors significantly associated with acceptability of “phone calls” to improve HIV treatment and care were as follows: living in Bouaké, the second town of Côte d’Ivoire (adjusted odds ratio [aOR]=6.24, 95% CI 3.98-9.79;  $P<.001$ ), in Lomé (aOR=0.57, 95% CI 0.37-0.88;  $P<.01$ ), or in Bobo-Dioulasso (aOR=3.06, 95% CI 2.11-4.43;  $P<.001$ ); having a primary school education or no education (aOR=7.15, 95% CI 5.05-10.12;  $P<.001$ ), and waiting over 1 hour before meeting a medical doctor on the appointment day (aOR=1.84, 95% CI 1.30-2.62;  $P=.01$ ; [Table 3](#)).

**Table 3.** Factors associated with acceptability of “phone calls” among patients living with HIV in the three participating countries (N=1111).

Factor	Total, n (%)	Bivariate analysis			Multivariate analysis		
		OR <sup>a</sup>	95% CI	P value	OR	95% CI	P value
<b>Town</b>							
Abidjan (Côte d'Ivoire)	478 (43.0)	1			1		
Bouaké (Côte d'Ivoire)	150 (13.5)	7.95	5.29-11.95	<.001	6.24	3.98-9.79	<.001
Lomé (Togo)	234 (21.1)	0.95	0.63-1.40	0.8	0.56	0.36-0.87	0.011
Bobo-Dioulasso (Burkina Faso)	249 (22.4)	5.28	3.77-7.39	<.001	3.06	2.11-4.43	<.001
<b>Age (years)</b>							
18-34	136 (12.2)	1					
35-49	659 (59.3)	1.44	0.95-2.17	0.08	N/A <sup>b</sup>	N/A	N/A
≥50	316 (28.4)	1.53	0.98-2.38	0.06	N/A	N/A	N/A
<b>Sex</b>							
Male	264 (23.8)	1					
Female	847 (76.2)	1.84	1.34-2.51	<.001	N/A	N/A	N/A
<b>Marital status</b>							
Single	566 (50.9)	1					
In a relationship	545 (49.0)	1.05	0.82-1.34	0.69	N/A	N/A	N/A
<b>Level of education</b>							
Secondary school and university	480 (43.2)	1			1		
No education and primary school	631 (56.8)	8.83	6.3-12.22	<.001	7.14	5.04-10.10	<.001
<b>Time to go to the health center (hours)</b>							
≤1	765 (68.8)	1					
>1	346 (31.1)	1.15	0.88-1.50	0.29	N/A	N/A	N/A
<b>Time to meet a doctor (hours)</b>							
≤1	330 (29.7)	1			1		
>1	781 (70.3)	1.53	1.16-2.03	0.003	1.84	1.29-2.62	<.001
<b>Monthly income (in FCFA<sup>c,d</sup>)</b>							
<60,000	824 (74.2)	1					
60,000-199,999	215 (19.3)	0.43	0.30-0.61	<.001	N/A	N/A	N/A
>200,000	72 (6.5)	0.18	0.08-0.40	<.001	N/A	N/A	N/A

<sup>a</sup>OR: odds ratio.<sup>b</sup>N/A: not applicable.<sup>c</sup>US \$1=576.90 FCFA; 60,000 FCFA=minimum wage in most West African countries.<sup>d</sup>FCFA:Franc Communauté Financière Africaine.

## Discussion

### Principal Findings

This study on PLHIV is the first one conducted in francophone sub-Saharan Africa that explores the acceptability of introducing an mHealth solution to better manage HIV treatment and care. In this study, almost all PLHIV reported their acceptability to introduce mHealth into their treatment and care. In addition, acceptability of mHealth was not dependent on sex, age, and level of literacy despite majority of the sample being illiterate. The level of literacy was not a barrier to the acceptability of

mHealth, as demonstrated by the fact that almost all indicated their acceptability of mHealth, although the level of literacy conditioned the choice of mHealth services.

Consistent with our findings, previous studies conducted in East and Southern Africa have found the use of mobile phones in the framework of HIV treatment and care acceptable. A 2007 study in South Africa among 300 PLHIV explored the patterns of mobile phone use to improve adherence to ART and reported that 81% had a mobile phone, 99% indicated their acceptability for verbal clinic contact through their mobile phones, and 96% agreed to contact through text messages [27]. The same trend

was observed in Kenya, where 99% of the sample supported the use of mobile technology in the management of HIV infection, with calls being the preferred method [28]. Those results suggest that the integration of mobile phone into the treatment and care of PLHIV in sub-Saharan Africa is an opportunity that should be considered.

Our study reported a mobile phone possession rate of 97.9% despite an unequal mobile network coverage across the three countries and confirmed important disparities in access to mobile internet between the countries, as defined by the Global System for Mobile Communications Association mobile connectivity index [21]. In Burkina Faso, a mobile internet-based solution would reach less than 1 of 10 people, while it would reach 3 of 10 people in Côte d'Ivoire and Togo. In these three countries, introducing a non-Web-based solution seems to be the most appropriate intervention. In addition, in a regional context where patients' fees have been pointed out as a main barrier to access HIV treatment [29,30], our study reported a readiness to pay for better quality health care or at least to support the additional out-of-pocket expenditures generated by the use of their mobile phone to improve the quality of their care, with no significant relationship between financial acceptability and monthly income. Even if our study did not focus on willingness to pay (WTP) as a measure used in the field of health economics, our findings are aligned to the conclusion of a randomized controlled trial carried out in Bangladesh, demonstrating patients' WTP to benefit from an SMS-based service for managing their type 2 diabetes [31]. In the same country, a previous study already showed patients' WTP for an improved patient-clinician relationship [32]. However, as there is a scarcity of studies conducted in sub-Saharan Africa in the field of HIV, which target the WTP for digital health services, it seems to be preferable to implement an mHealth solution with a business model that does not increase the user fees, but rather generates benefits to the patients in reducing transportation time, improves waiting time at the clinic, and allows savings on out-of-pocket expenditures.

In our study, participants generally indicated a preference for phone calls, and the acceptability for phone calls was significantly associated with levels of education, such that participants with no education or a primary school level education were more likely to prefer phone calls compared to text messages. This finding of preference for phone calls, and its association with education levels, has been corroborated by other studies in the sub-Saharan region [28,25]. Another study among PLHIV reported that almost all participants indicated their preference for a two-way communication with doctors or nurses, rather than SMS reminders [33]. These findings suggest that mHealth could respond to the need of restoring or strengthening the interface patient-clinician through an improved and personalized relationship. In this way, mHealth could be an appropriate tool to address the issue raised by Tran et al [34] who showed that listening to patients should lead decision makers to improve the quality of health services and decrease the burden of HIV treatment [34]. There is also a need for adaptation of types of communication (text messages or phone calls) to the target PLHIV according to their literacy level. Tailoring text messages interventions to fit with those specific

needs of PLHIV is an example of how this adaptation could be implemented. The preference for phone calls could also be explained by several factors including the issue of confidentiality. A qualitative study on PLHIV indicated that although SMS reminders are helpful for treatment adherence, the main concerns are on the delivery mode and the need for confidentiality [35]. This could suggest the importance of using strategies that would protect privacy, ensure confidentiality, and respect human rights for PLHIV.

### Strengths and Limitations

This study was conducted in three countries of West Africa, with a sample of more than 1000 PLHIV from six HIV treatment and care programs, including approximately one-third (34.7%, n=393) who could not read nor write. Findings from this study thus suggest a very good feasibility and acceptability of the use of mHealth in all types of HIV populations. In addition, financial acceptability of mHealth among PLHIV, rarely explored in the literature in West Africa and documented in our study, confer to it an important strength. Furthermore, this study was patient centered, with the involvement of PLHIV from the design and elaboration of the study to the interpretation of the results, as recommended for the design of digital health tool [36]. Most PLHIV organizations of Cote d'Ivoire including the Réseau Ivoirien des organisations de Personnes vivant avec le VIH/Sida (RIP+) and the Coalition des Organisations des Femmes vivant avec le VIH/Sida (COFCI) have been involved in this study since the beginning. In addition, focus groups and interviews were conducted as part of this study among 20 PLHIV and revealed that although messages are highly acceptable, message contents need to be further explored. In fact, PLHIV indicated that contents should be short, highly coded, and customized based on literacy levels.

However, this study has a few limitations. First, it occurred mainly in economic and administrative capitals where mobile network coverage is close to 100% and sharing of mobile phone is limited, even if in these settings, a significant number of people cannot not read or write. If an intervention were to be set up in those cities, further studies would be needed before scaling up to rural and remote areas. Second, only PLHIV already receiving ART were included in the study, of which nearly 70% were stable patients who initiated their treatment over 5 years ago. This could lead to underestimation of the demand for the introduction of an mHealth solution and the willingness to financially contribute, as the targeted population is already adherent and can self-perceive the benefits of the treatment. A bias could have also been introduced with the cross-sectional nature of the study and the fact that recruitment of participants occurred in HIV care and treatment center, systematically including those who are naturally inclined to be retained in care and to adhere to treatment.

Further research, such as qualitative studies, are needed to better understand expectations and perceptions of PLHIV on the contents of text messages or phone calls, especially in terms of confidentiality, privacy, and respect of human rights.

Qualitative studies should also explore health care workers' perception of the introduction of mHealth into care to assess

whether it would be considered a tool to ease their workload or hinder their performance.

Finally, considering that implication of end users is a key factor of success when designing a digital health tool and the need to prevent any potential harm that may result from the use of an mHealth solution in a stigmatizing environment, further qualitative studies shall be conducted in these francophone cities to consider frequency, timing, and content of the SMS in the design of mHealth interventions, as these factors could influence efficacy [36,37].

## Conclusions

Based on a considerable sample size, this study demonstrated that mobile technology is accessible in francophone low-income countries and could be used as a tool to improve the quality of HIV care and treatment. In such a context, mHealth could potentially help francophone West and Central African countries to bridge their treatment gap in the aim to achieve the 90-90-90 goal.

## Acknowledgments

We would like to acknowledge the people living with HIV who participated in this study and networks of PLHIV such as RIP+ and COF-CI who actively contributed.

## Conflicts of Interest

None declared.

## References

1. Joint United Nations Programme on HIV/AIDS (UNAIDS). Joint United Nations Programme on HIV/AIDS (UNAIDS). Geneva, Switzerland; 2017. UNAIDS Data 2017 URL: [https://www.unaids.org/sites/default/files/media\\_asset/20170720\\_Data\\_book\\_2017\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf) [accessed 2018-11-12]
2. Joint United Programme on HIV/AIDS (UNAIDS). Geneva, Switzerland; 2014. Fast-Track: Ending the AIDS epidemic by 2030 URL: [https://www.unaids.org/sites/default/files/media\\_asset/JC2686\\_WAD2014report\\_en.pdf](https://www.unaids.org/sites/default/files/media_asset/JC2686_WAD2014report_en.pdf) [accessed 2018-11-19]
3. Samji H, Cescon A, Hogg RS, Modur SP, Althoff KN, Buchacz K, North American AIDS Cohort Collaboration on ResearchDesign (NA-ACCORD) of IeDEA. Closing the gap: increases in life expectancy among treated HIV-positive individuals in the United States and Canada. PLoS One 2013 Dec;8(12):e81355 [FREE Full text] [doi: [10.1371/journal.pone.0081355](https://doi.org/10.1371/journal.pone.0081355)] [Medline: [24367482](https://pubmed.ncbi.nlm.nih.gov/24367482/)]
4. Teeraananchai S, Kerr SJ, Amin J, Ruxrungtham K, Law MG. Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. HIV Med 2017 Apr;18(4):256-266 [FREE Full text] [doi: [10.1111/hiv.12421](https://doi.org/10.1111/hiv.12421)] [Medline: [27578404](https://pubmed.ncbi.nlm.nih.gov/27578404/)]
5. World Health Organization (WHO). Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach. Geneva: World Health Organization; 2016.
6. World Health Organization (WHO). mHealth: New horizons for health through mobile technologies. In: Global observatory for ehealth series (Volume 3). Geneva: World Health Organization; 2011.
7. Rassi C, Gore-Langton GR, Gidudu Walimbwa B, Strachan CE, King R, Basharat S, et al. Improving health worker performance through text messaging: A mixed-methods evaluation of a pilot intervention designed to increase coverage of intermittent preventive treatment of malaria in pregnancy in West Nile, Uganda. PLoS One 2018 Sep;13(9):e0203554 [FREE Full text] [doi: [10.1371/journal.pone.0203554](https://doi.org/10.1371/journal.pone.0203554)] [Medline: [30188956](https://pubmed.ncbi.nlm.nih.gov/30188956/)]
8. Hirsch-Moverman Y, Daftary A, Yuengling KA, Saito S, Ntoane M, Frederix K, et al. Using mHealth for HIV/TB Treatment Support in Lesotho. JAIDS 2017;74:S37-S43. [doi: [10.1097/qai.0000000000001202](https://doi.org/10.1097/qai.0000000000001202)]
9. Jia K, Mohamed K. Evaluating the use of cell phone messaging for community Ebola syndromic surveillance in high risk settings in Southern Sierra Leone. Afr Health Sci 2015 Sep;15(3):797-802 [FREE Full text] [doi: [10.4314/ahs.v15i3.13](https://doi.org/10.4314/ahs.v15i3.13)] [Medline: [26957967](https://pubmed.ncbi.nlm.nih.gov/26957967/)]
10. Mangam K, Fiekowsky E, Bagayoko M, Norris L, Belemvire A, Longhany R, et al. Feasibility and Effectiveness of mHealth for Mobilizing Households for Indoor Residual Spraying to Prevent Malaria: A Case Study in Mali. Glob Health Sci Pract 2016 Jun 27;4(2):222-237. [doi: [10.9745/ghsp-d-15-00381](https://doi.org/10.9745/ghsp-d-15-00381)]
11. Mannik J, Figol A, Churchill V, Aw J, Francis S, Karino E, et al. Community-based screening for cardiovascular risk using a novel mHealth tool in rural Kenya. J Innov Health Inform 2018 Oct 31;25(3):176-182 [FREE Full text] [doi: [10.14236/jhi.v25i3.1012](https://doi.org/10.14236/jhi.v25i3.1012)] [Medline: [30398461](https://pubmed.ncbi.nlm.nih.gov/30398461/)]
12. Kingue S, Angandji P, Menanga AP, Ashuntantang G, Sobngwi E, Dossou-Yovo RA, et al. Efficiency of an intervention package for arterial hypertension comprising telemanagement in a Cameroonian rural setting: The TELEMED-CAM study. Pan Afr Med J 2013;15:153 [FREE Full text] [doi: [10.11604/pamj.2013.15.153.2655](https://doi.org/10.11604/pamj.2013.15.153.2655)] [Medline: [24396559](https://pubmed.ncbi.nlm.nih.gov/24396559/)]
13. Ekouevi DK, Balestre E, Ba-Gomis F, Eholie SP, Maiga M, Amani-Bosse C, IeDEA West Africa Collaboration. Low retention of HIV-infected patients on antiretroviral therapy in 11 clinical centres in West Africa. Trop Med Int Health 2010 Jun;15 Suppl 1:34-42 [FREE Full text] [doi: [10.1111/j.1365-3156.2010.02505.x](https://doi.org/10.1111/j.1365-3156.2010.02505.x)] [Medline: [20586958](https://pubmed.ncbi.nlm.nih.gov/20586958/)]



14. Plazy M, Orne-Gliemann J, Dabis F, Dray-Spira R. Retention in care prior to antiretroviral treatment eligibility in sub-Saharan Africa: a systematic review of the literature. *BMJ Open* 2015 Jun 24;5(6):e006927 [FREE Full text] [doi: [10.1136/bmjopen-2014-006927](https://doi.org/10.1136/bmjopen-2014-006927)] [Medline: [26109110](https://pubmed.ncbi.nlm.nih.gov/26109110/)]
15. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. *J Med Internet Res* 2014 Oct 17;16(10):e222 [FREE Full text] [doi: [10.2196/jmir.3442](https://doi.org/10.2196/jmir.3442)] [Medline: [25326646](https://pubmed.ncbi.nlm.nih.gov/25326646/)]
16. Mbuagbaw L, Thabane L, Ongolo-Zogo P, Lester RT, Mills EJ, Smieja M, et al. The Cameroon Mobile Phone SMS (CAMPS) trial: a randomized trial of text messaging versus usual care for adherence to antiretroviral therapy. *PLoS One* 2012 Dec;7(12):e46909 [FREE Full text] [doi: [10.1371/journal.pone.0046909](https://doi.org/10.1371/journal.pone.0046909)] [Medline: [23236345](https://pubmed.ncbi.nlm.nih.gov/23236345/)]
17. Orrell C, Cohen K, Mauff K, Bangsberg DR, Maartens G, Wood R. A Randomized Controlled Trial of Real-Time Electronic Adherence Monitoring With Text Message Dosing Reminders in People Starting First-Line Antiretroviral Therapy. *JAIDS* 2015;70(5):495-502. [doi: [10.1097/qai.0000000000000770](https://doi.org/10.1097/qai.0000000000000770)]
18. Joseph Davey D, Nhavoto JA, Augusto O, Ponce W, Traca D, Nguimfack A, et al. SMSaúde: Evaluating mobile phone text message reminders to improve retention in HIV care for patients on antiretroviral therapy. *JAIDS* 2016;73(2):e23-e30. [doi: [10.1097/qai.0000000000001115](https://doi.org/10.1097/qai.0000000000001115)]
19. Katz IT, Kaplan R, Fitzmaurice G, Leone D, Bangsberg DR, Bekker L, et al. Treatment guidelines and early loss from care for people living with HIV in Cape Town, South Africa: A retrospective cohort study. *PLoS Med* 2017 Nov;14(11):e1002434 [FREE Full text] [doi: [10.1371/journal.pmed.1002434](https://doi.org/10.1371/journal.pmed.1002434)] [Medline: [29136014](https://pubmed.ncbi.nlm.nih.gov/29136014/)]
20. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC Public Health* 2014 Feb 21;14:188 [FREE Full text] [doi: [10.1186/1471-2458-14-188](https://doi.org/10.1186/1471-2458-14-188)] [Medline: [24555733](https://pubmed.ncbi.nlm.nih.gov/24555733/)]
21. GSMA. 2018. The mobile connectivity index 2018 URL: <https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2018/09/GSMA-Mobile-Connectivity-Index-GLOBAL-Fcous.pdf> [accessed 2018-11-02]
22. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? *Glob Health Action* 2014 Oct;7:25606 [FREE Full text] [doi: [10.3402/gha.v7.25606](https://doi.org/10.3402/gha.v7.25606)] [Medline: [25361730](https://pubmed.ncbi.nlm.nih.gov/25361730/)]
23. Nachega J, Skinner D, Jennings L, Magidson J, Altice F, Burke J, et al. Acceptability and feasibility of mHealth and community-based directly observed antiretroviral therapy to prevent mother-to-child HIV transmission in South African pregnant women under Option B+: an exploratory study. *PPA* 2016 Apr:683. [doi: [10.2147/ppa.s100002](https://doi.org/10.2147/ppa.s100002)]
24. Cormick G, Ciganda A, Cafferata ML, Ripple MJ, Sosa-Estani S, Buekens P, et al. Text message interventions for follow up of infants born to mothers positive for Chagas disease in Tucumán, Argentina: a feasibility study. *BMC Res Notes* 2015 Sep 29;8:508 [FREE Full text] [doi: [10.1186/s13104-015-1498-9](https://doi.org/10.1186/s13104-015-1498-9)] [Medline: [26419230](https://pubmed.ncbi.nlm.nih.gov/26419230/)]
25. Kim J, Zhang W, Nyonyitono M, Lourenco L, Nanfuka M, Okobo S, et al. Feasibility and acceptability of mobile phone short message service as a support for patients receiving antiretroviral therapy in rural Uganda: a cross-sectional study. *Journal of the International AIDS Society* 2015 Dec 10;18(1):20311. [doi: [10.7448/ias.18.1.20311](https://doi.org/10.7448/ias.18.1.20311)]
26. Berry G, Smith CL, Macaskill P, Irwig L. Analytic methods for comparing two dichotomous screening or diagnostic tests applied to two populations of differing disease prevalence when individuals negative on both tests are unverified. *Statist Med* 2002 Mar 30;21(6):853-862. [doi: [10.1002/sim.1066](https://doi.org/10.1002/sim.1066)]
27. Crankshaw T, Corless IB, Giddy J, Nicholas PK, Eichbaum Q, Butler LM. Exploring the patterns of use and the feasibility of using cellular phones for clinic appointment reminders and adherence messages in an antiretroviral treatment clinic, Durban, South Africa. *AIDS Patient Care STDS* 2010 Nov;24(11):729-734. [doi: [10.1089/apc.2010.0146](https://doi.org/10.1089/apc.2010.0146)] [Medline: [21039181](https://pubmed.ncbi.nlm.nih.gov/21039181/)]
28. Kinyua F, Kiptoo M, Kikui G, Mutai J, Meyers AFA, Muiruri P, et al. Perceptions of HIV infected patients on the use of cell phone as a tool to support their antiretroviral adherence; a cross-sectional study in a large referral hospital in Kenya. *BMC Public Health* 2013 Oct 21;13:987 [FREE Full text] [doi: [10.1186/1471-2458-13-987](https://doi.org/10.1186/1471-2458-13-987)] [Medline: [24143931](https://pubmed.ncbi.nlm.nih.gov/24143931/)]
29. Le prix de l'oubli. In: Médecins Sans Frontières. Bruxelles, Belgique: Médecins sans frontières; 2016.
30. The Western and Central Africa Catch-up plan: Putting HIV treatment on the fast-track by 2018. In: Joint United Nations Programme on HIV/AIDS (UNAIDS). Geneva: UNAIDS; 2018.
31. Shariful Islam SM, Lechner A, Ferrari U, Seissler J, Holle R, Niessen LW. Mobile phone use and willingness to pay for SMS for diabetes in Bangladesh. *J Public Health (Oxf)* 2016 Mar;38(1):163-169. [doi: [10.1093/pubmed/fdv009](https://doi.org/10.1093/pubmed/fdv009)] [Medline: [25687131](https://pubmed.ncbi.nlm.nih.gov/25687131/)]
32. Pavel MS, Chakrabarty S, Gow J. Assessing willingness to pay for health care quality improvements. *BMC Health Serv Res* 2015 Feb 01;15:43 [FREE Full text] [doi: [10.1186/s12913-015-0678-6](https://doi.org/10.1186/s12913-015-0678-6)] [Medline: [25638204](https://pubmed.ncbi.nlm.nih.gov/25638204/)]
33. Georgette N, Siedner MJ, Zandoni B, Sibaya T, Petty CR, Carpenter S, et al. The Acceptability and Perceived Usefulness of a Weekly Clinical SMS Program to Promote HIV Antiretroviral Medication Adherence in KwaZulu-Natal, South Africa. *AIDS Behav* 2016 Nov;20(11):2629-2638 [FREE Full text] [doi: [10.1007/s10461-016-1287-z](https://doi.org/10.1007/s10461-016-1287-z)] [Medline: [26781866](https://pubmed.ncbi.nlm.nih.gov/26781866/)]
34. Tran V, Messou E, Mama Djima M, Ravaud P, Ekouevi DK. Patients' perspectives on how to decrease the burden of treatment: a qualitative study of HIV care in sub-Saharan Africa. *BMJ Qual Saf* 2019 Apr;28(4):266-275 [FREE Full text] [doi: [10.1136/bmjqs-2017-007564](https://doi.org/10.1136/bmjqs-2017-007564)] [Medline: [29706594](https://pubmed.ncbi.nlm.nih.gov/29706594/)]

35. Sherman E, Clauson KA, Elrod S, Eckardt P, Alkhateeb F, Jacobs R. Patient Perceptions of Text Messaging to Improve Antiretroviral Therapy Adherence: A Qualitative Study of Patients in a Ryan White Clinic. *J Int Assoc Provid AIDS Care* 2018 Feb;17:2325958218759209. [doi: [10.1177/2325958218759209](https://doi.org/10.1177/2325958218759209)] [Medline: [29473483](https://pubmed.ncbi.nlm.nih.gov/29473483/)]
36. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. *Acad Emerg Med* 2015 Jun;22(6):754-756 [FREE Full text] [doi: [10.1111/acem.12692](https://doi.org/10.1111/acem.12692)] [Medline: [25997375](https://pubmed.ncbi.nlm.nih.gov/25997375/)]
37. Finitis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. *PLoS One* 2014 Feb;9(2):e88166 [FREE Full text] [doi: [10.1371/journal.pone.0088166](https://doi.org/10.1371/journal.pone.0088166)] [Medline: [24505411](https://pubmed.ncbi.nlm.nih.gov/24505411/)]

## Abbreviations

**ART:** antiretroviral treatment

**mHealth:** mobile health

**PLHIV:** People Living with HIV

**SMS:** short messages service

**UNAIDS:** United Nations Programme on HIV/AIDS

**WHO:** World Health Organization

*Edited by G Eysenbach; submitted 21.02.19; peer-reviewed by R Garfein, K Stasiak, D Mendes; comments to author 18.04.19; revised version received 07.06.19; accepted 19.06.19; published 13.11.19.*

*Please cite as:*

*Lepère P, Touré Y, Bitty-Anderson AM, Boni SP, Anago G, Tchounga B, Touré P, Minga A, Messou E, Kanga G, Koule S, Poda A, Calmy A, Ekouevi DK, Coffie PA*

*Exploring the Patterns of Use and Acceptability of Mobile Phones Among People Living With HIV to Improve Care and Treatment: Cross-Sectional Study in Three Francophone West African Countries*

*JMIR Mhealth Uhealth* 2019;7(11):e13741

URL: <https://mhealth.jmir.org/2019/11/e13741>

doi: [10.2196/13741](https://doi.org/10.2196/13741)

PMID: [31719023](https://pubmed.ncbi.nlm.nih.gov/31719023/)

©Phillipe Lepère, Yélamikan Touré, Alexandra M Bitty-Anderson, Simon P Boni, Gildas Anago, Boris Tchounga, Pendadiago Touré, Albert Minga, Eugène Messou, Guillaume Kanga, Serge Koule, Armel Poda, Alexandra Calmy, Didier K Ekouevi, Patrick A Coffie. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 13.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# A Model for Assessing Necessary Conditions for Rural Health Care's Mobile Health Readiness: Qualitative Assessment of Clinician-Perceived Barriers

Bryan Weichelt<sup>1</sup>, MBA, PhD; Casper Bendixsen<sup>1</sup>, PhD; Timothy Patrick<sup>2</sup>, PhD

<sup>1</sup>Marshfield Clinic Research Institute, National Farm Medicine Center, Marshfield, WI, United States

<sup>2</sup>University of Wisconsin-Milwaukee, Health Informatics, Milwaukee, WI, United States

**Corresponding Author:**

Bryan Weichelt, MBA, PhD

Marshfield Clinic Research Institute

National Farm Medicine Center

1000 N Oak Ave

Marshfield, WI, 54449

United States

Phone: 1 2217276

Email: [weichelt.bryan@marshfieldresearch.org](mailto:weichelt.bryan@marshfieldresearch.org)

## Abstract

**Background:** Mobile health (mHealth) technology dissemination has penetrated rural and urban areas alike. Yet, health care organization oversight and clinician adoption have not kept pace with patient use. mHealth could have a unique impact on health and quality of life for rural populations. If organizations are prepared to manage mHealth, clinicians may improve the quality of care for their patients, both rural and urban. However, many organizations are not yet prepared to prescribe or prohibit third-party mHealth technologies.

**Objective:** This study explored organizational readiness for rural mHealth adoption, the use of patient-reported data by clinical care teams, and potential impact on improving rural health care delivery.

**Methods:** Semistructured, open-ended interviews were used to investigate clinicians' current practices, motivators, and perceived barriers to their use of mHealth technologies in rural settings.

**Results:** A total of 13 clinicians were interviewed, and 53.8% (7/13) reported encouraging use of mHealth apps or wearable devices with rural patients. Perceived barriers to adoption were categorized into three primary themes: (1) personal (clinician), (2) patient, and (3) organizational. Organizational was most prominent, with subcodes of time, uniformity, and policy or direction. Thematic analysis revealed code-category linkages that identify the complex nature of a rural health care organization's current climate from a clinician's perspective. A thematic map was developed to visualize the flow from category to code. Identified linkages guided the development of a refined rural mHealth readiness model.

**Conclusions:** Clinicians (including physicians) have limited time for continuing education, research, or exploration of emerging technologies. Clinicians are motivated to learn more, but they need guidance through organization-led directives. Rural health care institutions should consider investing in mHealth analysis, tool development, and formal recommendations of sanctioned tools for clinicians to use with patients.

(*JMIR Mhealth Uhealth* 2019;7(11):e11915) doi:[10.2196/11915](https://doi.org/10.2196/11915)

**KEYWORDS**

mHealth; clinician; physician; rural; patient; mobile health; health care

## Introduction

**Background and Significance**

Out of an unprecedented adoption of mobile communication technologies and the progressive advancement of their

application to personal and population health management, a new field of science, research, and health care has emerged—the study of mobile health (mHealth). The World Health Organization defines mHealth as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs),

and other wireless devices” [1]. Evolving in recent years, the Knowledge for Health project describes the discipline by stating, “mobile health, or mHealth, is broadly the use of mobile and wireless technologies to support the delivery and utilization of health care services.” mHealth is a young field, and a limited evidence base exists for demonstrating its efficacy, effectiveness, and comparative effectiveness, especially its cost-effectiveness [2]. Both these definitions imply an organizationally driven or practice-driven approach to health care. In addition, the definitions suggest that the process is merely enhanced by the use of these new technologies.

Applications (apps) are software programs. Mobile apps are software developed specifically for a mobile device, such as a mobile phone or tablet. Native apps are software programs developed to be installed directly onto a device, typically downloaded through an app store. In this paper, mHealth apps include software accessible through websites and native mobile apps.

mHealth also includes wearable activity monitors (WAMs). Hundreds of WAMs are on the market and being used by consumers across a wide range of industries and occupations [3]. WAMs include a multitude of devices that are worn in and on various body parts or clothing, such as the wrist or pocket. Some companies offering these devices include Apple, Fitbit, Under Armour, Garmin, Jawbone, Pebble Time, LG, and Misfit [3].

The treatment of many chronic conditions takes place primarily outside the purview and physical environment of a doctor’s office. Although clinicians often require patients to recall detailed information about their symptoms and condition, the appointments may be spread apart, and recollection of specific situations can be challenging for patients [4]. Innovative mHealth apps provide the opportunity to enhance and improve the data collection and reporting processes, the patient-clinician interaction, and health outcomes [5-9]. Furthermore, the technology may make self-reporting easier for the patient and possibly more accurate and complete as well [10,11].

Some clinicians and organizations may argue that they should not base clinical decisions on patient-reported data. This study sought to uncover barriers to the use of patient-reported data and also motivators for clinicians who are already practicing in this manner. The opportunity exists for patient-driven health management and is readily available through a number of consumer software. The number of mHealth apps in publicly available app stores has grown exponentially in recent years to

more than 325,000, with Android as the leading platform in 2017 [12,13].

iTunes and Google Play stores are the 2 largest in terms of apps available [14], as well as in number of available mHealth apps. iTunes is a publicly available app store where Apple device users can access software apps and directly install them onto their personal devices. Some apps are free, whereas others charge a fee for the installation, and some have a monthly or annual subscription [15]. Google Play is very similar, although the apps available in that store are for Android users. Although the public has access to an increasingly large number of apps, there appears to be limited traction among rural health care organizations to actively prescribe or engage patients with mHealth technologies, much less adopt and integrate patient-reported data into an existing electronic medical record (EMR). Therefore, despite the public’s easy access to hundreds of thousands of mHealth apps, they are not widely implemented by rural health care providers.

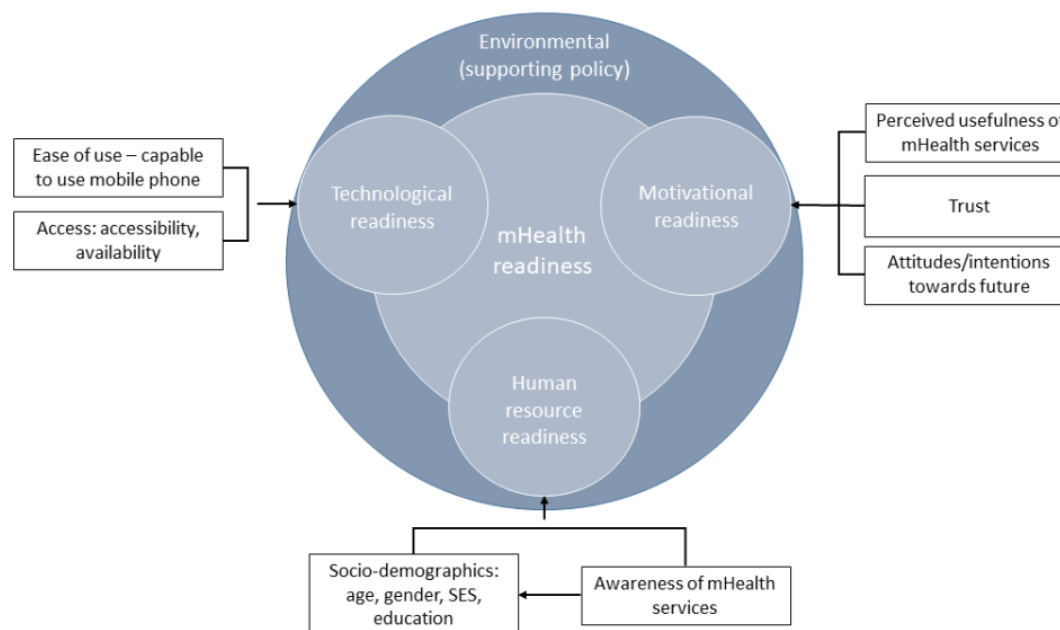
### A Framework for Readiness

In a rural community of Bangladesh, Khatun et al [16] surveyed 4915 randomly selected household members aged 18 years and older. The research team found that only 5% of participants had internet connectivity, only 50% were aware of SMS apps, and only 37% generally read them. Literacy was the primary barrier. In addition, 21% needed to charge their phones at someone else’s home, as there was no electricity in their own homes. Despite these barriers, the majority (73%) showed an interest in using mHealth technology in the future.

The team developed a framework for assessing community readiness for mHealth. The developed framework, further described herein and seen in Figure 1, has led to some interesting findings that have helped guide other studies, including this paper. The framework identified 3 high-level areas of readiness that are described here in greater detail. Furthermore, a model assessing clinician readiness was developed.

Previous literature is encouraging, but it should be noted that many of those studies were single interventions, often focused on isolated communities. There is currently a gap in the knowledge and literature around clinicians’ unofficial use of these types of technologies with rural patients. It may be that clinicians doing so are acting in silos with limited or no organizational direction, oversight, or approval processes. The purpose of this study was to create a model for assessing the necessary conditions for rural health care’s mHealth readiness through a qualitative assessment of clinician-perceived barriers.

**Figure 1.** Conceptual model for mHealth readiness, developed by Khatun et al [16]. mHealth: mobile health; SES: socioeconomic status.



## Methods

### Overview

mHealth technologies, such as apps, are widely available, along with the personal devices needed to run them. Clinicians and researchers have had successes with mHealth in urban and rural settings at various locations worldwide. However, limited organizational adoption has emerged in rural Wisconsin.

As a result, there are rarely formal policies and procedures to evaluate when assessing the use and impact of mHealth technologies. Therefore, this study team conducted semistructured, open-ended interviews with clinicians, as they were seen as the primary decision makers as to whether mHealth was used in patient care. Clinicians were recruited based on availability and not whether they were actively using mHealth technology. The open-ended nature of the interview allowed clinicians to explore their thoughts about using mHealth without assuming any prior beliefs or commitments on the part of the researcher [17]. These open-ended discussions provided qualitative data that were later used to refine a technology readiness model.

The 2-part methodology included consultations and interviews as the primary data collection and refinement of an mHealth readiness model using this study's findings. Clinicians who see patients are particularly busy and may be unlikely to complete an electronic or paper-based survey without an incentive [18-20].

### Theory

The Diffusion of Innovation theory is one of the oldest social science theories and was developed by E.M. Rogers in 1962 [21]. The theory originated in communications, and the concept guided explanations of how ideas may gain in popularity and momentum to spread through a population or social system. This theory informed and guided the project and will be further assessed in its relation to the newly refined conceptual model

for mHealth readiness, particularly as a framework for clinician and organizational adoption of mHealth technologies.

### Recruitment

Interviews and consultations were conducted at Marshfield Clinic Health System (MCHS), Wisconsin. Interviews were one-on-one interactions. Clinicians were recruited via phone and email, and interviews were scheduled at their convenience.

### Data Collection

Data collection was carried out between February and April 2016. Sessions were recorded using a battery-powered Olympus DM-620 audio recorder. Handwritten notes were also taken.

Consultation interviews were conducted as a way to refine the interview schedule. A total of 3 MD clinicians were interviewed as consultations for the project. Each represented a different clinical specialty and background.

The principal method of investigation was open-ended, semistructured interviews. The interviews were loosely bound around questions that asked informants to reflect on their participation in mHealth. Open-ended, semistructured interviews are largely conversation driven, and clinician time was a constraint. Thus, it was difficult to predict exactly which questions may come up with individual subjects and at what point in the flow and under what context. The initial proposal called for 10 interviews or until data collection reached the point of theme saturation.

### Coding: Identifying Themes and Categories

Following the consultations and interviews, additional notes were taken, notes were reviewed, and the audio recording was reviewed for additional note-taking and thematic analysis. The primary method of analysis was audio coding [22]. On the basis of the literature review and the first author's experience, codes were identified. These experiences also formed the basis of the study and the research questions. Deductive reasoning

techniques, coding qualitative data with pre-existing codes already in mind, were used to construct theme and category linkages.

The following existing theme categories, identified by Khatun et al in their study of barriers to rural community readiness to adopt mHealth [16], were used as a starting point: (1) technological readiness, (2) motivational readiness, and (3) human resource readiness. These were a natural fit for the evaluation of community readiness, particularly a rural community. The researchers in a study by Khatun et al were also interested in the individuals within that community and had surveyed residents, using door-to-door, in-person data collection methods.

The abovementioned categories were modified for this study to (1) clinician readiness, (2) patient readiness, and (3) organizational readiness. As noted, these 3 categories were identified through the first author's experience in the field and through the literature review, including a study by Khatun et al [16]. The themes themselves were identified as relationships between the codes and the categories, such as the relationship between the organization and the legal/liability barriers to mHealth adoption, which is described in further detail in the Results section.

The review of the audio recordings allowed for inductive reasoning [23]—the possibility that other themes and subthemes would emerge. This process was used to identify additional themes.

### Mobile Health Readiness Model Refinement

The description of the work by Khatun et al [16] was also used here as the basis for the readiness model refinement. The refinement of the model was assessed before the interviews, and several new categories were identified. The refined model is focused on clinician readiness with data collected from clinician interviews. In addition to modifying category titles and influencers, each category was further assessed for thematic linkages to codes. Those linkages or relationships encompass

the core of this study, and the recommendations are built into a newly refined model.

### Institutional Review Board Approval

This study was approved as minimal risk after review by the institutional review board of the University of Wisconsin-Milwaukee. The protocol has been granted Exempt Status under Category 2 as governed by 45 CFR 46.101(b). This study was also reviewed by the Office of Research and Integrity of Marshfield Clinic Research Institute and approved as minimal risk and exempt from further review (45 CFR 46.101(b)(2)).

## Results

The patient base of this study was primarily Marshfield, Wisconsin, and the surrounding area. Generally, their patients lived within a 20-mile radius of the medical center, with some exceptions.

### Interviews

The interviews uncovered a strong personal interest and passion for the participants' patients, and the conversations exposed uses of the technology that are already in their practice or that they are considering for future patient care (see Table 1). There were 12 different clinical specialties represented. Given the small sample and the other reported demographics, for the protection of the participants' anonymity, clinical specialties have been omitted.

Of the 13 clinicians interviewed, 12 were physically located at the Marshfield center, Marshfield, Wisconsin, and 1 was located in Eau Claire, Wisconsin, but they saw patients remotely through telehealth and in person in Marshfield and Eau Claire. Additional 10 clinicians were contacted during recruitment, but they were unable to participate. There were 5 female and 8 male participants. At the time of the interviews, 100% of participants owned tablets, and 92% owned mobile phones, and the 1 clinician who did not own a mobile phone was planning on purchasing one within the next few weeks. In addition, 77% were parents, and 30% of those were grandparents.

**Table 1.** Descriptive demographics of interview participants.

Process	Age (years)	Gender	Mobile phone owner	Tablet owner	Patients seen per month (n)	Percent of patients viewed as rural (%)	Prescribing/encouraging mHealth app/technology	Examples used or plan to use
Consultation	61	Female	Yes	Yes	Question not asked	Question not asked	Yes	Fitbit and My-FitnessPal
Consultation	62	Male	Yes	Yes	14-22	100	No	— <sup>a</sup>
Consultation	37	Male	Yes	Yes	200	10-20	No; but hopes to	Fitness
Interview	58	Male	Yes	Yes	200	90	No	—
Interview	37	Male	Yes	Yes	14-22	Most	Yes	MyFitnessPal
Interview	47	Female	Yes	Yes	100	95-100	Yes	Fitbit
Interview	43	Male	Yes	Yes	40	100	Yes	MyFitnessPal
Interview	47	Female	Yes	Yes	200	75	Yes	Omitted
Interview	53	Female	Yes	Yes	40	75	No; but hopes to	Blood pressure
Interview	40	Male	Yes	Yes	10	60	No; but hopes to	Physical Rehab
Interview	40	Male	Yes	Yes	200	100	Yes	Apple iWatch and Fitbit
Interview	46	Female	Yes	Yes	48	100	Yes	Blood sugar, fitness, and others omitted
Interview	37	Male	Yes	Yes	200	90	No; but plans to	Fitness

<sup>a</sup>Not applicable.

## Demographics

Although all participants were clinicians, few had an understanding of the term *mHealth* or its scope, with only 23% responding that they were familiar with the term before the interview. Although once described and examples given, all participants knew of other cases and gave examples of mobile devices or wearables used for personal health-related activities.

## Barriers

The analyses and coding of the data showed that barriers to adoption were one of the most important factors affecting clinician readiness to adopt. At this stage of the model refinement, we have focused on the issues of clinicians' perceived barriers, further described in the sections below.

**Table 2** describes the top barriers identified during the interviews and consultations combined. This question specifically asked, "Now that we've talked for a while, I want to ask again—what do you think are the top three barriers to clinicians using mHealth technologies with rural patients?" Several clinicians listed more than 3 barriers: this table only shows the top 3 responses tallied (the remaining answers were omitted). *Clinician familiarity* and *clinician time* accounted for 38.4% of identified barriers (15/39) to using mHealth apps or technologies with rural patients. Of the next 10 barriers, 8 were organizationally based, including EMR/data with 12.8% (5/39), and Health Insurance Portability and Accountability Act/Protected Health Information (HIPAA/PHI) with 10.2% (4/39) of total responses. Patient connectivity was combined with technology adoption and accounted for 10.2% (4/39) of the total responses, which was higher than anticipated.

**Table 2.** Top three barriers to mobile health adoption, identified during interviews and consultations (n=13; 39 responses).

Barriers	Responses, n (%)
Clinician familiarity	9 (23.1)
Clinician time	6 (15.4)
Electronic medical record/data	5 (12.8)
Health Insurance Portability and Accountability Act/Protected Health Information	4 (10.1)
Patient connectivity/technology	4 (10.1)
Organizational direction	2 (5.1)
Patient acceptance	2 (5.1)
Patient affordability	2 (5.1)
Uniformity of use	2 (5.1)
Hindering patient-provider communication	1 (2.6)
Technology reliance and limited patient face time	1 (2.6)
Usability of the app/technology	1 (2.6)

### Clinician Adoption

As shown in Table 1, 53.8% of participants reported encouraging the use of apps or WAMs with patients. When discussing MyFitnessPal, one clinician said the following:

*I've had a fair amount of success by cueing people into some things like My Fitness Pal. If I've assessed a patient and they have some readiness to change, I'll do some motivational interviewing around that and get people to buy into it and committed to making a lifestyle change. I'll specifically call out a tool to help them. I've had the patient pull it up in the app store and download it in the office and kind of get them started in terms of what you need to do. Both I and my spouse have used that tool in our practices and have had patients that have had tremendous results because of it.*

According to the participant, the couple had heard a speaker giving a presentation on this mHealth tool at a continuing medical education event, and also noted the following:

*...it wasn't driving patients to alter drug therapy; it was a way to track, record, and provide feedback. So, it was just a great way to engage patients. My spouse and I started using it, and then started recommending it to patients.*

Although the participants did note limitations and that it is not a one size fits all approach, they described notable successes and a positive outlook on the technology and its use in patient care:

*Does it help everybody? No. But a much higher percentage of patients had a positive change as a result of using that, than I ever saw had a positive change by me speaking with them and giving them a pamphlet to take home. It gave them something actionable, something that they had to report into. And when they would come back in, they would actually pull it up on their smartphones. We could kind of walk through and I could see they were*

*actually doing it and we were seeing the results, in terms of what their numbers looked like.*

This participant compared the data collection with a patient's home blood sugar monitoring:

*I don't go through all of those [data], I look at their average or I look at a range and say their blood sugar over the past two weeks ranged from fasting morning sugars of 100 up to 150 and I record things more globally that way.*

Although noting that he/she does put some of the patient-reported data into the EMR, he/she did prefer to have some type of automatic feed.

### Clinicians' Perceived Barriers Category 1: Personal (Clinician)

This category was established as a division of themes relating back to the clinician's personal barriers and not necessarily influenced by the organization or the patient. Of the 3 categories, this category yielded the smallest number of responses.

#### Personal/Clinician: Familiarity With Mobile Health Options

Although a majority of the participants reported engagement with mHealth on a personal level, a lack of familiarity with appropriate and effective clinical apps of mHealth apps and technologies was common. This also was the leader in the *top three barriers* question, with 23% of all possible answers being *familiarity*. In addition, 66.6% (2/3) of consultations also reported this barrier, including one who stated that "taking the time to learn about what apps might be useful" is a challenge for them.

Unless they were a personal adopter, mHealth technologies were not something these clinicians were actively looking for, including trying to find new tools that might help with their clinical practice. Of those that were not currently using mHealth or WAM technologies with patients (46.1%), familiarity was one of the biggest barriers. One clinician was quoted saying:

*How do you learn about it? To me that's a barrier. It's a time issue. I'm not going to spend my free time*



*looking for it. And if the clinic doesn't offer it to me I'm probably not going to learn a lot about it. If someone put on a continuing education course on mobile apps that you can use with your patients, I would probably go, but I have never seen one. There's never been one offered here. I'm not sitting around at my desk a lot in my medical practice doing nothing, where I would have time to look these things up.*

### **Clinicians' Perceived Barriers Category 2: Patient**

One clinician who spoke at length about patient barriers was also an adopter of mHealth and frequently promoted the use of several technologies with her patients. The majority of the patient barriers she identified during the interview were based on experience, not speculation. Those barriers are described in this section, although these barriers were not highlighted across all interviews or specialties. The data are not exclusively described in a way in which the patient barriers are assumptions based on speculation of the clinician or based on actual attempts to implement mHealth technologies. The results described here are a mix of both.

#### **Patient: Affordability**

Patient affordability emerged early on but was limited, and it was mentioned only once in the *top three barriers* of clinicians' question, although it was discussed during 2 other interviews as a potential barrier. Affordability does not appear to hold a strong argument for barriers to adoption. An assumption is that this was likely a more significant barrier 5 to 10 years ago as the mobile phone adoption rates were just starting to climb.

#### **Patient: Willingness and Adherence**

Coded as a patient acceptance response to the *top three barriers* question, the willingness and adherence of patients were discussed with 2 participants. One of those participants did use mHealth technologies with his/her patients, but he/she also saw a unique patient population not representative of the whole patient population. Although, as noted above, the specialty is omitted to protect the identity of the participant, in this case, further research with regard to barriers related specifically to the participants' specialty is warranted.

#### **Patient: Access and Connectivity**

Access and connectivity was a barrier identified on 10% of the responses to the *top three barriers* question. It also emerged throughout the interviews. Although conversations touched on this topic at some point, it was not always in a way that access and connectivity were viewed as barriers. For example, one participant talked about access and connectivity as a barrier when they had first started to promote MyFitnessPal with patients. This was, in part, because many of the patients still had flip phones at the time and were encouraged to use their home computer to keep track of activity. Use of home computers had limited success. However, since the increased adoption of mobile phones among that participant's patient population, the participant has seen this particular theme of access and connectivity fade as a barrier. Others also noted that access and connectivity is no longer a hurdle because cell coverage is so well spread across Marshfield Clinic's service area.

### **Clinicians' Perceived Barriers Category 3: Organizational**

#### **Organizational: Limited Time and Existing Knowledge**

All 3 consultations noted time limitation as a barrier to understanding, locating, evaluating, or implementing mHealth technologies with patients. Time also accounted for 15.3% of the responses to the *top three barriers* question. Clinicians, particularly physicians, have tight schedules with very limited time for continuing education, research, or exploration into new technologies. One participant was concerned about the potential *time sink* that an app, which facilitates patient-to-physician interaction, might impose. When asked what he thought other clinicians' barriers might be, he said:

*Once I've initiated the use of the app, will I be overwhelmed by the amount of information that's provided to me? More than I can handle, and at times disappoint the expectations of my patients because I can't? You know, they're emailing me or texting me, or doing whatever it is that this app is going to do, and I don't have time to respond to all my patients who are using the app. That's another potential barrier I think.*

#### **Organizational: Uniformity**

A lack of uniformity across clinicians and departments emerged from the interviews. Uniformity in this context is described as a common and guided practice among different clinicians to use similar mHealth tools. Clinicians who commented on uniformity noted that it should be considered in any type of organizational model moving forward. Several participants noted they did not have any type of departmental process in place but felt they would benefit from some type of uniformity or common footing. This concept links closely with the next section, which more specifically defines an organizational piece that is missing.

#### **Organizational: Lack of Policy/Direction**

The tone of the interviews highlighted a number of different concepts and ideas that were not necessarily identified in the participants' *top three barriers*. Lack of policy/direction emerged from every interview and consultation as a point of discussion. Although when asked the question of *top three barriers*, only 23% mentioned this as a top barrier. If asked, none of the participants were aware of a policy or specific process, either departmental or organizational, in regard to the review or screening of a new mHealth technology that a clinician would like to put into practice with his/her patients. As it stands, those that are currently encouraging patient use of some type of mHealth technology have done so with no authoritative oversight other than their own or sometimes peer recommendation and review. Nevertheless, of all the technologies mentioned, the majority were health and fitness related (eg, MyFitnessPal and Fitbit), with others that were more specialty focused (data omitted to protect participant anonymity).

Clinicians also discussed additional challenges of understanding organizational policy and procedures, especially regarding instances where precedence has not already been established.

One conversation that followed a question regarding the process for integrating a new mobile technology led to one clinician stating that:

*I came here and I now understand Dilbert cartoons, because you can't just do it. In fact, in my first year I innocently tried to do a few things and got my hand slapped because I was going outside of channels and I didn't even know channels existed yet and that wasn't mHealth sorts of things, but other things.*

The participant did note that this was in regard to another type of issue, not mHealth, but still held weighted and influenced the subjects' perception of approval processes. Another participant noted that the act of "trying to get all the different people on board, and even knowing who to ask sometimes, is really difficult."

In the absence of a formal written policy, the policy section can be further broken down into the following sections: (1) information systems security and related policies, (2) legal concerns with liabilities, and (3) app quality and patient safety. Surprisingly, none of these 3 sections were discussed by participants with any significance. Policy was discussed from an organizational direction perspective. Security was discussed, but from a patient data and patient privacy perspective, and legal concerns and liabilities were discussed and are further noted below. There were just 2 mentions of app/WAM quality, which are mentioned later in the paper, and there was no discussion about patient safety.

#### **Organizational: Health Insurance Portability and Accountability Act/Protected Health Information**

Responses that were combined into this code included *HIPAA*, *PHI*, *privacy*, *security*, *confidentiality*, and *patient privacy*. Clinicians were outspoken about this as a barrier to how they practice or how they would like to practice medicine, with 10.2% of responses in the *Top 3 barriers* question. One clinician noted that "HIPAA turns out to be a barrier, because everyone is so afraid of messing it up," and another clinician referred to the general concept as "HIPAA-phobia" and often a prohibitive element within the health care industry. Another participant referred to HIPAA regulations as "very confusing and difficult to interpret."

#### **Organizational: Data and the Electronic Medical Record**

Almost all the interviews and consultations voiced their desire to integrate patient-reported data into the EMR. Clinicians also noted that the current system would not be able to handle it in an efficient way. None of the participants commented in detail on the constraints and technological challenges of mapping

patient-reported data to clinical systems or how best to store that data for later retrieval and how it would be used with clinical decision support. In summary, clinicians are already using that data and would like to have them somehow integrated into the EMR.

When talking with one of the participants about the current Cattails MD and My Marshfield Clinic System that is used at MCHS, the clinician was quoted saying:

*We have to learn how to leverage tools like that, in concert with some of these apps and other things, to say how can we exchange information out of your Fitbit that helps connect your care team with how you're doing? Are you getting your steps every day? Are you on track with your caloric intake and all of that stuff? So, I think there's some opportunity to marry those through a tool like the Portal.*

Another participant who was interested in using that data noted he/she often would make a note of using the MyFitnessPal with the patient, but there is no structured or uniform method of capturing such data. When asked about the process of recording it, the clinician noted that he/she did not transfer detailed data such as weights, steps, and activities, "mostly because it's just an onerous task to transcribe those over."

#### **Organizational: Legal and Liability**

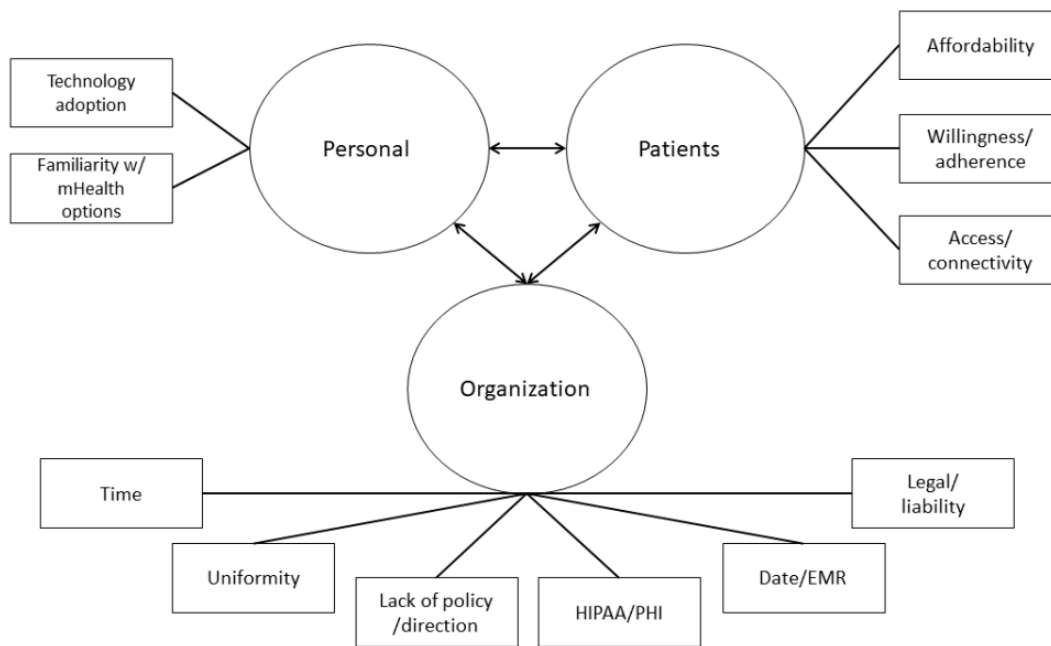
Legal refers to the Legal Department within the organization. As just a code, legal/liability holds little meaning and could be easily linked to a clinician barrier, as personal liability or malpractice; that particular linkage was only identified as a barrier during the interviews once. One clinician noted that "some are afraid that by recommending something, if anything ever went wrong, that they would somehow be personally liable." Meanwhile, the linkage between the *organization* and legal/liability was more frequent.

All 3 consultations noted specific organizational constraints in regard to using mHealth apps and technologies with patients. Of the 3 consultations, 2 cited *legal* as one of the barriers in trying to get an organizational decision. One of the consultations suggested departmental consensus and recommendations as a way to move new techniques through legal and into practice.

#### **Thematic Map**

Figure 2 displays the coded responses of the top barriers identified throughout the interview process as well as the thematic placement in identified categories of personal, patients, and organizational.

**Figure 2.** Thematic map of clinician-identified barriers. EMR: electronic medical record; HIPAA/PHI: Health Insurance Portability and Accountability Act/Protected Health Information; mHealth: mobile health.



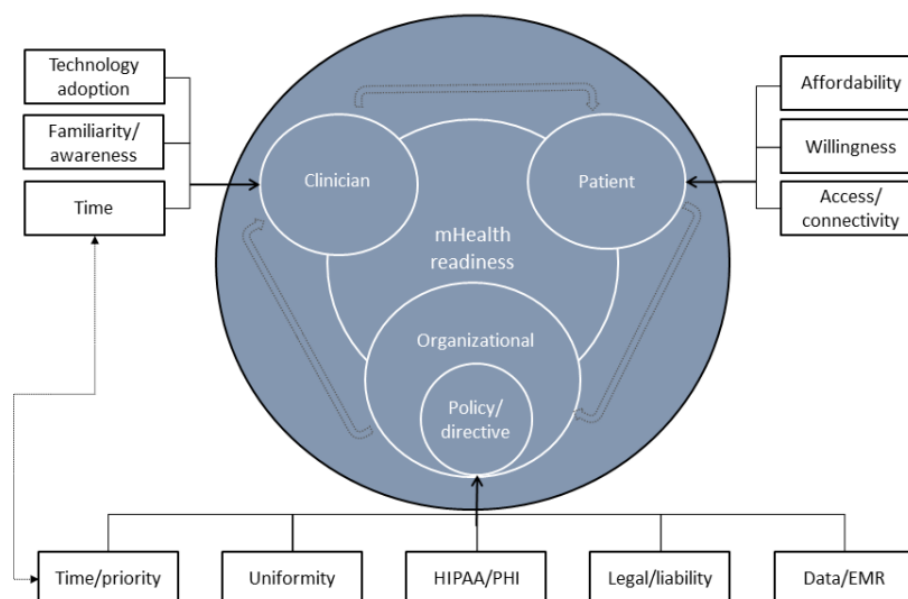
**Mobile Health Readiness Model Refinement**

The original model of mHealth readiness was based on a study by Khatun et al [24]. As the refinement process began, 3 pillars and primary dimensions of readiness were formed: technological, motivational, and clinician. *Technological* is primarily the technical ability for the clinician and his/her patients to be able to connect and engage with an mHealth app. Both would need to have a device (eg, mobile phone or tablet), connectivity (home for patients, office/examination room for clinicians) through mobile broadband, and/or connectivity through broadband (eg, Wi-Fi). This study did not seek to assess the connectivity and overall readiness of the patients, just the

perceived patient readiness in the opinion of clinicians. *Motivational* refers to the willingness of clinicians to engage with patients in mHealth. Clinician and organizational readiness are the key areas of inquiry for this project; thus, the refinement began.

As the interviews unveiled thematic linkages between codes and categories, the framework refinements began to take shape as the model displayed in Figure 3. The technological pillar was eliminated early, as the data showed this is a very small barrier, if at all. The motivational barrier was also eliminated from the framework once enough data showed that clinicians who participated in this study were adopters themselves and promoters of mHealth technologies with patients.

**Figure 3.** Conceptual model for assessing necessary conditions for rural health care’s mHealth readiness, with an emphasis on clinician-perceived barriers. EMR: electronic medical record; HIPAA/PHI: Health Insurance Portability and Accountability Act/Protected Health Information; mHealth: mobile health.



This is a framework that has not yet been tested. The model should be used with the assumption that patients and clinicians are already motivated and that technology is not a barrier to adoption. As both those pillars were eliminated from the original model, they are not to be included in assessments using the new model.

The coded barriers form a basis of decision making and readiness for each of the 3 categorical pillars, now labeled as *clinician*, *patient*, and *organizational*. These 3 primary dimensions have been defined as pinnacle and foundational elements of assessing overall readiness to adopt mHealth in patient care. The double-wide arrows represent an informational flow from organization to clinician to patient and back to the organization in the form of feedback and patient experience. This framework was developed for MCHS specifically and for other health care organizations generally.

There is an arrow connecting *time* from the clinician and the organizational categories. This represents a unique barrier that can be placed in either or both of the categories. For example, the clinician's time was a clear barrier to adoption, but it is not known if that is an organizational barrier related to not enough time allowed in the schedules for research/education. The other complexity is that time and priority are potential barriers for the organization itself. Owing to the current cultural climate and the competitive expansion into new markets, including the acquisition of a hospital and its thousands of staff members, the organizational priority and available administrative time to dedicate to mHealth may be limited.

Although created based on international work, the model may not be as easily adaptable to international health care organizations. It could possibly serve as a framework, but it would have to be thoroughly tested, as there is so much variation in external variables and the potential for different barriers in each organization, region, or country.

## Discussion

### Principal Findings

This project sought to understand the barriers to adoption, with a strong assumption that some innovative clinicians were already using mHealth tools with rural patients. This assumption was confirmed during the interview process, and in fact, an even higher rate of adoption was found than anticipated. The small sample size of our study is not statistically significant, and it is likely not representative of the entire MCHS clinician body. Nevertheless, our results strongly suggest that health care organizations should consider investing in mHealth analysis, tool development, and the promotion/recommendation of sanctioned tools for clinicians to use with their patients.

Many studies have investigated effectiveness and progress of mHealth adoption, particularly in developing and low-income countries [25-29]. Yet, much work remains to be done in the United States, particularly among rural health care organizations, which is a rapidly changing landscape as mergers, acquisitions, and hospital closures have shaken up the industry in recent decades [30,31].

Organizational mHealth readiness depends on a number of factors, not all of them addressed in this study. However, the readiness of clinicians was identified, as well as barriers preventing them from furthering their mHealth practices. A follow-on study could help to ascertain the perceived relative importance of the barriers identified. For example, it will be useful to discover whether *time* and *familiarity* fall under personal/clinician or under organizational category. It could be argued that the organization is responsible for the education, training, and communication about new technologies for its physicians and staff. It could as well be argued that the clinician is responsible for keeping abreast of new opportunities, including mHealth, for improving patient care.

One of the initial consultation interviews was with a participant who was very interested in the project and what our team might learn about mHealth adoption. Near the end of the interview, the author described more about some of what researchers and clinicians have been using, as described in the literature. The participant was very interested, saying, "you've been holding that back this whole time?" The interview schedule, as described earlier, was semistructured. Later in the interview process, as in this case, there was often additional discussion about what falls within the realm of mHealth and what some clinicians and organizations have had success with, as available in the scientific literature. The clinician also said, "these are things that we should start talking about as a group, immediately. These are inspiring."

The same clinician continued to describe some of the possibilities where mHealth would be useful. The clinician then added, "I think about maybe kids that are in that tween to teenager years that are having challenges to get more active and curb their weight, lose weight, do something like that. I'm inspired; I think I should start something like this. That would be great." The author had not planned on discussing mHealth as a way to influence the participant in any way. The tone of any information was neutral, as much as possible, so as to not persuade or motivate the participant to consider mHealth technologies. Limitations, challenges, and barriers were discussed, as well as some of the interventions in the literature. This particular physician's reaction is another example that suggests health care organizations should consider investing in mHealth adoption.

Other interviews yielded similar results and further showed that the organization is not yet prepared for innovative clinicians. mHealth as a tool for improving patient outcomes is already being embraced by some, and others may be very interested in its potential. One clinician's perceived value was described as the time between visits, which is otherwise unmonitored and loosely reported with unknown accuracy during the next appointment. That clinician commented:

*Where I see the real value, is in between visits. So, I may only see a patient with diabetes every 6 months or see a patient with high cholesterol once a year. So, you might cue them in to My Fitness Pal and talk about lifestyle modifications. But do I bring them back in two months specifically to say, are you using your app? No. If there's no clinical reason to bring them*

*back and it's just preventative care, you might not see that person for a year. But if that was then connected and you have the ability to make that link you could actually have some tracking of patients' in-between time and provide some positive reinforcements. So let's say the patient starts using a tool like My Fitness Pal and they hit a ten pound weight loss. Well, that would be great if that could somehow trigger onto my worklist so that our practice could do an outbound communication to that person as a "thumbs up" you hit the ten pound weight loss! That's awesome! Let us know how we can help. Let people know that you're involved in their care and you're recognizing and applauding them for their efforts and helping to fuel the fire.*

Although there are some risks that would need to be managed and mitigated, if organizations were prepared to manage mHealth, it is very likely that clinicians could improve the quality of care for their patients. Our preliminary results suggest that many physicians may see such benefits of mHealth technologies. However, many organizations, including MCHS, are not yet prepared to prescribe or prohibit the use of mHealth technologies. Having said that, there are some recommendations that come from this study, particularly the formation and promotion of a Resource menu. Those details are further described in the section Unexpected Outcomes.

### Unexpected Outcomes

The project yielded a number of unexpected outcomes from the interviews. These outcomes suggest potential future projects as well as immediately practical actions to take based on recommendations from the participants in this study. Outcomes are also discussed in the section below.

#### Mobile Health Resource Menu

Participants showed a desire for guidance and instruction, especially with regard to recommendations on what to use with patients, what others are having success with, and what has positively affected patient outcomes. Several participants noted that a seminar, webinar, or presentation on mHealth apps would be helpful, and they would attend. Participants were interested in a *menu* that offers the organization's recommended mobile apps or technologies. One of the participant's responses included:

*I would love to have some sort of databank of vetted apps that we could recommend for our patients, without each and every one of us just playing around and trying to figure it out for ourselves. To say that, if you're going to motivate patients for weight loss, here are some good apps that have been looked at by the organization, are as free of bias as you can be, and safe. I know there has been some discussion, and once in a while I hear bits and pieces, but I don't think there's anything formal in place.*

After several subjects requested this type of document during interviews, it was added to the schedule as a question for others. Of those who were asked if they would be interested in seeing

such a list, all the participants said they would. One was quoted saying:

*I think having some of that stuff sanctioned and clarified for our providers who are generally mystified by all of this, because most of them only know what they've taken initiative to learn on their own. And as an organization, I think we can do a lot better job of steering and helping people to go through those thousands of different options and say here's what we consider to be the cream of the crop. And give providers, and maybe even patients, a library of things that they could go to that might be helpful for them.*

Although app and technology quality never arose as a solidified theme, some clinicians hinted to it in ways such as this participant's comment:

*I don't think we want to promote an app that is a 30-day grapefruit diet. So who's going to vet all that stuff when the world of apps is just exploding. I think it's a real issue for us.*

A quality review of mHealth technologies certainly needs to be a part of any type of formal vetting and recommendation process.

### Accuracy and Honesty in the Data

The honesty of participants is often critiqued in qualitative research, questioning the integrity of participant responses, and thus the data. This study was able to eliminate some of that critique through the context that bookended said responses. Participants were honest about experiences that they would have been behooved to lie about. For example, the use of unsanctioned apps and technologies in patient care was described at length by clinicians who knowingly were practicing outside the purview of higher authority, guidance, or direction. Therefore, it seems as though all participants were providing candid responses.

### Limitations

Clinicians noted they were using various mHealth technologies with patients, and they also expressed excitement of those interventions' results. Several examples were intentionally omitted from this paper because of the possibility of identifying participants through their clinical specialty and the specificity of the mHealth apps being used with patients. Multiple institutions and a broader geographic reach for participant recruitment would be ideal. Not only would data be more representative of the population, we would also more descriptively report the findings. Specifically, we could have described participants' clinical specialties and the unique technologies and apps that are being used therein. Those identifiers were stricken from this paper.

### Sample Size and Time of Clinicians

Recruitment was a challenge, as is often the case in health care research [32]. Clinicians, particularly physicians, have a demanding schedule and are often tasked with seeing as many patients as possible throughout their workday. Even the nonphysician staff who were contacted (eg, physical therapists, clinical psychologists, and nurse practitioners) were either

nonresponsive or not able to accommodate a 45-min to 1-hour interview. Those who did participate were interested in shortening it during the scheduling process. Several interviews were interrupted for a patient care issue that needed to be addressed, but they were reconvened minutes later. The only participants who were not interrupted had dedicated research time as part of their contract, were currently traveling, or were meeting at the end of their workday.

## Conclusions

Mobile technologies are not often restricted by geographic boundaries or distance, traffic conditions, or weather. On the basis of these factors alone, rural populations may have more to gain through the use of these technologies. Beyond rural, the application of these technologies may also be critical to other populations, such as parents with two jobs who have a difficult time taking vacation for appointments.

Research needs to further assess organizational/administrative perceptions, which may include interview with legal teams and decision-making administrative leaders of health care organizations. What has been found through this research and a review of the literature is that clinicians are interested in or already using mHealth technologies with patients. In addition, patients are interested in or already using mHealth technologies, and some research has shown positive outcomes from mHealth interventions. Meanwhile, health care organizations have not yet actively embraced or supported its use. Organizational leadership should review these studies and the results so that they can make more informed decisions to proceed with formalized and sanctioned adoption of secure and safe evidence-based mHealth tools. Removing the identified barriers is necessary for adoption, but not sufficient for successful implementation or readiness.

## Acknowledgments

The authors would like to thank the clinicians who participated in this study.

## Conflicts of Interest

None declared.

## References

1. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies. Switzerland: World Health Organization; 2011.
2. mHealth Evidence. About mHealth Evidence URL: <https://www.mhealthevidence.org/about> [accessed 2019-10-21]
3. WebCite. The Wearables Database URL: <https://www.webcitation.org/71ZHrfURI> [accessed 2018-08-10] [[WebCite Cache ID 71ZHrfURI](#)]
4. Domingo-Salvany A. The science of real-time data capture: self-reports in health research. *J Epidemiol Commun Health* 2008;62(5):471. [doi: [10.1136/jech.2007.068551](https://doi.org/10.1136/jech.2007.068551)]
5. Sakakibara BM, Chakrabarti S, Krahn A, Mackay MH, Sedlak T, Singer J, et al. Delivery of peer support through a self-management mHealth intervention (Healing Circles) in patients with cardiovascular disease: protocol for a randomized controlled trial. *JMIR Res Protoc* 2019 Jan 11;8(1):e12322 [[FREE Full text](#)] [doi: [10.2196/12322](https://doi.org/10.2196/12322)] [Medline: [30635261](https://pubmed.ncbi.nlm.nih.gov/30635261/)]
6. Silva EH, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. *J Cancer Surviv* 2019 Feb;13(1):97-107. [doi: [10.1007/s11764-018-0730-8](https://doi.org/10.1007/s11764-018-0730-8)] [Medline: [30635865](https://pubmed.ncbi.nlm.nih.gov/30635865/)]
7. Steinhubl SR, Muse ED, Topol EJ. Can mobile health technologies transform health care? *J Am Med Assoc* 2013 Dec 11;310(22):2395-2396. [doi: [10.1001/jama.2013.281078](https://doi.org/10.1001/jama.2013.281078)] [Medline: [24158428](https://pubmed.ncbi.nlm.nih.gov/24158428/)]
8. Ben-Zeev D. Mobile health for all: public-private partnerships can create a new mental health landscape. *JMIR Ment Health* 2016 Jun 6;3(2):e26 [[FREE Full text](#)] [doi: [10.2196/mental.5843](https://doi.org/10.2196/mental.5843)] [Medline: [27268283](https://pubmed.ncbi.nlm.nih.gov/27268283/)]
9. Bauer AM, Iles-Shih M, Ghomi RH, Rue T, Grover T, Kincler N, et al. Acceptability of mHealth augmentation of Collaborative Care: a mixed methods pilot study. *Gen Hosp Psychiatry* 2018;51:22-29 [[FREE Full text](#)] [doi: [10.1016/j.genhosppsy.2017.11.010](https://doi.org/10.1016/j.genhosppsy.2017.11.010)] [Medline: [29272712](https://pubmed.ncbi.nlm.nih.gov/29272712/)]
10. Rönby S, Lundberg O, Fagher K, Jacobsson J, Tillander B, Gauffin H, et al. mHealth self-report monitoring in competitive middle- and long-distance runners: qualitative study of long-term use intentions using the technology acceptance model. *JMIR Mhealth Uhealth* 2018 Aug 13;6(8):e10270 [[FREE Full text](#)] [doi: [10.2196/10270](https://doi.org/10.2196/10270)] [Medline: [30104183](https://pubmed.ncbi.nlm.nih.gov/30104183/)]
11. Thompson D, Mackay T, Matthews M, Edwards J, Peters NS, Connolly SB. Metadata correction: direct adherence measurement using an ingestible sensor compared with self-reporting in high-risk cardiovascular disease patients who knew they were being measured: prospective intervention. *JMIR Mhealth Uhealth* 2018 Apr 27;6(4):e13 [[FREE Full text](#)] [doi: [10.2196/mhealth.8317](https://doi.org/10.2196/mhealth.8317)] [Medline: [30578223](https://pubmed.ncbi.nlm.nih.gov/30578223/)]
12. research2guidance. 325,000 mobile health apps available in 2017 – Android now the leading mHealth platform URL: <https://research2guidance.com/325000-mobile-health-apps-available-in-2017/> [accessed 2019-10-21] [[WebCite Cache ID 71ZIAzZe7](#)]

13. Business Wire. 2015. IMS Health Study: Patient Options Expand as Mobile Healthcare Apps Address Wellness and Chronic Disease Treatment Needs URL: <https://www.businesswire.com/news/home/20150917005044/en/IMS-Health-Study-Patient-Options-Expand-Mobile> [accessed 2019-10-20] [WebCite Cache ID 71ZIfip5U]
14. Statista. Number of Apps Available in Leading App Stores as of 3rd Quarter 2019 URL: <http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/> [accessed 2019-10-20] [WebCite Cache ID 71ZIpV40V]
15. CVi Connect. URL: <https://cviconnect.co/> [accessed 2019-10-23]
16. Khatun F, Heywood AE, Ray PK, Hanifi SM, Bhuiya A, Liaw S. Determinants of readiness to adopt mHealth in a rural community of Bangladesh. *Int J Med Inform* 2015 Oct;84(10):847-856. [doi: [10.1016/j.ijmedinf.2015.06.008](https://doi.org/10.1016/j.ijmedinf.2015.06.008)] [Medline: [26194141](https://pubmed.ncbi.nlm.nih.gov/26194141/)]
17. Bernard HR. *Research Methods in Anthropology: Qualitative and Quantitative Methods*. Third Edition. Walnut Creek, CA: AltaMira Press; 2002.
18. Klabunde CN, Willis GB, McLeod CC, Dillman DA, Johnson TP, Greene SM, et al. Improving the quality of surveys of physicians and medical groups: a research agenda. *Eval Health Prof* 2012 Dec;35(4):477-506. [doi: [10.1177/0163278712458283](https://doi.org/10.1177/0163278712458283)] [Medline: [22947596](https://pubmed.ncbi.nlm.nih.gov/22947596/)]
19. Leung GM, Johnston JM, Saing H, Tin KY, Wong IO, Ho L. Prepayment was superior to postpayment cash incentives in a randomized postal survey among physicians. *J Clin Epidemiol* 2004 Aug;57(8):777-784. [doi: [10.1016/j.jclinepi.2003.12.021](https://doi.org/10.1016/j.jclinepi.2003.12.021)] [Medline: [15485729](https://pubmed.ncbi.nlm.nih.gov/15485729/)]
20. James KM, Ziegenfuss JY, Tilburt JC, Harris AM, Beebe TJ. Getting physicians to respond: the impact of incentive type and timing on physician survey response rates. *Health Serv Res* 2011 Feb;46(1 Pt 1):232-242 [FREE Full text] [doi: [10.1111/j.1475-6773.2010.01181.x](https://doi.org/10.1111/j.1475-6773.2010.01181.x)] [Medline: [20880042](https://pubmed.ncbi.nlm.nih.gov/20880042/)]
21. Roger G. *Diffusion Of Innovations*. Third Edition. New York: Free Press; 2003.
22. Wainwright M, Russell A. *Social Research Update. Using NVivo Audio-Coding: Practical, Sensorial and Epistemological Considerations* URL: <http://sru.soc.surrey.ac.uk/SRU60.pdf> [accessed 2018-08-10] [WebCite Cache ID 71ZKMxyZR]
23. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *Am J Eval* 2006;27(2):237-246. [doi: [10.1177/1098214005283748](https://doi.org/10.1177/1098214005283748)]
24. Epstein RM, Street RL. The values and value of patient-centered care. *Ann Fam Med* 2011;9(2):100-103 [FREE Full text] [doi: [10.1370/afm.1239](https://doi.org/10.1370/afm.1239)] [Medline: [21403134](https://pubmed.ncbi.nlm.nih.gov/21403134/)]
25. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmarts 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](https://doi.org/10.1093/jamia/ocv052)] [Medline: [26078410](https://pubmed.ncbi.nlm.nih.gov/26078410/)]
26. Khoja S, Scott RE, Casebeer AL, Mohsin M, Ishaq AF, Gilani S. e-Health readiness assessment tools for healthcare institutions in developing countries. *Telemed J E Health* 2007 Aug;13(4):425-431. [doi: [10.1089/tmj.2006.0064](https://doi.org/10.1089/tmj.2006.0064)] [Medline: [17848110](https://pubmed.ncbi.nlm.nih.gov/17848110/)]
27. Leon N, Daviaud E, Schneider H. *BMC Medical Informatics and Decision Making. Applying a Framework for Assessing the Health System Challenges to Scaling Up mHealth in South Africa* URL: <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-12-123> [accessed 2019-10-23]
28. Cho YM, Lee S, Islam SM, Kim SY. Theories applied to m-Health interventions for behavior change in low- and middle-income countries: a systematic review. *Telemed J E Health* 2018 Oct;24(10):727-741 [FREE Full text] [doi: [10.1089/tmj.2017.0249](https://doi.org/10.1089/tmj.2017.0249)] [Medline: [29437546](https://pubmed.ncbi.nlm.nih.gov/29437546/)]
29. Mengesha W, Steege R, Kea AZ, Theobald S, Datiko DG. Can mHealth improve timeliness and quality of health data collected and used by health extension workers in rural Southern Ethiopia? *J Public Health (Oxf)* 2018 Dec 1;40(suppl\_2):ii74-ii86 [FREE Full text] [doi: [10.1093/pubmed/fdy200](https://doi.org/10.1093/pubmed/fdy200)] [Medline: [30551131](https://pubmed.ncbi.nlm.nih.gov/30551131/)]
30. Williams DR, Spaulding TJ. An empirical study of the determinants of location of hospital mergers and acquisitions: 1997-2016. *J Health Care Finance* 2018;45(1) [FREE Full text]
31. Ricketts TC. The changing nature of rural health care. *Annu Rev Public Health* 2000;21:639-657. [doi: [10.1146/annurev.publhealth.21.1.639](https://doi.org/10.1146/annurev.publhealth.21.1.639)] [Medline: [10884968](https://pubmed.ncbi.nlm.nih.gov/10884968/)]
32. Miller K, McKeever P, Coyte PC. Recruitment issues in healthcare research: the situation in home care. *Health Soc Care Commun* 2003 Mar;11(2):111-123. [doi: [10.1046/j.1365-2524.2003.00411.x](https://doi.org/10.1046/j.1365-2524.2003.00411.x)]

## Abbreviations

- EMR:** electronic medical record
- HIPAA:** Health Insurance Portability and Accountability Act
- MCHS:** Marshfield Clinic Health System
- mHealth:** mobile health
- PHI:** Protected Health Information
- WAM:** wearable activity monitor

*Edited by G Eysenbach; submitted 10.08.18; peer-reviewed by R Wynn, L Ranandeh, A Bairagi; comments to author 08.01.19; revised version received 06.06.19; accepted 04.09.19; published 08.11.19.*

*Please cite as:*

*Weichelt B, Bendixsen C, Patrick T*

*A Model for Assessing Necessary Conditions for Rural Health Care's Mobile Health Readiness: Qualitative Assessment of Clinician-Perceived Barriers*

*JMIR Mhealth Uhealth 2019;7(11):e11915*

*URL: <http://mhealth.jmir.org/2019/11/e11915/>*

*doi: [10.2196/11915](https://doi.org/10.2196/11915)*

*PMID: [31702564](https://pubmed.ncbi.nlm.nih.gov/31702564/)*

©Bryan Weichelt, Casper Bendixsen, Timothy Patrick. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 08.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Making Sense of Negative Findings from Mobile Attention Bias Modification Interventions for Individuals with Addictive Disorders: Quantitative Feasibility Study

Melvyn Zhang<sup>1</sup>, MBBS, MRCPsych; Jiangbo Ying<sup>1</sup>, MBBS, MRCPsych; Syidda B Amron<sup>1</sup>, Bsc(Hons); Zaakira Mahreen<sup>1</sup>, Bsc(Hons); Guo Song<sup>1</sup>, PhD; Daniel SS Fung<sup>2</sup>, MBBS, MMED; Helen E Smith<sup>3</sup>, DM, FFPHM

<sup>1</sup>National Addictions Management Service, Institute of Mental Health, Singapore, Singapore

<sup>2</sup>Department of Developmental Psychiatry, Institute of Mental Health, Singapore, Singapore

<sup>3</sup>Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore

**Corresponding Author:**

Melvyn Zhang, MBBS, MRCPsych  
National Addictions Management Service  
Institute of Mental Health  
10 Buangkok Green Medical Park  
Singapore, 539747  
Singapore  
Phone: 65 63802504  
Email: [melvynzhangweibin@gmail.com](mailto:melvynzhangweibin@gmail.com)

## Abstract

**Background:** Advances in experimental psychology have led to a better understanding of unconscious, automatic processes that result in individuals relapsing into their substance-using habits. While some reviews have demonstrated the effectiveness of bias retraining of these unconscious biases, there have been other reviews that have highlighted that bias retraining is not always effective. Other studies have revealed there was no baseline biases among some participants. An examination of mobile bias retraining interventions has also revealed mixed results, with some reporting effectiveness and others null findings. A recent feasibility and acceptability study, done by the authors, revealed that 53% of participants have had no baseline biases and 21% of those with positive baseline biases did not have a positive change in magnitude following intervention.

**Objective:** The aim of this paper was to explore potential variables (demographic and clinical) that could account for the negative baseline biases in the prior feasibility and acceptability study, and to discuss some of the factors that could account for the absence of baseline biases. We also explored potential reasons for why there was no reduction in the magnitude of attentional biases among individuals with baseline biases.

**Methods:** Participants who were in the rehabilitation phase of their treatment were invited to participate. During the study they had to complete a set of baseline questionnaires, and on each day that they were on the ward they had to complete an attention bias assessment and modification task and rate their cravings using a visual analogue scale. Attention bias was deemed to be present if individuals had a positive score.

**Results:** In our study, 53% (16/30) of individuals did not present with baseline attentional biases, and among those with positive baseline biases a total of 21% (3/14) of participants did not have a reduction in the overall magnitude of attentional biases. Chi-square analyses undertaken to compare the demographic characteristics of participants with and without baseline biases did not reveal any significant findings. However, with respect to clinical characteristics, those who had positive baseline biases had experimented with more substances.

**Conclusions:** Our study is one of the first to have explored negative findings in attention bias modification interventions for individuals with addictive disorders. We postulate that several factors could account for the absence of baseline biases and there being no changes following bias retraining. Future research ought to take into consideration these factors.

(*JMIR Mhealth Uhealth* 2019;7(11):e16325) doi:[10.2196/16325](https://doi.org/10.2196/16325)

**KEYWORDS**

attention bias; cognitive bias; psychiatry

## Introduction

Advances in experimental psychology over the years have led to a better understanding of unconscious, automatic processes [1] that result in individuals relapsing into their substance-using habits. Such processes include attentional biases and approach biases. Attentional biases refer to the automatic tendencies for attention to be preferentially allocated towards substance-related stimuli [2], whereas approach biases refer to the automated tendencies that result in individuals reaching out for substance stimuli [3]. The dual-process theoretical model explains that automatic biases develop as the chronic use of a substance results in increased automatic processing of the substance cue, with a resultant inhibition in the normal conscious cognitive control processes [4]. Tasks, such as the visual probe task and the Stroop task, are used in the assessment and modification of these automatic biases. The effectiveness of bias modification has been extensively evaluated. In a previous review, which considered studies involving participants with alcohol and tobacco use problems, bias modification interventions were deemed to be effective with a moderate effect size [5]. In a recent review from Heitman et al [6], they reported that only 10 of the identified 18 studies provided evidence for symptom change and effectiveness of attention bias modification. Of these 18 studies, 9 were focused on alcohol use, 6 were on nicotine use, and 3 were on opiate use [6]. Overall, the review found that multi-session interventions were more effective, especially for those who had adopted the alcohol attention control program [5].

Conventionally, attention bias modification interventions have been delivered in the laboratory, but advances in electronic and mobile health technologies have transformed the delivery of these interventions [7]. Zhang et al [8], in a review of a mobile cognitive bias intervention, reported that there had been at least 7 studies reporting that using the intervention with new technologies was more effective. However, one study (Robinson et al [9]) examining attention bias modification among smokers reported bias training to be ineffective, as the mobile intervention did not produce any changes in attentional biases. Similarly, Zhu et al [10], in their study examining the effectiveness of a mobile-based, computerized, cognitive addiction–therapy intervention, reported that their intervention led to no reduction in overall attention biases despite there being improvements in cognitive impairment and impulse control. In our recently published feasibility and acceptability study of a mobile attention bias intervention for inpatients with various addictive disorders (alcohol, opioids, cannabis, or stimulants) [11], we found that 16/30 participants (53%) did not have baseline attentional biases. Among those with baseline biases, attentional retraining did reduce the magnitude of the biases in all but three of the participants [11].

Negative findings are common in attention bias modification interventions. Negative findings are found in two contexts: at baseline assessment where there is an absence of biases, and when bias retraining fails to reduce the magnitude of attentional biases. Negative findings (the absence of effectiveness of bias modification) were highlighted in Heitman et al's [6] review of the effectiveness of bias retraining for individuals with opioid

use disorders. Sharpe [12], in his commentary, highlighted a recent study involving attention bias modification for pain in children which also had negative findings. In this study [13], participants did not present with baseline biases. One of the factors that Sharpe [12] highlighted was that the experimental procedure was not appropriate for children, and prior meta-analysis has highlighted the need to present stimulus of longer presentation rates [14] for these biases to be detected. Factors like the quantity and frequency of substance use and cravings could also affect the magnitude of attentional biases [15].

In our prior feasibility and acceptability study, there were at least 53% (16/30) of individuals who did not present with baseline biases, and of the participants with baseline biases, 21% (3/14) did not have a reduction of biases following bias intervention [10]. Thus, the aim of this paper is to explore potential variables (demographic and clinical) that could account for negative baseline biases, and to discuss some of the factors that could account for the absence of baseline biases. We also explored potential reasons for there to be no reduction in the magnitude of attentional biases among individuals with baseline biases. An exploration of these negative findings is pertinent, as it has a consequential implication on future attention bias research.

## Methods

### Overview

The methods for this study have been previously published as a study protocol [16] and in Zhang's published feasibility and acceptability study [11]. Participants, who were in the rehabilitation phase of their treatment program at the National Addictions Management Service, Institute of Mental Health, were invited to participate. Participants who agreed to participate completed an informed consent form. The minimum recruitment target for the study was set at 30 participants.

### Ethical Approval

This study has been approved by the National Healthcare Group's Domain Specific Research Board with the following reference number (2018/00316) on May 2, 2018.

### Inclusion and Exclusion Criteria

Patients were included in the study if they were aged between 21–65 years old, were diagnosed with a primary psychiatric disorder of alcohol, opioid, cannabis, stimulants, or polysubstance dependence, were diagnosed with polysubstance dependence, with alcohol, opioid, cannabis or stimulants as the main substance of use, were able to read and write in English, and were capable of using a smartphone or tablet device. Patients were excluded from the study if they had a known history of cognitive impairment or dementia, if they had a history of seizures or a prior history of withdrawal seizures, if they had a history of migraines triggered by flashing lights, and if they had any moderate to severe comorbid psychiatric disorders based on clinical assessment.

## Measures

Upon recruitment into the study, baseline demographic and clinical information were acquired from the participants, including: nationality, gender, marital status, race, religion, highest level of education, housing conditions, current substance use, previous treatment history, chronic diseases (psychiatric or physical disorders), and any current psychiatric medication use. Participants also completed the modified Addiction Severity Index (ASI)–Lite, the Severity of Drug Dependence Scale (SDS), and the Short Form (SF-12) questionnaires. The ASI-Lite collects information about the participants' drug and alcohol use. All participants were asked about their alcohol and substance use in the last 30 days, in the last last month, and over their lifetime. The SDS questionnaire has 5 questions, all of which are focused on the psychological components of dependence. The SF-12 assessed the participants' self-reported quality of life.

Upon the completion of these questionnaires, participants were oriented on how they should use the application. On the first day of the intervention, participants were required to undertake an assessment task, then an intervention task, and then rest for 15 minutes before completing another assessment task. On the subsequent days of the intervention, participants were asked to complete an intervention task and then rest for 10 minutes before completing the assessment task. All participants were also asked to complete a visual-analogue craving scale before and upon the completion of each of the bias modification tasks. Participants were to undertake the tasks only if it was a weekday and were exempted from undertaking these tasks on weekends or public holidays.

In the assessment task, participants were presented with a central fixation cross for 500 milliseconds. Following the disappearance of the fixation cross, they were presented with a set of images, with one image being a substance image and the other a neutral image. The images would then disappear, and a probe would replace the position of one of these images. Participants were instructed to indicate the position of the probe by pressing on buttons within the application as rapidly as they could. In the assessment task, the probe would replace the substance-related image and the neutral image equally. In the intervention task, the probe would always replace the neutral image for the successful retraining of the attentional processes.

## Statistical Analyses

Data collated was analyzed using SPSS Version 22 (IBM Corp, Armonk, New York, United States). Baseline demographic information of the subjects was summarized using descriptive statistics. The presence of attentional biases was determined based on the mean reaction times taken to respond to the position of the probes that replaced either the drug or neutral stimuli. The formula used for the computation of attentional biases was

$$\frac{T_1 - T_0}{T_1 + T_0}$$
, where  $T_1$  refers to the time for probes that replaced the

neutral stimulus,  $n_1$  refers to the number of trials for probes that replaced the neutral stimulus,  $T_0$  refers to the time for probes that replaced the substance stimulus, and  $N_0$  refers to the number of trials for probes that replaced the substance stimulus. A chi-square test was conducted to compare the demographic characteristics of those with and without baseline attentional biases.

## Results

### Summary

40 participants were screened and invited to participate in the study, of which 10 declined, thus leaving 30 who participated. A total of 11 of these participants failed to complete all the planned interventions, 10 of whom elected for premature discharge from the ward and 1 who withdrew from the study following the initial intervention. [Table 1](#) (previously published in Zhang et al's feasibility and acceptability study) [11] provides an overview of the baseline attention bias scores for each of the participants and their change in scores across time.

### Baseline Attentional Biases

It is evident from [Table 1](#) that only 14/30 participants presented with positive attentional biases at baseline. The remaining 16 participants did not have any underlying attentional biases.

### Changes in Attentional Biases Over Time for Individuals With Positive Baseline Biases

Among those with baseline attention biases, there was a general decrease in their attentional bias scores from baseline till the end of the intervention, except for three participants (numbers 8, 19, and 20). The changes in attentional biases scores ranged from 12.0 to 409.5 milliseconds.

Participants with and without baseline attentional biases were compared. [Table 2](#) provides an overview of the baseline demographic characteristics of participants with and without attentional biases at baseline. The mean age of those with positive baseline attention bias was 46.3 years old, while those without attention bias had a mean age of 45.8 years. In terms of substance use, among those with positive baseline attention bias, 3 were diagnosed with alcohol use, 9 with opioid use, and 2 with stimulant use. Among those without baseline attention bias, 3 were diagnosed with alcohol use, 8 with opioids use, 2 with cannabis use, and 3 with stimulant use. Most of the participants were Singaporean. There was one female in the positive group and three females in the negative group. Most of the participants had a primary or secondary school education, and most were unemployed. Physical disorders were more common (43.8%) among those with negative baseline attention bias. The mean scores for the severity of substance dependence were comparable (10.93 for the positive group and 10.31 for the negative group).

**Table 1.** Overview of baseline attention bias scores and changes in scores over time.

Participant	Drug	Baseline	Post first session	Post second session	Post third session	Post fourth session	Post fifth session	Overall change in attentional bias
001	Stimulants	30.3	70.6	36.3	9.3	-23.6	13.3	17.0
002 <sup>a</sup>	Stimulants	-22.4	-23.4	-11.7	— <sup>b</sup>	—	—	10.7 (increased)
003	Stimulants	6.7	-3.6	-28.9	-11.3	4.1	-7.3	14.0
004 <sup>c</sup>	Opioid	32.1	28.7	12.2	31.2	20.1	—	12.0
005 <sup>c</sup>	Alcohol	91.2	-23.3	-37.4	-5.3	-33.2	—	124.4
006 <sup>c</sup>	Opioid	98.9	33.4	14.5	-9.7	-31.5	—	130.4
007	Stimulants	-30.5	-13.5	-23.4	-27.7	-28.2	-7.6	22.9 (increased)
008 <sup>a</sup>	Opioids	58.7	85.1	—	—	—	—	26.4 (increased)
009 <sup>a</sup>	Opioids	25.8	13.5	—	—	—	—	12.3
010	Cannabis	-9.9	-20.7	-54.6	-14.1	-22.9	-48.4	38.5
011	Opioids	-30.9	2.4	-15.4	-7.4	14.2	7.3	38.2 (increased)
012 <sup>a</sup>	Opioids	0.7	34.8	-15.3	-52.4	—	—	53.1
013	Alcohol <sup>d</sup>	-20.9	-12.3	—	—	-75.8	-48.6	27.7
014 <sup>a</sup>	Alcohol	397.7	44.2	45.3	-32.0	-11.7	—	409.4
015	Cannabis	-7.7	-40.5	8.6	-33.6	-31.8	-50.3	42.6
016 <sup>a</sup>	Opioids	-27.4	—	—	—	—	—	—
017	Opioids	-42.5	-64.8	63.4	8.9	-15.5	26.7	69.2 (increased)
018	Opioids	27.9	-17.8	3.2	-22.6	-104.6	-9.2	37.1
019 <sup>a</sup>	Opioids	3.8	35.1	13.7	32.4	—	—	28.6 (increased)
020	Opioids	10.1	9.4	105.2	54.1	-1.7	20.3	10.2 (increased)
021	Alcohol	224.5	61.5	73.3	176.7	130.3	107.0	117.4
022	Opioids	-52.9	10.9	5.6	39.1	36.5	76.6	129.6 (increased)
023	Opioids	-36.4	18.0	45.0	74.8	35.9	41.3	77.7 (increased)
024 <sup>a</sup>	Opioids	-16.9	45.2	1.49	3.8	33.9	—	50.8 (increased)
026 <sup>a</sup>	Opioids	-77.1	-82.5	—	—	—	—	5.4
027	Stimulants	-15.2	-11.6	10.1	-27.0	-25.7	-1.6	13.6 (increased)
028	Alcohol	-33.3	-29.4	-48.3	-10.9	-15.6	-11.6	21.8 (increased)
029 <sup>a</sup>	Alcohol	-41.4	-18.0	20.6	28.2	—	—	69.6 (increased)
030	Opioids	1.3	38.6	7.9	11.5	13.0	-8.9	10.2
031 <sup>a</sup>	Opioids	-38.4	-282.7	-166.4	-190.1	—	—	151.6

<sup>a</sup>Participants who did not complete the study as they left the voluntary program.

<sup>b</sup>Not applicable.

<sup>c</sup>There was a holiday during the assessment period and thus the maximum number of sessions completed was 4.

<sup>d</sup>Due to a technical issue, Participant 13 was not administered an assessment task following the second intervention so the participant took another intervention task instead. Attentional bias assessment was performed only after the fourth session.

**Table 2.** Baseline demographic characteristics (n=30).

Demographic characteristics	Positive baseline attention bias (n=14)	Negative baseline attention bias (n=16)
Age (years), mean (SD)	46.3 (12.7)	45.8 (10.7)
<b>Substance use, n (%)</b>		
Alcohol	3 (21)	3 (19)
Opioids	9 (64)	8 (50)
Cannabis	0 (0)	2 (13)
Stimulants	2 (14)	3 (19)
<b>Nationality, n (%)</b>		
Singaporean	12 (86)	15 (94)
Others	2 (14)	1 (6)
<b>Gender, n (%)</b>		
Male	13 (93)	13 (81)
Female	1 (7)	3 (19)
<b>Race, n (%)</b>		
Chinese	4 (29)	3 (19)
Malay	6 (43)	7 (38)
Indian	3 (21)	6 (38)
Others	1 (7)	1 (6)
<b>Education, n (%)</b>		
Primary education	4 (29)	3 (19)
Secondary education	7 (50)	9 (56)
Junior college or polytechnic/technical studies	1 (8)	4 (25)
Undergraduate studies	2 (14)	0 (0)
<b>Employment, n (%)</b>		
Unemployed	10 (71)	13 (81)
Part-time employment	1 (8)	2 (13)
Full-time employment	3 (21)	1 (6)
<b>Housing, n (%)</b>		
Homeless	4 (29)	2 (13)
1 room	4 (29)	4 (25)
2 rooms	2 (14)	0 (0)
3 rooms	1 (7)	2 (13)
4 rooms	0 (0)	7 (44)
5 rooms	2 (14)	0 (0)
Others	1 (7)	1 (6)
Psychiatric disorders, n (%)	2 (14)	1 (6)
Physical disorders, n (%)	4 (29)	7 (44)
Severity of substance dependence scores, mean (SD)	10.93 (3.33)	10.31 (3.28)

Using chi-square ( $\chi^2$ ) analysis to compare the categorical demographical variables amongst those with and without baseline attentional biases, we found no significant differences between the two groups with respect to nationality ( $\chi^2_1=0.536$ ;  $P=.46$ ), gender ( $\chi^2_1=0.871$ ;  $P=.35$ ), substances ( $\chi^2_3=1.875$ ;

$P=.60$ ), race ( $\chi^2_3=1.014$ ;  $P=.80$ ), education status ( $\chi^2_3=4.078$ ;  $P=.25$ ), employment ( $\chi^2_2=1.598$ ;  $P=.45$ ), accommodation/housing ( $\chi^2_6=11.92$ ;  $P=.06$ ), psychiatric

diagnoses ( $\chi^2_1=0.536$ ;  $P=.46$ ) and physical diagnosis ( $\chi^2_1=1.916$ ;  $P=.38$ ).

Table 3 and Table 4 provide an overview of participants with and without baseline attentional biases and their reported drug use, recorded on the modified ASI-Lite questionnaire. As evident from Table 2, Table 3, and Table 4, both the two

participants with cannabis use disorder did not present with any baseline attentional biases. Most of the participants had started using their substance of abuse at a young age, and most of them were using daily prior to admission. Individuals in the group with positive baseline attention bias had experimented with more substances, as compared to individuals in the group without baseline attention bias.

**Table 3.** Substance use history of participants with positive baseline attention bias.

Participant	Age	Baseline bias	Current drug	Drug history	Age of onset	Age of problematic use	Frequency prior to admission
001	46	30.3	Stimulants	Polysubstance (9) <sup>a</sup>	45	45	Daily (1-2 grams, once per day)
003	35	6.7	Stimulants	Polysubstance (7)	31	31	Daily (0.5 grams, once per day)
004	46	32.1	Opioids	Polysubstance (12)	14	17	Daily (30-40 mL, 1-2 times per day)
005	26	91.2	Alcohol	Alcohol (1)	20	21	Daily (750 mL, 2-3 times per day)
006	64	98.9	Opioids	Polysubstance (11)	20	22	Daily (2 Straws, once daily)
008	60	58.7	Opioids	Polysubstance (4)	13	13	Daily (1 Straw, 2 times daily)
009	28	25.8	Opioids	Polysubstance (5)	17	19	Daily (1.5 straws, 6 times daily)
012	28	0.7	Opioids	Polysubstance (5)	18	25	Daily (5 tablets, 5 times daily)
014	61	397.7	Alcohol	Polysubstance (5)	13	27	Daily (3 tall cans, 2 times daily)
018	54	27.9	Opioids	Polysubstance (5)	34	34	*Patient did not fill up
019	57	3.8	Opioids	Polysubstance (3)	40	40	Daily (90 mL, once per day)
020	45	10.1	Opioids	Polysubstance (4)	23	23	Daily (1 straw, 3 times daily)
021	50	224.5	Alcohol	Alcohol & tobacco (2)	15	50	Daily (8 cans, once daily)
030	48	1.3	Opioids	Polysubstance (3)	16	16	Daily (4 Straws, 4 times daily)

<sup>a</sup>(n): This refers to the total number of substances the individual has experimented with.

**Table 4.** Substance use history of participants with negative baseline attention bias

Participant	Age	Baseline bias	Current drug	Drug history	Age of onset	Age of problematic use	Frequency prior to admission
002	44	-22.4	Stimulants	Polysubstance (4) <sup>a</sup>	38	40	Daily (Unknown quantity, 6 times daily)
007	32	-30.5	Stimulants	Polysubstance (3)	29	30	Daily (0.3 grams, 5 times daily)
010	59	-9.9	Cannabis	Polysubstance (5)	59	59	Daily (2 joints once per day)
011	50	-30.9	Opioids	Polysubstance (3)	24	24	Daily (4 straws, 4 times per day)
013	42	-20.9	Alcohol	Alcohol and tobacco (2)	18	18	Daily (3-4 bottles, 3 times per day)
015	57	-7.7	Cannabis	Polysubstance (7)	57	57	Daily (1 joint, 20 times per day)
016	47	-27.4	Opioids	Polysubstance (6)	17	17	Daily (1 joint, 6-7 times per day)
017	63	-42.5	Opioids	Polysubstance (4)	16	17	Daily (0.5 straws, 3 times per day)
022	43	-52.9	Opioids	Polysubstance (6)	16	17	Daily (2 straws, 2 times per day)
023	54	-36.4	Opioids	Polysubstance (6)	15	16	Daily (3 straws, 10 times per day)
024	46	-16.9	Opioids	Tobacco & heroin (2)	17	21	Daily (7 straws, 4 times per day)
026	57	-77.1	Opioids	Tobacco & heroin (2)	17	28	Daily (1 straw, 2 times per day)
027	31	-15.2	Stimulants	Polysubstance (9)	17	25	Daily (0.3 grams throughout the day)
028	44	-33.3	Alcohol	Alcohol and tobacco (2)	17	28	Daily (750 mL, 16-18 times per day)
029	39	-41.4	Alcohol	Polysubstance (3)	14	30	Daily (9-10 cans, 9 times per day)
031	25	-38.4	Opioids	Polysubstance (8)	20	20	Daily (1 straw, 2 times per day)

<sup>a</sup>(n): This refers to the total number of substances the individual has experimented with.

## Discussion

### Primary Findings

We are unaware of any other study that has examined negative findings in attention bias modification intervention for individuals with addictive disorders. In our study, 53% (16/30) of individuals did not present with baseline attentional biases, and among those with positive baseline biases, 21% (3/14) of participants did not have a reduction in the overall magnitude of attentional biases. Chi-square analyses undertaken to compare the demographic characteristics of participants with and without baseline biases did not reveal any significant findings. However, with respect to clinical characteristics, those who had positive baseline biases had experimented with more substances. We postulate that there are four factors that could account for the absence of baseline biases: the period of abstinence, the images

used in the mobile application, the stimulus timings, and the clinical characteristics of the participants. With regards to the lack of retraining effects, we theorize that this might be due to practice effects or to the novelty of the other neutral stimulus.

The lack of baseline biases we observed may be due firstly to the fact that our participants had received treatment and had been abstinent for at least seven days prior to being recruited into the study. This period of abstinence might have affected their baseline attentional biases. In most of the published studies [17,18], participants who had been recruited were in the detoxification phase of treatment. In our study, participants were considered for the intervention only upon the successful completion of their detoxification program. Participants who were in their detoxification phase were not recruited, as we anticipated that them having withdrawal symptoms might affect their ability to undertake the required interventions. The existing

medication-assisted detoxification was highly effective in ameliorating attentional biases among these participants. Psychotropic medications prescribed in the acute detoxification phase might have affected or reduced attentional biases. Dopaminergic and serotonergic agents (such as antipsychotics and antidepressants) could have reduced attention biases in the acute phase, as highlighted in Zhang et al's [19] prior review. Some of these medications might have been used for symptomatic relief of psychotic or affective symptoms in these individuals during the withdrawal phase. Unfortunately, in this current study we did not request ethical approval to extract information from the medical records, and thus the association between the prescription of medications and attentional biases cannot be determined.

Secondly, the nature of the images/stimuli used in our application may have contributed to a null finding at baseline assessment. The images used in the application may not be good representations of the substances participants were familiar with, and thus did not manage to capture their attention. This is in line with Field et al's [20] report that one of the key factors leading to the poor reliability of the visual probe task is that of the nature of the stimulus used. Field et al [20] highlighted the importance of personalization of the stimulus presented to the participants, as it is postulated that a stimulus that is relevant and identifiable to the participant could increase their baseline attentional bias score and provide evidence of greater change in the magnitude of attentional biases. Many of the images included in the mobile application were extracted from the internet via the United States Drug Enforcement Agency media library. Fewer images came from Singapore's Central Narcotics Bureau's website. It might be possible that the images included do not approximate the real object and are not realistic enough for participants.

Thirdly, as highlighted by Sharpe [12], the experimental procedure that we used in our study might have affected the detection of baseline biases, in regard to the duration of the presentation of the images. In our study, we presented the stimulus for 500 milliseconds, like most prior studies which had evaluated the reliability of the visual probe task [21,22]. We are aware that there have been other studies, such as those by Constantinou et al, Frankland et al, and two by Garland et al [21-24], that have presented the stimulus for as little as 200 milliseconds to as long as 1500 to 2000 milliseconds. It has been previously postulated in the literature that the short stimulus interval helped in the evaluation of automatic orientating tendencies, whereas the long stimulus interval helped in the evaluation of controlled attentional processing. In some of these previous studies [23-26], the stimulus was being presented to individuals at both a short and long stimulus interval. The fact that we had a fixed stimulus timing interval of 500 milliseconds might have resulted in us not capturing

potentially controlled attentional processes, which might account for individuals not having baseline biases. It is thus of importance for there to be a review of stimulus timings that have been previously used and to correlate the timings with the effectiveness of bias detection and modification.

Lastly, with regards to the absence of baseline biases, Dean et al's [15] research has highlighted that clinical characteristics of participants modulate their baseline attentional biases. In our study, we did not manage to find any demographic variables that could account for the differences among those with baseline biases and those without. There were, however, more individuals in the group with positive baseline biases that had previously abused a larger quantity of substances. This observation is congruent with previous research, which reported that the frequency and quantity of drug use do moderate the magnitude of the attentional biases. Field et al [21], in their previous study, reported that there were more robust attentional biases among individuals who were heavy drinkers when the alcohol images were presented for 500 and 2000 milliseconds, respectively. In another study by Noel et al [27], the authors reported that there was an absence of attentional biases among those who were abstinent from alcohol use. It is evident from these studies that the amount of substance use affects the attentional biases, and this should be an important consideration when planning for future attention bias modification interventions for individuals with addictive disorders since there may be individuals who are abstinent and without baseline biases, so attention bias modification might not be appropriate for them. It is important that future studies consider evaluating individuals for baseline biases before administration of bias modification.

Regarding the negative findings among those with positive baseline attentional biases (ie, no impact of attentional retraining), we postulate that the absence of effects could be accounted for by practice effects. Thus, for some individuals, over the course of the intervention they have learnt to focus only on the position of the probe and not on the stimulus. There is also the possibility that the neutral stimulus might have appeared to be more novel and appealing to some individuals, which lead to an increased focus on the neutral stimulus and a resultant more rapid response in identifying probes that replace the neutral stimulus.

## Conclusions

Our study is one of the first to have explored negative findings in attention bias modification interventions for individuals with addictive disorders. We postulate that there could be several factors that could account for the absence of baseline biases and there being no changes following bias retraining. Future research ought to take into consideration these factors, and it is especially pertinent for clinical studies to evaluate for baseline bias prior to administration of bias retraining interventions.

---

## Acknowledgments

We would like to thank the National Addictions Management Service for providing finance support for this project.

---



## Authors' Contributions

MZ, JY, SBA, ZM, GS, DSSF, HES jointly conceptualized the study. MZ, SBA, ZM were involved in data collation. MZ and JY coded and analyzed the initial data. MZ, JY, SBA, ZM, GS, DSSF, HES reviewed the final dataset. MZ and HES jointly worked on and wrote up the initial draft of the manuscript. HES provided further comments, which aided MZ in formulating the second draft of the manuscript. All authors proofread and approved of the manuscript prior to publication.

## Conflicts of Interest

None declared.

## References

1. Zhang M, Ying J, Song G, Ho RC, Fung DS, Smith H. Attention Bias in Individuals with Addictive Disorders: Systematic Review Protocol. *JMIR Res Protoc* 2018 Feb 08;7(2):e41 [FREE Full text] [doi: [10.2196/resprot.9385](https://doi.org/10.2196/resprot.9385)] [Medline: [29422449](https://pubmed.ncbi.nlm.nih.gov/29422449/)]
2. Field M, Marhe R, Franken IHA. The clinical relevance of attentional bias in substance use disorders. *CNS Spectr* 2014 Jun;19(3):225-230. [doi: [10.1017/S1092852913000321](https://doi.org/10.1017/S1092852913000321)] [Medline: [23663386](https://pubmed.ncbi.nlm.nih.gov/23663386/)]
3. Wiers RW, Rinck M, Kordts R, Houben K, Strack F. Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 2010 Feb;105(2):279-287. [doi: [10.1111/j.1360-0443.2009.02775.x](https://doi.org/10.1111/j.1360-0443.2009.02775.x)] [Medline: [20078486](https://pubmed.ncbi.nlm.nih.gov/20078486/)]
4. Wiers R, Teachman B, De Houwer J. Implicit cognitive processes in psychopathology: an introduction. *J Behav Ther Exp Psychiatry* 2007 Jun;38(2):95-104. [doi: [10.1016/j.jbtep.2006.10.002](https://doi.org/10.1016/j.jbtep.2006.10.002)] [Medline: [17112462](https://pubmed.ncbi.nlm.nih.gov/17112462/)]
5. Cristea I, Kok R, Cuijpers P. The Effectiveness of Cognitive Bias Modification Interventions for Substance Addictions: A Meta-Analysis. *PLoS ONE* 2016 Sep 9 2016 Sep;11(9):A.
6. Heitmann J, Bennik EC, van Hemel-Ruiter ME, de Jong PJ. The effectiveness of attentional bias modification for substance use disorder symptoms in adults: a systematic review. *Syst Rev* 2018 Oct 13;7(1). [doi: [10.1186/s13643-018-0822-6](https://doi.org/10.1186/s13643-018-0822-6)]
7. Zhang M, Ying J, Ho RC. Gamified Cognitive Bias Interventions for Psychiatric Disorders: Protocol of a Systematic Review. *JMIR Res Protoc* 2018 Oct 16;7(10):e10154 [FREE Full text] [doi: [10.2196/10154](https://doi.org/10.2196/10154)] [Medline: [30327292](https://pubmed.ncbi.nlm.nih.gov/30327292/)]
8. Zhang M, Ying J, Song G, Fung D, Smith H. Attention and Cognitive Bias Modification Apps: Review of the Literature and of Commercially Available Apps. *JMIR Mhealth Uhealth* 2018 May 24;6(5):e10034 [FREE Full text] [doi: [10.2196/10034](https://doi.org/10.2196/10034)] [Medline: [29793899](https://pubmed.ncbi.nlm.nih.gov/29793899/)]
9. Robinson C, Muench C, Brede E, Endrighi R, Szeto E, Sells J, et al. Effect of attentional retraining on cognition, craving, and smoking in African American smokers. *Psychol Addict Behav* 2017 Aug;31(5):636-646 [FREE Full text] [doi: [10.1037/adb0000286](https://doi.org/10.1037/adb0000286)] [Medline: [28627913](https://pubmed.ncbi.nlm.nih.gov/28627913/)]
10. Zhu Y, Jiang H, Su H, Zhong N, Li R, Li X, et al. A Newly Designed Mobile-Based Computerized Cognitive Addiction Therapy App for the Improvement of Cognition Impairments and Risk Decision Making in Methamphetamine Use Disorder: Randomized Controlled Trial. *JMIR Mhealth Uhealth* 2018 Jun 20;6(6):e10292. [doi: [10.2196/10292](https://doi.org/10.2196/10292)]
11. Zhang M, Ying J, Amron S, Mahreen Z, Song G, Fung D, et al. A Smartphone Attention Bias App for Individuals With Addictive Disorders: Feasibility and Acceptability Study. *JMIR Mhealth Uhealth* 2019;7(9):e15465. [doi: [10.2196/preprints.15465](https://doi.org/10.2196/preprints.15465)]
12. Sharpe L. Attention bias modification for children. *PAIN* 2018;159(2):191-192. [doi: [10.1097/j.pain.0000000000001107](https://doi.org/10.1097/j.pain.0000000000001107)]
13. Heathcote LC, Jacobs K, Van Ryckeghem DM, Fisher E, Eccleston C, Fox E, et al. Attention bias modification training for adolescents with chronic pain. *PAIN* 2018;159(2):239-251. [doi: [10.1097/j.pain.0000000000001084](https://doi.org/10.1097/j.pain.0000000000001084)]
14. Cristea IA, Mogoia e C, David D, Cuijpers P. Practitioner Review: Cognitive bias modification for mental health problems in children and adolescents: a meta-analysis. *J Child Psychol Psychiatr* 2015 Jan 30;56(7):723-734. [doi: [10.1111/jcpp.12383](https://doi.org/10.1111/jcpp.12383)]
15. Dean AC, Nurmi EL, Moeller SJ, Amir N, Rozenman M, Ghahremani DG, et al. No effect of attentional bias modification training in methamphetamine users receiving residential treatment. *Psychopharmacology* 2018 Nov 10;236(2):709-721. [doi: [10.1007/s00213-018-5100-8](https://doi.org/10.1007/s00213-018-5100-8)]
16. Zhang M, Ying J, Amron SB, Mahreen Z, Song G, Fung DS, et al. A Smartphone Attention Bias Intervention for Individuals With Addictive Disorders: Protocol for a Feasibility Study. *JMIR Res Protoc* 2018 Nov 19;7(11):e11822. [doi: [10.2196/11822](https://doi.org/10.2196/11822)]
17. Manning V, Staiger PK, Hall K, Garfield JB, Flaks G, Leung D, et al. Cognitive Bias Modification Training During Inpatient Alcohol Detoxification Reduces Early Relapse: A Randomized Controlled Trial. *Alcohol Clin Exp Res* 2016 Aug 04;40(9):2011-2019. [doi: [10.1111/acer.13163](https://doi.org/10.1111/acer.13163)]
18. Eberl C, Wiers RW, Pawelczack S, Rinck M, Becker ES, Lindenmeyer J. Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work best? *Developmental Cognitive Neuroscience* 2013 Apr;4:38-51. [doi: [10.1016/j.dcn.2012.11.002](https://doi.org/10.1016/j.dcn.2012.11.002)]
19. Zhang MW, Ying JB, Song G, Fung DSS, Smith HE. Recent Advances in Attention Bias Modification for Substance Addictions. *Int J Environ Res Public Health* 2018 Apr 04;15(4) [FREE Full text] [doi: [10.3390/ijerph15040676](https://doi.org/10.3390/ijerph15040676)] [Medline: [29617325](https://pubmed.ncbi.nlm.nih.gov/29617325/)]
20. Field M, Christiansen P. Commentary on Ataya et al. (2012), 'Internal reliability of measures of substance-related cognitive bias'. *Drug and Alcohol Dependence* 2012 Aug;124(3):189-190. [doi: [10.1016/j.drugalcdep.2012.02.009](https://doi.org/10.1016/j.drugalcdep.2012.02.009)]

21. Field M. ; Mogg, K. ; Bradley, B. P. Cognitive bias and drug craving in recreational cannabis users. *Drug Alcohol Depend*,74, 105 2004:111.
22. Field M. Cannabis 'dependence' and attentional bias for cannabis-related words. *Behav. Pharmacol*,16, 473? 2005:476.
23. Constantinou N, Morgan CJA, Battistella S, O'Ryan D, Davis P, Curran HV. Attentional bias, inhibitory control and acute stress in current and former opiate addicts. *Drug Alcohol Depend* 2010 Jun 01;109(1-3):220-225. [doi: [10.1016/j.drugalcdep.2010.01.012](https://doi.org/10.1016/j.drugalcdep.2010.01.012)] [Medline: [20172662](https://pubmed.ncbi.nlm.nih.gov/20172662/)]
24. Frankland L, Bradley BP, Mogg K. Time course of attentional bias to drug cues in opioid dependence. *Psychol Addict Behav* 2016 Aug;30(5):601-606. [doi: [10.1037/adb0000169](https://doi.org/10.1037/adb0000169)] [Medline: [27031089](https://pubmed.ncbi.nlm.nih.gov/27031089/)]
25. Garland E, Howard MO. Opioid attentional bias and cue-elicited craving predict future risk of prescription opioid misuse among chronic pain patients. *Drug Alcohol Depend* 2014 Nov 01;144:283-287 [FREE Full text] [doi: [10.1016/j.drugalcdep.2014.09.014](https://doi.org/10.1016/j.drugalcdep.2014.09.014)] [Medline: [25282309](https://pubmed.ncbi.nlm.nih.gov/25282309/)]
26. Garland E, Froeliger BE, Passik SD, Howard MO. Attentional bias for prescription opioid cues among opioid dependent chronic pain patients. *J Behav Med* 2013 Dec;36(6):611-620 [FREE Full text] [doi: [10.1007/s10865-012-9455-8](https://doi.org/10.1007/s10865-012-9455-8)] [Medline: [22968666](https://pubmed.ncbi.nlm.nih.gov/22968666/)]
27. Noël X, Colmant M, Van Der Linden M, Bechara A, Bullens Q, Hanak C, et al. Time course of attention for alcohol cues in abstinent alcoholic patients: the role of initial orienting. *Alcohol Clin Exp Res* 2006 Nov;30(11):1871-1877. [doi: [10.1111/j.1530-0277.2006.00224.x](https://doi.org/10.1111/j.1530-0277.2006.00224.x)] [Medline: [17067351](https://pubmed.ncbi.nlm.nih.gov/17067351/)]

## Abbreviations

**ASI:** Addiction Severity Index

**SDS:** Severity of Drug Dependence Scale

**SF-12:** Short Form 12-item questionnaire

*Edited by G Eysenbach; submitted 19.09.19; peer-reviewed by R Ho, T Klose; comments to author 09.10.19; revised version received 09.10.19; accepted 17.10.19; published 12.11.19.*

*Please cite as:*

*Zhang M, Ying J, Amron SB, Mahreen Z, Song G, Fung DSS, Smith HE*

*Making Sense of Negative Findings from Mobile Attention Bias Modification Interventions for Individuals with Addictive Disorders: Quantitative Feasibility Study*

*JMIR Mhealth Uhealth* 2019;7(11):e16325

URL: <http://mhealth.jmir.org/2019/11/e16325/>

doi: [10.2196/16325](https://doi.org/10.2196/16325)

PMID: [31714248](https://pubmed.ncbi.nlm.nih.gov/31714248/)

©Melvyn Wb Zhang, Jiangbo Ying, Syidda B Amron, Zaakira Mahreen, Guo Song, Daniel SS Fung, Helen E Smith. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 12.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Development of a Deep Learning Model for Dynamic Forecasting of Blood Glucose Level for Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial

Syed Hasib Akhter Faruqui<sup>1</sup>, MS; Yan Du<sup>2</sup>, RN, MPH, PhD; Rajitha Meka<sup>1</sup>, MIE; Adel Alaeddini<sup>1</sup>, PhD; Chengdong Li<sup>2</sup>, PhD; Sara Shirinkam<sup>3</sup>, PhD; Jing Wang<sup>2</sup>, RN, MPH, PhD

<sup>1</sup>Department of Mechanical Engineering, University of Texas at San Antonio, San Antonio, TX, United States

<sup>2</sup>Center on Smart and Connected Health Technologies, University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

<sup>3</sup>Department of Mathematics and Statistics, University of the Incarnate Word, San Antonio, TX, United States

**Corresponding Author:**

Jing Wang, RN, MPH, PhD

Center on Smart and Connected Health Technologies

University of Texas Health Science Center at San Antonio

7703 Floyd Curl Drive

San Antonio, TX

United States

Phone: 1 210 450 8561

Email: [wangj1@uthscsa.edu](mailto:wangj1@uthscsa.edu)

## Abstract

**Background:** Type 2 diabetes mellitus (T2DM) is a major public health burden. Self-management of diabetes including maintaining a healthy lifestyle is essential for glycemic control and to prevent diabetes complications. Mobile-based health data can play an important role in the forecasting of blood glucose levels for lifestyle management and control of T2DM.

**Objective:** The objective of this work was to dynamically forecast daily glucose levels in patients with T2DM based on their daily mobile health lifestyle data including diet, physical activity, weight, and glucose level from the day before.

**Methods:** We used data from 10 T2DM patients who were overweight or obese in a behavioral lifestyle intervention using mobile tools for daily monitoring of diet, physical activity, weight, and blood glucose over 6 months. We developed a deep learning model based on long short-term memory-based recurrent neural networks to forecast the next-day glucose levels in individual patients. The neural network used several layers of computational nodes to model how mobile health data (food intake including consumed calories, fat, and carbohydrates; exercise; and weight) were progressing from one day to another from noisy data.

**Results:** The model was validated based on a data set of 10 patients who had been monitored daily for over 6 months. The proposed deep learning model demonstrated considerable accuracy in predicting the next day glucose level based on Clark Error Grid and  $\pm 10\%$  range of the actual values.

**Conclusions:** Using machine learning methodologies may leverage mobile health lifestyle data to develop effective individualized prediction plans for T2DM management. However, predicting future glucose levels is challenging as glucose level is determined by multiple factors. Future study with more rigorous study design is warranted to better predict future glucose levels for T2DM management.

(*JMIR Mhealth Uhealth* 2019;7(11):e14452) doi:[10.2196/14452](https://doi.org/10.2196/14452)

**KEYWORDS**

type 2 diabetes; long short-term memory (LSTM)-based recurrent neural networks (RNNs); glucose level prediction; mobile health lifestyle data

## Introduction

Diabetes mellitus is a serious health condition resulting from defects of insulin secretion and/or insulin action [1]. Patients with type 2 diabetes mellitus (T2DM) need to maintain strict glycemic control to avoid the risk of hypoglycemia, hyperglycemia, and consequential complications [2]. T2DM, characterized by the combination of insufficient insulin secretion and insulin resistance, accounts for approximately 90% to 95% of all diabetes cases [3]. It has become a major public health concern as it is burdensome for individuals, health systems, and society [4]. Self-management of diet, physical activity, weight, and medication and self-monitoring of blood glucose are essential for glycemic control [5,6]. However, it is very challenging to adhere to this self-management regimen [7].

Emerging evidence has demonstrated that mobile technologies can promote a healthy lifestyle and medication adherence and improve diabetes outcomes [8,9]. The underlying mechanisms might include the frequent reminder for blood glucose monitoring [10], self-awareness and control of diabetes [11,12], or behavior adjustment based on tracked behaviors [13]. For instance, Padhye et al [11] reported that in patients with T2DM, smartphone users are more likely to adhere to self-monitoring of diet, physical activity, blood glucose, and body weight when compared with paper diary users. Many studies have evidenced that compliance with self-monitoring of diet and physical activity can lead to weight loss [14] and hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>) improvement through behavior adjustment [8,15]. However, digital diabetes care has shown only modest HbA<sub>1c</sub> improvement in multiple studies [16]. Despite the modest effects of digital self-monitoring on HbA<sub>1c</sub>, recorded lifestyle data may shed light on improving glycemic control through predicting blood glucose level.

There are several algorithms, such as Biostator (Miles-Ames), for automated insulin delivery in order to improve blood glucose control [2,17]. Meanwhile, with an ever-growing amount of data, several machine learning techniques are being developed to understand patterns and develop models that predict the health conditions of patients [18]. For instance, Plis et al [19] described a generic physiological model of blood glucose dynamics to extract informative features to train a support vector regression model on patient-specific data [20-22]. Model predictive control

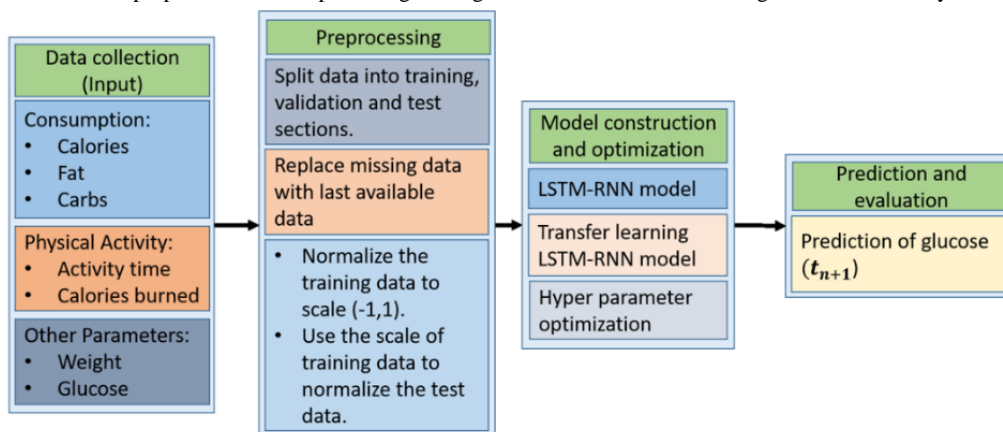
is also used to avoid long delays and open-loop characteristics of the control algorithms [23]. As the relation between input features and glucose levels is nonlinear, dynamic, interactive, and patient-specific, nonlinear regression models are used to build the predictive models [24]. Specifically, neural networks have increasingly been used to model glucose levels using multilayer perceptrons [25,26], time series convolution neural networks, recurrent neural networks [27], convolutional recurrent neural networks [28], and deep convolutional neural networks [29]. Quchani et al [30] compared multilayer perceptron neural networks with Elman recurrent neural networks for predicting the glucose level in patients with type 1 diabetes mellitus (T1DM) to show improvement in the accuracy of the model using recurrent neural networks.

However, to what extent self-monitoring data on health behaviors and weight can help predict blood glucose level in T2DM patients has rarely been studied. Available literature exploring glucose prediction in T2DM mainly focuses on glucose responses to nutrition [31]. However, glucose level is determined by a variety of factors [31-33], and a prediction model incorporating multiple lifestyle factors in a real-world setting is needed. In addition, most of the existing machine learning models predict glucose level for a very short interval (ie, a few minutes [34]), which makes it difficult to plan for effective control strategies. By using long short-term memory (LSTM)-based recurrent neural networks (RNNs), this study aimed to dynamically forecast the next-day glucose levels in individuals with T2DM based on their daily mobile health lifestyle data on diet, physical activity, weight, and previous glucose levels. The study also developed a transfer learning strategy to cope with data scarcity and improve prediction accuracy for individual patients. Additionally, the study used the advanced design of experiments to optimize the hyperparameters of the LSTM-based RNN model.

## Methods

Forecasting the glucose level of a T2DM patient is critical in planning for future medication and food habits. This study was a secondary analysis of data collected by a randomized controlled trial (RCT) consisting of several steps including data collection, data preprocessing, model construction and optimization, and prediction and evaluation (Figure 1).

**Figure 1.** General scheme of the proposed method of predicting blood glucose level. LSTM-RNN: long short-term memory recurrent neural networks.



## Data Collection

Our study used data collected from a smartphone group based in a pilot randomized trial [11]. The details of the original pilot RCT were published elsewhere [9]. In the randomized trial, overweight/obese adults (BMI >25 kg/m<sup>2</sup>) who were literate in English and diagnosed with T2DM for at least 6 months were eligible to participate. A total of 26 participants aged between 21 and 75 years were enrolled and randomly assigned to a smartphone group (n=11), paper diary group (n=9), and control group (n=6). Participants in the smartphone group received a standard behavioral lifestyle education. The Lose It! (FitNow, Inc) smartphone app was used in this group to self-monitor physical activity, diet, and weight. Blood glucose levels were collected using MyGlucoHealth, a Bluetooth-enabled glucometer (Entra Health Systems) and the DiabetesConnect app (PHRQL, Inc). Informed consent was obtained from each participant, and the study was approved by the Committee for the Protection of Human Subjects at the University of Texas Health Science Center at Houston. One participant in the smartphone group withdrew and did not complete the study.

The data in the smartphone group included an abundance of dynamically monitored lifestyle and health information that has not been fully explored and deserves further mining and analysis to generate study results and provide suggestions and directions for future studies and practices to improve health outcomes. The data collected from the clinical trial was a good fit for our

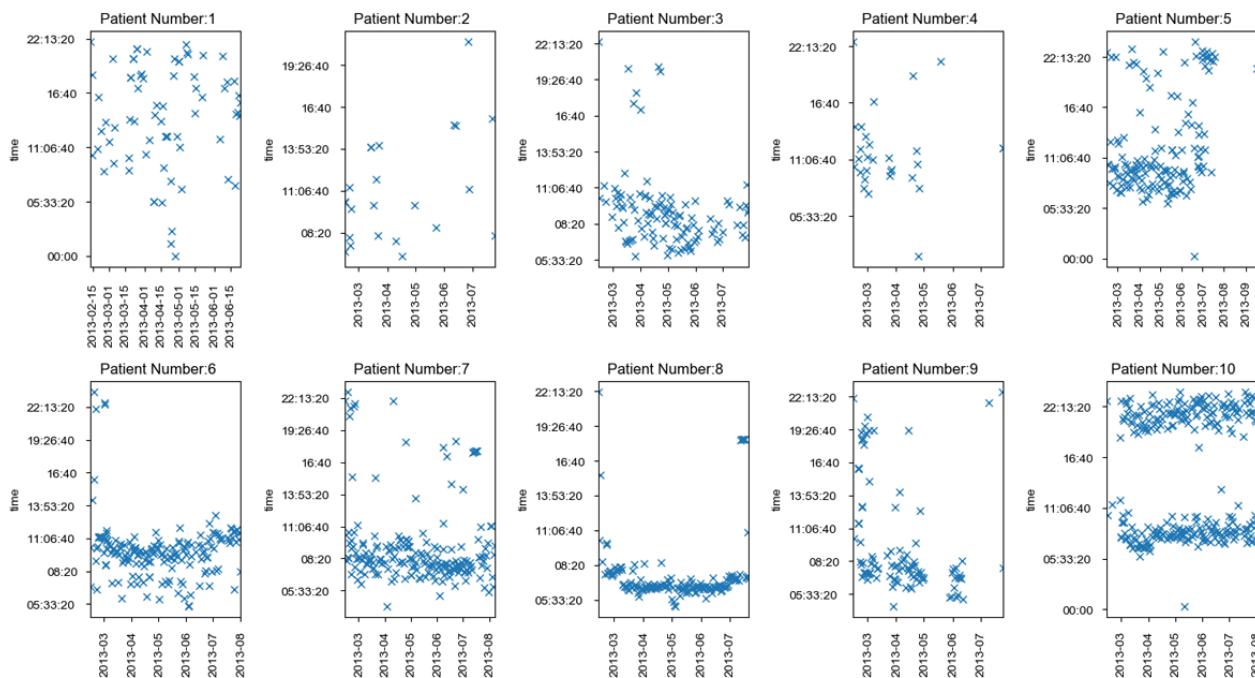
study objective of predicting glucose levels. The 10 participants who were in the smartphone group and recorded at least 150 days of self-monitored data were included in this study. The data for each participant include daily diet information, where collected food intake data (breakfast, lunch, dinner, snacks) is discretized into calories, macronutrient content (carbohydrates and fat), physical activity (where exercise time is translated into calories burned from standard food nutrient charts), weight, and glucose levels. A descriptive summary of the data is presented in Table 1. From the table, it can be seen that patients 1, 2, 4, and 9 have the highest number of missing values in terms of self-reported blood glucose values.

Patients were not required to take glucose readings at a fixed time of day but were asked to be consistent in terms of collecting blood glucose readings every day. Figure 2 shows the distribution of each patient's blood glucose recording time. For patients 3, 5, 6, 7, 8, and 9, the recorded times are generally between 8:00 am to 11:00 am. However, for patient 10, the recorded times are divided between 8:00 am and 10:00 am or 8:00 pm to 10:00 pm. For this patient, we considered the readings taken from 8:00 am to 10:00 am. For patients 1, 2, and 4, the number of recorded instances were fewer and scattered throughout the day. Figure 3 shows the self-monitored collected data for patient 1. While the patient has recorded their food intake for the day, they haven't adhered to a daily exercise regimen, as can be seen from the calories burned (cb) subfigure.

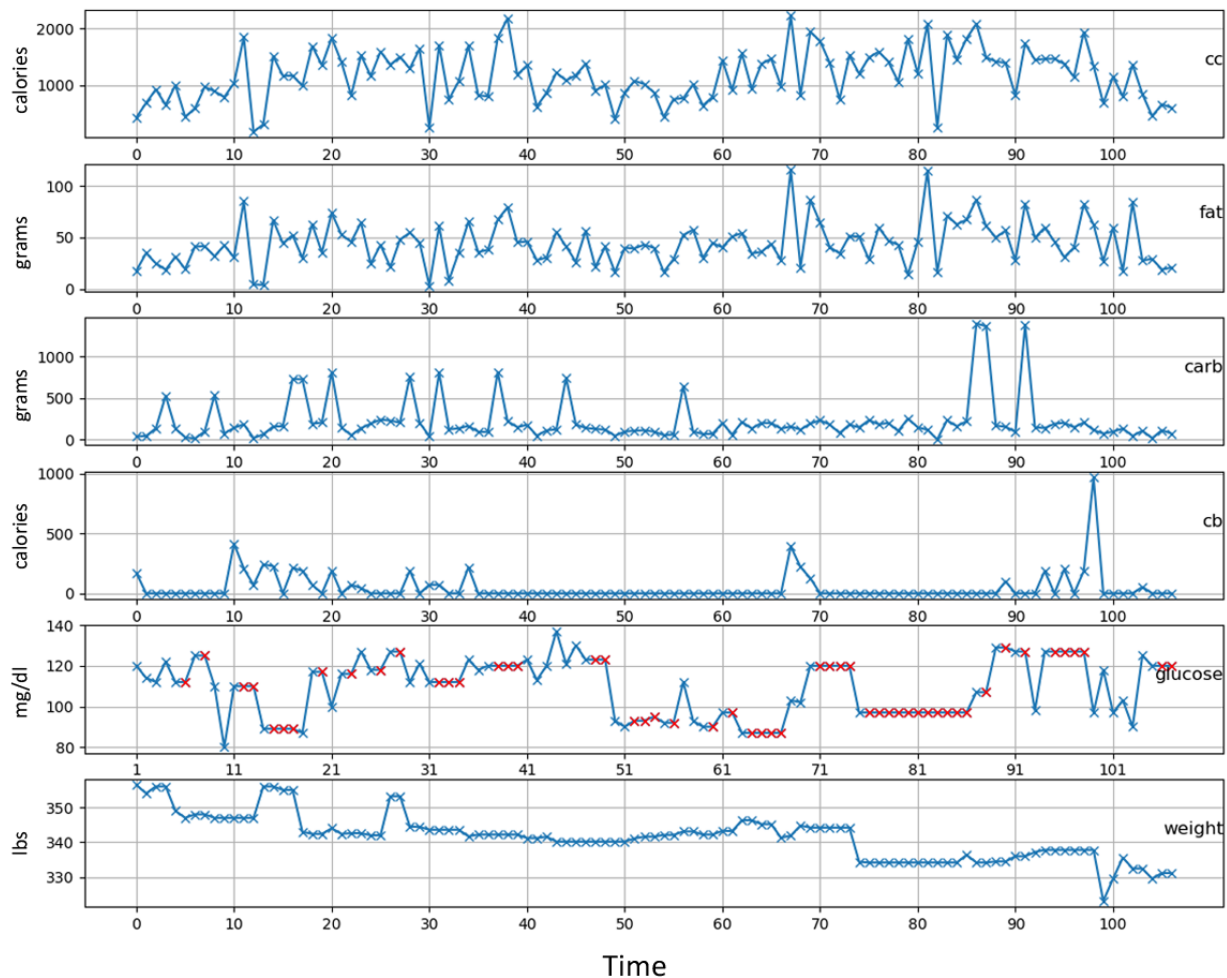
**Table 1.** Descriptive statistics of glucose and weight levels for the patients in the study.

Patient #	Number of missing observations	Glucose level (mg/dL)			Weight (lbs)		
		Min	Max	Mean (SD)	Min	Max	Mean (SD)
1	118	80	137	110.99 (13.97)	323	356	341.60 (6.39)
2	161	108	171	136.15 (18.95)	225	241	230.42 (4.94)
3	73	275	87	151.00 (27.64)	144	148	145.62 (0.83)
4	154	137	192	155.97 (17.27)	170	183	178.38 (2.88)
5	48	73	287	132.79 (29.81)	304	312	30.06 (3.88)
6	29	85	306	196.62 (53.43)	178	191	184.39 (4.02)
7	16	72	200	119.95 (27.26)	273	248	260.40 (6.17)
8	26	111	168	138.10 (13.15)	150	155	152.97 (1.41)
9	102	74	1770	118.33 (39.98)	274	285	279.65 (3.15)
10	53	86	147	104.82 (10.80)	222	229	226.97 (3.17)

**Figure 2.** Time distribution of self-monitored blood glucose level collection by the patients over the clinical trial. The x-axis represents the date and the y-axis represents the time of the day that data has been collected. If the same date has two recorded times, that means the patient has collected their blood glucose twice in the same day.



**Figure 3.** Patient 1 self-recorded data including calories consumed (cc), fat consumed (fat), carbohydrates consumed (carb), calories burned (cb), recorded blood glucose level for the day (glucose), and recorded weight for the day (weight), where blue markers represent original data points and red markers show the imputed data points.



### Data Preprocessing

The study dataset is based on patient data entry with several complicating factors, including missing values, (possible) wrong entries, calculated features, and irregular sampling. Therefore, major preprocessing steps are needed to clean the data and make it compatible with the proposed deep learning model. The preprocessing steps considered for this study include handling of missing values, data scaling, and data splitting.

### Handling Missing Values

The measurements in the database are sparse, irregularly sampled, and presented with missing data points. To handle missing values, assuming there is less chance of abrupt change in glucose level on the following day, the missing values of data were handled by replacing them with the last available data (last observation). Meanwhile, we noticed that many patients had a

considerable number of missing values that could potentially affect the performance of the methods. To address this problem, we developed a transfer learning strategy to leverage the similarity between the information of patients to improve the predictions when dealing with data scarcity.

### Data Scaling

The range of values for each feature in the dataset varies extensively. Thus, the performance of the learning algorithm might be dominated by features with a wider range of values. The goal of this step was to scale the values of each feature within a predefined limit without losing the inherent information. For this purpose, we used data scaling based on min-max normalization [35] (Figure 4), where  $X$  denotes the original value of the feature of interest,  $X_{min}$  denotes the minimum value of the feature,  $X_{max}$  denotes maximum value of the feature, and  $R$  denotes the desired range of the scaled features, namely  $[-1, 1]$ .

**Figure 4.** Equation for data scaling based on min-max normalization.

$$\bar{X} = \frac{(X - X_{min})}{(X_{max} - X_{min})}$$

$$X_{scaled} = \bar{X}(R_{max} - R_{min}) + R_{min}$$

**Data Splitting**

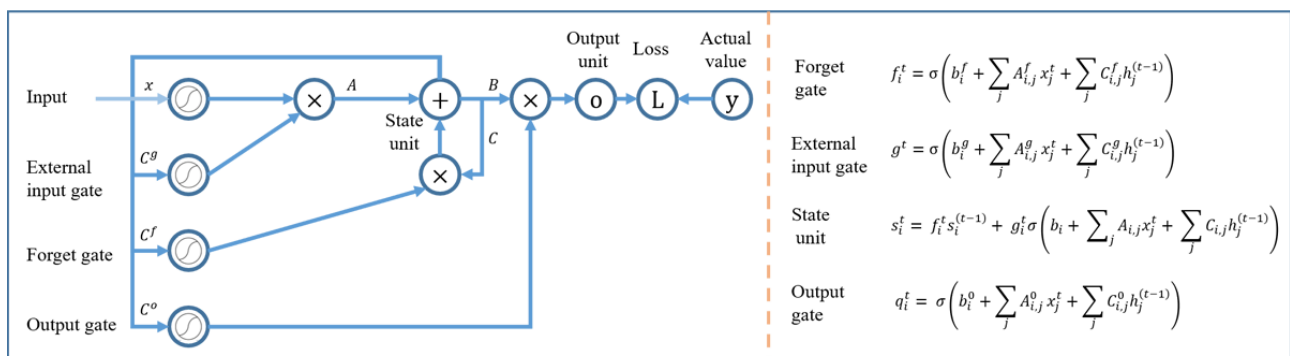
When making a dietary, physical activity, or medication plan for a patient, it is important to consider the time it takes for those changes to affect the patient. In order to provide enough time to observe patient behavior and test the model, we consider approximately 120 days of data for training, 30 days of data for validation, and 30 days of data for testing, where possible. For the patient with a smaller number of available records, we reduced the size of training, validation, and test sets proportionally. For example, for a patient with 41 days of entries, we considered 27 days of data for training, 7 days of data for validation, and 7 days of data for testing.

**Model Construction and Optimization: Long Short-Term Memory–Based Recurrent Neural Networks**

We constructed a specialized RNN known as LSTM for predictive modeling of daily glucose levels using mobile health

time series data. RNNs use the concept of parameter sharing across different layers and can effectively model data sequences of different lengths. However, classical RNNs suffer from the vanishing (often) or exploding (rarely) gradient information problem. Here, we consider an LSTM network that is explicitly designed to avoid the vanishing gradient problem by regulating the information flow using three distinct gates: forget gate, external input gate, and output gate. The forget gate ( $f_i^t$ ) is a linear self-loop weight that decides which information to keep and which to drop. The external input gate ( $g_i^t$ ) helps with deciding which new information to update in an LSTM memory unit/cell. The output gate controls the extent to which the values in the cell are to be used to compute the output activation. The state unit ( $s_i^t$ ) is calculated based on the forget gate and external input gate. The output unit ( $q_i^t$ ) then provides the necessary information to predict the output ( $y^t$ ), which is the predicted glucose level for the next day (Figure 5).

**Figure 5.** Block diagram of the LSTM neural network, where the left portion of the figure shows how an LSTM regulates the information flow using the three control gates and the right portion provides the mathematical equations for the key gates and units of the LSTM model. LSTM: long short-term memory.



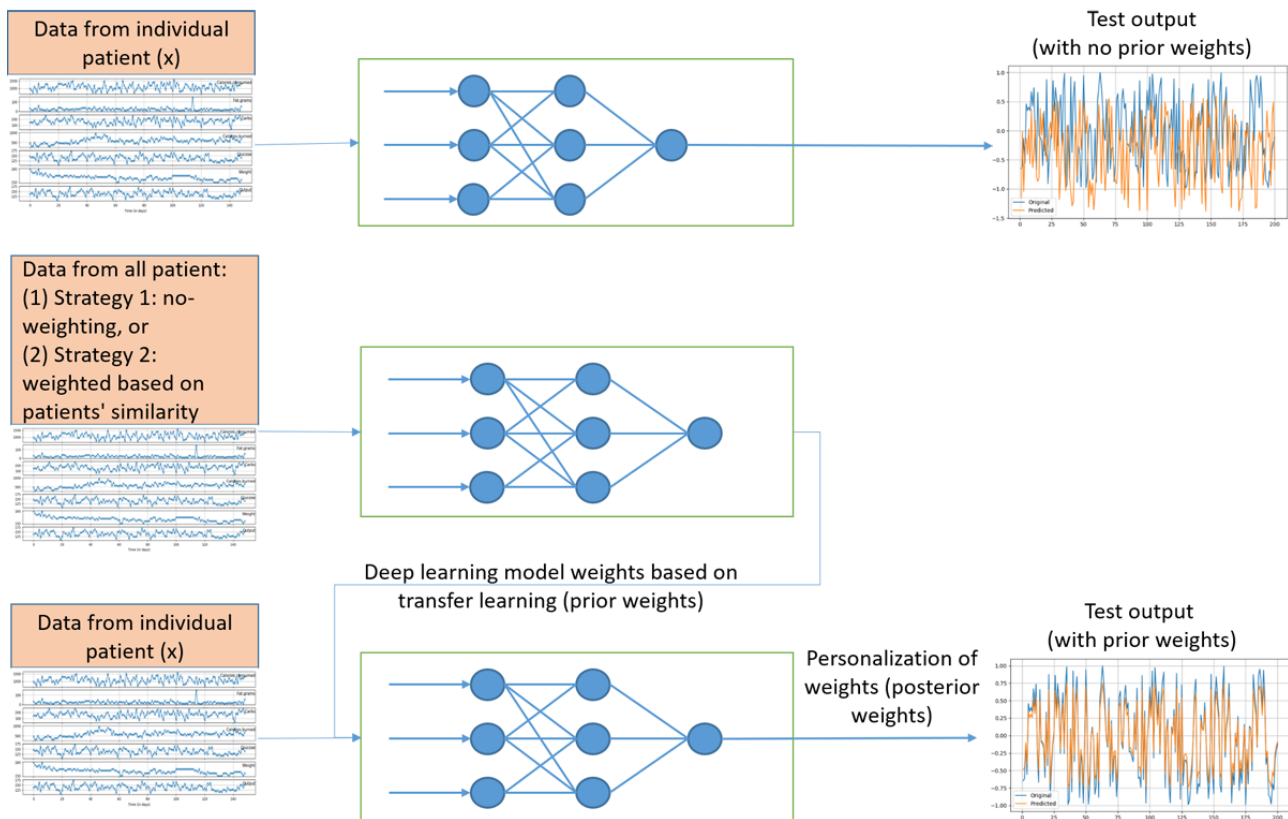
**Knowledge Transfer Across Patients**

It has been shown that transfer learning is useful in learning tasks when the data are scarce, contain many missing/imputed values, and/or suffer from complex patterns [36,37]. Here we developed two transfer learning strategies for coping with data scarcity and improving the predictions for individual patients. The first strategy used all the patient data (training data) to create the transfer learning dataset to pretrain a global LSTM model. The global model was then personalized for each patient based on their individual records. The second strategy used the similarity in glucose patterns between patients to create a transfer learning dataset for each patient. For that, a similarity matrix was created for each patient comparing their glucose patterns with all other patients using dynamic time warping

(DTW) [38]. DTW is often used to compare two dynamic patterns and calculate their similarity by calculating the minimum distance between the two time series and aligning the significant patterns [39]. Next, it creates a transfer learning dataset for each patient by sampling records from other patients according to their similarity. It then uses the sampled data to train a deep learning model for the patient of interest, where the deep learning model weights of the trained model will be used as the prior. Finally, we personalized the deep learning model weights to the patient of interest using their own data. Figure 6 shows a visual representation of the proposed transfer learning strategies. In the results section, we compare the performance of the two transfer learning strategies to the no-transfer learning strategy.



**Figure 6.** General scheme of the proposed transfer learning strategy. For patients with fewer observations, we pretrain a model with observations sampled from all the patients in the dataset (based on either no weighting or weighting strategy). The pretrained model is then fine-tuned with the data of the patient of interest.



## Model Selection and Parameter Tuning

The proposed LSTM model for prediction of glucose level has three hyperparameters to be optimized to achieve the best predictive performance. These three hyperparameters are dropout rate, number of neurons in LSTM layers, and number of neurons in the feed-forward neural network layers [40]. We considered lower bounds of (0.10, 5, 5) and upper bounds of (0.45, 60, 40) for the three hyperparameters, respectively. Optimizing the hyperparameters involved building, training, and validating many versions of the LSTM network based on various choices of the hyperparameters.

Considering an allowable unit change of 0.01 for the dropout rate parameter and 1 for the number of neurons in LSTM and feed-forward layers, we had to test a total of ( $35 \times 55 \times 35 = 67,375$ ) combinations before finding the optimal hyperparameters, which was time-consuming. Thus, to identify the optimal value of the hyperparameters with a minimum number of trials and errors, we used an advanced design of experiments method based on Bayesian optimization [41]. The advanced design of experiment process began by generating a small sample of 15 experimental settings based on the three hyperparameters of the LSTM using a Latin hypercube design [42,43]. Next, for each of the initial set of 15 experimental settings, an LSTM network was built and tested over the validation dataset based on the root mean squared error (RMSE) of the actual versus predicted glucose level. Afterward, using the hyperparameters of the LSTM models as the input and the respective RMSE as the output, a surrogate model was fitted based on a Gaussian process. Then,

the expected improvement criterion was used to identify the optimal point of the Gaussian process which represented the estimated optimal hyperparameter setting of the LSTM. The estimate of the optimal hyperparameter was used as the next hyperparameter setting to experiment [44]. This procedure continued until no improvement in the RMSE was observed. For the LSTM, expected improvement methods converged to the global optimum point after five iterations or 24 total evaluations (15 initial evaluations + 9 additional evaluations).

## Results

### Overview

We considered three variants of the proposed deep learning models for evaluation. The three variants were (1) the base LSTM-NN (without transfer learning), where the model was trained using only the respective patient's data, (2) LSTM-NN-TF-ALL (first transfer learning strategy), where a general model was trained using all patients' data and then personalized based on the target patient's records, and (3) LSTM-NN-TF-DTW (second transfer learning strategy), where separate transfer learning datasets were created for training individualized models for each patient using similarity-based sampling from other patients' records. We evaluated the performance of the deep learning models along with several baseline machine learning methods including an ANN [45], k-nearest neighbors (KNN) regression, ridge regression, kernel ridge regression with Gaussian kernel, and a moving average model. The validation dataset was used for tuning the hyperparameters of the comparing models, such as the optimal

number of nearest neighbors in KNN (found at  $k=3$ ), the sample size in the moving average ( $n=3$ ), and the optimal value of the penalty term in the (kernel) ridge regression.

### Evaluation Criteria

Mean squared error and mean absolute error are commonly used to evaluate the performance of prediction models. However, these criteria do not consider the clinical impact of the prediction error and how it might affect medical decision making. Here, we considered two criteria which were related to the mean squared error and provided information about the clinical impact. The first criterion was the Clark Error Grid [46], which determined the acceptable error for the accuracy of blood glucose prediction in comparison with the actual observation. The second criterion, based on the prescription point of care [47], was the prediction accuracy within the range of  $\pm 10\%$  of the actual value.

### Prediction Accuracy Based on the Clark Error Grid

The Clark Error Grid [46] is one of the most widely used tools to assess the clinical accuracy of blood glucose estimation. The Clark Error Grid is a plot with five major zone of attention (zone A, B, C, D, and E) for interpretation of the predicted glucose levels. Zone A represents those values within 20% of the reference value that generally lead to the appropriate treatment

of patients. Zone B represents those values that are outside zone A, yet do not lead to inappropriate treatment of the patients. Prediction values falling in zone C lead to inappropriate treatment but without any dangerous consequences for the patient. Prediction values on zone D lead to failure in detecting hypoglycemia or hyperglycemia. Finally, prediction values in zone E lead to the inappropriate treatment of hyperglycemia instead of hypoglycemia and vice versa depending on the zone location.

Table 2 summarizes the percentage of prediction points falling in various zones of the Clark Error Grid for each of the comparing methods. As shown in the table, the proposed LSTM-NN-TF-DTW model has the highest percentage of predicted values in zone A (84.12%), followed by the kernel ridge regression (83.03%), and the moving average (82.01%). On the other hand, the moving average and kernel ridge regression have the lowest percentage of predicted values in zone C, D, and E, followed by LSTM and artificial neural network (ANN) models. Overall, ANN provides the lowest performance among all methods, which may be attributed to the large amount of data that it requires and the problem with vanishing gradient in RNN. Multimedia Appendix 1 (Figure C.1) complements Table 2 with visual illustrations of the Clark Error Grid for each of the comparing methods.

**Table 2.** Percentage of prediction points on the Clark Error Grid zones.

Clark Error Grid zone	LSTM-NN-TF-DTW <sup>a</sup> (with transfer learning) (%)	LSTM-NN-TF-ALL <sup>b</sup> (with transfer learning) (%)	LSTM-NN <sup>c</sup> (without transfer learning) (%)	ANN <sup>d</sup> (%)	KNN <sup>e</sup> regression (%)	Ridge regression (%)	Kernel ridge regression (%)	Moving average (last 3 days) (%)
A	84.12	78.17	75.81	69.31	71.12	83.03	76.9	82.01
B	15.16	21.17	23.47	29.96	26.71	16.25	23.1	17.99
C	0	0	0	0	0	0	0	0
D	0.72	0.65	0.72	0.72	2.17	0.72	0	0
E	0	0	0	0	0	0	0	0

<sup>a</sup>LSTM-NN-TF-DTW: second transfer learning strategy.

<sup>b</sup>LSTM-NN-TF-ALL: first transfer learning strategy.

<sup>c</sup>LSTM-NN: without transfer learning.

<sup>d</sup>ANN: artificial neural network.

<sup>e</sup>KNN: k-nearest neighbors.

### Prediction Accuracy Based on the $\pm 10\%$ Range

Table 3 provides the predictive accuracy of the comparing methods based on the  $\pm 10\%$  range of the actual values. As demonstrated in the table, LSTM neural networks generally outperform other methods by a margin of significance. Also, transfer learning strategies provide meaningful improvements

to the LSTM network, with DTW transfer learning (weighted strategy) delivering better results. Meanwhile, there are a couple of exceptions, such as patients 5 and 8, where the moving average method makes better predictions. Further investigation of such cases reveals that those patients suffer from many (adjacent) missing values over a long period of time (see also Multimedia Appendix 1, Part D).

**Table 3.** Prediction accuracy of the proposed deep learning models along with other comparing methods based on the  $\pm 10\%$  range of the actual glucose level value.

Patient #	LSTM-NN-TF-DTW <sup>a</sup> (with transfer learning) (%)	LSTM-NN-TF-ALL <sup>b</sup> (with transfer learning) (%)	LSTM-NN <sup>c</sup> (without transfer learning) (%)	ANN <sup>d</sup> (%)	KNN <sup>e</sup> regression (%)	Ridge regression (%)	Kernel ridge regression (%)	Moving average (last 3 days) (%)
1	76.67	73.33	73.33	73.33	46.67	76.67	43.33	73.33
2	86.67	86.67	83.33	86.67	53.33	86.67	50.00	76.67
3	60.00	60.00	60.00	53.33	50.00	53.33	43.33	53.33
4	85.71	85.71	85.71	71.43	42.86	71.43	71.43	71.43
5	63.33	46.67	16.67	13.33	36.66	66.67	66.67	76.66
6	46.00	36.67	33.33	26.67	10.00	36.67	26.67	43.33
7	33.33	26.67	33.33	30.00	33.33	26.67	26.67	20.00
8	63.33	56.67	66.67	56.67	46.67	60.00	43.33	80.00
9	60.00	16.67	60.00	20.00	26.67	40.00	20.00	40.00
10	73.33	70.00	73.33	73.33	56.67	56.67	63.33	63.33

<sup>a</sup>LSTM-NN-TF-DTW: second transfer learning strategy.

<sup>b</sup>LSTM-NN-TF-ALL: first transfer learning strategy.

<sup>c</sup>LSTM-NN: without transfer learning.

<sup>d</sup>ANN: artificial neural network

<sup>e</sup>KNN: k-nearest neighbors.

## Discussion

### Principal Findings

The objective of this study was to dynamically forecast the next-day glucose levels in patients with T2DM based on their daily mobile health data including diet, physical activity, weight, and glucose levels from the day before. To achieve this objective, we developed an LSTM-based RNN that leverages these data and finds the pattern of glucose level change. We also developed two transfer learning strategies to deal with the issue of data scarcity and/or when a new patient starts using our model. The numerical results show the transfer learning model provided better prediction accuracy, especially in cases where there weren't enough data available (and for patients with high variability). This provided the intuition for building an initial model that worked as a prior while collecting more data to personalize the predictions. Additionally, we used advanced design of experiments to optimize the hyperparameters of the proposed deep learning models with minimum effort. The proposed deep learning models performed well in comparison with the baseline models such as Kernel ridge regression and KNN. This pilot investigation has significant implications for future research studies in using real-world patient-generated lifestyle data to predict blood glucose changes to achieve optimal diabetes management.

The modeling of our study was closer to daily life in the real world involving dynamic data of physical activity, food intake, calories burned, body weight, and blood glucose generated by patients for 6 months. Previous studies have extensively focused on predicting glucose level for T1DM by engineering an artificial pancreas and simulating its insulin delivery to assist with the glycemic control of T1DM, and most predicting

approaches are building upon physiological modeling [48]. Unlike T1DM, which is characterized as absolute inadequate insulin secretion of the body, the management of T2DM is largely determined by lifestyle [49-51]. Meanwhile, unlike the existing literature on forecasting glucose for T2DM based solely on food intake or glucose [52,53], our proposed model forecasts future glucose level more comprehensively by considering dietary habits, physical activity, weight, and previous glucose levels for patients and might provide a practical guide to T2DM management. We admit that blood glucose level can be affected instantly by an extreme lifestyle event (eg, a large amount of carbohydrate consumption or high levels of intensive physical activity without sufficient carbohydrate supplementation) [54,55]. However, our goal is to guide a patient through lifestyle changes (or choices) steadily, considering changing lifestyle choices takes time. Thus, the intention of this model was not to predict short-term blood glucose level variation throughout the day. Instead, it was designed to manage routine lifestyle in T2DM patients, and it is important to help patients understand how their lifestyle behaviors may change their blood glucose level in the very next day.

Glucose prediction was personalized in this model. Even though we have not counted all variations of each individual (demographic conditions, family history of diseases, etc), the prediction model does consider previous blood glucose level, which is the result of the interaction of lifestyle included in the model and other unexamined factors (eg, genes) [56]. In particular, the current glucose level may improve our understanding of glucose dynamics in patients with diabetes and serve as a crucial predicting factor for a future glucose level. For example, in T1DM patients, continuous glucose monitoring (CGM) is evidenced to predict future glucose level with high accuracy [49,57]. While CGM is generally not available and is

not currently recommended for all T2DM patients [58], it is promising that CGM will be available with low-cost and noninvasive devices in the near future [59]. Using our model, as more data are recorded by the patient, the model will become more personalized to them, thus attaining higher accuracy in terms of predicting glucose levels, especially short-term glucose levels, throughout the day. Together with future advancements in characterizing biological traits, a more personalized and proactive diabetes management program will likely become practical. Our study has provided a promising piece of precision health, integrating dynamic lifestyle and daily glucose monitoring. Moreover, by using the prediction model to assimilate the massive amount of lifestyle data, health care providers can provide T2DM management guidance to patients without personally reviewing the data collected by mobile health technologies. This further makes precision health more feasible.

However, it should be noted that the prediction accuracy was low or modest for some participants. There are several possible explanations. First, glucose levels are multifactorial, and it might be hard to predict with limited input. Adding other patient features (eg, age, genetic profiling, demographic conditions, medication usage) [60,61] may increase the accuracy of prediction. Second, individual variability, such as genes and individual differences in glycemic response to lifestyle, further complicate the prediction. Studies are needed to explore models to mimic the interactions among predisposed traits (eg, genes), new input (eg, lifestyle), and the interactions between the two. Third, it could also be that blood glucose has a stronger dependency on short-term lifestyle choices than on long-term choices. For example, if a patient decides to consume a lot of carbohydrates or consume carbs at irregular times or work out irregularly, the glucose value would be unpredictable because there is no such information in the model. This further reinforces the importance of our study trying to predict future glucose levels and guide individual lifestyle choices. We would suggest that a more rigorous study design is needed to help identify the right model to predict future glucose levels. In particular, behaviors such as food consumption, physical activity, or medication use performed right before the predicting glucose level used to test prediction accuracy will need to be considered in the model. A model developed from a rigorous study design and data collection with high prediction accuracy may provide significant clinical implications to manage T2DM.

## Limitations

There are several limitations to this study. First, this study has a relatively small sample size, which may limit our study generalizability. Future studies with larger sample sizes are needed. Second, there are substantial variations in terms of accuracy when predicting blood glucose. Some underlying mechanisms, such as individual variations of age, gender, gut microbiota, and genetic traits [60-62], which are beyond the scope of this study, may have contributed to the variations. Future studies incorporating these factors are warranted. Third, given the nature of using secondary data from a previous trial, the data collection was not designed to predict future glucose levels. For example, the time of glucose level testing and diabetic medication use were not well documented and hard to include in the model. Last, several study participants had a large amount of missing data on glucose monitoring, and data were imputed using standard imputation methods widely used in the literature. Nevertheless, this is one of the first attempts of using digital monitored lifestyle data, weight, and previous glucose levels to predict future glucose levels in T2DM. It provides important information for future studies regarding data collection, model selection, and the implications of glucose prediction for individuals living with T2DM.

## Conclusion

In this work, we proposed a personalized dynamic forecasting model for glucose levels in T2DM patients based on LSTM-based RNN. We developed a transfer learning strategy based on weighted sampling from all patients to improve predictions, especially when dealing with data scarcity. We also used an advanced design of experiments based on Bayesian optimization and expectation maximization for efficient optimization of deep neural network hyperparameters with the minimum number of experiments. We tested our model using a longitudinal mobile health lifestyle dataset of 10 patients who provided self-monitoring data over 6 months on food intake (carbohydrates, fats, and calories), physical activity (exercise time and calories burned), weight, and previous glucose levels. Predicting future glucose levels is challenging as glucose level is determined by multiple factors. Future research with a more rigorous study design is warranted to help identify a model or models to predict future glucose levels.

## Acknowledgments

The Center on Smart and Connected Health Technologies at University of Texas Health San Antonio provided pilot funding for this study. AA is partially supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number 1SC2GM118266-01. YD is supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant TL1 TR002647. JW's contribution is partially supported by the Hugh Roy Cullen Professorship.

## Authors' Contributions

SHAF, AA, and SS developed the proposed algorithms. SHAF preprocessed the data, coded the algorithms, and conducted the numerical studies. RM conducted the advanced design of experiments. JW contributed to the study design and data collection. SHAF, YD, AA, CL, and JW reviewed the results. SHAF, YD, and RM wrote the manuscript; all authors reviewed and edited the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Additional information on models, training procedure, and data analysis.

[PDF File (Adobe PDF File), 1799 KB - [mhealth\\_v7i11e14452\\_app1.pdf](#)]

## References

- Mougiakakou S, Prountzou A, Iliopoulou D, Nikita K, Vazeou A, Bartsocas C. Neural network based glucose—insulin metabolism models for children with type 1 diabetes. *Conf Proc IEEE Eng Med Biol Soc* 2006;1:3545-3548. [doi: [10.1109/IEMBS.2006.260640](#)] [Medline: [17947036](#)]
- Kovatchev B, Patek S, Dassau E, Doyle FJ, Magni L, De Nicolao G, et al. Control to range for diabetes: functionality and modular architecture. *J Diabetes Sci Technol* 2009 Sep;3(5):1058-1065. [doi: [10.1177/193229680900300509](#)] [Medline: [20144419](#)]
- Olefsky JM. Prospects for research in diabetes mellitus. *JAMA* 2001 Feb 07;285(5):628-632. [doi: [10.1001/jama.285.5.628](#)] [Medline: [11176871](#)]
- Narayan KM, Gregg EW, Fagot-Campagna A, Engelgau MM, Vinicor F. Diabetes—a common, growing, serious, costly, and potentially preventable public health problem. *Diabetes Res Clin Pract* 2000 Oct;50 Suppl 2:S77-S84. [doi: [10.1016/s0168-8227\(00\)00183-2](#)] [Medline: [11024588](#)]
- van den Arend I, Stolk R, Krans H, Grobbee D, Schrijvers A. Management of type 2 diabetes: a challenge for patient and physician. *Patient Educ Couns* 2000 May;40(2):187-194. [doi: [10.1016/s0738-3991\(99\)00067-1](#)] [Medline: [10771372](#)]
- Nyenwe EA, Jerkins TW, Umpierrez GE, Kitabchi AE. Management of type 2 diabetes: evolving strategies for the treatment of patients with type 2 diabetes. *Metabolism* 2011 Jan;60(1):1-23 [FREE Full text] [doi: [10.1016/j.metabol.2010.09.010](#)] [Medline: [21134520](#)]
- Blonde L. Current challenges in diabetes management. *Clin Cornerstone* 2005;7 Suppl 3:S6-S17. [Medline: [16545737](#)]
- Hunt CW. Technology and diabetes self-management: an integrative review. *World J Diabetes* 2015 Mar 15;6(2):225-233 [FREE Full text] [doi: [10.4239/wjcd.v6.i2.225](#)] [Medline: [25789104](#)]
- Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. *JMIR Mhealth Uhealth* 2018 Apr 10;6(4):e92 [FREE Full text] [doi: [10.2196/mhealth.4478](#)] [Medline: [29636320](#)]
- Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. *Diabetes Technol Ther* 2009 Feb;11(2):99-106. [doi: [10.1089/dia.2008.0022](#)] [Medline: [19848576](#)]
- Padhye NS, Jing Wang. Pattern of active and inactive sequences of diabetes self-monitoring in mobile phone and paper diary users. *Conf Proc IEEE Eng Med Biol Soc* 2015;2015:7630-7633. [doi: [10.1109/EMBC.2015.7320159](#)] [Medline: [26738059](#)]
- Nundy S, Dick JJ, Solomon MC, Peek ME. Developing a behavioral model for mobile phone-based diabetes interventions. *Patient Educ Couns* 2013 Jan;90(1):125-132 [FREE Full text] [doi: [10.1016/j.pec.2012.09.008](#)] [Medline: [23063349](#)]
- Arsand E, Tatara N, Østengen G, Hartvigsen G. Mobile phone-based self-management tools for type 2 diabetes: the few touch application. *J Diabetes Sci Technol* 2010 Mar;4(2):328-336 [FREE Full text] [Medline: [20307393](#)]
- Wang J, Sereika SM, Chasens ER, Ewing LJ, Matthews JT, Burke LE. Effect of adherence to self-monitoring of diet and physical activity on weight loss in a technology-supported behavioral intervention. *Patient Prefer Adherence* 2012;6:221-226 [FREE Full text] [doi: [10.2147/PPA.S28889](#)] [Medline: [22536058](#)]
- Kebede MM, Zeeb H, Peters M, Heise TL, Pischke CR. Effectiveness of digital interventions for improving glycemic control in persons with poorly controlled type 2 diabetes: a systematic review, meta-analysis, and meta-regression analysis. *Diabetes Technol Ther* 2018 Nov;20(11):767-782. [doi: [10.1089/dia.2018.0216](#)] [Medline: [30257102](#)]
- Faruque LI, Wiebe N, Ehteshami-Afshar A, Liu Y, Dianati-Maleki N, Hemmelgarn BR, Alberta Kidney Disease Network. Effect of telemedicine on glycated hemoglobin in diabetes: a systematic review and meta-analysis of randomized trials. *CMAJ* 2017 Mar 06;189(9):E341-E364 [FREE Full text] [doi: [10.1503/cmaj.150885](#)] [Medline: [27799615](#)]
- Clemens AH, Chang PH, Myers RW. The development of Biostator, a Glucose Controlled Insulin Infusion System (GCIIS). *Horm Metab Res* 1977;Suppl 7:23-33. [Medline: [873440](#)]
- Aczon M, Ledbetter D, Ho L, Gunny A, Flynn A, Williams J. Dynamic mortality risk predictions in pediatric critical care using recurrent neural networks. 2017 Jan 23. URL: <https://arxiv.org/abs/1701.06675v1> [accessed 2019-04-19]
- Plis K, Bunescu R, Marling C, Shubrook J, Schwartz F. A machine learning approach to predicting blood glucose levels for diabetes management. 2014. URL: <https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/view/8737> [accessed 2019-04-19]
- Begg R, Kamruzzaman J, Sarker R. *Neural Networks in Healthcare: Potential and Challenges*. Hershey: IGI Global; 2001.

21. Ma F, Chitta R, Zhou J, You Q, Sun T, Gao J. Dipole: diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. 2017 Presented at: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2017; New York p. 1903-1911. [doi: [10.1145/3097983.3098088](https://doi.org/10.1145/3097983.3098088)]
22. Che C, Xiao C, Liang J, Jin B, Zho J, Wang F. An RNN architecture with dynamic temporal matching for personalized predictions of Parkinson's disease. 2017 Presented at: Proceedings of the SIAM International Conference on Data Mining; Society for Industrial and Applied Mathematics; 2017; Houston p. 198-206. [doi: [10.1137/1.9781611974973.23](https://doi.org/10.1137/1.9781611974973.23)]
23. Hovorka R, Canonico V, Chassin LJ, Haueter U, Massi-Benedetti M, Orsini Federici M, et al. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol Meas* 2004 Aug;25(4):905-920. [Medline: [15382830](https://pubmed.ncbi.nlm.nih.gov/15382830/)]
24. Georga E, Protopappas V, Fotiadis D. Glucose prediction in type 1 and type 2 diabetic patients using data driven techniques. In: Funatsu K, editor. *Knowledge-Oriented Applications in Data Mining*. London: IntechOpen; 2011.
25. Kok P. Predicting blood glucose levels of diabetics using artificial neural networks [Thesis]. Delft: Delft University of Technology; 2004.
26. Zitar R, Al-Jabali A. Towards neural network model for insulin/glucose in diabetics-II. 2005. URL: <https://pdfs.semanticscholar.org/ca91/3f222c97e37029fddd8c1214b0786be92f42.pdf> [accessed 2019-10-09]
27. Tresp V, Briegel T, Moody J. Neural-network models for the blood glucose metabolism of a diabetic. *IEEE Trans Neural Netw* 1999;10(5):1204-1213. [doi: [10.1109/72.788659](https://doi.org/10.1109/72.788659)] [Medline: [18252621](https://pubmed.ncbi.nlm.nih.gov/18252621/)]
28. Li K, Daniels J, Herrero-viñas P, Liu C, Georgiou P. Convolutional recurrent neural networks for blood glucose prediction. 2018. URL: <https://arxiv.org/pdf/1807.03043> [accessed 2019-10-09]
29. Mhaskar HN, Pereverzyev SV, van der Walt MD. A deep learning approach to diabetic blood glucose prediction. *Front Appl Math Stat* 2017 Jul 14;3. [doi: [10.3389/fams.2017.00014](https://doi.org/10.3389/fams.2017.00014)]
30. Quchani S, Tahami E. Comparison of MLP and Elman neural network for blood glucose level prediction in type 1 diabetics. In: Ibrahim F, Osman NA, Usman J, Kadri NA, editors. *3rd Kuala Lumpur International Conference on Biomedical Engineering*. Berlin: Springer; 2007:54-58.
31. Albers DJ, Levine M, Gluckman B, Ginsberg H, Hripcsak G, Mamykina L. Personalized glucose forecasting for type 2 diabetes using data assimilation. *PLoS Comput Biol* 2017 Apr;13(4):e1005232 [FREE Full text] [doi: [10.1371/journal.pcbi.1005232](https://doi.org/10.1371/journal.pcbi.1005232)] [Medline: [28448498](https://pubmed.ncbi.nlm.nih.gov/28448498/)]
32. Hordern MD, Cooney LM, Beller EM, Prins JB, Marwick TH, Coombes JS. Determinants of changes in blood glucose response to short-term exercise training in patients with type 2 diabetes. *Clin Sci (Lond)* 2008 Nov;115(9):273-281. [doi: [10.1042/CS20070422](https://doi.org/10.1042/CS20070422)] [Medline: [18254721](https://pubmed.ncbi.nlm.nih.gov/18254721/)]
33. Gibson BS, Colberg SR, Poirier P, Vancea DMM, Jones J, Marcus R. Development and validation of a predictive model of acute glucose response to exercise in individuals with type 2 diabetes. *Diabetol Metab Syndr* 2013 Jul 01;5(1):33. [doi: [10.1186/1758-5996-5-33](https://doi.org/10.1186/1758-5996-5-33)] [Medline: [23816355](https://pubmed.ncbi.nlm.nih.gov/23816355/)]
34. Aliberti A, Pupillo I, Terna S, Macii E, Di Cataldo S, Patti E, et al. A multi-patient data-driven approach to blood glucose prediction. *IEEE Access* 2019;7:69311-69325. [doi: [10.1109/access.2019.2919184](https://doi.org/10.1109/access.2019.2919184)]
35. Rakthanmanon T, Campana B, Mueen A, Batista G, Westover B, Zhu Q, et al. Addressing big data time series: Mining trillions of time series subsequences under dynamic time warping. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 2013 Sep 01;7(3):10. [doi: [10.1145/2513092.2500489](https://doi.org/10.1145/2513092.2500489)]
36. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010 Oct;22(10):1345-1359. [doi: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191)]
37. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? 2014 Presented at: NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2; 2014; Montreal p. 3320-3328 URL: <http://papers.nips.cc/paper/5347-how-transferable-are-features-in-deep-neural-networks.pdf>
38. Rakthanmanon T, Campana B, Mueen A, Batista G, Westover B, Zhu Q, et al. Addressing big data time series. *ACM Trans Knowl Discov Data* 2013 Sep 01;7(3):1-31. [doi: [10.1145/2513092.2500489](https://doi.org/10.1145/2513092.2500489)]
39. Salvador S, Chan P. Toward accurate dynamic time warping in linear time and space. *J Intel Data Anal* 2007 Oct 10;11(5):561-580. [doi: [10.3233/ida-2007-11508](https://doi.org/10.3233/ida-2007-11508)]
40. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014 Jan;15(1):1929-1958.
41. Jones D, Schonlau M, Welch W. Efficient global optimization of expensive black-box functions. *J Global Optimiz* 1998 Dec 01;13(4):455-492. [doi: [10.1023/A:1008306431147](https://doi.org/10.1023/A:1008306431147)]
42. McKay MD, Beckman RJ, Conover WJ. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 1979 May;21(2):239-245. [doi: [10.1080/00401706.1979.10489755](https://doi.org/10.1080/00401706.1979.10489755)]
43. Pronzato L, Müller WG. Design of computer experiments: space filling and beyond. *Stat Comput* 2011 Apr 1;22(3):681-701. [doi: [10.1007/s11222-011-9242-3](https://doi.org/10.1007/s11222-011-9242-3)]
44. Rasmussen C, Nickisch H. Gaussian processes for machine learning (GPML) toolbox. *J Mach Learn Res* 2010 Nov;11:3011-3015. [doi: [10.7551/mitpress/3206.001.0001](https://doi.org/10.7551/mitpress/3206.001.0001)]

45. Chatterjee S, Dutta K, Xie H, Byun J, Pottathil A, Moore M. Persuasive and pervasive sensing: a new frontier to monitor, track and assist older adults suffering from type-2 diabetes. 2013 Presented at: 46th Hawaii international conference on system sciences; 2013; Washington p. 2636-2645. [doi: [10.1109/hicss.2013.618](https://doi.org/10.1109/hicss.2013.618)]
46. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987;10(5):622-628. [doi: [10.2337/diacare.10.5.622](https://doi.org/10.2337/diacare.10.5.622)] [Medline: [3677983](https://pubmed.ncbi.nlm.nih.gov/3677983/)]
47. Klonoff DC. Point-of-care blood glucose meter accuracy in the hospital setting. *Diabetes Spectr* 2014 Aug;27(3):174-179 [FREE Full text] [doi: [10.2337/diaspect.27.3.174](https://doi.org/10.2337/diaspect.27.3.174)] [Medline: [26246776](https://pubmed.ncbi.nlm.nih.gov/26246776/)]
48. McGrath T, Murphy KG, Jones NS. Quantitative approaches to energy and glucose homeostasis: machine learning and modelling for precision understanding and prediction. *JR Soc Interface* 2018 Jan 24;15(138):20170736. [doi: [10.1098/rsif.2017.0736](https://doi.org/10.1098/rsif.2017.0736)]
49. Taylor PJ, Thompson CH, Brinkworth GD. Effectiveness and acceptability of continuous glucose monitoring for type 2 diabetes management: a narrative review. *J Diabetes Investig* 2018 Jul;9(4):713-725 [FREE Full text] [doi: [10.1111/jdi.12807](https://doi.org/10.1111/jdi.12807)] [Medline: [29380542](https://pubmed.ncbi.nlm.nih.gov/29380542/)]
50. Sami W, Ansari T, Butt NS, Hamid MRA. Effect of diet on type 2 diabetes mellitus: a review. *Int J Health Sci (Qassim)* 2017;11(2):65-71 [FREE Full text] [Medline: [28539866](https://pubmed.ncbi.nlm.nih.gov/28539866/)]
51. Look AHEAD Research Group, Wing RR. Long-term effects of a lifestyle intervention on weight and cardiovascular risk factors in individuals with type 2 diabetes mellitus: four-year results of the Look AHEAD trial. *Arch Intern Med* 2010 Sep 27;170(17):1566-1575 [FREE Full text] [doi: [10.1001/archinternmed.2010.334](https://doi.org/10.1001/archinternmed.2010.334)] [Medline: [20876408](https://pubmed.ncbi.nlm.nih.gov/20876408/)]
52. Albers DJ, Levine M, Gluckman B, Ginsberg H, Hripcsak G, Mamykina L. Personalized glucose forecasting for type 2 diabetes using data assimilation. *PLoS Comput Biol* 2017 Apr;13(4):e1005232 [FREE Full text] [doi: [10.1371/journal.pcbi.1005232](https://doi.org/10.1371/journal.pcbi.1005232)] [Medline: [28448498](https://pubmed.ncbi.nlm.nih.gov/28448498/)]
53. Martinsson J, Schliep A, Eliasson B, Meijner C, Persson S, Mogren O. Automatic blood glucose prediction with confidence using recurrent neural networks. 2018. URL: <http://ceur-ws.org/Vol-2148/paper10.pdf> [accessed 2019-04-19]
54. Esposito K, Ciotola M, Carleo D, Schisano B, Sardelli L, Di Tommaso D, et al. Post-meal glucose peaks at home associate with carotid intima-media thickness in type 2 diabetes. *J Clin Endocrinol Metab* 2008 Apr;93(4):1345-1350. [doi: [10.1210/jc.2007-2000](https://doi.org/10.1210/jc.2007-2000)] [Medline: [18198229](https://pubmed.ncbi.nlm.nih.gov/18198229/)]
55. Younk LM, Mikeladze M, Tate D, Davis SN. Exercise-related hypoglycemia in diabetes mellitus. *Expert Rev Endocrinol Metab* 2011 Jan 01;6(1):93-108 [FREE Full text] [doi: [10.1586/eem.10.78](https://doi.org/10.1586/eem.10.78)] [Medline: [21339838](https://pubmed.ncbi.nlm.nih.gov/21339838/)]
56. Lee PG, Halter JB. The pathophysiology of hyperglycemia in older adults: clinical considerations. *Diabetes Care* 2017 Apr;40(4):444-452. [doi: [10.2337/dc16-1732](https://doi.org/10.2337/dc16-1732)] [Medline: [28325795](https://pubmed.ncbi.nlm.nih.gov/28325795/)]
57. Frandes M, Timar B, Timar R, Lungeanu D. Chaotic time series prediction for glucose dynamics in type 1 diabetes mellitus using regime-switching models. *Sci Rep* 2017 Jul 24;7(1):6232. [doi: [10.1038/s41598-017-06478-4](https://doi.org/10.1038/s41598-017-06478-4)] [Medline: [28740090](https://pubmed.ncbi.nlm.nih.gov/28740090/)]
58. Sakurai K, Kawai Y, Yamazaki M, Komatsu M. Prediction of lowest nocturnal blood glucose level based on self-monitoring of blood glucose in Japanese patients with type 2 diabetes. *J Diabetes Complications* 2018 Dec;32(12):1118-1123. [doi: [10.1016/j.jdiacomp.2018.09.007](https://doi.org/10.1016/j.jdiacomp.2018.09.007)] [Medline: [30293932](https://pubmed.ncbi.nlm.nih.gov/30293932/)]
59. Cappon G, Acciaroli G, Vettoretti M, Facchinetti A, Sparacino G. Wearable continuous glucose monitoring sensors: a revolution in diabetes treatment. *Electronics* 2017 Sep 05;6(3):65. [doi: [10.3390/electronics6030065](https://doi.org/10.3390/electronics6030065)]
60. Rich SS, Cefalu WT. The impact of precision medicine in diabetes: a multidimensional perspective. *Diabetes Care* 2016 Nov;39(11):1854-1857 [FREE Full text] [doi: [10.2337/dc16-1833](https://doi.org/10.2337/dc16-1833)] [Medline: [27926886](https://pubmed.ncbi.nlm.nih.gov/27926886/)]
61. Florez J. Precision medicine in diabetes: is it time? *Diabetes Care* 2016 Jul 01;39(7):1085-1088. [doi: [10.1016/b978-0-12-800685-6.00014-x](https://doi.org/10.1016/b978-0-12-800685-6.00014-x)] [Medline: [27289125](https://pubmed.ncbi.nlm.nih.gov/27289125/)]
62. Utzschneider KM, Kratz M, Damman CJ, Hullar M. Mechanisms linking the gut microbiome and glucose metabolism. *J Clin Endocrinol Metab* 2016 Apr;101(4):1445-1454 [FREE Full text] [doi: [10.1210/jc.2015-4251](https://doi.org/10.1210/jc.2015-4251)] [Medline: [26938201](https://pubmed.ncbi.nlm.nih.gov/26938201/)]

## Abbreviations

- ANN:** artificial neural network
- CGM:** continuous glucose monitoring
- DTW:** dynamic time warping
- HbA<sub>1c</sub>:** hemoglobin A1c
- KNN:** k-nearest neighbors
- LSTM:** long short-term memory
- LSTM-NN:** without transfer learning
- LSTM-NN-TF-ALL:** first transfer learning strategy
- LSTM-NN-TF-DTW:** second transfer learning strategy
- RCT:** randomized controlled trial
- RMSE:** root mean squared error
- RNN:** recurrent neural networks
- T1DM:** type 1 diabetes mellitus

**T2DM:** type 2 diabetes mellitus

*Edited by G Eysenbach; submitted 22.04.19; peer-reviewed by J Martinsson, DCJ( Wu, D Goldner; comments to author 31.05.19; revised version received 26.08.19; accepted 24.09.19; published 01.11.19.*

*Please cite as:*

*Faruqui SHA, Du Y, Meka R, Alaeddini A, Li C, Shirinkam S, Wang J*

*Development of a Deep Learning Model for Dynamic Forecasting of Blood Glucose Level for Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial*

*JMIR Mhealth Uhealth 2019;7(11):e14452*

*URL: <https://mhealth.jmir.org/2019/11/e14452>*

*doi: [10.2196/14452](https://doi.org/10.2196/14452)*

*PMID: [31682586](https://pubmed.ncbi.nlm.nih.gov/31682586/)*

©Syed Hasib Akhter Faruqui, Yan Du, Rajitha Meka, Adel Alaeddini, Chengdong Li, Sara Shirinkam, Jing Wang. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 01.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Mental Health Apps in China: Analysis and Quality Assessment

Jie Shang<sup>1</sup>, MSc; Shaoming Wei<sup>2</sup>, MPH; Jianbo Jin<sup>2</sup>, BSc; Puhong Zhang<sup>1</sup>, PhD

<sup>1</sup>The George Institute for Global Health, Peking University Health Science Center, Beijing, China

<sup>2</sup>School of Public Health, Peking University Health Science Center, Beijing, China

**Corresponding Author:**

Puhong Zhang, PhD

The George Institute for Global Health  
Peking University Health Science Center  
18F Jinqiu Tower B, No 6 Zhichun Rd  
Haidian District  
Beijing, 100088  
China

Phone: 86 10 8280 0577

Email: [zpuhong@georgeinstitute.org.cn](mailto:zpuhong@georgeinstitute.org.cn)

## Abstract

**Background:** Mental disorders have been a great burden on health care systems, affecting the quality of life of millions of people worldwide. Developing countries, including China, suffer from the double burden of both the increasing mental health issues in population and the deficiency in mental health care resources. The use of mobile health technologies, especially for mobile phone apps, can be a possible solution.

**Objective:** This review aimed to describe the features and assess the quality of mental health apps in major mobile phone app markets in China and further discuss the priorities for mental health app development.

**Methods:** Keywords including *psychology, psychological health, psychological hygiene, psychological health service(s), mental, mental health, mental hygiene, mental health service(s), depression, and anxiety* were searched in Chinese in 3 Android app markets (*Baidu Mobile Assistant, Tencent MyApp, and 360 Mobile Assistant*) and iOS App Store independently. Mental health apps were then selected according to established criteria for in-depth analysis and quality assessment by the Mobile App Rating Scale.

**Results:** In total, 63 of 997 mental health apps were analyzed in depth, of which 78% (49/63) were developed by commercial entities for general population, 17% (11/63) were for patients or clients of specialized psychiatric hospitals or counseling agencies, 3% (2/63) were by government or local Centers for Disease Control and Prevention for general information, and 2% (1/63) for students of a university. Major built-in features of the apps included counseling services, mental health education, and self-assessment of mental health status by validated self-rating scales. The overall quality score of the MH apps was *acceptable*.

**Conclusions:** Mental health apps are emerging in the area of mobile health in China. Popular mental health apps usually provide a synthetic platform organizing resources of information, knowledge, counseling services, self-tests, and management for the general population with mental health-related inquiries. The quality of the apps was rated as *acceptable* on average, suggesting some space for improvement. Official guidelines and regulations are urgently required for the field in the future.

(*JMIR Mhealth Uhealth* 2019;7(11):e13236) doi:[10.2196/13236](https://doi.org/10.2196/13236)

**KEYWORDS**

mental health; mental disorder; quality assessment; mobile health; digital health; innovative health; smartphone application

## Introduction

**Background**

The global burden of mental disorders has been estimated to account for 32.4% of years lived with disability and 13.0% of disability-adjusted life years [1]. The global lifetime prevalence of any mental disorder ranges from 47.4% in the United States

to 12% in Nigeria [2]. In China, the lifetime prevalence of any mental disorder is 16.6% [3].

Challenges, however, are critical for mental health care in China. Clinical resources are limited, but the service demands are increasing fast for mental health-related inquiries, including anxiety and depression [4-6]. The lifetime contact rate of patients with common mental disorders, such as depression,

with any mental health care was around 2.7% in Beijing and 3.1% in Shanghai, which was strikingly low even in major metropolitan areas, whereas for the treated patients with any type of mental disorder, only 37.5% measured up to the minimum treatment adequacy [7]. The gap in human resources is also alarming, given that there are 1.7 clinical psychiatrists per 100,000 people in China, compared with 12 to 15 per 100,000 people in countries such as Australia and the United States [7,8]. The mental health resource distribution is also uneven, as most mental health professionals and mental health care facilities are concentrated in the Eastern China, having the more developed regions [9]. Although the management of severe mental disorders has been included in the basic public health service (BPHS) package by the Chinese government, the workload of frontline community health care workers is critical [10,11]. In addition, mental disorders are still strongly stigmatized, and people are reluctant to seek professional help [12-15]. Supportive social environment is also lacking for effective mental health education and early intervention [16]. Therefore, it is urgent to improve the access and delivery of mental health services and reallocate the resources in China.

### Objective

On the other hand, driven essentially by the development of novel technology and the demands of mental health services, mobile phone apps have given birth to the era of digitalized mental health care globally [17-20]. Mental health apps also expand exponentially, especially in developing countries [21-23]. China has a gigantic and growing market of both internet and mobile phone usage [24]. By the end of 2017, there were 753 million Chinese people connected to the internet through the mobile phone, accounting for 97.5% of all internet users in the country [24]. Recently, fruitful research has been done exploring contents and features of mobile phone apps in mental health-related areas globally, for example, smoking cessation, suicide prevention, and bipolar disorder treatment [25-28]. There is a study investigating the features and characteristics of apps on maternal and child health in China [29]. However, no academic literature discussing mental health-relevant apps in China was identified.

Therefore, this study aimed to investigate the features and further evaluate the quality of mental health apps in major mobile phone app markets in China and further discuss the priorities for mental health app development in the future.

## Methods

### Overall Procedures

We adopted the following 4 steps to identify relevant mental health apps and analyze the data: preselection, selection, data extraction, and app quality evaluation. All the 4 steps or procedures were conducted by 2 independent investigators following established criteria and procedures. Any discrepancy that occurred was resolved through discussion under the supervision of a senior researcher.

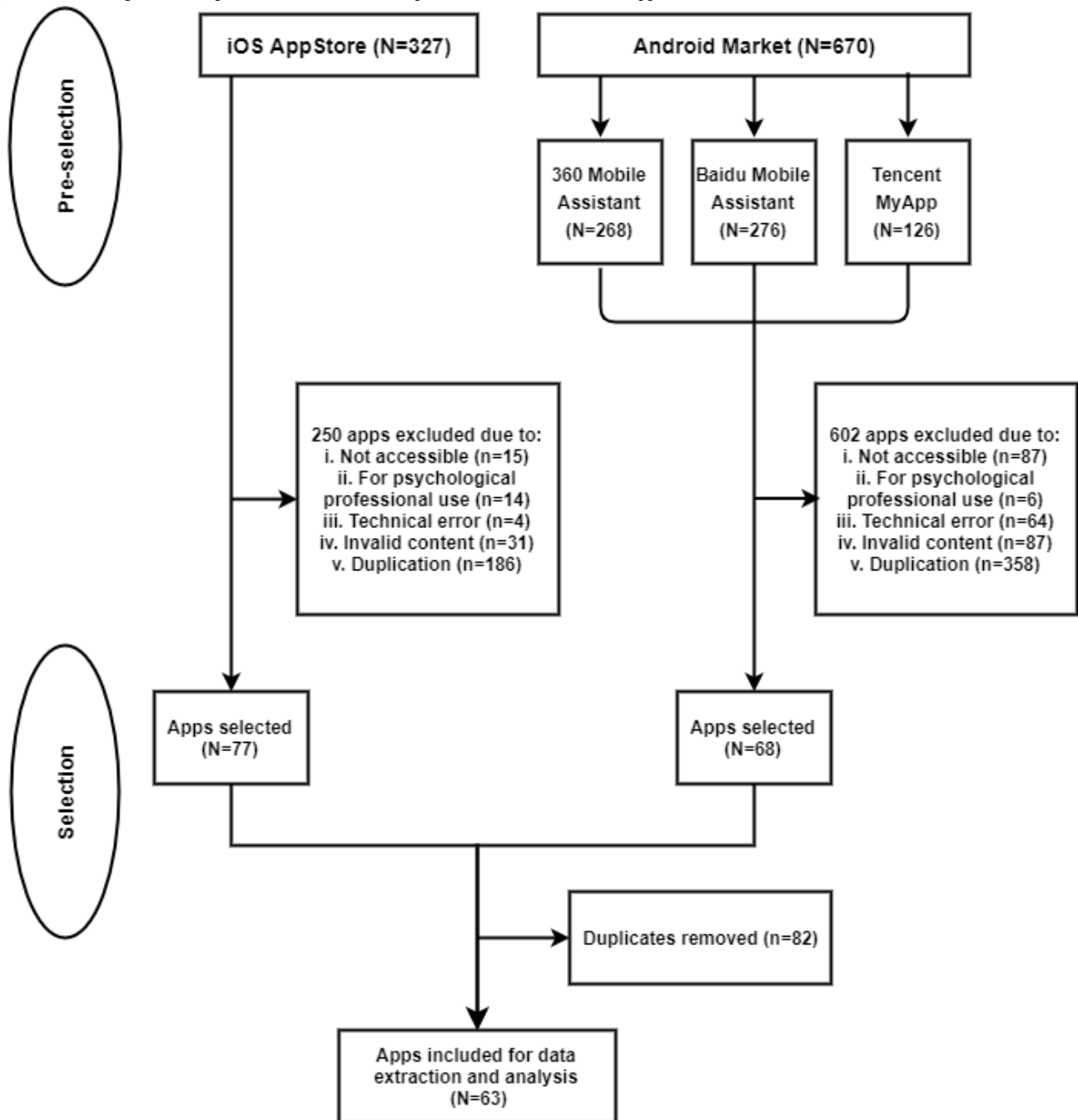
### Preselection of Mental Health Apps

We chose the top 3 Android app markets, *Baidu Mobile Assistant*, *Tencent MyApp*, and *360 Mobile Assistant*, with the biggest market shares in China, and iOS App Store to identify mental health apps [30].

Keywords *psychology* () , *psychological health* () , *psychological hygiene* () , *psychological health service(s)* () , *mental* () , *mental health* () , *mental hygiene* () , *mental health service(s)* () , *depression* () , and *anxiety* () were searched in mandarin Chinese on May 14, 2019. Mental health apps were preselected using the following process: (1) for each keyword, the top 50 apps (if available) appeared were reviewed based on their names and app description to identify the relevant apps; and (2) for each app market, all the relevant apps under the 10 keywords were collected, which gave the number of preselected apps. Apps preselected from the 3 Android markets were combined into 1 group, whereas apps from the iOS App Store were collected as a separate group (Figure 1).

The number of apps listed in the results of keyword search is unknown to customers in the iOS App Store, so we needed to confirm a range for selection. Moreover, we tested and found that the number of relevant apps appeared under each mental health-related keyword was less than 50 when counting downward from the top of the list. As searching through the 10 keywords, duplication was increasing. Therefore, we adopted the strategy of searching for the top 50 apps for each of the 10 keywords in every app store and then conducting deduplication. It is thus very likely to exhaust the relevant mental health apps in the China market.

Figure 1. Flow diagram of the preselection and selection procedures of mental health apps.



**Selection of Mental Health Apps**

All preselected apps were downloaded, installed, and registered on an iPhone 8s Plus and a Huawei P30 Pro for further selection. We accessed each app to further examine whether it was

adequate for feature analysis and quality assessment with regard to the inclusion and exclusion criteria (Textbox 1). After removing the inadequate and duplicate apps, the remaining apps were combined for data extraction, analysis, and quality evaluation (Figure 1).

**Textbox 1.** Inclusion and exclusion criteria for mental health app selection.

<p>Inclusion criteria</p> <p>Mental health apps would be included if they were:</p> <ul style="list-style-type: none"> <li>• accessible;</li> <li>• targeted at the individual mental health consumer, excluding mental health professionals and companies;</li> <li>• usable;</li> <li>• focused on relevant contents for people seeking professional assistance; and</li> <li>• developed in Chinese or has a Chinese version.</li> </ul> <p>Exclusion criteria</p> <p>Mental health apps would be excluded if they were:</p> <ul style="list-style-type: none"> <li>• not accessible (requiring enterprise/school identification to access);</li> <li>• end products for doctors or commercial enterprises;</li> <li>• not usable due to technical errors;</li> <li>• focused on irrelevant contents such as entertainment and gaming, training and education, social and communication, business and commercial, romance and relationship, and e-books, which are irrelevant to mental health service; and</li> <li>• duplication of apps with identical internal design or contents in the market.</li> </ul>
--

**Data Extraction and Analysis**

A table sheet was generated and saved in a Microsoft Excel file with specific variables to record the basic information and potential features of different aspects of each of the included apps. We extracted the data of apps regarding the themes and contents of mental health–informative articles (emotion control, depression, anxiety, stress management, and familial relationship), self-assessment of mental health status (different self-rating scales testing the level of depression, anxiety, autism, sleep quality, and personality), counseling services (types of counselors and ways of communication), mental health courses, and meditation. Data calculation and analysis were performed using SAS 9.4 on Microsoft Windows.

**Quality Evaluation Using Mobile App Rating Scale  
App Quality Criteria**


The quality of included apps was evaluated using Mobile App Rating Scale (MARS) [31]. MARS contains 23 individual items sorted into the following 5 criteria categories: engagement, functionality, aesthetics, information, and subjective qualities. Each was rated by 2 independent investigators. The score of each individual item was rated from 1 to 5 points (1: inadequate, 2: poor, 3: acceptable, 4: good, and 5: excellent). Mean scores were calculated for the following 4 objective quality criteria: engagement, functionality, aesthetics, and information. In addition, a mean total score was calculated across the 4 objective categories. The 2 investigators were trained to use MARS by scoring a small group of apps individually and adjusting for discrepancies in understanding of the criteria. Disagreements were resolved by thorough discussions between the 2 investigators. We did not use the MARS app classification because of the fact that all the apps included were classified as mental health apps, and our focus was on app quality evaluation.

**Results****Basic Information of Included Mental Health Apps**

In total, 997 preselected mental health apps from Android and iOS markets were found through keyword search in the preselection (Figure 1). We found 327 relevant mental health apps from the iOS App Store, 268 apps from 360 Mobile Assistant, 276 apps from Baidu Mobile Assistant, and 126 from Tencent MyApp. Through further selection with respect to the inclusion and exclusion criteria, 63 unique apps were included for quality assessment and in-depth analyses (Figure 1).

Of the included apps, 78% (49/63) were developed by commercial entities providing mental health services, for example, Web-based mental health articles, psychological counseling, and mental health status assessment by self-rating scales. There were also 17% (11/63) apps developed by specialized psychiatric hospitals or psychological counseling agencies, featured by relevant services including appointment booking and consultation, for individual patients or clients of the entity. The rest were 2 apps provided by government or district-level Centers for Disease Control and Prevention (CDC) in Beijing, and 1 app developed for the students of Chongqing University.

**Features Built in Mental Health Apps**

According to the findings above, the mental health apps in China were aimed to provide as broad range of mental health services as possible to nonspecific users. We only found 1 app purely for communication between depression patients (*Snail* ). No apps were found purely for other common mental disorders, for examples, substance abuse and anxiety. Simultaneously, there has not been an app designed for the BPHS delivery in China that involves extensive procedures for the management of patients with severe mental disorders, including patient identification and registration, home visits, and drug

dissemination [10]. Therefore, the following features were summarized with respect to the detailed services provided by the mental health apps found in the China market: the common features found in the mental health apps included were mental health education (67%, 42/63) either by informative articles or relevant courses, counseling services (65%, 41/63), self-assessment of mental health status (44%, 28/63), and question and answer (Q&A) module (40%, 25/63) providing quick answer to mental health–relevant questions that were available to all users. Other features included meditation (11%, 7/63) and self-management tools (16%, 10/63) such as drug description and intake alert and hotline for general mental health inquiries.

### **Mental Health Education**

The mental health education component in the mental health apps can be accessed by either reading informative articles (38/63, 60%) or taking relevant courses (23/63, 37%). Informative articles addressed the causes, risk factors, symptoms, and coping skills of common psychological disorders or conditions, for instance, depression (33/63, 52%), anxiety (25/63, 40%), and emotion coping (34/63, 54%). For the apps providing mental health–related courses, 18 apps offered lectures with slides or recordings on the concise scientific knowledge of common mental disorders. In addition, 7 apps contained courses that were designed based on verified psychological therapy or training, for example, cognitive behavioral therapy (CBT) [32]. In total, 14 apps provided free courses. For paid courses, fees ranged largely according to different types across the apps.

### **Mental Health Counseling Services**

Mental health counseling services were an essential feature in the included apps. A total of 80% (33/41) of the counseling services in these apps were provided by mental health professionals with verification. Counselors were verified by obtaining personally identifiable information or photocopy of their Counselor Certificate, and clinical psychiatrists were verified by obtaining the following information: name, title, and the hospital they work with. The users can purchase one-to-one counseling sessions through the apps. Web-based counseling sessions were a common method, including audio phone calls and video connection, available in 90% (37/41) of apps with the feature. Users can also write reviews and rate the counseling service, which was available to other app users. The fee for an mental health counseling session varied from less than 50 *yuan* to over 1000 *yuan* per hour.

### **Self-Assessment of Mental Health Status**

Self-assessment of mental health status was conducted through self-rating scales. There were 28 apps providing at least 1 verified scale for depression or anxiety. Zung Self-rating Scale for Depression (21/28, 75%) and Zung Self-Rating Scale for Anxiety (SAS; 22/28, 79%) were the most frequently used tests for depression and anxiety detection, respectively [33,34]. Other scales used for depression self-diagnosis included Symptom Checklist 90 (15/28, 54%), Beck Depression Inventory (10/28, 36%), and Patient Health Questionnaire (PHQ) 9 (8/28, 29%) [35,36]. Scales for anxiety assessment also included Generalized Anxiety Disorder Screener (GAD-7; 4/28, 14%) and Beck Anxiety Inventory (6/28, 21%) [37,38]. Moreover, half (14/28, 50%) of the apps provided validated self-rating scales on sleep quality and somatization symptoms, including Athens Insomnia Scale and PHQ-15 [39,40]. In total, 89% of these apps (25/28) provided results for the self-rating scales.

### **Other Features of Mental Health Apps**

Q&A section was available in 40% (25/63) of all the included mental health apps. Users can post questions relevant to mental health issues. Both mental health professionals and general users can respond to the questions, and all answers were open to be viewed by all users. Self-management–related features included tools for self-monitoring and management, such as mood diary, medication record, and reminder. There were also 11% apps (7/63) designed with modules on meditation training for everyday practice that aimed to improve stress coping, emotion regulation, and sleep quality.

### **Quality Assessment Scores by Mobile App Rating Scale**

The mean overall score of MARS was 2.96 for the total of 63 mental health apps (Table 1). Engagement was rated lowest of the 4 objective app quality criteria, which was 2.72. The highest was the mean function score, which was 0.53 higher than engagement. The aesthetics and information were similar, and both did not exceed the mean overall score.

In total, 56 apps can be classified as *acceptable*, and 8 apps as *good* according to the mean overall MARS scores given by the criteria (Table 2). However, there were still 7 apps rated as *poor* quality. For apps in the engagement category, 42 apps were classified into *acceptable* or above, whereas as high as 21 apps were classified into the *poor* group. The aesthetics category was similar to the mean overall group; the majority (55/63, 87%) of apps were *acceptable* or above, leaving 8 apps in the *poor* group. In the information category, 81% of the apps were *acceptable* or above, but still 12 apps were in the *poor* group. The function category was, in general, rated higher, given that 97% of all apps were classified into *acceptable* or *good* groups.

**Table 1.** Mean Mobile App Rating Scale scores of mental health apps (N=63).

Mobile App Rating Scale category	Minimum	Maximum	Mean (SD)
Engagement	1.8	4.0	2.72 (0.54)
Function	2.3	4.3	3.25 (0.06)
Aesthetics	2.0	4.0	2.99 (0.06)
Information	2.0	4.0	2.88 (0.54)
Overall	2.1	4.0	2.96 (0.40)

**Table 2.** Classification of mental health apps according to the mean Mobile App Rating Scale scores (N=63).

Mobile App Rating Scale category	Inadequate, n (%)	Poor, n (%)	Acceptable, n (%)	Good, n (%)	Excellent, n (%)
Engagement	— <sup>a</sup>	21 (33)	36 (57)	6 (10)	—
Function	—	2 (3)	38 (60)	23 (37)	—
Aesthetics	—	8 (13)	44 (70)	11 (17)	—
Information	—	12 (19)	42 (67)	9 (14)	—
Overall	—	7 (11)	48 (76)	8 (13)	—

<sup>a</sup>No app fell into this category.

## Discussion

### Summary and Discussion of Major Findings

Through the investigation into 63 of 997 mental health apps from iOS and 3 Android markets in China, major features and characteristics have been summarized from informative details. Features, including mental health education, counseling services, self-assessment of mental health status, and Q&A section, were commonly found in the apps across stores. The apps played the role of a platform gathering the users and the mental health professionals, allowing services to be delivered through the innovative platform. Moreover, the mean of overall MARS scores has shown that the quality was acceptable for the mental health apps. Quality scores of the 4 objective categories of MARS, engagement, aesthetics, function, and information, were similar to the overall rating. However, the apps, in general, rated better in the function category than others.

The rise of mental health apps in the China market might be because of deficient knowledge and negative perception toward mental disorders [12,41]. It has been a cultural stigma that mental illnesses in China are often considered as abnormal, and patients with mental disorders are shamed, especially in less developed rural area [10,41]. Therefore, people more tend to refer mental issues to physical symptoms, which may prevent people from actively seeking professional help [10]. Mobile health technologies, such as mental health apps, have the potential to protect patients' privacy and reduce the impact of social stigma toward mental illnesses.

There were common features of mental health apps in China compared with apps in the global market, which include real-time communication through Web-based services and mental health education [27]. From the analyses of this study, most of the mental health apps provided Web-based counseling services. People can make real-time communication with their preferred mental health professionals in terms of mental health inquiries. Moreover, mental health information offered by the app can be an efficient tool to improve mental health literacy [42]. On the other side, differences were also found between mental health apps in China and global market. Apps in global market were commonly seen to be designed for a specific mental disorder or condition, for example, anxiety and depression [43,44]. Contrastingly, mental health apps in China were very likely to include information and services for a broad range of mental disorders or conditions in 1 single app. They often included a group of features in 1 app, such as counseling services, self-tests, and other mental health education on a broad

range of psychological issues, instead of dealing with 1 particular disorder.

Moreover, most of the apps included in this study were developed by commercial entities for commercial purpose. Public health impact has hardly been stressed. More apps are expected to pay more attention to the common advantage from the public health perspective. For example, the patient management of severe mental disorders is a critical component of the BPHS in China, together with health records creation for every resident, health education, immunization, chronic disease (hypertension and diabetes) management, and so forth [11]. In China, BPHS has been provided through a ladder system of county hospital and CDC, township health care centers, and village clinics [11]. The frontline community health care workers, who are more often called village doctors working at village clinics in rural China, have critical workload of patient finding and registration, home visit, and drug dissemination [5,20,45,46]. However, no mental health app was found to assist in such public health care services. Innovative health technologies are promising in mental health care practice in community and the whole mental health system of the country.

### Implications From Features of the Mental Health Apps

The popularity of counseling services reflects the increasing demand in mental health interventions of the general population. The apps act as a platform bridging the gap between mental health professionals and the app users, rather than directly engage in interventions for patients with the need of mental health care. Therefore, the mental health apps in China did not involve much core psychological methodologies but more incorporate functions of coordination and communication.

Web-based counseling outperforms traditional face-to-face counseling in a few ways. One aspect is that it allows reorganization of psychiatric resources, especially for available professionals. This can indeed benefit the Chinese population that is facing unbalanced and limited resources in mental health care [9]. However, concerns have also been raised regarding the quality of service provided by those psychological professionals. There is no standardized guidelines or official regulations for conducting Web-based counseling [47,48]. Moreover, it would be difficult for the mental health apps, as commercial entities, to effectively supervise counselors for their service.

Mental health apps can promote mental health education to lay public by offering informative articles and relevant courses. Low mental health literacy has long been considered as a major

obstruction in help-seeking and treatment delivery for people with mental disorders in China [42,49]. However, quality concern has also been raised toward the mental health educational materials provided by mental health apps. It is necessary to establish more official guidance and promotion to improve mental health literacy.

Only 2 apps (*Alpha psychology* [ ] and *Jianxinjiayuan* [ ]) were found to be based on formal psychotherapeutic techniques of CBT [32]. CBT has been found effective in a variety of mental disorders and conditions, including affective disorders, psychosis, gambling behavior, and substance abuse [50-53]. Many apps from global market were also designed to incorporate the core techniques of CBT to elicit a cognitive and behavioral change of mental disorder and conditions, such as substance abuse and affective disorders [26,54]. More mental health techniques and psychotherapy components, including techniques used in CBT, are expected to be involved in designing the feature modules.

Self-rating scales of common mental disorders, such as depression and anxiety, were prevalent in mental health apps. Verified self-rating scales were convenient tools for quick assessment. Some apps even provided multiple rating scales for 1 disorder that allows users to compare GAD-7 and SAS for anxiety testing. However, there were drawbacks in using the Web-based version of self-rating scales. These include errors in the questions, as well as misunderstanding and misinterpretation of the test results. Most self-rating scales were written originally in English and translated directly into Chinese; therefore, the questions were asked in the way more suitable for Western lifestyle and culture [55]. Translation errors and discrepancies always occurred. Moreover, there are special requirements for doing self-tests; for instance, some questions may require intuitive answers, but some may involve precise recalling. Therefore, it is better to have the presence of a trained mental health professional when scoring the scales. In addition, lay users may feel to be overlabeled for a mental disorder, and unnecessary panic can occur [27,56].

In addition, features of self-management were found in some apps, including mood diary, drug description, and medication reminder. Those features can be practical tools to support regular treatment for clinical patients already diagnosed with mental disorders [20]. However, they were overall less *smart* and not relevant to mental health, for example, a drug reminder is a simple clock. Considering gaps currently in the mental health

services in China, the developers might put more efforts to combine the potential of smart technologies and mental health relevance [5]. Meditation is another growing aspect of mental health app features. The apps provided simple guidance by recorded audio clips. Therefore, it is 1-direction demonstration, and no feedback was required from the users. However, the users did not have the opportunity to correct or improve in everyday practice.

### Quality Evaluation by Mobile App Rating Scale

This study used MARS quality criteria to evaluate the 63 included mental health apps [31]. Both mean overall score and individual score of the 4 objective categories had shown that the majority of the apps were *acceptable*, which was the middle of the 5-point scale. The quality did not vary much across the 4 trajectories of engagement, aesthetics, function, and information, but the function category was rated to be slightly higher. The mean scores of MARS suggested that there would be space for improvement in the above trajectories. For instance, mental health apps can pay more attention to the visual appeal, content engagement, and display of app-relevant information in the market.

### Strengths and Limitations

This is the first academic review on the characteristics and common features of mental health apps in China. Limitations might include that we searched mental health apps by keyword, following the default algorithm of each app market. This is to simulate the real-world situation when mobile phone users search for mental health apps in common app markets. Moreover, built-in features that required in-app purchase to be accessed were out of the scope of analyses. For example, one should pay for a counseling session to communicate with a counselor in an app with that feature. However, the analyses of this study were purely descriptive. Further research on experiences of the paid features is expected by qualitative study in the future.

### Conclusions

In conclusion, this study has systematically reviewed and investigated the built-in features of mental health apps in current China market. Moreover, the quality of the apps was also evaluated using valid assessment criteria. Overall, the area of mental health-related innovative technologies in China is experiencing fast development, and appropriate guidance would be beneficial to the growth of the field.

---

### Conflicts of Interest

None declared.

---

### References

1. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *Lancet Psychiatry* 2016 Feb;3(2):171-178. [doi: [10.1016/S2215-0366\(15\)00505-2](https://doi.org/10.1016/S2215-0366(15)00505-2)] [Medline: [26851330](https://pubmed.ncbi.nlm.nih.gov/26851330/)]
2. Kessler RC, Angermeyer M, Anthony JC, de Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* 2007 Oct;6(3):168-176 [FREE Full text] [Medline: [18188442](https://pubmed.ncbi.nlm.nih.gov/18188442/)]

3. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yan J, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry* 2019 Mar;6(3):211-224. [doi: [10.1016/S2215-0366\(18\)30511-X](https://doi.org/10.1016/S2215-0366(18)30511-X)] [Medline: [30792114](https://pubmed.ncbi.nlm.nih.gov/30792114/)]
4. Phillips MR. Mental health services in China. *Epidemiol Psychiatr Soc* 2011 Oct 11;9(2):84-88. [doi: [10.1017/s1121189x00008253](https://doi.org/10.1017/s1121189x00008253)]
5. Liu J, Ma H, He YL, Xie B, Xu YF, Tang HY, et al. Mental health system in China: history, recent service reform and future challenges. *World Psychiatry* 2011 Oct;10(3):210-216 [FREE Full text] [doi: [10.1002/j.2051-5545.2011.tb00059.x](https://doi.org/10.1002/j.2051-5545.2011.tb00059.x)] [Medline: [21991281](https://pubmed.ncbi.nlm.nih.gov/21991281/)]
6. Phillips MR, Chen H, Diesfeld K, Xie B, Cheng HG, Mellso G, et al. China's new mental health law: reframing involuntary treatment. *Am J Psychiatry* 2013 Jun;170(6):588-591. [doi: [10.1176/appi.ajp.2013.12121559](https://doi.org/10.1176/appi.ajp.2013.12121559)] [Medline: [23625139](https://pubmed.ncbi.nlm.nih.gov/23625139/)]
7. Patel V, Xiao S, Chen H, Hanna F, Jotheeswaran AT, Luo D, et al. The magnitude of and health system responses to the mental health treatment gap in adults in India and China. *Lancet* 2016 Dec 17;388(10063):3074-3084. [doi: [10.1016/S0140-6736\(16\)00160-4](https://doi.org/10.1016/S0140-6736(16)00160-4)] [Medline: [27209149](https://pubmed.ncbi.nlm.nih.gov/27209149/)]
8. World Health Organization. 2015. Mental Health Atlas 2014. URL: [https://apps.who.int/iris/bitstream/handle/10665/178879/9789241565011\\_eng.pdf?sequence=1#targetText=PROJECT%20TEAM%20AND%20PARTNERS,published%20in%202005%20and%202011](https://apps.who.int/iris/bitstream/handle/10665/178879/9789241565011_eng.pdf?sequence=1#targetText=PROJECT%20TEAM%20AND%20PARTNERS,published%20in%202005%20and%202011) [accessed 2019-01-01]
9. Liu C, Chen L, Xie B, Yan J, Jin T, Wu Z. Number and characteristics of medical professionals working in Chinese mental health facilities. *Shanghai Arch Psychiatry* 2013 Oct;25(5):277-285 [FREE Full text] [doi: [10.3969/j.issn.1002-0829.2013.05.003](https://doi.org/10.3969/j.issn.1002-0829.2013.05.003)] [Medline: [24991167](https://pubmed.ncbi.nlm.nih.gov/24991167/)]
10. Ran MS, Zhang T, Wong IY, Yang X, Liu C, Liu B, CMHP Study Group. Internalized stigma in people with severe mental illness in rural China. *Int J Soc Psychiatry* 2018 Feb;64(1):9-16. [doi: [10.1177/0020764017743999](https://doi.org/10.1177/0020764017743999)] [Medline: [29183250](https://pubmed.ncbi.nlm.nih.gov/29183250/)]
11. Guan L, Liu J, Wu XM, Chen D, Wang X, Ma N, et al. Unlocking patients with mental disorders who were in restraints at home: a national follow-up study of China's new public mental health initiatives. *PLoS One* 2015;10(4):e0121425 [FREE Full text] [doi: [10.1371/journal.pone.0121425](https://doi.org/10.1371/journal.pone.0121425)] [Medline: [25848800](https://pubmed.ncbi.nlm.nih.gov/25848800/)]
12. Angermeyer MC, Matschinger H. The stigma of mental illness: effects of labelling on public attitudes towards people with mental disorder. *Acta Psychiatr Scand* 2003 Oct;108(4):304-309. [doi: [10.1034/j.1600-0447.2003.00150.x](https://doi.org/10.1034/j.1600-0447.2003.00150.x)] [Medline: [12956832](https://pubmed.ncbi.nlm.nih.gov/12956832/)]
13. Lin TY, Lin MC. Love, denial and rejection: responses of chinese families to mental illness. In: Kleinman A, Lin TY, editors. *Normal and Abnormal Behavior in Chinese Culture*. New York City: Springer; 1981:387-401.
14. Ng CH. The stigma of mental illness in Asian cultures. *Aust N Z J Psychiatry* 1997 Jun;31(3):382-390. [doi: [10.3109/00048679709073848](https://doi.org/10.3109/00048679709073848)] [Medline: [9226084](https://pubmed.ncbi.nlm.nih.gov/9226084/)]
15. Alonso J, Buron A, Bruffaerts R, He Y, Posada-Villa J, Lepine J, World Mental Health Consortium. Association of perceived stigma and mood and anxiety disorders: results from the World Mental Health Surveys. *Acta Psychiatr Scand* 2008 Oct;118(4):305-314 [FREE Full text] [doi: [10.1111/j.1600-0447.2008.01241.x](https://doi.org/10.1111/j.1600-0447.2008.01241.x)] [Medline: [18754833](https://pubmed.ncbi.nlm.nih.gov/18754833/)]
16. Kleinman A, Lin TY, editors. *Normal and Abnormal Behaviour in Chinese Culture*. Dordrecht, Holland: Springer; 2010.
17. Hill C, Martin JL, Thomson S, Scott-Ram N, Penfold H, Creswell C. Navigating the challenges of digital health innovation: considerations and solutions in developing online and smartphone-application-based interventions for mental health disorders. *Br J Psychiatry* 2017 Aug;211(2):65-69. [doi: [10.1192/bjp.bp.115.180372](https://doi.org/10.1192/bjp.bp.115.180372)] [Medline: [28522435](https://pubmed.ncbi.nlm.nih.gov/28522435/)]
18. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS Med* 2013;10(1):e1001362 [FREE Full text] [doi: [10.1371/journal.pmed.1001362](https://doi.org/10.1371/journal.pmed.1001362)] [Medline: [23349621](https://pubmed.ncbi.nlm.nih.gov/23349621/)]
19. Zhang MW, Ho CS, Cheok CC, Ho RC. Smartphone apps in mental healthcare: the state of the art and potential developments. *BJPsych Adv* 2018 Jan 2;21(5):354-358. [doi: [10.1192/apt.bp.114.013789](https://doi.org/10.1192/apt.bp.114.013789)]
20. Ben-Zeev D, Davis KE, Kaiser S, Krzysos I, Drake RE. Mobile technologies among people with serious mental illness: opportunities for future services. *Adm Policy Ment Health* 2013 Jul;40(4):340-343 [FREE Full text] [doi: [10.1007/s10488-012-0424-x](https://doi.org/10.1007/s10488-012-0424-x)] [Medline: [22648635](https://pubmed.ncbi.nlm.nih.gov/22648635/)]
21. Kahn JG, Yang JS, Kahn JS. 'Mobile' health needs and opportunities in developing countries. *Health Aff (Millwood)* 2010 Feb;29(2):252-258. [doi: [10.1377/hlthaff.2009.0965](https://doi.org/10.1377/hlthaff.2009.0965)] [Medline: [20348069](https://pubmed.ncbi.nlm.nih.gov/20348069/)]
22. Poushter J. Pew Research Center. 2016. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. URL: <https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed 2019-05-10]
23. Silva BM, Rodrigues JJ, de la Torre ID, 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](https://doi.org/10.1016/j.jbi.2015.06.003)] [Medline: [26071682](https://pubmed.ncbi.nlm.nih.gov/26071682/)]
24. China Internet Network Information Center. 2018. Statistical Report on Internet Development in China. URL: <https://cnnic.com.cn/IDR/ReportDownloads/201807/P020180711391069195909.pdf> [accessed 2019-01-01]
25. Nicholas J, Larsen M, Christensen H, Proudfoot J. Systematic assessment of mobile apps for bipolar disorder: features and content. *Bipolar Disord* 2015;17:129 [FREE Full text]



26. Heffner JL, Vilardaga R, Mercer LD, Kientz JA, Bricker JB. Feature-level analysis of a novel smartphone application for smoking cessation. *Am J Drug Alcohol Abuse* 2015 Jan;41(1):68-73 [FREE Full text] [doi: [10.3109/00952990.2014.977486](https://doi.org/10.3109/00952990.2014.977486)] [Medline: [25397860](https://pubmed.ncbi.nlm.nih.gov/25397860/)]
27. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Ment Health* 2016 Mar 1;3(1):e7 [FREE Full text] [doi: [10.2196/mental.4984](https://doi.org/10.2196/mental.4984)] [Medline: [26932350](https://pubmed.ncbi.nlm.nih.gov/26932350/)]
28. Larsen ME, Nicholas J, Christensen H. A systematic assessment of smartphone tools for suicide prevention. *PLoS One* 2016;11(4):e0152285 [FREE Full text] [doi: [10.1371/journal.pone.0152285](https://doi.org/10.1371/journal.pone.0152285)] [Medline: [27073900](https://pubmed.ncbi.nlm.nih.gov/27073900/)]
29. Zhang P, Dong L, Chen H, Chai Y, Liu J. The rise and need for mobile apps for maternal and child health care in china: survey based on app markets. *JMIR Mhealth Uhealth* 2018 Jun 8;6(6):e140 [FREE Full text] [doi: [10.2196/mhealth.9302](https://doi.org/10.2196/mhealth.9302)] [Medline: [29884601](https://pubmed.ncbi.nlm.nih.gov/29884601/)]
30. Erdenebileg Z. China Business Review. 2017. At Your Fingertips: Approaching China's Mobile App Store Market. URL: <https://www.chinabusinessreview.com/at-your-fingertips-approaching-chinas-mobile-app-store-market/> [accessed 2019-01-01]
31. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
32. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clin Psychol Rev* 2006 Jan;26(1):17-31. [doi: [10.1016/j.cpr.2005.07.003](https://doi.org/10.1016/j.cpr.2005.07.003)] [Medline: [16199119](https://pubmed.ncbi.nlm.nih.gov/16199119/)]
33. Zung WW. A self-rating depression scale. *Arch Gen Psychiatry* 1965 Jan;12:63-70. [doi: [10.1001/archpsyc.1965.01720310065008](https://doi.org/10.1001/archpsyc.1965.01720310065008)] [Medline: [14221692](https://pubmed.ncbi.nlm.nih.gov/14221692/)]
34. Zung WW. SAS. Self-rating anxiety scale. In: Guy W, editor. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, Maryland: US National Institute of Health; 1976:337-340.
35. Aben I, Verhey F, Lousberg R, Lodder J, Honig A. Validity of the beck depression inventory, hospital anxiety and depression scale, SCL-90, and Hamilton depression rating scale as screening instruments for depression in stroke patients. *Psychosomatics* 2002;43(5):386-393. [doi: [10.1176/appi.psy.43.5.386](https://doi.org/10.1176/appi.psy.43.5.386)] [Medline: [12297607](https://pubmed.ncbi.nlm.nih.gov/12297607/)]
36. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
37. Beck AT. *Great Plains Health*. 1988. Beck Anxiety Inventory (BAI). URL: <https://www.gphealth.org/media/1087/anxiety.pdf> [accessed 2019-09-17]
38. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
39. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. *J Psychosom Res* 2000 Jun;48(6):555-560. [doi: [10.1016/s0022-3999\(00\)00095-7](https://doi.org/10.1016/s0022-3999(00)00095-7)] [Medline: [11033374](https://pubmed.ncbi.nlm.nih.gov/11033374/)]
40. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;64(2):258-266. [doi: [10.1097/00006842-200203000-00008](https://doi.org/10.1097/00006842-200203000-00008)] [Medline: [11914441](https://pubmed.ncbi.nlm.nih.gov/11914441/)]
41. Xu Z, Rüschen N, Huang F, Kösters M. Challenging mental health related stigma in China: systematic review and meta-analysis. I. Interventions among the general public. *Psychiatry Res* 2017 Sep;255:449-456. [doi: [10.1016/j.psychres.2017.01.008](https://doi.org/10.1016/j.psychres.2017.01.008)] [Medline: [28780127](https://pubmed.ncbi.nlm.nih.gov/28780127/)]
42. Jorm AF. Mental health literacy: empowering the community to take action for better mental health. *Am Psychol* 2012 Apr;67(3):231-243. [doi: [10.1037/a0025957](https://doi.org/10.1037/a0025957)] [Medline: [22040221](https://pubmed.ncbi.nlm.nih.gov/22040221/)]
43. Proudfoot J, Parker G, Pavlovic DH, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. *J Med Internet Res* 2010 Dec 19;12(5):e64 [FREE Full text] [doi: [10.2196/jmir.1475](https://doi.org/10.2196/jmir.1475)] [Medline: [21169174](https://pubmed.ncbi.nlm.nih.gov/21169174/)]
44. BinDhim NF, Shaman AM, Trevena L, Basyouni MH, Pont LG, Alhawassi TM. Depression screening via a smartphone app: cross-country user characteristics and feasibility. *J Am Med Inform Assoc* 2015 Jan;22(1):29-34 [FREE Full text] [doi: [10.1136/amiajnl-2014-002840](https://doi.org/10.1136/amiajnl-2014-002840)] [Medline: [25326599](https://pubmed.ncbi.nlm.nih.gov/25326599/)]
45. Charlson FJ, Baxter AJ, Cheng HG, Shidhaye R, Whiteford HA. The burden of mental, neurological, and substance use disorders in China and India: a systematic analysis of community representative epidemiological studies. *Lancet* 2016 Jul;388(10042):376-389. [doi: [10.1016/s0140-6736\(16\)30590-6](https://doi.org/10.1016/s0140-6736(16)30590-6)]
46. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 2003 Feb;60(2):184-189. [doi: [10.1001/archpsyc.60.2.184](https://doi.org/10.1001/archpsyc.60.2.184)] [Medline: [12578436](https://pubmed.ncbi.nlm.nih.gov/12578436/)]
47. Leung SA, Guo L, Lam MP. The development of counseling psychology in higher educational institutions in China. *Couns Psychol* 2016 Jun 30;28(1):81-99. [doi: [10.1177/0011000000281005](https://doi.org/10.1177/0011000000281005)]
48. Hou ZJ, Zhang N. Counseling psychology in China. *Applied Psychol* 2007 Jan;56(1):33-50. [doi: [10.1111/j.1464-0597.2007.00274.x](https://doi.org/10.1111/j.1464-0597.2007.00274.x)]
49. Angermeyer MC, Matschinger H, Riedel-Heller SG. Whom to ask for help in case of a mental disorder? Preferences of the lay public. *Soc Psychiatry Psychiatr Epidemiol* 1999 Apr;34(4):202-210. [doi: [10.1007/s001270050134](https://doi.org/10.1007/s001270050134)] [Medline: [10365626](https://pubmed.ncbi.nlm.nih.gov/10365626/)]

50. Palmier-Claus JE, Ainsworth J, Machin M, Barrowclough C, Dunn G, Barkus E, et al. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BioMed Central Psychiatry* 2012 Oct 17;12(1):172. [doi: [10.1186/1471-244x-12-172](https://doi.org/10.1186/1471-244x-12-172)]
51. Leung KS, Cottler LB. Treatment of pathological gambling. *Curr Opin Psychiatry* 2009 Jan;22(1):69-74. [doi: [10.1097/YCO.0b013e32831575d9](https://doi.org/10.1097/YCO.0b013e32831575d9)] [Medline: [19122538](https://pubmed.ncbi.nlm.nih.gov/19122538/)]
52. Morgenstern J, Morgan TJ, McCrady BS, Keller DS, Carroll KM. Manual-guided cognitive-behavioral therapy training: a promising method for disseminating empirically supported substance abuse treatments to the practice community. *Psychol Addict Behav* 2001 Jun;15(2):83-88. [doi: [10.1037//0893-164x.15.2.83](https://doi.org/10.1037//0893-164x.15.2.83)] [Medline: [11419234](https://pubmed.ncbi.nlm.nih.gov/11419234/)]
53. Okajima I, Komada Y, Inoue Y. A meta - analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia. *Sleep Biol Rhythm* 201 Nov 28;9(1):24-34. [doi: [10.1111/j.1479-8425.2010.00481.x](https://doi.org/10.1111/j.1479-8425.2010.00481.x)]
54. Dagö J, Asplund RP, Bsenko HA, Hjerling S, Holmberg A, Westh S, et al. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. *J Anxiety Disord* 2014 May;28(4):410-417. [doi: [10.1016/j.janxdis.2014.02.003](https://doi.org/10.1016/j.janxdis.2014.02.003)] [Medline: [24731441](https://pubmed.ncbi.nlm.nih.gov/24731441/)]
55. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000 Dec 15;25(24):3186-3191. [doi: [10.1097/00007632-200012150-00014](https://doi.org/10.1097/00007632-200012150-00014)] [Medline: [11124735](https://pubmed.ncbi.nlm.nih.gov/11124735/)]
56. Rosenfield S. Labeling mental illness: the effects of received services and perceived stigma on life satisfaction. *Am Sociol Rev* 1997 Aug;62(4):660-672. [doi: [10.2307/2657432](https://doi.org/10.2307/2657432)]

## Abbreviations

**BPHS:** Basic Public Health Service  
**CBT:** cognitive behavioral therapy  
**CDC:** Centers for Disease Control and Prevention  
**GAD:** Generalized Anxiety Disorder Screener  
**MARS:** Mobile App Rating Scale  
**PHQ:** Patient Health Questionnaire  
**Q&A:** question and answer  
**SAS:** Self-Rating Scale for Anxiety

*Edited by G Eysenbach; submitted 24.12.18; peer-reviewed by J Car, J Han, Y Zhou, L Ranandeh; comments to author 08.04.19; revised version received 18.06.19; accepted 31.08.19; published 07.11.19.*

*Please cite as:*

Shang J, Wei S, Jin J, Zhang P  
*Mental Health Apps in China: Analysis and Quality Assessment*  
*JMIR Mhealth Uhealth* 2019;7(11):e13236  
URL: <https://mhealth.jmir.org/2019/11/e13236>  
doi: [10.2196/13236](https://doi.org/10.2196/13236)  
PMID: [31697245](https://pubmed.ncbi.nlm.nih.gov/31697245/)

©Jie Shang, Shaoming Wei, Jianbo Jin, Puhong Zhang. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 07.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Mediators of Intervention Effects on Depressive Symptoms Among People Living With HIV: Secondary Analysis of a Mobile Health Randomized Controlled Trial Using Latent Growth Curve Modeling

Mengting Zhu<sup>1\*</sup>, MS; Weiping Cai<sup>2\*</sup>, MD; Linghua Li<sup>2\*</sup>, MD; Yan Guo<sup>1,3,4\*</sup>, PhD; Aliza Monroe-Wise<sup>5\*</sup>, MD, MSc; Yiran Li<sup>1\*</sup>, MS; Chengbo Zeng<sup>6,7</sup>, MS; Jiaying Qiao<sup>1</sup>, MS; Zhimeng Xu<sup>1</sup>, MS; Hanxi Zhang<sup>8</sup>, MS; Yu Zeng<sup>1</sup>, MS; Cong Liu<sup>2</sup>, MSN

<sup>1</sup>Department of Medical Statistics and Epidemiology, School of Public Health, Sun Yat-sen University, Guangzhou, China

<sup>2</sup>Department of Infectious Diseases, Guangzhou Eighth People's Hospital, Guangzhou, China

<sup>3</sup>Center for Migrant Health Policy, Sun Yat-sen University, Guangzhou, China

<sup>4</sup>Sun Yat-sen Global Health Institute, Sun Yat-sen University, Guangzhou, China

<sup>5</sup>Department of Global Health, University of Washington, Seattle, WA, United States

<sup>6</sup>South Carolina SmartState Center of Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

<sup>7</sup>Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

<sup>8</sup>National Center of AIDS/STD Control and Prevention, China Center for Disease Control, Beijing, China

\*these authors contributed equally

**Corresponding Author:**

Yan Guo, PhD

Department of Medical Statistics and Epidemiology, School of Public Health

Sun Yat-sen University

74 Zhong Shan 2nd Road

Guangzhou

China

Phone: 86 1 350 150 2582

Email: [guoy8@mail.sysu.edu.cn](mailto:guoy8@mail.sysu.edu.cn)

## Abstract

**Background:** Although several studies have investigated the effects of mobile health (mHealth) interventions on depression among people living with HIV, few studies have explored mediators of mHealth-based interventions to improve mental health in people living with HIV. Identifying influential mediators may enhance and refine effective components of mHealth interventions to improve mental health of people living with HIV.

**Objective:** This study aimed to examine mediating factors of the effects of a mHealth intervention, *Run4Love*, designed to reduce depression among people living with HIV using 4 time-point measurement data.

**Methods:** This study used data from a randomized controlled trial of a mHealth intervention among people living with HIV with elevated depressive symptoms in Guangzhou, China. A total of 300 patients were assigned to receive either the mHealth intervention (n=150) or a waitlist control group (n=150) through computer-generated block randomization. Depressive symptoms, coping, and HIV-related stigma were measured at baseline, 3-, 6-, and 9-month follow-ups. The latent growth curve model was used to examine the effects of the intervention on depressive symptoms via potential mediators. Mediating effects were estimated using bias-corrected 95% bootstrapped CIs (BCIs) with resampling of 5000.

**Results:** Enhanced positive coping and reduced HIV-related stigma served as effective treatment mediators in the mHealth intervention. Specially, there was a significant indirect effect of the mHealth intervention on the slope of depressive symptoms via the slope of positive coping (beta=-2.86; 95% BCI -4.78 to -0.94). The indirect effect of the mHealth intervention on the slope of depressive symptoms via the slope of HIV-related stigma was also statistically significant (beta=-1.71; 95% BCI -3.03 to -0.40). These findings indicated that enhancement of positive coping and reduction of HIV-related stigma were important mediating factors of the mHealth intervention in reducing depression among people living with HIV.

**Conclusions:** This study revealed the underlying mediators of a mHealth intervention to reduce depression among people living with HIV using latent growth curve model and 4 time-point longitudinal measurement data. The study results underscored the importance of improving positive coping skills and mitigating HIV-related stigma in mHealth interventions to reduce depression among people living with HIV.

(*JMIR Mhealth Uhealth* 2019;7(11):e15489) doi:[10.2196/15489](https://doi.org/10.2196/15489)

## KEYWORDS

mobile health; depression; HIV; randomized controlled trial; longitudinal studies

## Introduction

### Background

Depression is highly prevalent among people living with HIV (PLWH) [1]. Data from a systematic review in China indicate that the pooled prevalence of depressive symptoms is 50.8% in PLWH [2]. In contrast, the prevalence of depressive symptoms in the general population is 17.1% [3]. Depression is consistently associated with impaired role functioning, worsened antiretroviral therapy adherence, elevated risks of HIV-related morbidity and mortality, and increased health care costs [4-6]. Although depression is highly prevalent and disabling, it remains greatly undertreated worldwide with more than 90% of those diagnosed with depression in China and India not on treatment and more than 50% untreated even in many high-resource settings [7,8].

Literature has shown that interventions such as cognitive behavioral stress management (CBSM) are effective at reducing depression in various populations including PLWH [9-11]. However, few studies have explored mediators of how such interventions reduce depression. Mediators are statistical constructs that constitute 1 type of mechanism for how an intervention affects outcomes [12,13]. Understanding influential mediators is critical for understanding how interventions may achieve effective outcomes and help enhance and refine intervention components for future implementation or scaling up to improve mental health of PLWH [14].

Previous studies have indicated that factors such as coping and stigma may be strong predictors of depression. Coping strategies have been defined as individuals' emotional, cognitive, and behavioral attempts in response to stressful events [15,16]. HIV-related stigma is defined as processes that devalue, label, negatively stereotype, or treat unfairly objects or people associated with HIV [17,18]. Higher positive coping was associated with lower depressive symptoms [19]. Higher levels of HIV-related stigma and negative coping were associated with higher depressive symptoms [19,20]. CBSM can instruct patients to modify maladaptive cognitions such as HIV-related stigma, acquire social support, and engage in more adaptive behaviors such as positive coping to reduce depression and facilitate adjustment [21,22]. In the Stigma and HIV Disparities Model proposed by Earnshaw, enhancing positive coping could improve resilience to stigma and ultimately improve health outcomes among PLWH [23]. Existing studies have suggested that coping and stigma might be important factors in influencing the effects of psychological treatments on depression among PLWH [19,24,25]. For example, Tshabalala and colleagues observed that HIV-positive women randomized to cognitive

behavioral intervention group reported greater reduction in depression than the control group. In addition, a significant reduction in HIV-related stigma and negative coping and improvement in positive coping were observed in comparison with the control group [25]. Utilizing qualitative analysis, the same study found that reasons for the effectiveness of the cognitive behavioral intervention to reduce depression might be the enhancement of participants' coping and assertiveness skills and reduction in HIV-related stigma during the intervention. Similarly, another trial among HIV-seropositive men who have sex with men in the United States also reported that CBSM intervention was efficacious in improving cognitive coping strategies and improved coping might be an important determinant of both depression and anxiety reduction [21].

However, the few studies that explored intervention mediators of depression reduction were mostly conducted within interventions that occurred in clinic settings and were delivered face-to-face. To the best of the knowledge of the authors, no study has been done based on mobile health (mHealth) interventions [21,25]. It is not clear that factors effective in mediating intervention effects on depression reduction in traditional delivery settings remain effective in mHealth interventions. With the wide coverage of smart mobile phones and emerging literature on the initial effectiveness of mHealth interventions, it is important to understand the underlying mechanisms and mediating factors for mHealth interventions, especially in comparison with traditional face-to-face interventions [26-29]. As mHealth interventions have the potential to reach a large population with less stigma, lower cost, and increased convenience, more research is needed for greater understanding of mediators of effective mHealth interventions [30-33].

In addition, existing studies mostly used qualitative analysis or pre-post measurements to examine mediators of interventions for depression reduction; longitudinal studies with multiple waves are lacking [21,25]. As pre and post measurements contain minimal information on individual changes, longitudinal data with multiple ( $\geq 3$ ) waves allow investigation on how factors change and how these changes are related to health outcomes over time [12,34]. Mediators should temporally precede the outcomes to demonstrate causal temporal relationships [12,13].

### Objectives

To bridge the gap in the existing literature, this study aimed to examine mediating factors of the effects of an mHealth intervention, Run4Love, designed for depression reduction among PLWH based on CBSM with 4 time-point measurement data. We hypothesized that both coping and HIV-related stigma

would play important roles in mediating the effects of the Run4Love mHealth intervention. Specifically, we hypothesized that the intervention would improve positive coping and decrease HIV-related stigma, which in turn would lead to reduced depression among PLWH.

## Methods

### Design and Procedure

This study used data from a randomized controlled trial (RCT) of an mHealth intervention, Run4Love, for depression reduction among PLWH in China. This study is a secondary analysis of the Run4Love RCT (ChiCTR-IPR-17012606) [35]. Participants were recruited by trained research staff at the outpatient clinic of a large hospital designated for HIV treatment. The study was conducted from September 9, 2017, to October 1, 2018, at Guangzhou Eighth People's Hospital in Guangzhou, China. Details of study design and procedures can be found in the study protocol [35]. Briefly, patients who met the eligibility criteria were provided with a study pamphlet that described research procedures and were invited to join the study. A total of 300 patients were enrolled in the study. Following enrollment, the patients were assigned to either the Run4Love mHealth intervention (n=150) or the waitlist control group (n=150) through computer-generated block randomization with a block size of 4 and an allocation ratio of 1:1. By the nature of the trial design, neither research staff nor participants were blinded to the Run4Love mHealth intervention. The allocation sequence was not concealed from the research staff. The duration of the intervention was 3 months, and participants were followed up to 6 months after the intervention. Participants were assessed at baseline, 3-month follow-up, 6-month follow-up, and 9-month follow-up using electronic questionnaires on tablets. These assessments were conducted during face-to-face sessions with the (nonblinded) research staff at the outpatient clinic. The primary outcome was depressive symptoms assessed by the Center for Epidemiological Studies Depression Scale (CES-D) [36]. Written informed consent was obtained from each participant before data collection. Participants who completed each of the study assessments received 50 RMB (ie, about US \$8) or gifts of equivalent value for completion of each survey. The Consolidated Standards of Reporting Trials EHEALTH checklist is shown in [Multimedia Appendix 1](#).

### Participants

Patients were eligible to participate if they were aged 18 years or older, HIV seropositive, having elevated depressive symptoms (CES-D $\geq$ 16), willing to provide hair samples, and using WeChat, the most popular application for instant communication in China [37]. Hair samples were collected to test the cortisol content as a biomarker of chronic stress. Patients were excluded if they were taking psychiatric drugs, unable to finish questionnaires because of mental or other illnesses or other reasons, unable to read or listen to the materials sent via WeChat (ie, short articles, audios, and posters), unable to engage in physical activities because of medical reasons, or had hair permed or dyed in the past 3 months. Those who refused to participate provided information on specific reasons for refusal.

This study was approved by the institutional review board of Sun Yat-sen University in Guangzhou, China.

### Run4Love Mobile Health Intervention Program

Participants in the intervention group participated in a 3-month mHealth intervention delivered by the enhanced WeChat platform, consisting of the adapted CBSM course and regular physical activity promotion [37]. The adapted CBSM course consisted of 9 sessions and 3 review sessions on coping skills and stress reduction management such as practices of effective coping skills, cognitive distortions, meditation, and breathing. Sessions were in multiple formats, including audio clips, articles, and posters. On average, the articles were 1300 words and took about 5 min to read; the audios were 5 to 10 min. The physical activity promotion program included information about benefits of and guidance on regular exercise and a healthy diet. Participants in the intervention group received CBSM and physical activity promotion information on their WeChat account 3 to 5 times a week. Participants could review the materials they had received at any time.

The Run4Love intervention also included 5 phone calls from research staff at the first week and 1, 2, 5, and 8 months after enrollment. The purpose of the phone calls was to confirm participation and proper use of the platform at the first week and to offer social support, facilitate intervention implementation, identify barriers to adherence, and provide feedback on mental health in subsequent months. All calls had a script for reference. The calls lasted for an average of 10 min during the first week and 15 min during subsequent weeks. The intervention is described in detail elsewhere [35].

### Control Program

Participants in the control group received a brochure on nutrition and healthy living in addition to usual care for HIV treatment. Moreover, they were offered to receive the Run4Love intervention as soon as the study ended (ie, 9 months after enrollment).

### Measures

#### Depressive Symptoms

Depressive symptoms were assessed by the CES-D with good reliability and validity [36]. The 20-item measure included 4 subscales: depressed affect, positive affect, interpersonal relationship, and somatic and retarded activity. Higher scores indicated higher levels of depressive symptoms. All items used a 4-point Likert scale from 0 ("Rarely or none of the time") to 3 ("Most or all of the time"). The total scores ranged from 0 to 60, with scores of 16 or above being considered as elevated depressive symptoms. The CES-D scale demonstrated good internal consistency, with Cronbach alphas ranging from .88 to .93 across the 4 waves of assessment.

#### Coping

The 20-item Simplified Ways of Coping Questionnaire (SWCQ) was used to assess coping with proven reliability and validity [38]. The SWCQ assessed different attitudes and measures of coping that people adopted in their daily lives using a 4-point Likert scale from 0 ("Not used at all") to 3 ("Used frequently"). The SWCQ measurement consisted of 2 subscales: positive

coping (12 items) and negative coping (8 items). The positive coping scores ranged from 0 to 36, with higher scores indicating better positive coping. Cronbach alphas for positive coping ranged from .85 to .93 across 4 time points. The negative coping scores ranged from 0 to 24. Higher scores indicated higher levels of negative coping. Cronbach alphas for negative coping ranged from .65 to .74 across 4 time points.

### **HIV-Related Stigma**

HIV-related stigma was measured by 14 items derived from the HIV Stigma Scale with good reliability and validity among Chinese PLWH [39,40]. The measurement included 2 subscales: perceived stigma (6 items) and internalized stigma (8 items). Each item used a 4-point Likert scale from 1 (“strongly disagree”) to 4 (“strongly agree”). The total scores ranged from 14 to 56, with higher scores indicating higher levels of HIV-related stigma. An example of the items assessing perceived stigma was “People with HIV lose their jobs when their employers learn.” A sample item of internalized stigma was “I feel guilty because I have HIV.” The HIV-related stigma scale demonstrated good internal consistency, with Cronbach alphas ranging from .92 to .96 across the 4 waves of assessment.

### **Demographic Variables**

Demographic variables included age, gender, educational level, sexual orientation, marital status, employment status, family monthly income, and duration since HIV diagnosis.

### **Statistical Analysis**

All analyses were conducted based on intention-to-treat principle. Baseline characteristics were compared between the intervention and control groups using *t* tests or Wilcoxon rank-sum tests for numeric outcomes and chi-square tests for categorical variables. All statistical tests were 2-sided, and *P* value <.05 was considered statistically significant.

The latent growth curve model (LGCM) was used to examine mediating factors of the effects of the mHealth intervention on depressive symptoms among PLWH using 4 time-point measurement data. As an extension of structural equation modeling (SEM), LGCM allowed simultaneous analysis of multiple time points, thus potentially providing more accurate estimation of changes over time [41]. LGCM was suited in this study because it accommodated longitudinal data with multiple waves where both the mediator and the outcome changed simultaneously over time, therefore allowing researchers to address longitudinal mediation [42,43]. Though the framework of LGCM was similar to that of SEM, LGCM allowed for estimation of inter- and intraindividual variation over time and exploration of predictors of these individual differences [34,43]. Repeated measurements collected at 4 time points (ie, at baseline and 3-, 6-, and 9-month follow-ups) were included as observed indicators, with the latent intercept (ie, initial status) and slope (ie, rate of change) factors being estimated. With the longitudinal data, the parallel-process LGCM allowed exploration of the effects of the intervention on changes of variables of interest, or the growth rates (ie, slopes) of the variables of interest.

A stepwise approach was used to examine the mediating effects, which were widely used with longitudinal data [43]. First,

unconditional parallel-process LGCM of HIV-related stigma, coping, and depressive symptoms were specified to estimate growth trajectories and each construct’s temporal stability for both groups. Factor loadings for the intercept at each time point were set to 1. Factor loadings of the latent slope were set to a model with an unspecified shape (ie, 0, 1, \* and \*) such that the third and fourth factor loadings of the slope could be freely estimated. Instead of assuming linear growth of each factor from the baseline to 9-month follow-up, it was more reasonable to have a model with an unspecified shape without such a strong assumption [44].

Second, conditional LGCM was conducted to examine the effects of the intervention on the outcome and potential mediators. Intervention condition (ie, mHealth intervention vs control) was explored as a predictor of changes in depressive symptoms, coping, and HIV-related stigma across time (ie, predicting the slope factor). A dummy variable was created to represent group assignment. The mHealth intervention group was coded as 1 and the control group as 0. A significant path from the intervention condition to the slope of the variable of interest would indicate a significantly larger change in that variable over time in intervention group than in control group.

Third, a longitudinal mediation model was used to investigate whether the mHealth intervention was effective in reducing depressive symptoms via the potential mediators by examining changes in the slopes of the outcome (ie, depressive symptoms) and mediators (ie, coping and HIV-related stigma). The mediating effects were estimated using bias-corrected 95% bootstrapped CIs (BCIs) with resampling of 5000 [45-47].

All LGCMs were conducted using maximum likelihood estimation. Model fit was assessed using the comparative fit index (CFI) and the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the relative chi-square ratio (chi-square/df). A LGCM model with a good model fit met the following criteria: CFI>0.90, RMSEA<0.08, SRMR<0.08, and relative chi-square ratio<3.0 [48,49]. Preliminary statistical analyses were performed using R software, version 3.5 (R foundation for Statistical Computing). LGCM analyses were performed using Mplus software, version 7 (Muthén & Muthén).

## **Results**

### **Participant Enrollment**

A total of 1555 PLWH were assessed for eligibility, among whom 1255 were excluded or withdrawn before enrollment. A total of 1017 were excluded because of lower CES-D scores (ie, CES-D<16), and 538 were further screened. Among the 538 participants, 164 declined to participate; 24 refused eligibility interview, and 50 were excluded because of other reasons such as currently taking psychotropic medication, participating in other studies, or unable to read because of eye problems. The RCT included 300 participants, with 150 in the intervention group and 150 in the control group. Details of the recruitment process are described in the study protocol [35]. Dropout rates for 3-, 6-, and 9-month follow-ups were 8.7% (26/300), 11.7% (35/300), and 13.3% (40/300), respectively.

All demographic characteristics displayed in Table 1 were examined, and characteristics of dropouts were not statistically different from those who completed the study, aside from being slightly older. Reasons for dropping out included nonresponse (n=29), refusing to continue (n=9), transferring to another hospital (n=1), and imprisonment (n=1).

### Descriptive Analyses

Descriptive statistics for baseline data are presented in Table 1 by group. There was no significant difference between the intervention and control groups except for sexual orientation, where more heterosexual participants were allocated in the control group. The participants were primarily male (277/300, 92.3%), employed (251/300, 83.7%), and with a median age (interquartile range) of 27.5 years (24.5-31.3). The majority

(245/300, 81.7%) were homosexual or bisexual or uncertain of their sexual orientation. On average, participants in the intervention group completed 55% (33/60) of the 12 total Run4Love sessions at 3 months. Repeated measurements of the outcome variable (ie, depressive symptoms) and potential mediators (ie, HIV-related stigma and coping) at 4 time points for the intervention and control groups are shown in Table 2. The observed growth trajectories of mean depressive symptoms, coping, and HIV-related stigma scores for the intervention group and control group over 4 time points are presented in Figure 1. Our RCT indicated that the Run4Love mHealth intervention significantly reduced depressive symptoms and HIV-related stigma and improved positive coping compared with the control group at 3-, 6-, and 9-month follow-ups.

**Table 1.** Participants' characteristics for the intervention and control groups at baseline.

Characteristics	Total (n=300)	Intervention (n=150)	Control (n=150)	P value
Age (years), median (interquartile range, IQR)	27.5 (24.5-31.3)	27.4 (24.3-31.1)	27.8 (24.6-32.2)	.40 <sup>a</sup>
Male, n (%)	277 (92.3)	142 (94.7)	135 (90.0)	.19 <sup>b</sup>
Educational level > high school, n (%)	182 (60.7)	98 (65.3)	84 (56.0)	.12 <sup>b</sup>
Homosexual/bisexual/uncertain, n (%)	245 (81.7)	130 (86.7)	115 (76.7)	.04 <sup>b</sup>
Married, n (%)	38 (12.7)	18(12.0)	20 (13.3)	.73 <sup>b</sup>
Employed, n (%)	251 (83.7)	123 (82.0)	128 (85.3)	.53 <sup>b</sup>
Family monthly income ≥7000 (yuan), n (%)	124 (41.3)	68 (45.3)	56 (37.3)	.20 <sup>b</sup>
Duration since HIV diagnosis (years), median (IQR)	1.7 (0.6-3.7)	1.7 (0.6-4.0)	1.8 (0.6-3.9)	.62 <sup>c</sup>
Center for Epidemiological Studies Depression Scale, mean (SD)	24.1 (6.6)	23.9 (6.4)	24.3 (6.9)	.68 <sup>a</sup>
SWCQ <sup>d</sup> , positive coping, mean (SD)	18.4 (5.8)	18.4 (5.5)	18.3 (6.2)	.92 <sup>a</sup>
SWCQ, negative coping, mean (SD)	11.8 (3.9)	11.8 (3.9)	11.8 (3.9)	.94 <sup>a</sup>
HIV Stigma Scale, mean (SD)	37.5 (7.6)	37.1 (7.7)	38.0 (7.5)	.31 <sup>a</sup>

<sup>a</sup>Based on *t* test.

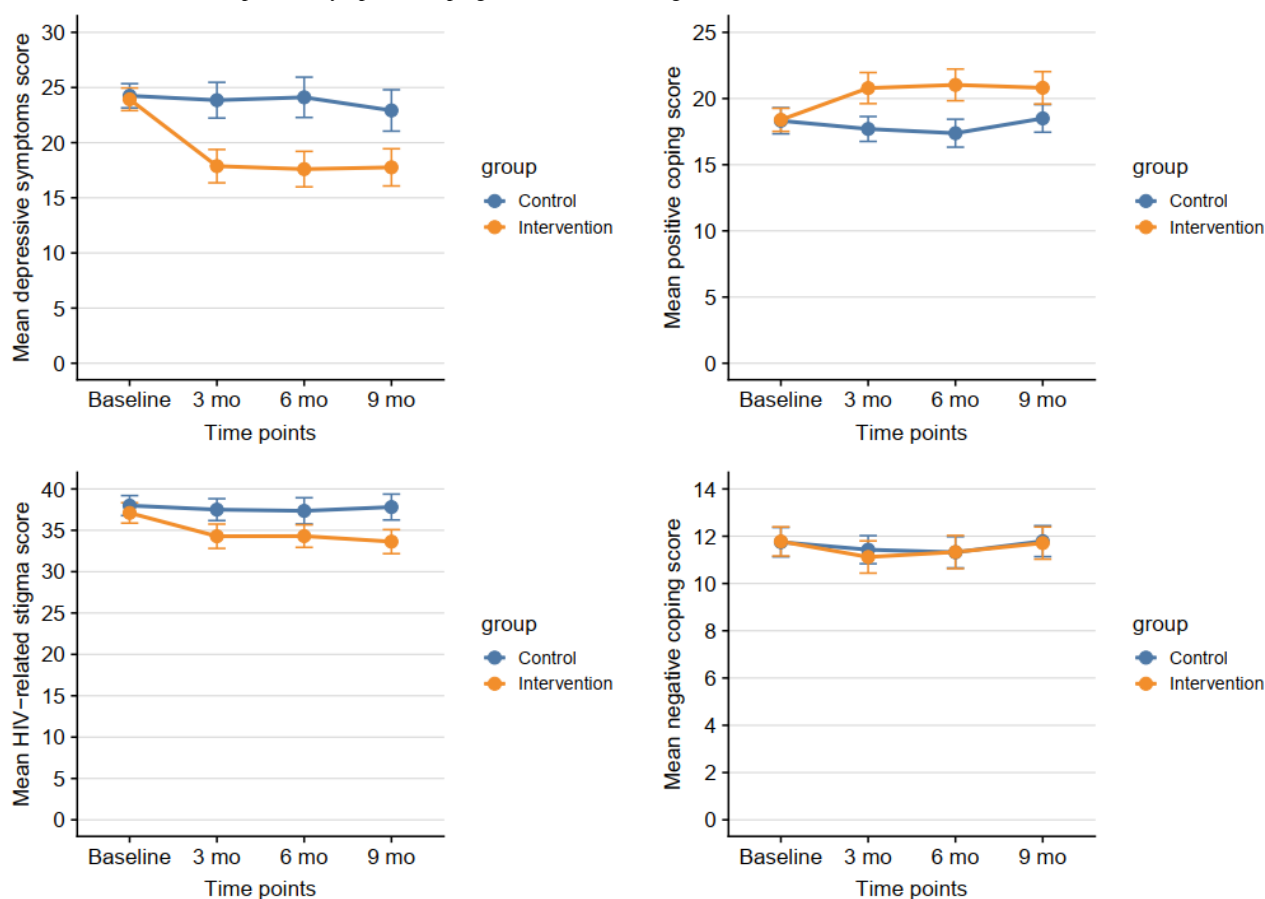
<sup>b</sup>Based on chi-square test, the Fisher exact *P* values were used.

<sup>c</sup>Based on Wilcoxon rank-sum test.

<sup>d</sup>SWCQ: Simplified Ways of Coping Questionnaire.

**Table 2.** Repeated measurements of depressive symptoms and potential mediators in the Run4Love randomized controlled trial.

Variables, group	Baseline, mean (SD)	3-month follow-up, mean (SD)	6-month follow-up, mean (SD)	9-month follow-up, mean (SD)
<b>Depressive symptoms</b>				
Intervention	23.93 (6.39)	17.87 (9.44)	17.60 (10.06)	17.86 (10.72)
Control	24.25 (6.86)	23.85 (10.11)	24.11 (11.42)	23.43 (11.45)
<b>Positive coping</b>				
Intervention	18.39 (5.45)	20.79 (7.33)	21.03 (7.48)	20.95 (7.75)
Control	18.32 (6.15)	17.70 (5.88)	17.38 (6.59)	18.31 (6.41)
<b>Negative coping</b>				
Intervention	11.78 (3.85)	11.12 (4.26)	11.33 (4.38)	11.71 (4.09)
Control	11.75 (3.88)	11.43 (3.71)	11.32 (4.14)	11.87 (4.09)
<b>HIV-related stigma</b>				
Intervention	37.10 (7.67)	34.28 (9.19)	34.30 (8.52)	33.98 (9.01)
Control	37.99 (7.54)	37.50 (8.27)	37.35 (9.92)	37.79 (9.99)

**Figure 1.** Measurements of depressive symptoms, coping, and HIV-related stigma over time. Error bars indicate 95% confidence intervals.

## Latent Growth Curve Modeling

### Intervention Effects

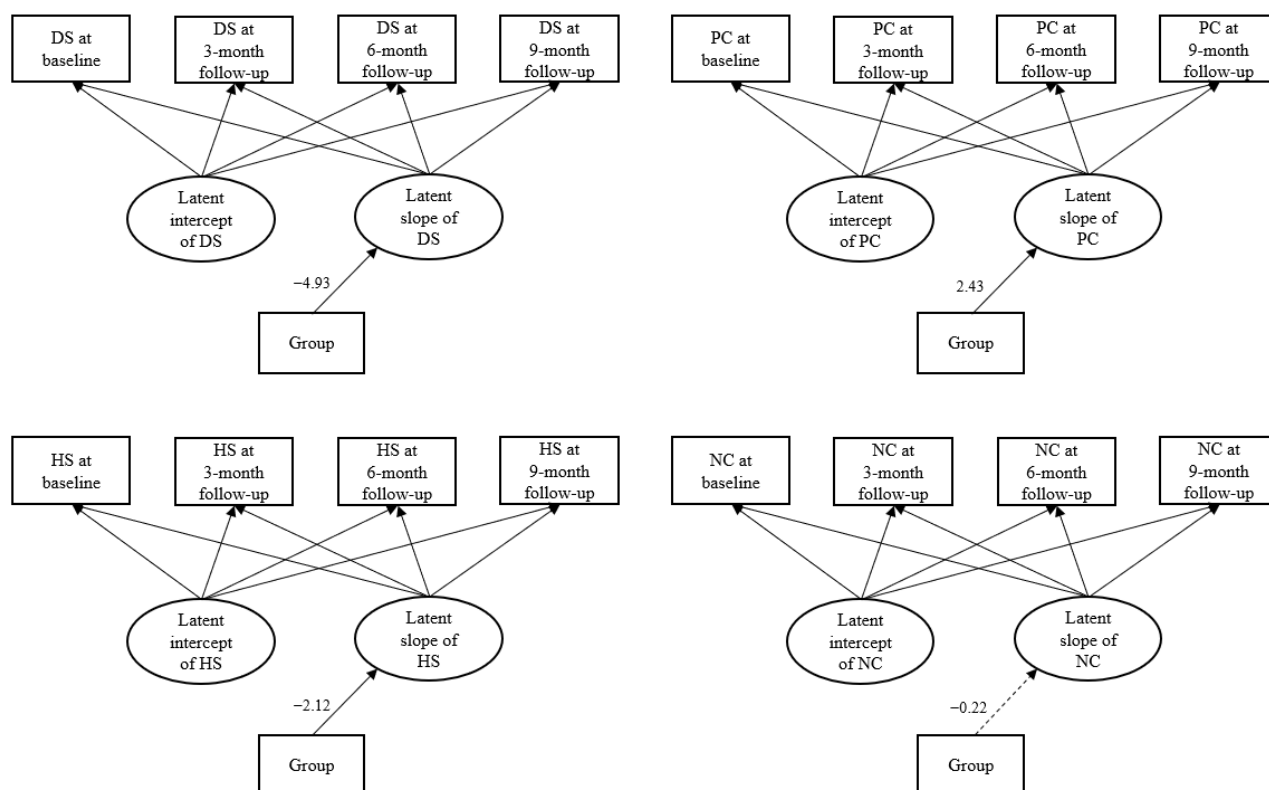
Conditional LGCM supported the beneficial effects of the intervention on the outcome and potential mediators. Figure 2 presents the path diagram of each conditional LGCM. There were significant effects of the intervention on the slopes of depressive symptoms ( $\beta = -4.93$ ;  $P < .001$ ), positive coping ( $\beta = 2.43$ ;  $P < .001$ ), and HIV-related stigma ( $\beta = -2.12$ ;

$P < .001$ ). The results indicated that there was significantly more reduction in depressive symptoms and HIV-related stigma and significantly more improvement in positive coping over time in the intervention group than in the control group. The LGCMs for the abovementioned 3 variables showed good model fit. Table 3 presents model fit indices for all LGCMs. In addition, the intervention did not have a significant intervention effect on the slope of negative coping ( $\beta = -0.22$ ;  $P = .49$ ), indicating that there was no significant reduction in negative coping over



time in the intervention group compared with the control group. Therefore, negative coping was not included in the final analyses of mediating effects.

**Figure 2.** Conditional latent growth curve modeling examining the effects of the mobile health intervention on the outcome and potential mediators. Continuous lines with arrows indicate statistically significant paths. Dotted lines indicate nonsignificant paths. The first and second factor loadings of the latent slope of all models were set to 1, the third and fourth factor loadings of the latent slope of all models were freely estimated. Group: intervention or control group; DS: depressive symptoms; HS: HIV-related stigma; PC: positive coping; NC: negative coping.



**Table 3.** Model fit indices of all latent growth curve models.

Model	CFI <sup>a</sup>	RMSEA <sup>b</sup>	SRMR <sup>c</sup>	Relative chi-square ratio (df)
Reference	>0.90	<0.08	<0.08	<3.0
LGCM <sup>d</sup> for depressive symptoms	1.00	0.04	0.02	1.4 (7)
LGCM for positive coping	1.00	0.00	0.03	0.8 (7)
LGCM for negative coping	1.00	0.02	0.04	1.1 (7)
LGCM for HIV-related stigma	1.00	0.00	0.02	0.8 (7)
Final LGCM	0.98	0.05	0.04	1.7 (56)

<sup>a</sup>CFI: comparative fit index.

<sup>b</sup>RMSEA: root mean square error of approximation.

<sup>c</sup>SRMR: standardized root mean square residual.

<sup>d</sup>LGCM: latent growth curve model.

### Mediating Effects of Positive Coping and HIV-Related Stigma

The path diagram of LGCM in Figure 3 shows how the mHealth intervention reduced depressive symptoms via mediators of positive coping and HIV-related stigma. Pathways and BCIs of the final parallel-process LGCM are presented in Table 4. The final model indicated good model fit (CFI=0.98, RMSEA=0.05, SRMR=0.04, relative chi-square ratio (df)=1.7 (56)). The results of the parallel-process LGCM indicated significant mediating

effects of positive coping and HIV-related stigma on depression reduction in the mHealth intervention.

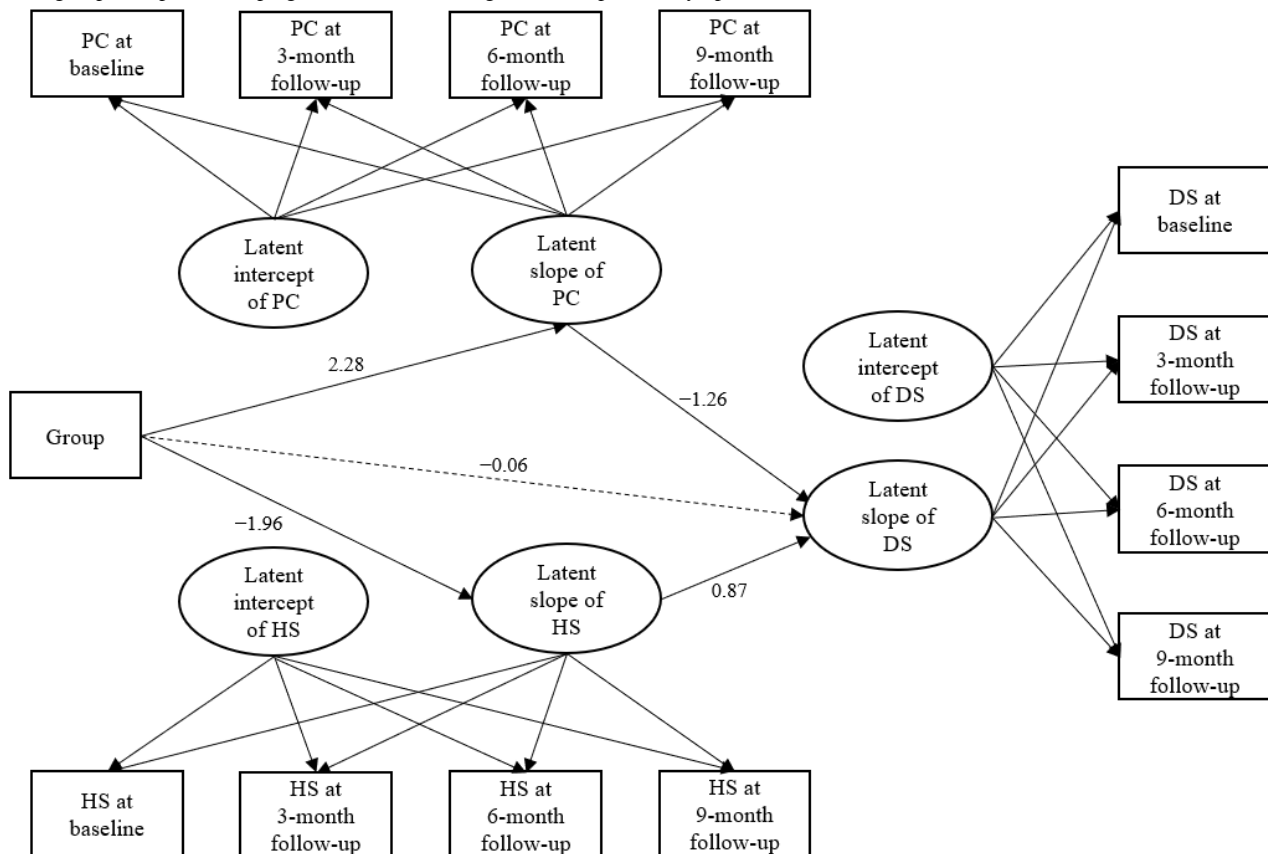
The results indicated a significantly indirect effect of the intervention on the slope of depressive symptoms via the slope of positive coping (beta=2.275×(-1.257)=-2.86; 95% BCI -4.78 to -0.94). This means that the mHealth intervention significantly improved participants' positive coping over time, which in turn significantly reduced depressive symptoms of the participants over time. Similarly, there was also a significantly indirect effect of the intervention on the slope of depressive symptoms via the

slope of HIV-related stigma ( $\beta = -1.962 \times 0.873 = -1.71$ ; 95% BCI  $-3.03$  to  $-0.40$ ), indicating significant intervention effects in reducing HIV-related stigma, which in turn significantly reduced depressive symptoms of the participants over time.

The direct effect of the intervention on the slope of depressive symptoms was not statistically significant ( $\beta = -0.06$ ; 95%

BCI  $-2.15$  to  $2.03$ ) when the mediators were added, indicating no direct effect of the intervention on depressive symptoms of the participants. Therefore, the effects of the mHealth intervention Run4Love in reducing depressive symptoms of the participants might be largely explained by the indirect effects of the intervention on enhancing positive coping and reducing HIV-related stigma among PLWH.

**Figure 3.** Latent growth curve modeling examining mediating effects between the mobile health intervention and changes in depressive symptoms. Continuous lines with arrows indicate statistically significant paths. Dotted lines indicate nonsignificant paths. The first and second factor loadings of the latent slope of the model were set to 1, the third and fourth factor loadings of the latent slope of the model were freely estimated. Group: intervention or control group; PC: positive coping; HS: HIV-related stigma; DS: depressive symptoms.



**Table 4.** Coefficients and bootstrapping CIs of the final parallel-process latent growth curve modeling.

Effect	Estimate	95% BCI <sup>a</sup>	Standardized estimate
<b>Direct effect on the slope of depressive symptoms</b>			
Slope of positive coping	-1.26 <sup>b</sup>	-1.89 to -0.62	-0.59
Slope of HIV-related stigma	0.87 <sup>b</sup>	0.48 to 1.27	0.52
Group	-0.06	-2.15 to 2.03	-0.00
Intercept of positive coping	-0.22	-0.47 to 0.04	-0.15
Intercept of HIV-related stigma	0.08	-0.06 to 0.22	0.08
<b>Direct effect on the slope of positive coping</b>			
Group	2.28 <sup>b</sup>	1.06 to 3.50	0.37
Intercept of depressive symptoms	0.20 <sup>b</sup>	0.04 to 0.36	0.37
Intercept of HIV-related stigma	-0.10	-0.22 to 0.02	-0.21
<b>Direct effect on the slope of HIV-related stigma</b>			
Group	-1.96 <sup>b</sup>	-3.22 to -0.70	-0.25
Intercept of positive coping	-0.35 <sup>b</sup>	-0.57 to -0.13	-0.40
Intercept of depressive symptoms	-0.14 <sup>b</sup>	-0.28 to -0.01	-0.21
<b>Indirect effect of group on the slope of depressive symptoms</b>			
Via the slope of positive coping	-2.86 <sup>b</sup>	-4.78 to -0.94	-0.22
Via the slope of HIV-related stigma	-1.71 <sup>b</sup>	-3.03 to -0.40	-0.13
Total indirect effect	-4.57 <sup>b</sup>	-7.01 to -2.14	-0.35

<sup>a</sup>BCI: bootstrapped CIs.

<sup>b</sup>CI does not contain zero.

## Discussion

### Principal Findings

Our study is among the first to examine mediating factors of the effects of an mHealth intervention designed for depression reduction among PLWH. In addition, this study is the first study that has utilized latent growth curve modeling for examining mediators in an mHealth intervention study using 4 time-point measurement data among PLWH. We found that enhancement of positive coping and reduction of HIV-related stigma were important mediating factors of the mHealth intervention in reducing depression among PLWH.

Previous studies using qualitative analysis have demonstrated a mediating effect of positive coping on achieving depression reduction in face-to-face interventions [25,50]. Results of the LGCM in our quantitative analysis affirmed this finding in the Run4Love mHealth intervention [25,50]. This finding underscored the critical role of positive coping in depression reduction and suggested that enhancing participants' positive coping may be an effective treatment strategy to reduce depression among PLWH in mHealth-based interventions.

One reason for the significant enhancement of positive coping among PLWH might be that the Run4Love mHealth intervention was adapted from evidence-based CBSM program, with an important component of training in coping skills [21]. In the

intervention, participants were instructed to practice positive coping skills by listening to audio recordings and reading short essays delivered via our enhanced mHealth platform [35]. The short essays and audio recordings on effective coping skills included problem- and emotion-focused coping and relaxation exercises. As reported in the previous studies of face-to-face CBSM interventions, problem- and emotion-focused coping and relaxation exercises were effective ways of improving positive coping among PLWH [51,52].

One advantage of the Run4Love mHealth intervention over traditional face-to-face interventions may be that the materials (eg, audio clips and essays) can be repeated, read, or heard at any time or location of participants' choice. The tracking and monitoring functions of the Run4Love intervention can provide timely feedback to both researchers and participants, such as whether and for how long each participant read or listened to the materials sent via the mHealth platform. Instead of recalling what is learned in face-to-face sessions or seeking clinicians' suggestions in traditional interventions, participants in mHealth interventions are able to read, listen to, and review materials related to positive coping skills whenever they encounter challenges in their daily lives [53,54]. Therefore, as opposed to traditional interventions, mHealth interventions such as Run4Love have the advantage of delivering psychological materials and support with increased accessibility, lower cost, and increased privacy for participants. These characteristics, in

turn, increase the potential to scale-up the intervention for widespread implementation, especially in resource-limited settings [28].

The results of this study also revealed that depression reduction in the Run4Love mHealth intervention was significantly mediated by reduction of HIV-related stigma. Previous studies found that face-to-face cognitive behavioral interventions were effective in reducing HIV-related stigma and depression in PLWH [25,55,56]. For example, 1 study showed that a face-to-face cognitive behavioral intervention was effective in reducing stigma, which in turn led to alleviated emotional symptoms in outpatients with anxiety and depressive symptoms in a pre- and postintervention assessment design [57]. Our study contributed to existing literature by demonstrating that an mHealth intervention can significantly reduce HIV-related stigma and also that reduced HIV-related stigma served as an important mediator in the overall effect of the mHealth intervention on depression reduction.

Although the specific features of the Run4Love intervention that might explain the significant reduction in HIV-related stigma are not clear, 2 possible elements might be (1) the important component of HIV stigma reduction messages in the CBSM and (2) the additional 5 phone calls from research staff that occurred throughout the study. The phone calls may have served as additional social support for the participants as research staff helped to facilitate participation, improve intervention adherence, and provide feedback and guidance on participants' mental health status. Previous research indicates that social support is an important factor in reducing HIV-related stigma [58]. Our study findings provided empirical evidence that the Run4Love mHealth intervention decreased the experience of HIV-related stigma, which in turn effectively reduced depressive symptoms among PLWH.

### Limitations

There were several limitations in this study. First, although our study focused on positive coping and HIV-related stigma, other factors such as stress, self-efficacy, physical activity, patient

satisfaction, guidance (working alliance), emotion regulation skills, and expectations might also serve as potential mediating factors for depression reduction among PLWH in the Run4Love mHealth intervention. Social support that the research staff may have provided with the phone calls might also serve as potential mediators, but we did not measure social support in this study. Future studies should further assess and explore the potential mediating effects of these factors in mHealth interventions. Second, data in this study were self-reported, which might introduce recall and social desirability biases. More objective measures such as biomarkers could be incorporated in the future studies. Third, as the data in this study were collected in an urban setting, caution should be exercised when generalizing the results to other places such as rural areas. Fourth, as the temporal sequence could not be identified between the mediators and outcome in this study, a causal relationship between the mediators and outcome cannot be confirmed. Fifth, selection bias in enrollment may limit the generalizability of the findings. Sixth, mediators only statistically narrow the mechanisms of change and might not necessarily be congruent [12,13]. Despite these limitations, this study provided additional empirical and quantitative evidence for a better understanding of the effects of an mHealth intervention on depression reduction among PLWH through mediating factors.

### Conclusions

In conclusion, this study revealed several mediating effects of an mHealth intervention using latent growth curve modeling and 4 time-point longitudinal measurement data. Positive coping and HIV-related stigma were important mediating factors of the Run4Love mHealth intervention in reducing depression among PLWH. The study's findings provided empirical evidence for future research to enhance positive coping and reduce HIV-related stigma in mHealth interventions to reduce depression among PLWH. In addition, future interventions and policies aimed at reducing depression among PLWH should be designed with specific features that address positive coping and stigma to maximize intervention efficacy.

---

### Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant no 71573290) and China Medical Board open competition funding (grant no 17-271). The funders provided grants to implement this program but had no role in the design of the study, data collection, analysis, interpretation of the data, and preparation of the manuscript.

---

### Authors' Contributions

YG designed the study. WC, LL, and CL were important collaborators. MZ, YL, ZX, and JQ analyzed the data. MZ wrote the first draft of the manuscript. YG and AMW made significant revisions of the manuscript. MZ, YL, CZ, ZX, JQ, and YZ reviewed data analysis and revised the manuscript. All authors reviewed and approved the final manuscript for publication.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 2178 KB - mhealth\\_v7i11e15489\\_app1.pdf](#)]

## References

1. Nanni MG, Caruso R, Mitchell AJ, Meggiolaro E, Grassi L. Depression in HIV infected patients: a review. *Curr Psychiatry Rep* 2015 Jan;17(1):530. [doi: [10.1007/s11920-014-0530-4](https://doi.org/10.1007/s11920-014-0530-4)] [Medline: [25413636](https://pubmed.ncbi.nlm.nih.gov/25413636/)]
2. Wang T, Fu H, Kaminga AC, Li Z, Guo G, Chen L, et al. Prevalence of depression or depressive symptoms among people living with HIV/AIDS in China: a systematic review and meta-analysis. *BMC Psychiatry* 2018 May 31;18(1):160 [FREE Full text] [doi: [10.1186/s12888-018-1741-8](https://doi.org/10.1186/s12888-018-1741-8)] [Medline: [29855289](https://pubmed.ncbi.nlm.nih.gov/29855289/)]
3. Yu B, Zhu Q, Meng G, Gu Y, Zhang Q, Liu L, et al. Habitual yoghurt consumption and depressive symptoms in a general population study of 19,596 adults. *Eur J Nutr* 2018 Oct;57(7):2621-2628. [doi: [10.1007/s00394-017-1532-x](https://doi.org/10.1007/s00394-017-1532-x)] [Medline: [28856430](https://pubmed.ncbi.nlm.nih.gov/28856430/)]
4. Tao J, Vermund SH, Qian H. Association between depression and antiretroviral therapy use among people living with HIV: a meta-analysis. *AIDS Behav* 2018 May;22(5):1542-1550. [doi: [10.1007/s10461-017-1776-8](https://doi.org/10.1007/s10461-017-1776-8)] [Medline: [28439754](https://pubmed.ncbi.nlm.nih.gov/28439754/)]
5. Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, DCP MNS Author Group. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. *Lancet* 2016 Apr 16;387(10028):1672-1685. [doi: [10.1016/S0140-6736\(15\)00390-6](https://doi.org/10.1016/S0140-6736(15)00390-6)] [Medline: [26454360](https://pubmed.ncbi.nlm.nih.gov/26454360/)]
6. Gonzalez JS, Batchelder AW, Psaros C, Safren SA. Depression and HIV/AIDS treatment nonadherence: a review and meta-analysis. *J Acquir Immune Defic Syndr* 2011 Oct 1;58(2):181-187 [FREE Full text] [doi: [10.1097/QAI.0b013e31822d490a](https://doi.org/10.1097/QAI.0b013e31822d490a)] [Medline: [21857529](https://pubmed.ncbi.nlm.nih.gov/21857529/)]
7. Patel V, Xiao S, Chen H, Hanna F, Jotheeswaran AT, Luo D, et al. The magnitude of and health system responses to the mental health treatment gap in adults in India and China. *Lancet* 2016 Dec 17;388(10063):3074-3084. [doi: [10.1016/S0140-6736\(16\)00160-4](https://doi.org/10.1016/S0140-6736(16)00160-4)] [Medline: [27209149](https://pubmed.ncbi.nlm.nih.gov/27209149/)]
8. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet* 2007 Sep 8;370(9590):841-850 [FREE Full text] [doi: [10.1016/S0140-6736\(07\)61414-7](https://doi.org/10.1016/S0140-6736(07)61414-7)] [Medline: [17826169](https://pubmed.ncbi.nlm.nih.gov/17826169/)]
9. Nwobi UA, Eseadi C, Emeka O, Ekwealor N, Ogbonnaya KA, Oboegbulem AI, et al. A stress management intervention for adults living with HIV in Nigerian community settings: An effects study. *Medicine (Baltimore)* 2018 Nov;97(44):e12801 [FREE Full text] [doi: [10.1097/MD.00000000000012801](https://doi.org/10.1097/MD.00000000000012801)] [Medline: [30383633](https://pubmed.ncbi.nlm.nih.gov/30383633/)]
10. Berger S, Schadt T, von Wyl V, Ehlert U, Zellweger C, Furrer H, et al. Effects of cognitive behavioral stress management on HIV-1 RNA, CD4 cell counts and psychosocial parameters of HIV-infected persons. *AIDS* 2008 Mar 30;22(6):767-775. [doi: [10.1097/QAD.0b013e318282f511dc](https://doi.org/10.1097/QAD.0b013e318282f511dc)] [Medline: [18356607](https://pubmed.ncbi.nlm.nih.gov/18356607/)]
11. McGregor BA, Dolan ED, Murphy KM, Sannes TS, Highland KB, Albano DL, et al. Cognitive behavioral stress management for healthy women at risk for breast cancer: a novel application of a proven intervention. *Ann Behav Med* 2015 Dec;49(6):873-884 [FREE Full text] [doi: [10.1007/s12160-015-9726-z](https://doi.org/10.1007/s12160-015-9726-z)] [Medline: [26290001](https://pubmed.ncbi.nlm.nih.gov/26290001/)]
12. Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annu Rev Clin Psychol* 2007;3:1-27. [doi: [10.1146/annurev.clinpsy.3.022806.091432](https://doi.org/10.1146/annurev.clinpsy.3.022806.091432)] [Medline: [17716046](https://pubmed.ncbi.nlm.nih.gov/17716046/)]
13. Lemmens LH, Müller VN, Arntz A, Huibers MJ. Mechanisms of change in psychotherapy for depression: an empirical update and evaluation of research aimed at identifying psychological mediators. *Clin Psychol Rev* 2016 Dec;50:95-107. [doi: [10.1016/j.cpr.2016.09.004](https://doi.org/10.1016/j.cpr.2016.09.004)] [Medline: [27770716](https://pubmed.ncbi.nlm.nih.gov/27770716/)]
14. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 2002 Oct;59(10):877-883. [doi: [10.1001/archpsyc.59.10.877](https://doi.org/10.1001/archpsyc.59.10.877)] [Medline: [12365874](https://pubmed.ncbi.nlm.nih.gov/12365874/)]
15. Folkman S, Moskowitz JT. Coping: pitfalls and promise. *Annu Rev Psychol* 2004;55:745-774. [doi: [10.1146/annurev.psych.55.090902.141456](https://doi.org/10.1146/annurev.psych.55.090902.141456)] [Medline: [14744233](https://pubmed.ncbi.nlm.nih.gov/14744233/)]
16. Nakano K. Coping strategies and psychological symptoms in a Japanese sample. *J Clin Psychol* 1991 May;47(3):346-350. [doi: [10.1002/1097-4679\(199105\)47:3<346::aid-jclp2270470304>3.0.co;2-h](https://doi.org/10.1002/1097-4679(199105)47:3<346::aid-jclp2270470304>3.0.co;2-h)] [Medline: [2066401](https://pubmed.ncbi.nlm.nih.gov/2066401/)]
17. Logie C, James L, Tharao W, Loutfy M. Associations between HIV-related stigma, racial discrimination, gender discrimination, and depression among HIV-positive African, Caribbean, and Black women in Ontario, Canada. *AIDS Patient Care STDS* 2013 Feb;27(2):114-122. [doi: [10.1089/apc.2012.0296](https://doi.org/10.1089/apc.2012.0296)] [Medline: [23373665](https://pubmed.ncbi.nlm.nih.gov/23373665/)]
18. Mahajan AP, Sayles JN, Patel VA, Remien RH, Sawires SR, Ortiz DJ, et al. Stigma in the HIV/AIDS epidemic: a review of the literature and recommendations for the way forward. *AIDS* 2008 Aug;22(Suppl 2):S67-S79 [FREE Full text] [doi: [10.1097/01.aids.0000327438.13291.62](https://doi.org/10.1097/01.aids.0000327438.13291.62)] [Medline: [18641472](https://pubmed.ncbi.nlm.nih.gov/18641472/)]
19. Kiene SM, Dove M, Wanyenze RK. Depressive symptoms, disclosure, HIV-related stigma, and coping following HIV testing among outpatients in Uganda: a daily process analysis. *AIDS Behav* 2018 May;22(5):1639-1651 [FREE Full text] [doi: [10.1007/s10461-017-1953-9](https://doi.org/10.1007/s10461-017-1953-9)] [Medline: [29081046](https://pubmed.ncbi.nlm.nih.gov/29081046/)]
20. Cherenack EM, Sikkema KJ, Watt MH, Hansen NB, Wilson PA. Avoidant coping mediates the relationship between self-efficacy for HIV disclosure and depression symptoms among men who have sex with men newly diagnosed with HIV. *AIDS Behav* 2018 Oct;22(10):3130-3140 [FREE Full text] [doi: [10.1007/s10461-018-2036-2](https://doi.org/10.1007/s10461-018-2036-2)] [Medline: [29372454](https://pubmed.ncbi.nlm.nih.gov/29372454/)]
21. Lutgendorf SK, Antoni MH, Ironson G, Starr K, Costello N, Zuckerman M, et al. Changes in cognitive coping skills and social support during cognitive behavioral stress management intervention and distress outcomes in symptomatic human immunodeficiency virus (HIV)-seropositive gay men. *Psychosom Med* 1998;60(2):204-214. [doi: [10.1097/00006842-199803000-00017](https://doi.org/10.1097/00006842-199803000-00017)] [Medline: [9560871](https://pubmed.ncbi.nlm.nih.gov/9560871/)]

22. Allen N. Cognitive therapy of depression. Aaron T Beck, A John Rush, Brian F Shaw, Gary Emery. New York: Guilford Press, 1979. Aust N Z J Psychiatry 2002 Apr;36(2):275-278. [doi: [10.1046/j.1440-1614.2002.t01-5-01015.x](https://doi.org/10.1046/j.1440-1614.2002.t01-5-01015.x)] [Medline: [11982561](https://pubmed.ncbi.nlm.nih.gov/11982561/)]
23. Earnshaw VA, Bogart LM, Dovidio JF, Williams DR. Stigma and racial/ethnic HIV disparities: moving toward resilience. Am Psychol 2013;68(4):225-236 [FREE Full text] [doi: [10.1037/a0032705](https://doi.org/10.1037/a0032705)] [Medline: [23688090](https://pubmed.ncbi.nlm.nih.gov/23688090/)]
24. Tao J, Wang L, Kipp AM, Qian H, Yin L, Ruan Y, et al. Relationship of stigma and depression among newly HIV-diagnosed Chinese men who have sex with men. AIDS Behav 2017 Jan;21(1):292-299 [FREE Full text] [doi: [10.1007/s10461-016-1477-8](https://doi.org/10.1007/s10461-016-1477-8)] [Medline: [27376900](https://pubmed.ncbi.nlm.nih.gov/27376900/)]
25. Tshabalala J, Visser M. Developing a cognitive behavioural therapy model to assist women to deal with HIV and stigma. South Afr J Psychol 2011 Mar;41(1):17-28. [doi: [10.1177/008124631104100103](https://doi.org/10.1177/008124631104100103)]
26. Firth J, Torous J, Nicholas J, Carney R, Prapat A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
27. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: [10.1002/cpp.1855](https://doi.org/10.1002/cpp.1855)] [Medline: [23918764](https://pubmed.ncbi.nlm.nih.gov/23918764/)]
28. van Luenen S, Garnefski N, Spinhoven P, Kraaij V. Guided internet-based intervention for people with HIV and depressive symptoms: a randomised controlled trial in the Netherlands. Lancet HIV 2018 Sep;5(9):e488-e497. [doi: [10.1016/S2352-3018\(18\)30133-4](https://doi.org/10.1016/S2352-3018(18)30133-4)] [Medline: [30135045](https://pubmed.ncbi.nlm.nih.gov/30135045/)]
29. Schnall R, Cho H, Mangone A, Pichon A, Jia H. Mobile health technology for improving symptom management in low income persons living with HIV. AIDS Behav 2018 Oct;22(10):3373-3383 [FREE Full text] [doi: [10.1007/s10461-017-2014-0](https://doi.org/10.1007/s10461-017-2014-0)] [Medline: [29299790](https://pubmed.ncbi.nlm.nih.gov/29299790/)]
30. Corrigan P. How stigma interferes with mental health care. Am Psychol 2004 Oct;59(7):614-625. [doi: [10.1037/0003-066X.59.7.614](https://doi.org/10.1037/0003-066X.59.7.614)] [Medline: [15491256](https://pubmed.ncbi.nlm.nih.gov/15491256/)]
31. Kempf M, Huang C, Savage R, Safren SA. Technology-delivered mental health interventions for people living with HIV/AIDS (PLWHA): a review of recent advances. Curr HIV/AIDS Rep 2015 Dec;12(4):472-480 [FREE Full text] [doi: [10.1007/s11904-015-0292-6](https://doi.org/10.1007/s11904-015-0292-6)] [Medline: [26452644](https://pubmed.ncbi.nlm.nih.gov/26452644/)]
32. Schueller SM, Glover AC, Rufa AK, Dowdle CL, Gross GD, Karnik NS, et al. A mobile phone-based intervention to improve mental health among homeless young adults: pilot feasibility trial. JMIR Mhealth Uhealth 2019 Jul 2;7(7):e12347 [FREE Full text] [doi: [10.2196/12347](https://doi.org/10.2196/12347)] [Medline: [31267980](https://pubmed.ncbi.nlm.nih.gov/31267980/)]
33. Chan KL, Leung WC, Tiwari A, Or KL, Ip P. Using smartphone-based psychoeducation to reduce postnatal depression among first-time mothers: randomized controlled trial. JMIR Mhealth Uhealth 2019 May 14;7(5):e12794 [FREE Full text] [doi: [10.2196/12794](https://doi.org/10.2196/12794)] [Medline: [31094354](https://pubmed.ncbi.nlm.nih.gov/31094354/)]
34. Mansell G, Hill JC, Main CJ, Von Korff M, van der Windt D. Mediators of treatment effect in the back in action trial: using latent growth modeling to take change over time into account. Clin J Pain 2017 Sep;33(9):811-819 [FREE Full text] [doi: [10.1097/AJP.0000000000000463](https://doi.org/10.1097/AJP.0000000000000463)] [Medline: [27930393](https://pubmed.ncbi.nlm.nih.gov/27930393/)]
35. Guo Y, Hong YA, Qiao J, Xu Z, Zhang H, Zeng C, et al. Run4Love, a mHealth (WeChat-based) intervention to improve mental health of people living with HIV: a randomized controlled trial protocol. BMC Public Health 2018 Jun 26;18(1):793 [FREE Full text] [doi: [10.1186/s12889-018-5693-1](https://doi.org/10.1186/s12889-018-5693-1)] [Medline: [29940921](https://pubmed.ncbi.nlm.nih.gov/29940921/)]
36. Chhabria KS, Carnaby GD. Psychometric validation of the Center for Epidemiological Studies Depression Scale in Head and Neck Cancer patients. Oral Oncol 2017 Dec;75:158-162. [doi: [10.1016/j.oraloncology.2017.11.010](https://doi.org/10.1016/j.oraloncology.2017.11.010)] [Medline: [29224813](https://pubmed.ncbi.nlm.nih.gov/29224813/)]
37. Zeng F, Deng G, Wang Z, Liu L. WeChat: a new clinical teaching tool for problem-based learning. Int J Med Educ 2016 Apr 25;7:119-121 [FREE Full text] [doi: [10.5116/ijme.5708.e5c4](https://doi.org/10.5116/ijme.5708.e5c4)] [Medline: [27111920](https://pubmed.ncbi.nlm.nih.gov/27111920/)]
38. Lei H, Cheong CM, Li S, Lu M. The relationship between coping style and Internet addiction among mainland Chinese students: A meta-analysis. Psychiatry Res 2018 Dec;270:831-841. [doi: [10.1016/j.psychres.2018.10.079](https://doi.org/10.1016/j.psychres.2018.10.079)] [Medline: [30551332](https://pubmed.ncbi.nlm.nih.gov/30551332/)]
39. Berger BE, Ferrans CE, Lashley FR. Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. Res Nurs Health 2001 Dec;24(6):518-529. [doi: [10.1002/nur.10011](https://doi.org/10.1002/nur.10011)] [Medline: [11746080](https://pubmed.ncbi.nlm.nih.gov/11746080/)]
40. Xiao Z, Li X, Qiao S, Zhou Y, Shen Z, Tang Z. Using communication privacy management theory to examine HIV disclosure to sexual partners/spouses among PLHIV in Guangxi. AIDS Care 2015;27(Suppl 1):73-82 [FREE Full text] [doi: [10.1080/09540121.2015.1055229](https://doi.org/10.1080/09540121.2015.1055229)] [Medline: [26616128](https://pubmed.ncbi.nlm.nih.gov/26616128/)]
41. Byrne BM, Lam WW, Fielding R. Measuring patterns of change in personality assessments: an annotated application of latent growth curve modeling. J Pers Assess 2008 Nov;90(6):536-546. [doi: [10.1080/00223890802388350](https://doi.org/10.1080/00223890802388350)] [Medline: [18925494](https://pubmed.ncbi.nlm.nih.gov/18925494/)]
42. Selig JP, Preacher KJ. Mediation models for longitudinal data in developmental research. Res Hum Dev 2009;6(2-3):144-164. [doi: [10.1080/15427600902911247](https://doi.org/10.1080/15427600902911247)]
43. Cheong J, Mackinnon DP, Khoo ST. Investigation of mediational processes using parallel process latent growth curve modeling. Struct Equ Modeling 2003 Apr 1;10(2):238 [FREE Full text] [doi: [10.1207/S15328007SEM1002\\_5](https://doi.org/10.1207/S15328007SEM1002_5)] [Medline: [20157639](https://pubmed.ncbi.nlm.nih.gov/20157639/)]

44. Brown TA. Temporal course and structural relationships among dimensions of temperament and DSM-IV anxiety and mood disorder constructs. *J Abnorm Psychol* 2007 May;116(2):313-328. [doi: [10.1037/0021-843X.116.2.313](https://doi.org/10.1037/0021-843X.116.2.313)] [Medline: [17516764](https://pubmed.ncbi.nlm.nih.gov/17516764/)]
45. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods* 2008 Aug;40(3):879-891. [Medline: [18697684](https://pubmed.ncbi.nlm.nih.gov/18697684/)]
46. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr* 2009;76(4):408-420. [doi: [10.1080/03637750903310360](https://doi.org/10.1080/03637750903310360)]
47. Taylor AB, MacKinnon DP, Tein JY. Tests of the Three-Path Mediated Effect. *Organ Res Method* 2008;11(2):241-269. [doi: [10.1177/1094428107300344](https://doi.org/10.1177/1094428107300344)]
48. Kline RB. *Principles And Practice Of Structural Equation Modeling*. New York City: The Guilford Press; 2015.
49. Hu L, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Method* 1998;3(4):424-453. [doi: [10.1037/1082-989X.3.4.424](https://doi.org/10.1037/1082-989X.3.4.424)]
50. Antoni MH, Ironson G, Schneiderman N. *Cognitive-Behavioral Stress Management for Individuals Living with HIV*. Oxford, England, UK: Oxford University Press; 2007.
51. Chesney MA, Chambers DB, Taylor JM, Johnson LM, Folkman S. Coping effectiveness training for men living with HIV: results from a randomized clinical trial testing a group-based intervention. *Psychosom Med* 2003;65(6):1038-1046. [doi: [10.1097/01.psy.0000097344.78697.ed](https://doi.org/10.1097/01.psy.0000097344.78697.ed)] [Medline: [14645783](https://pubmed.ncbi.nlm.nih.gov/14645783/)]
52. Fife BL, Scott LL, Fineberg NS, Zwickl BE. Promoting adaptive coping by persons with HIV disease: evaluation of a patient/partner intervention model. *J Assoc Nurses AIDS Care* 2008;19(1):75-84 [FREE Full text] [doi: [10.1016/j.jana.2007.11.002](https://doi.org/10.1016/j.jana.2007.11.002)] [Medline: [18191771](https://pubmed.ncbi.nlm.nih.gov/18191771/)]
53. Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, et al. Mobile Health (mHealth) Versus Clinic-Based Group Intervention for People With Serious Mental Illness: A Randomized Controlled Trial. *Psychiatr Serv* 2018 Sep 1;69(9):978-985. [doi: [10.1176/appi.ps.201800063](https://doi.org/10.1176/appi.ps.201800063)] [Medline: [29793397](https://pubmed.ncbi.nlm.nih.gov/29793397/)]
54. Ben-Zeev D, Brian RM, Aschbrenner KA, Jonathan G, Steingard S. Video-based mobile health interventions for people with schizophrenia: Bringing the 'pocket therapist' to life. *Psychiatr Rehabil J* 2018 Mar;41(1):39-45 [FREE Full text] [doi: [10.1037/prj0000197](https://doi.org/10.1037/prj0000197)] [Medline: [27295133](https://pubmed.ncbi.nlm.nih.gov/27295133/)]
55. Tobin K, Davey-Rothwell MA, Nonyane BA, Knowlton A, Wissow L, Latkin CA. RCT of an integrated CBT-HIV intervention on depressive symptoms and HIV risk. *PLoS One* 2017;12(12):e0187180 [FREE Full text] [doi: [10.1371/journal.pone.0187180](https://doi.org/10.1371/journal.pone.0187180)] [Medline: [29240757](https://pubmed.ncbi.nlm.nih.gov/29240757/)]
56. Safren SA, Bedoya CA, O'Cleirigh C, Biello KB, Pinkston MM, Stein MD, et al. Cognitive behavioural therapy for adherence and depression in patients with HIV: a three-arm randomised controlled trial. *Lancet HIV* 2016 Nov;3(11):e529-e538 [FREE Full text] [doi: [10.1016/S2352-3018\(16\)30053-4](https://doi.org/10.1016/S2352-3018(16)30053-4)] [Medline: [27658881](https://pubmed.ncbi.nlm.nih.gov/27658881/)]
57. Shimotsu S, Horikawa N, Emura R, Ishikawa S, Nagao A, Ogata A, et al. Effectiveness of group cognitive-behavioral therapy in reducing self-stigma in Japanese psychiatric patients. *Asian J Psychiatr* 2014 Aug;10:39-44. [doi: [10.1016/j.ajp.2014.02.006](https://doi.org/10.1016/j.ajp.2014.02.006)] [Medline: [25042950](https://pubmed.ncbi.nlm.nih.gov/25042950/)]
58. Harper GW, Lemos D, Hosek SG. Stigma reduction in adolescents and young adults newly diagnosed with HIV: findings from the Project ACCEPT intervention. *AIDS Patient Care STDS* 2014 Oct;28(10):543-554 [FREE Full text] [doi: [10.1089/apc.2013.0331](https://doi.org/10.1089/apc.2013.0331)] [Medline: [25216106](https://pubmed.ncbi.nlm.nih.gov/25216106/)]

## Abbreviations

- BCI:** bootstrapped CI
- CBSM:** cognitive behavioral stress management
- CES-D:** Center for Epidemiological Studies Depression Scale
- CFI:** comparative fit index
- LGCM:** latent growth curve model
- mHealth:** mobile health
- PLWH:** people living with HIV
- RCT:** randomized controlled trial
- RMSEA:** root mean square error of approximation
- SEM:** structural equation modeling
- SRMR:** standardized root mean square residual
- SWCQ:** Simplified Ways of Coping Questionnaire

*Edited by G Eysenbach; submitted 16.07.19; peer-reviewed by C Simons, M Domhardt, N Karnik; comments to author 15.08.19; revised version received 25.09.19; accepted 20.10.19; published 15.11.19.*

*Please cite as:*

*Zhu M, Cai W, Li L, Guo Y, Monroe-Wise A, Li Y, Zeng C, Qiao J, Xu Z, Zhang H, Zeng Y, Liu C*

*Mediators of Intervention Effects on Depressive Symptoms Among People Living With HIV: Secondary Analysis of a Mobile Health Randomized Controlled Trial Using Latent Growth Curve Modeling*

*JMIR Mhealth Uhealth 2019;7(11):e15489*

URL: <http://mhealth.jmir.org/2019/11/e15489/>

doi: [10.2196/15489](https://doi.org/10.2196/15489)

PMID: [31730042](https://pubmed.ncbi.nlm.nih.gov/31730042/)

©Mengting Zhu, Weiping Cai, Linghua Li, Yan Guo, Aliza Monroe-Wise, Yiran Li, Chengbo Zeng, Jiaying Qiao, Zhimeng Xu, Hanxi Zhang, Yu Zeng, Cong Liu. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 15.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Assessment of a Mobile App by Adolescents and Young Adults With Cystic Fibrosis: Pilot Evaluation

Isa Rudolf<sup>1,2\*</sup>, MD; Katharina Pieper<sup>1\*</sup>; Helga Nolte<sup>3</sup>; Sibylle Junge<sup>1</sup>, MD; Christian Dopfer<sup>1,2</sup>, MD; Annette Sauer-Heilborn<sup>4</sup>, MD; Felix C Ringshausen<sup>2,4</sup>, MD; Burkhard Tümmler<sup>1,2</sup>, MD; Ute von Jan<sup>5</sup>, PhD; Urs-Vito Albrecht<sup>5</sup>, MD; Jan Fuge<sup>2</sup>; Gesine Hansen<sup>1,2</sup>, MD; Anna-Maria Dittrich<sup>1,2</sup>, MD

<sup>1</sup>Pediatric Pneumology, Allergology, and Neonatology, Hannover Medical School, Hannover, Germany

<sup>2</sup>Biomedical Research in End-Stage and Obstructive Lung Disease Hannover, German Centre for Lung Research, Hannover Medical University, Hannover, Germany

<sup>3</sup>Haus Schutzensel, Mukoviszidose eV, Hannover, Germany

<sup>4</sup>Department of Respiratory Medicine, Hannover Medical University, Hannover, Germany

<sup>5</sup>Peter L Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

\*these authors contributed equally

**Corresponding Author:**

Anna-Maria Dittrich, MD

Pediatric Pneumology, Allergology, and Neonatology

Hannover Medical School

Carl-Neuberg-Straße 1

Hannover, 30625

Germany

Phone: 49 17615329785

Email: [dittrich.anna-maria@mh-hannover.de](mailto:dittrich.anna-maria@mh-hannover.de)

## Abstract

**Background:** Cystic fibrosis (CF) continues to be the most common life-limiting chronic pulmonary disease in adolescents and young adults. Treatment of CF demands a high treatment time investment to slow the progression of lung function decline, the most important contributor to morbidity and mortality. Adherence is challenging in CF due to the high treatment burden and the lack of immediate health consequences in case of nonadherence. Lung function decline is particularly pronounced in the transition phase between 12 and 24 years of age. The improvement of self-management and self-responsibility and independence from parents and desire for normalcy are conflicting aspects for many adolescents with CF, which influence adherence to the time-consuming pulmonary therapy. Mobile health (mHealth) care apps could help to support self-management and independence and thereby reconcile seemingly conflicting goals to improve adherence, quality of life, and ultimately CF life expectancy.

**Objective:** This study aimed to (1) assess user behavior and satisfaction among adolescents and young adults with CF over an observation period of three months using an mHealth app; (2) identify areas of improvement for this mHealth app; and (3) compare overall and disease-specific satisfaction, lung function, and anthropometry before and after using the mHealth app.

**Methods:** A total of 27 adolescents and young adults with CF (age range 12-24 years, mean age 16 years, SD 3 years; 14 females, 11 males) used a free mHealth app for three months of whom 25 provided questionnaire data for analysis at the end of the study. Data collection was carried out using questionnaires on usage characteristics and life satisfaction, and standardized assessment of lung function and anthropometry.

**Results:** The use of the reminder function for medication declined from 70% (15/21) of the participants at week 4 to 65% (13/20) at week 8 of the observation period. At the end of the study, only 17% (4/23) of the participants wanted to continue using the app. Nevertheless, 56% (14/25) of participants saw the mobile app as a support for everyday life. Potential improvements targeting hedonistic qualities were identified to improve mHealth app adherence. Comparisons of satisfaction with different life aspects hinted at improvements or stabilization for the subitem respiration and the subitem lack of handicap by CF, suggesting that app use might stabilize certain CF-specific aspects of the weighted satisfaction with life. Lung function and anthropometry were not affected consistently.

**Conclusions:** Most of the patients did not want to continue using the app after the study period. Only a few CF-specific aspects of weighted life satisfaction were possibly stabilized by the mHealth app; clinical parameters were not affected. Adaptation of the functions to adolescent-specific needs could improve the long-term use and thus positively affect the disease course.

(*JMIR Mhealth Uhealth* 2019;7(11):e12442) doi:[10.2196/12442](https://doi.org/10.2196/12442)

## KEYWORDS

mobile phone; mobile phone app; mHealth; self-management; adolescence; cystic fibrosis

## Introduction

### Background

Cystic fibrosis (CF) is the most common systemic metabolic disease in Caucasians, affecting more than 70,000 patients worldwide, about 8000 of whom reside in Germany, where currently about 55% are in the adolescent or transition age, that is, between 12 and 29 years old [1]. Despite progress, CF remains life-limiting, mainly because of chronic pulmonary disease, with a high morbidity burden.

Treatment of CF demands one of the highest treatment time investments of all chronic diseases, which turns treatment adherence critical and vulnerable [2,3]. Treatment time investment in CF is high mainly because of the need for stringent, twice daily inhalational therapy coupled with physiotherapy techniques, but also because of the continuous need of enzyme replacement therapy with every fat-containing meal. Changes in everyday symptom load typically occur in the context of common colds and need to be judged by the patient for their severity, as they might necessitate treatment changes, that is, additional inhalations or antibiotics. Adherence in CF is particularly challenging, because of not only the high time investment but also the lack of immediate changes in health status in case of nonadherence, as disease progression is not immediately palpable in case of lung function decline.

In CF, as in all chronic diseases, the transition phase, that is, the period of life between approximately 12 and 24 years of age, constitutes the period in life when the most rapid loss of organ function can be observed [3], crucially setting the stage for future health [4]. A large proportion of loss of organ function is due to the psychological demands inherent to puberty. Adolescents afflicted by a chronic disease, which demands a high degree of treatment adherence, face the difficulty to meet the requirements of their disease, including the development of sufficient disease investment, in addition to the pubertal requirements to develop self-responsibility, self-determination, and independence [5]. During adolescence, patients with CF must learn to assess changes in everyday symptom load by themselves, to be able to take over this critical task from their parents. Disease investment because of a high load of therapy and the need to monitor health status versus parental independence and the desire for normalcy create conflicting situations for adolescents with chronic disease and turn adherence into a particularly vulnerable quality during this period of life [5-7]. Development of independent self-management techniques, a goal for many adolescents, is thus of an even higher priority in adolescents with chronic diseases such as CF.

The adolescent population, regardless of having a chronic disease, is difficult to approach, often coupled to a lack of attainability via conventional communication venues. Communication via mobile phones constitutes a promising venue in that respect. In Germany, in 2015, 90% of adolescents aged between 14 and 19 years owned a mobile phone, of whom 80% considered their mobile phone to be *very important* or *important* to them [8], providing proof to the common knowledge of high mobile phone affinity of this age group worldwide [9]. Currently available mobile apps offer novel approaches to support the daily therapy of patients with chronic diseases. On those lines, systematic reviews show promise in the use of mobile health (mHealth) apps to improve self-management for chronic diseases [10], also in the area of CF [7]. However, many of the existing mHealth apps are not disease specific [11], or do not provide all the desirable functions for the particular disease [12], although this is a typical demand in user-centered surveys among CF patients [9,11]. Development of CF-specific mHealth apps by pharmaceutical companies introduces a bias, possibly conflicting with patients' actual needs [13]. Some comprehensive mHealth apps have been designed by CF patients themselves [14] or doctors involved in CF patients' care [14-16], yet language and differences in medical systems provide barriers to extrapolation in different countries. Stringent scientific evaluation of mHealth apps, specifically for patients in the transition age, remains rare [9], even though studies have shown age-dependent differences in usage of mHealth apps [17].

### Hypothesis

The studies cited above indicate that rigorous analyses of mHealth app use in adolescents with CF are necessary when aiming at the development of supportive mHealth apps for this particular age group. We hypothesized that a positive assessment of a mobile phone app would lead to sustained use over time.

### Objectives

To address this hypothesis, we introduced a subgroup of our CF patients in our transition clinic to a mobile phone app, the KiOAPP [18], designed to support disease management. There are several assets in KiOAPP that influenced our choice of this specific app. As CF doctors, we consider self-management and communication essential to develop sustained adherence, regardless of patients' age. KiOAPP addresses these aspects by a diary, communication venue, a medication plan, and reminder function. Moreover, the developers of KiOAPP claim age-specific tailoring, for example, by using youth-specific, less formal language, colorful and adaptable design, placing emphasis on user autonomy [19], which might be essential to improve user affinity and, thereby, promoting sustained use.

We addressed our hypothesis in a pilot clinical trial that included 27 adolescent CF patients who were introduced to KiOAPP. Here, we present analyses on usage characteristics, product satisfaction, and clinical effects for 25 of them whose questionnaire data were available at the end of the study.

## Methods

### The KiOAPP Mobile Health App

The mHealth app (KiOAPP) was developed by a German nonprofit society for adolescents and young adults after solid organ transplantation [18]. The app is available free of charge in iOS and Android app stores. It offers the following features (also see [Multimedia Appendix 1](#)):

- Individualizable medication reminder, including a bar code scanner introduction aid and an audio reminder system (medication reminder function);
- Diary functions for recording daily vitals and personal observations (diary function);
- Communication platform for transmitting information to the physician (contact function); and
- Individualizable user surface (design function).

The integration of a bar code scanning function, allowing capturing bar codes on medication boxes as they are in use in Germany, facilitates the inclusion of a large number of different medications typical in CF into an individual mobile phone-based medication and medication reminder plan. It obviates the need for time-consuming preparation of an entry, which might deter users from this function. We considered the diary function and communication venue useful as they offer day-to-day documentation of symptom load, which adolescents need to learn to assess by themselves when taking over responsibility for their disease from their parents and offers the possibility to discuss changes with the attending physician. We considered the medication plan to provide independence from parental oversight and facilitate self-management of adolescents with CF, thereby providing support in the conflict between independence and disease investment.

### Study Population

Patients were eligible for inclusion into the interventional study via diagnosis of CF and age 12 to 24 years. The Hannover CF center routinely tends to approximately 400 patients with CF across all ages, of whom approximately 150 were of the required age during the recruitment period of the study. All CF patients at our center received a flyer explaining the app and the trial by email and handout upon their clinical appointments at the CF center. They were asked whether they wanted to participate in the study at their next and the consecutive center appointment, leading to a recruitment period of 9 months without randomization and recruitment of 27 patients. Of these patients, 2 were not willing to follow-up via questionnaires after the 3-month intervention period, leading to datasets of 25 participants for analysis at the completion of the study. Not all patients answered all items in all questionnaires. Numbers of patients who replied are indicated in the respective analyses and figures. For analyses on the effects of app use on forced expiratory volume in 1 second (FEV1) and body mass index

(BMI), which were available for 25 of the 27 participants who had utilized the app, a control group of 25 CF patients who had not participated in the app study was randomly selected from all CF patients at our center, 2 years after the study. For each study participant, we selected 1 CF patient of similar age, sex, FEV1, and BMI values at the time our interventional study had started.

### Study Design

The study was approved by the local ethics committee (#2826-2015). Upon consent, patients were introduced to the mHealth app, including an explanation of its different functions and the possible usefulness of these function to the everyday treatment burden and self-assessment necessary for patients with CF. At the same time point, all participants were assessed for life satisfaction via FLZ questionnaire (questionnaire on life satisfaction [Fragebogen zur Lebenszufriedenheit]; [Multimedia Appendices 2-4](#)), FEV1, and BMI. Within the next week, 4 and 8 weeks later, participants were contacted by phone to ensure installment and understanding of the app. At week 4 and 8 post installment, we systematically assessed app usage and satisfaction via questionnaire III ([Multimedia Appendix 5](#)). At these time points, not all participating patients were contactable via phone, leading to 21 (4 weeks) and 20 (8 weeks) phone interviews. Life satisfaction ([Multimedia Appendices 2-4](#)), app usage (questionnaire I, *Systems Usability Scale*, [Multimedia Appendix 6](#)) and satisfaction (questionnaire II, [Multimedia Appendix 7](#)), and product quality (AttrakDiff questionnaire) were consecutively assessed by questionnaires 3 months after recruitment at the observation endpoint, with patients answering paper and Web-based questionnaires during their waiting period for an ambulatory appointment, and interrogation about open questions or uncertainties by the physician during that appointment.

Lung function and BMI were assessed in 3-monthly intervals as part of routine care for the study group and the control group. We compared lung function and BMI of the 25 study subjects who provided questionnaire data at the end of the study and matched 25 control subjects after 3 months, 1 year, and 2 years post inclusion to investigate a possible time-dependent effect after using the app, as lung function and BMI are clinical parameters known to take some time to reflect changes induced by clinical interventions [20].

### Questionnaires

When possible, the authors chose well-established questionnaires which have been evaluated regarding reliability, validity, and reproducibility. Although the authors equally felt that for description of usage characteristics and satisfaction with the app, none of the established test systems offered appropriate content and thus self-designed 2 additional questionnaires, which thus lacked standardized assessment of their test qualities.

The established questionnaires were selected by 4 of the coauthors (HN, UVA, AMD, and KP; with a combined CF expertise of >95 years), each contributing unique backgrounds (social worker, medical doctor and informatician, CF specialist, and medical student within the age range of the target population). The same authors co-designed 2 questionnaires

([Multimedia Appendices 5](#) and [7](#)). After incorporation of feedback to these self-designed questionnaires by our multidisciplinary CF team, including another social worker, a psychologist, a data manager, a dietician, and a physiotherapist, the questionnaires were finalized.

### **App Usage and Satisfaction**

App usage and satisfaction were assessed by 3 questionnaires. Questionnaires I (standardized *system usability scale* questionnaire [21], see [Multimedia Appendix 6](#)) and II (self-designed questionnaire, see [Multimedia Appendix 7](#)) aimed at assessing user behavior. These questionnaires were handed out after the intervention, that is, 12 weeks after the introduction to the mHealth app, during a regular outpatient visit. Questionnaire III (self-designed questionnaire, see [Multimedia Appendix 5](#)) was used during telephone interviews with participants, 4 and 8 weeks after the introduction to the mHealth app.

### **Product Qualities of the KiOAPP**

Qualities of the KiOAPP were assessed by the AttrakDiff questionnaire, a Web-based questionnaire [22] 12 weeks after the introduction to the mHealth app during a regular outpatient visit. AttrakDiff is a well-established tool for assessing a product's usability and design. [23-25]. To this end, it differentiates between pragmatic and hedonic qualities. Usefulness and usability are aspects contributing to pragmatic quality, whereas features that help fulfill emotional needs such as identification with the product, or simply curiosity, are so-called hedonic factors. Both pragmatic and hedonic factors are important for a product's overall attractiveness. The assessment itself is done via many opposing word pairs, such as *technical* versus *human* or *harmless* versus *challenging*, with participants being asked to rate the product via these word pairs using a scale between -3 and +3.

### **Patients' Life Satisfaction**

Patients' life satisfaction was assessed by the FLZ questionnaire [26], a well-established questionnaire for the assessment of life satisfaction which offers the additional asset that it includes a CF-specific module and reference population. We assessed life satisfaction at the time of inclusion into the study and 12 weeks after the introduction to the mHealth app during regular outpatient visits. We assessed 3 modules (FLZ): *general life* (see [Multimedia Appendix 2](#)), *general health* (see [Multimedia Appendix 3](#)), and *CF life* (see [Multimedia Appendix 4](#)), each containing 8 categories [26]. The scores for the individual modules were summed up as global scores. Reference populations for the FLZ are CF patients aged 16 to 45 years,

n=251 [26]. Analysis of sum scores from study participants aged 12 to 15 years versus study participants aged 16 to 24 years were performed to assess age-specific differences.

### **Lung Function**

Spirometry was performed according to American Thoracic Society/European Respiratory Society criteria on a PowerCube Body Plethysmograph (Ganshorn). Absolute values were referenced according to Knudson providing percentages compared with a standard reference population and provided for the pertinent time points [27].

### **Body Mass Index**

BMI measurements were obtained by measurements of weight and height and provided as absolute values.

### **Statistical Analyses**

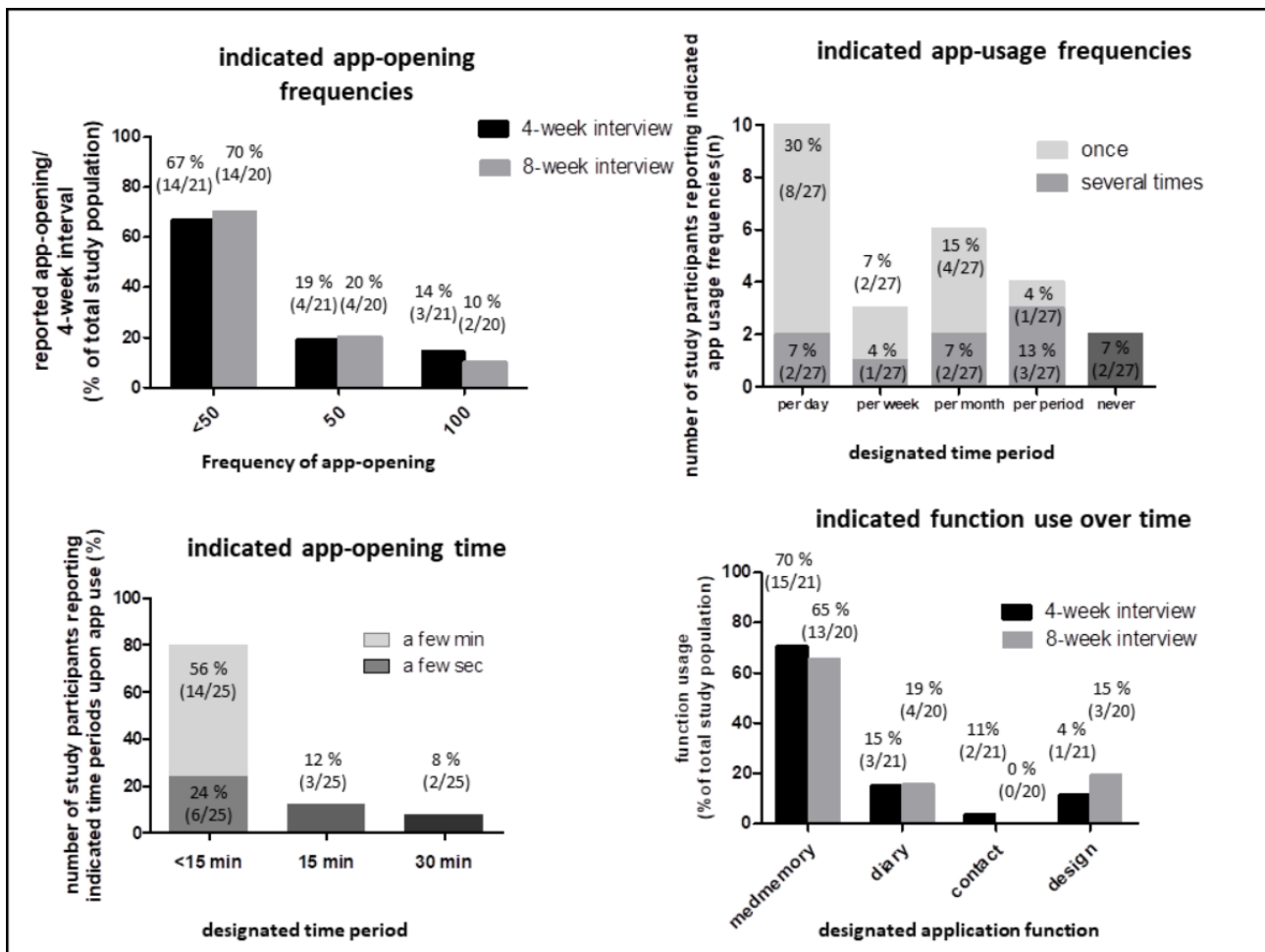
Where indicated, arithmetic mean (MW) and standard deviation (SD) were determined for all parameters. Normal distribution was tested with the Kolmogorov-Smirnov test. *P* values were calculated with a one-sample *t* test using the statistical software SPSS (version 23, IBM) where indicated throughout the manuscript. The significance level for all tests is indicated as actual values. Hedge *g* values were used as an alternative evaluation of effect size of the difference between the reference population and our study population by comparing FLZ sum scores between these 2 groups. Cohen *d* values were used as an alternative evaluation of effect size of the intervention by comparing lung function and anthropometry of our study group versus the age and gender-matched control population at given time points.

## **Results**

### **Mobile Health App Usage Characteristics**

Most of the patients used the app less than 50 times in a 4-week interval, followed by more frequent utilization times. These categories remained stable over the intervention ([Figure 1](#), assessment of question #2 "How often have you used the app in the past 4 weeks?", questionnaire [Multimedia Appendix 3](#)). After completion of the observation period, most of the participants reported that they had used the app several times or at least once per day ([Figure 1](#), assessment of question #2 "How often per day/per month/per three-month have you used the app?", questionnaire [Multimedia Appendix 7](#)). Most participants reported low usage times of a few seconds to minutes per opening of the app ([Figure 1](#), assessment of question #4 "How long have you used the app?", questionnaire [Multimedia Appendix 7](#)).

**Figure 1.** Usage characteristics of the mHealth app. Patients reported number of app-openings per 4 week-interval, estimated overall frequency of app use, reported opening times of app after study completion and use of different app functions. Numbers indicate percentages of participants and overall numbers.



The medication reminder function was the most frequently used (15/21, 70%, at 4 weeks), followed by the diary function (3/21, 15%, at 4 weeks), the contact function (1/21, 4%, at 4 weeks), and the design function (2/21, 11%, at 4 weeks) (Figure 1, assessment question #1 “Which function of the app have you used in the last 4 weeks?”, questionnaire Multimedia Appendix 5).

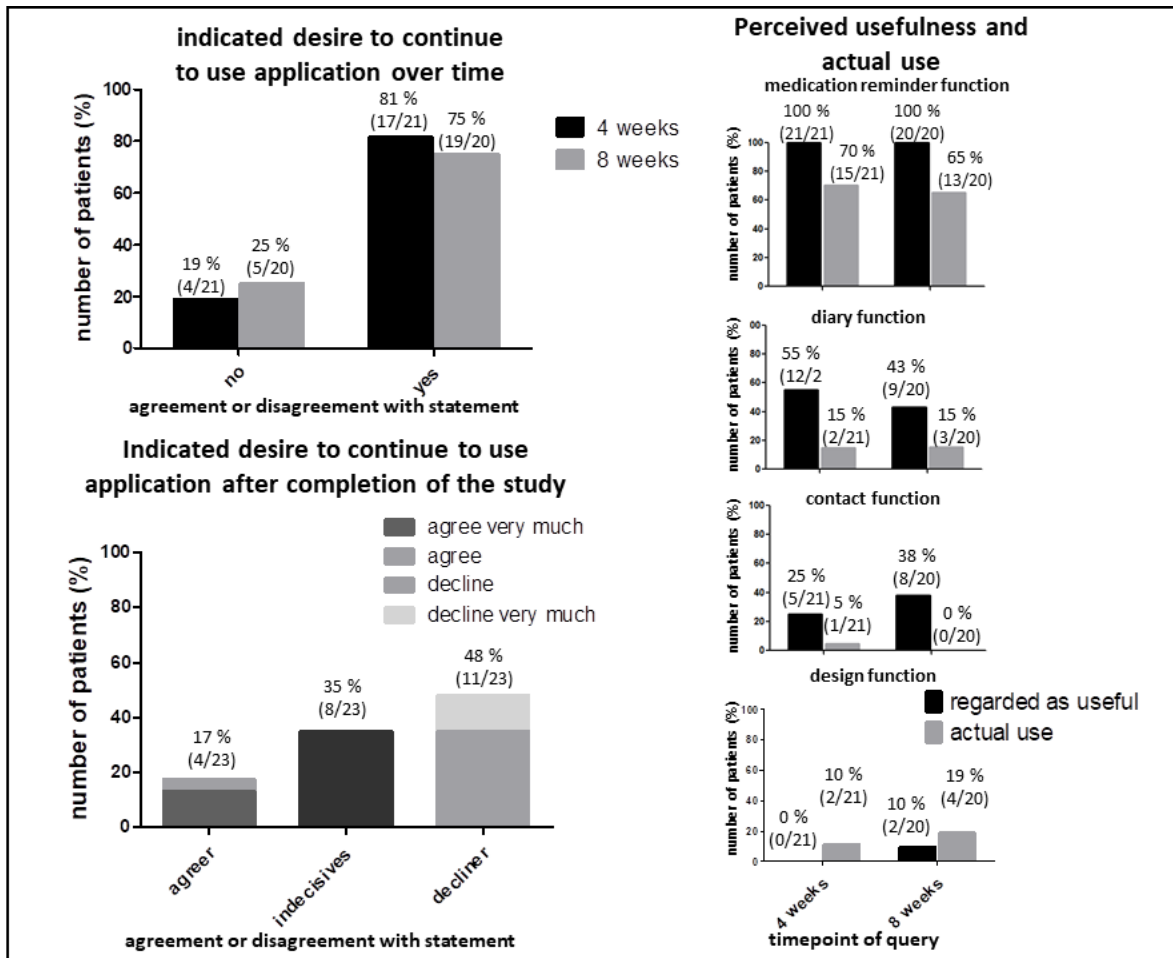
Use of the medication reminder function declined between time points from 70% (15/20) at 4 weeks to 65% (13/20) at 8 weeks (Figure 1). Use of the diary function was much lower but remained almost stable at 4 and 8 weeks (3/21 or rather 3/20). The contact function and the design function were both used by only a very limited proportion of participants (Figure 1).

Upon examining user patterns, we identified users who (1) used all app functions and reduced use of the medication reminder function over time; (2) used all apps at a steady level; and (3) used only the medication reminder function and reduced its use over time.

**User Ratings: Continuation of Use and Perceived Usefulness**

After 4 and 8 weeks, most of the participants wanted to continue to use the app (17/21, 81%, at 4 weeks; 15/20, 75%, after 8 weeks) (Figure 2, assessment question #10 “Can you imagine to use the app after finishing the study?”, questionnaire Multimedia Appendix 5).

**Figure 2.** Continuation of use and perceived usefulness versus actual use of mobile health (mHealth) app. Patients reported wish to continue to use app after 4 and 8 weeks and after completion of the study. Patients were queried for perceived usefulness and actual use of different mHealth functions: medication reminder function, diary function, contact function and design function. Perceived usefulness (black bars) was plotted versus reported usage (grey bars). Numbers indicate percentages of participants and overall numbers.



After the completion of the study, not all participants answered the items on their desire to continue to use the app, leading to 23 datasets. Also, the number of participants who wanted to continue to use the app declined to only 17% (4/23) of participants who wanted to continue using the app (Figure 2, sum of *agree very much* and *agree*, see assessment question #1 of the questionnaire in Multimedia Appendix 6) with a large proportion being indecisive (8/23, 35%), and 48% (11/23) of participants declining further usage (Figure 2, sum of *decline* and *decline very much*), suggesting a critical window of attrition between 8 and 12 weeks.

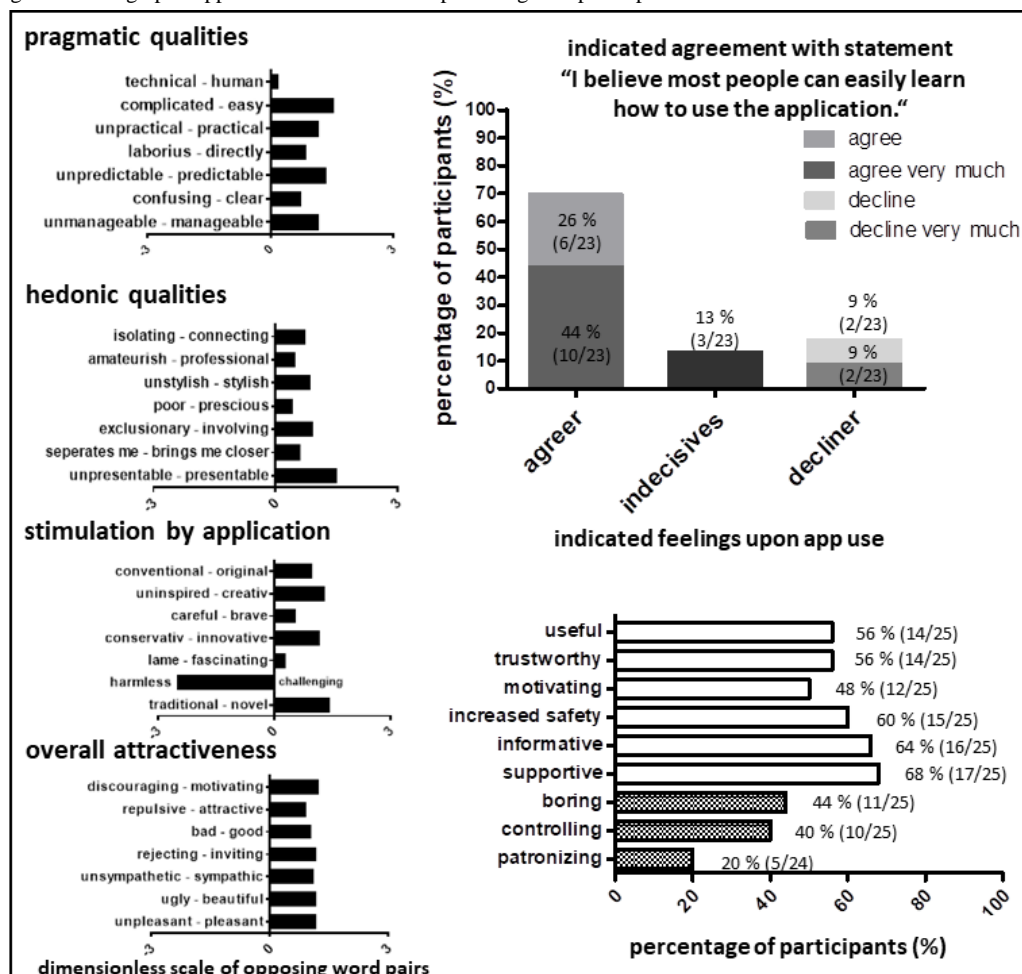
When asked which functions were regarded as useful, results for the medication reminder function were most similar to the actual use of this function (Figure 2, 100% (21/21) perceived usefulness vs. 70% (15/21) actual use at 4 weeks, 100% (20/20) perceived usefulness vs. 65% (13/20) actual use at 8 weeks, composite figure of 2 questions of questionnaire Multimedia Appendix 5: question #4=perceived usefulness, black bars and question #1=actual use, grey bars). Discrepancies between perceived usefulness and actual use were more pronounced for the diary function and the contact possibility as useful (Figure 2). For the design function, the picture was the opposite: more patients used this function than deeming it useful (Figure 2).

Still, the latter 2 functions were used and deemed useful only by a small proportion of participants. Perceived usefulness remained stable over the 8-week observation period, with only the design function possibly being regarded more useful after prolonged usage (Figure 2).

### User Ratings: Operability, Hedonistic Qualities, and Attractiveness

We used a subtest of the Web-based questionnaire, AttrakDiff (*opposing word pairs*), which assesses product operability (pragmatic quality, Figure 3), hedonic quality (Figure 3), *stimulation of the user by the app* (Figure 3), and overall attractiveness (Figure 3) of the mHealth app to further our understanding of the attitudes which might underlie the attrition we observed. Additionally, satisfaction of operability was assessed after 12 weeks of study participation by direct questionnaire. The overall good satisfaction with operability characteristics of the app by the system’s usability scale [21] (Figure 3, assessment of question #7; questionnaire, Multimedia Appendix 6) was confirmed by results from the AttrakDiff subtest *pragmatic quality* (Figure 3, first graph from top, mean scores for opposing word pairs “as indicated, from top to bottom: 0.2, 1.5, 1.2, 0.9, 1.4, 0.8, 1.2; mean overall score for all 7 word pairs =1.0).

**Figure 3.** User ratings assessing operability, hedonic qualities, and attractiveness of mobile health app. Pragmatic, hedonic qualities, and stimulation by app as well as overall attractiveness were assessed by the AttrakDiff test. Additionally, participants were queried by questionnaire for usability and for terms describing their feeling upon app use. Numbers indicate percentages of participants and overall numbers.



Comparing user ratings of hedonic qualities (Figure 3, second graph from top: mean scores for opposing word pairs as indicated, from top to bottom: 0.8, 0.5, 0.9, 0.4, 0.9, 0.6, 1.5; mean overall score for all 7 word pairs=0.8) and stimulation (Figure 3, third graph from top, mean scores for opposing word pairs as indicated, from top to bottom: 1.0, 1.3, 0.6, 1.2, 0.3, -2.5, 1.4, mean overall score for all 7 word pairs=0.5) and overall attractiveness (Figure 3, fourth graph from top, mean for opposing word pairs as indicated, scores from top to bottom: 1.3, 0.9, 1.1, 1.2, 1.2, 1.2, 1.2; mean overall score for all 7 word pairs=1.2), participants attributed more negatively connoted words for hedonic qualities and stimulation. It was most pronounced for the area of stimulation (Figure 3), attributable mainly to the negative ratings for the word pair *harmless* versus *challenging*, where participants viewed the app strongly *harmless* rather than *challenging*. Overall attractiveness was perceived most positively compared with the 3 other areas (Figure 3).

**Patients Feel Supported by Mobile Health App Use**

Upon completion of the 3-month observation period, we questioned the perceived usefulness of the app in supporting better therapeutic adherence, trustworthiness, motivation, promotion of safety, the informative quality and support by the app, boredom, perceived control, and paternalistic quality

exerted by the app (Figure 3, questions #7-14, questionnaire Multimedia Appendix 7). Again, not all participants completed all items, leading to 23 to 25 analyzable items, as indicated in Figure 3. The highest positive score was noted for feeling supported (17/25, 68%), followed by a feeling of being informed by the app (16/25, 64%), and a feeling of increased safety (15/25, 60%). Motivation (13/25, 52%), trust (14/25, 56%), and usefulness of the app for therapy adherence (14/25, 56%) were also perceived positively by more than half of the study participants. Fewer patients reported feelings of paternalism (5/24, 20%) or control (10/25, 40%) by the app, or boredom (11/25, 44%). These results suggest that most of the patients perceived the app positively and supportive of therapy adherence.

**Mobile Health App Use Fails to Modulate Weighted Satisfaction With Life and Health**

To obtain results on our study populations' satisfaction with life and health, we assessed the 3 FLZ modules *general life* (questionnaire Multimedia Appendix 2), *health life* (questionnaire Multimedia Appendix 3), and *CF life* (questionnaire Multimedia Appendix 4) in our study group (aged 12-24 years) before and after the mHealth app intervention. The sum scores for different age subgroups of our study group (12-15 years vs 16-24 years) for the 3 domains *general life*, *health life*, and *CF life* revealed no statistically important differences from

the results we observed for the overall study population of 12 to 24-year-olds (Table 1).

Our patients showed higher sum scores than the published reference population for weighted satisfaction in all 3 modules at the beginning and end of the observation period, particularly for general life and CF life. Similarly, our study population received higher scores in several subdomain scores. Comparison by Hedge  $g$  values supports these conclusions, as these attained medium to large effect sizes for most parameters calculated (Table 2). We attribute these differences to (1) inclusion of older patients who have worse health status, since CF is a chronic

progressive disease, and (2) the fact that this cohort was sampled >15 years ago [26], when CF patients had worse health status in general. Calculating  $P$  values for the observed differences never revealed  $P < .05$ , most likely because of the large standard deviations of the sum scores, and the subitem scores of the FLZ questionnaires [26]. However, the differences in these scores between the study population and the reference population were not a major research question of the study we present here, as we aimed to focus on changes induced by the usage of the mHealth app. In that respect, changes in the difference between our study population and the reference population might indicate effects of mHealth app usage, though, as we elaborate below.

**Table 1.** Questionnaire on life satisfaction (FLZ) sum scores weighted satisfaction for the modules general life, health life, and Cystic Fibrosis life for the complete study group aged 12 to 24 years, and subgroups of 12 to 25 years and 16 to 24 years.

Modules	Study population (age 12-24 years), sum scores		Study population (age 12-15 years), sum scores		Study population (age 16-24 years), sum scores	
	At inclusion	After completion	At inclusion	After completion	At inclusion	After completion
General life	60.04	58.4	62.33	63.6	58.57	54.09
Health life	73.86	72.38	79.89	74.5	70	70.45
Cystic fibrosis life	71.52	67.42	84.6	75.2	63.07	60.36

**Table 2.** Questionnaire on life satisfaction (FLZ) sum scores weighted satisfaction for the modules general life, health life, and Cystic Fibrosis (CF) life, the health life subitems ability to relax and audition/vision and the CF life subitems feeling of being needed/appreciated, understanding/integration of therapy, and lack of handicap by CF of our study population.

Modules and subitems	Reference population (age 16-45 years, n=251), sum scores	Study population (age 12-24 years, n=25), sum scores		Hedge $g$ value <sup>a</sup>		$P$ value <sup>a</sup>	
		At inclusion	After completion	At inclusion	After completion	At inclusion	After completion
<b>Modules</b>							
General life	46	60.04	58.4	0.44	0.38	.05	.09
Health life	61.67	73.86	72.38	0.30	0.25	.11	.17
CF <sup>b</sup> life	56.01	71.52	67.42	0.41	0.28	.08	.15
<b>Subitem for the module health life</b>							
Ability to relax	4.84	7.26	7.25	0.51	0.40	.02	.05
<b>Subitems for the module CF life</b>							
Respiration	4.56	6.5	8.1	0.24	0.44	.08	.05
Feeling of being needed/appreciated	5.63	9	8.3	0.61	0.48	.02	.12
Lack of handicap by CF	4.73	9.96	9.4	0.74	0.67	.02	.02
Understanding/integration of therapy	5.92	8.13	7.7	0.35	0.29	.06	.12

<sup>a</sup>Reference population versus study population.

<sup>b</sup>CF: cystic fibrosis.

Hedge  $g$  values were calculated to determine the effect size of the intervention by comparing the designated sum scores of the reference population versus the scores of our study population at inclusion and after completion.  $P$  values were calculated with a 1-sample  $t$  test between the sum scores of our study population (all participants, ie, age 12-24 years) versus the published reference population of the FLZ questionnaire (age 16-45 years).

The sum scores for *general life* remained the same during the observation period (60.04 vs 58.4 points), whereas the sum scores for *health life* and *CF life* declined during the observation period (73.86 vs 72.38 points and 71.52 vs 67.42 points, respectively, Table 2). These changes failed to attain statistical significance, suggesting that the mHealth app use did not modulate these general domains.



Compared with the published reference population, the *health life subitem ability to relax* was significantly higher in our study population at the beginning of the observation period, but not after the intervention (Table 2; *P* value .02 vs .05). In the domain *CF life*, the subitems *feeling of being needed/appreciated* and *lack of handicap by CF* were significantly higher at the beginning of the observation period, compared with the published reference population (*P* value .02 and .02, respectively), which remained significant after the study period for the subitem *lack of handicap by CF* (*P* value .01), suggesting stabilization of this subitem by the app use. The CF-specific subitem *respiration* was improved in our study population after the intervention, compared with values obtained by our study population before app use was assessed (6.5 points before intervention and 8.1 after intervention; Table 2). This increase was not statistically significant for our study population, yet the score was significantly better before and after the intervention, compared with the published reference population (4.56 points reference population vs 8.1 study population before intervention; *P* value .01; 4.56 points reference population vs 6.5 points study population after intervention; *P* value .05; Table 2), again suggesting stabilization of this value by the app use. The subitem *understanding/integration of therapy*, however, was not

significantly different from the published reference population either before or after study participation (Table 2).

### Lung Function and Body Mass Index Are Not Affected by Mobile Health App Use

Upon inclusion, the study participants had a mean age of 16 (SD 3) years, a mean FEV1 of 84% (SD 25%), and a mean BMI of 20 (SD 3) kg/m<sup>2</sup>. The control group had a mean age of 15 (SD 3) years, a mean FEV1 of 85% (SD 22%), and a mean BMI of 19 (SD 3) kg/m<sup>2</sup>, which did not differ significantly from the study participants (Table 3). Comparisons of our study group with the matched control group by *P* value or Cohen *d* value to gauge effect sizes of the intervention revealed no differences at 3 months (mean FEV1 for both groups 83%, SD 25% and 22%; Table 3), 1 year (mean FEV1 87%, SD 24% vs 86% SD 23%; Table 3), or 2 years (mean FEV1 79%, SD 25% vs 81%, SD 27%; Table 3). Comparison of the study group and the control group revealed a small effect size by Cohen *d* value on BMI directly after the study (study group 20 kg/m<sup>2</sup>, SD 3 kg/m<sup>2</sup>; control group 19 kg/m<sup>2</sup>, SD 3 kg/m<sup>2</sup>; Cohen *d* value 0.33) but no differences as per Cohen *d* values or *P* values at the other time points, indicating no consistent effect on somatic parameters.

**Table 3.** Age, forced expiratory volume in 1 second (% Knudson), and body mass index at inclusion of study subjects, and matched controls, 3 months, 1 year, and 2 years post inclusion.

Controls	Study subjects, mean (SD)	n <sup>a</sup>	Matched control group, mean (SD)	n <sup>a</sup>	<i>P</i> value	Cohen <i>d</i> value <sup>b</sup>
<b>Prestudy/inclusion (D0)</b>						
Age (years)	16 (3)	25	15 (3)	25	.93	— <sup>c</sup>
FEV1 (%)	84 (25)	25	85 (22)	25	.85	—
BMI <sup>d</sup> (kg/m <sup>2</sup> )	20 (3)	25	19 (3)	25	.37	—
<b>Poststudy (D0 plus 3 months)</b>						
FEV1 (%)	83 (23)	25	83 (23)	25	.73	0
BMI (kg/m <sup>2</sup> )	20 (3)	24	19 (3)	25	.42	0.33
<b>1-year follow-up (D0 plus 12 months)</b>						
FEV1 (%)	87 (24)	25	86 (23)	25	.91	0.042
BMI (kg/m <sup>2</sup> )	20 (3)	23	20 (2)	25	.77	0
<b>2-year follow-up (D0 plus 24 months)</b>						
FEV1 (%)	79 (25)	21	81 (27)	24	.38	0.08
BMI (kg/m <sup>2</sup> )	20 (3)	21	20 (2)	23	>.99	>.99

<sup>a</sup>Number of available measurements.

<sup>b</sup>Study subjects versus matched control group.

<sup>c</sup>Not applicable.

<sup>d</sup>BMI: body mass index.

Numbers of analyzable data points changed throughout the study because of missing values stemming from routine clinical care. Mean values were calculated for percentage FEV1 and BMI across the study group versus the control group. *P* values were calculated with a 1-sample *t* test between the mean scores of our study group (n=25) versus the control group (n=25).

Cohen *d* values were calculated to determine the effect size of the intervention by comparing the FEV1 and BMI values between the study group and the matched control group at the given time points.

## Discussion

### Principal Findings

The medication plan and reminder were the functions used most frequently and perceived most useful (Figures 1 and 2). The diary function and communication venue were used much less frequently, with a larger discrepancy between actual use and perceived usefulness (Figure 1). Finally, the design function was used most infrequently and perceived as of little use. These results might reflect the fact that the desire for independent self-management typical of adolescents was indeed addressed by the KiOAPP medication plan and reminder. Country-specific incorporation of a bar code scanning function, which allows quick incorporation of a large number of different medications typical for CF into the medication plan—an asset for a patient population under chronic time shortage—might have particularly supported the use of this app function. Lesser use of the diary and the communication functions could have been influenced by the fact that the adolescents did not consider these support measures important or that the app did not meet their demands; the discrimination of which probably needs a more qualitative approach to improve our understanding of which measures might further improve disease investment and communication for this age group.

The results indicate that the mHealth app is perceived attractive, useful, and supportive. Nevertheless, upon evaluation over time, we observed attrition in use, similar to other studies [9,10], and, more pronounced, low willingness for continued app usage (Figures 1 and 2). We can only speculate about the reasons for the sharp decline in continued willingness to use the app. Perceived usefulness did not match actual usage of some of the app's features, possibly because of competing support measures (ie, well-established phone and email contacts in place at our CF center vs the new app contact possibility). Parental (over-) involvement, which did not leave room for adolescents' involvement, and the well-known developmental challenge of low disease involvement during adolescence [4-6] might have also contributed. Finally, social desirability might have also played a role in the questionnaire responses, leading to overreporting of the perceived usefulness of certain features, with those features the participants actually found supportive likely achieving better matches of perceived usefulness and actual use. Moreover, the least popular function of the app, that is, the design function, only contains 3 options, which might have contributed to its low attractiveness, corroborated by the low ratings in hedonic and *stimulation* qualities we identified via the AttrakDiff questionnaires.

Although we initially thought that the design function could improve adolescents' acceptance and sustained use of the app, the unfavorable usage characteristics of this function suggest that the KiOAPP addressed these aspects only insufficiently. By assessing opposing word pairs in the AttrakDiff questionnaires, we identified the domain of hedonic qualities and *stimulation* as possible areas to address when aiming to reduce attrition. In the area of hedonic qualities, the AttrakDiff assesses the construct *self-identification* with the product assessed. This area achieved lower ratings compared with

pragmatic qualities and overall attractiveness (Figure 3). The least positive scores in the 3 subdomains of the word pair test of the AttrakDiff questionnaire were achieved in the area *stimulation*, particularly because of the app being perceived as *harmless* rather than *challenging* (Figure 3). On those lines, *challenges* might improve the motivation of our target population to continue using the app. Our results suggest that adding *gamification* approaches [28], more sophisticated design functions improving *stimulation*, or chat functions improving *self-identification* might be useful to increase app use adherence.

Finally, our results suggest that use of the mHealth app might stabilize satisfaction with the CF-specific item *lack of handicap by CF* and *respiration* in our study group (Table 2), but that it nevertheless has no consistent effect on the stringent clinical parameters FEV1 or BMI (Table 3). These exploratory findings are only insufficiently addressed by our approach and, therefore, do not allow definite conclusions on these interactions but instead require more detailed follow-up studies to address the complex interplay of disease support, disease investment, adherence, quality of life, and somatic function.

### Limitations

Our study and comparisons have several limitations. Methodological approaches such as the choice of questionnaires to assess usage characteristics and user experience, as well as the lack of randomization and blinding possibilities, in addition to recruitment difficulties and the short observation period, might also have precluded the identification of more significant effects by the app use.

The quantitative questionnaires we used might not be ideal to assess user experiences in the detail needed to explain some of the discrepancies of our results, for example, concerning the perceived usefulness and actual use of the app's features (Figure 2), where qualitative approaches, such as structured interviews, might be more informative. On the other hand, more objective measures, such as electronic usage read-outs, might provide more reliable usage characteristics than the patients' reported usage characteristics we chose to assess [9]. However, such an approach raises particular ethical concerns as it deeply affects patient privacy, possibly affecting recruitment possibilities, and this was therefore not considered worthwhile exploring at this early stage of the app evaluation. Also, about study design, the difficulties in recruitment precluded random assignment of patients willing to participate in the study into 2 groups (intervention vs no intervention), which might have induced a selection bias in participants. Refining our methodological approach in these areas could be important tools to improve results and reduce ambiguity of results in a future clinical trial.

Only one-sixth of the 150 eligible participants could be recruited into the study, despite intensive recruitment efforts, including hiring of additional personnel and a patient information flyer sent and handed out to all eligible patients. This strongly supports the well-known fact that adolescents and young adults are particularly challenging to draw into clinical studies and to motivate for higher disease investment [5,7], the primary reason for our study. The reasons for this low motivation are most likely multifactorial and correspond to similar findings by others pertaining to CF patients [29], with the overall difficulty in

motivation of this particular age group figuring prominently among possible causes [5], potentially even more pronounced in those affected by chronic diseases [4-6] because of a strong desire for normality [5,7]. Numerous competing studies, as well as numerous competing mHealth apps, might have also contributed.

Use of an mHealth app is not a blinded intervention. We chose a case-control design for the lung function and BMI comparisons as an attempt to arrive at meaningful comparisons despite the lack of blinding. Although for the FLZ questionnaires, assessing these questionnaires in an additional cohort at our center was not possible because of the amount of time and personnel necessary, as these questionnaires are not part of the clinical routine follow-ups as in the case of FEV1 and BMI. It left us with a reference population which was not age-matched, and thus we refrain from a comparison with the reference population for the changes we observed.

The short observation period, small study size population, and overall good health of our participants render a measurable improvement or stabilization of FEV1, BMI, and weighted life satisfaction due to any intervention difficult to attain [20]. To identify lung function improvements within a CF patient population with a well-preserved FEV1 as we had included, more sensitive lung function measurements, such as wash-out measurements (lung clearance index) [20], might be necessary, which, unfortunately for the age group we recruited, had not yet been implemented in everyday practice at our center at the time of the study. These shortcomings forego any conclusions about the effect of the mHealth app on somatic parameters, for which larger and longer interventional trials with more sensitive lung function methods are needed for this particular age group.

### Comparison With Prior Work

An improved version of the app might serve to improve app adherence. We chose the app we evaluated as we saw its functionalities address several aspects which might improve adherence in our target population. Still, several aspects might have influenced user affinity, and thereby sustained use, and could have thus impacted negatively on the support potential of the KiOAPP.

User-centered approaches to delineate CF-specific needs for mHealth app development suggest that a multifunctional app incorporating different CF-specific functions such education, enzyme dosage calculation, nutrition management, treatment organization, health diary, treatment follow-up, and practical guidelines for treatment [11], in addition to individuality and adaptability, is a key aspect agreed upon by CF patients [30]. Some freely available CF-specific apps [13,16], as well as current app developments [12,13,15,31], include CF-specific functions identified by a user-centered approach, which has been suggested to be crucial to improve acceptance and reduce attrition of app use [32,33]. The app we evaluated lacks CF-specific features which could reduce attrition. In a free-text item in one of our questionnaires, we did receive feedback that CF-specific features, such as a fat calculator or the possibility to receive reminders for physiotherapy and sports, were considered desirable additions to the existing app. However,

the number of answers to the free-text item were too few (n=3 total free-text answers) to address them statistically.

Communication with doctors and peers has also been identified as desirable features by others [30]. Considering that some of the cited studies included a significantly older study population compared with our study [11] and user preferences might reflect age-specific needs concerning mHealth apps [9,17], we suggest including adolescent-specific features into existing disease-specific apps to improve mHealth attachment for this difficult-to-attract age group. Our results on perceived product qualities give indications that, apart from the incorporation of CF-specific needs, additional (age-specific) adaptations, including the incorporation of *hedonistic* features typically desired by adolescents and young adults, might constitute a promising approach to provide adolescent-centered assets to improve app attachment in this age group. While the contribution of hedonistic features is not yet fully understood, this is a notion also supported in the literature (see for example the review presented by Diefenbach et al [34]).

Changes in quality of life have been perceived useful to assess effects of medical interventions in CF, as studies have shown that increases in adherence correlate with a more optimistic view of life in CF [33,35]. Improvements in weighted life satisfaction, measured by the FLZ questionnaire, could have thus indicated improved adherence by mHealth app use. However, as we did not observe such effects and, more importantly, did not include a direct measure of adherence, we cannot make such conclusions on this complex interaction.

Considering that the patient population between 12 and 29 years of age constitutes the largest patient population of Germany's current CF population [1], our results apply to a significant CF patient number and are probably also generalizable to CF patients in that age group beyond Germany, as CF care is very streamlined worldwide. While patients with other chronic diseases face similar challenges during adolescence, due to our results and those of others concerning disease-specific features of mHealth apps, advocating the development of user-centered approaches [32,33], we caution, however, against generalizing our results to the larger population of adolescents affected by chronic disease.

### Conclusions and Outlook

Low contact availability and disease involvement are challenges all physicians face when caring for adolescent patients with chronic diseases [4-6]. We had postulated that offering support in these areas via the mobile phone might improve self-management for this target age, given their high mobile phone affinity. We must acknowledge that at least in the current form, the app we used might not be sufficient to increase disease involvement of adolescents. However, given the relationship of life satisfaction, optimism, and adherence [6,35], we suggest that an improved version of the mHealth app might constitute a promising tool to improve medication and therapy adherence (via the medication reminder function), disease management (via a fat calculator), and disease investment (via an improved diary function, including at-home lung function monitoring), ultimately improving life satisfaction and also clinical well-being.

Our results suggest that only an improved mHealth app, along with an improved clinical trial design, could add proof to the hypothesis that mHealth apps can constitute a suitable tool to improve self-management and therapy adherence in adolescent patients with CF.

---

## Acknowledgments

The authors thank all patients, their parents, and the nonmedical personnel from the Christiane-Herzog-Transitionsambulanz at Hannover Medical School for supporting this study. This work was supported by the Christiane-Herzog-Foundation, the Vertex Circle-of-Care-Program, and the German Center for Lung Research.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Functions of the mHealth app (KiOAPP).

[PDF File (Adobe PDF File), 241 KB - [mhealth\\_v7i11e12442\\_app1.pdf](#) ]

---

### Multimedia Appendix 2

FLZ module on general life satisfaction.

[PDF File (Adobe PDF File), 309 KB - [mhealth\\_v7i11e12442\\_app2.pdf](#) ]

---

### Multimedia Appendix 3

FLZ module on health life satisfaction.

[PDF File (Adobe PDF File), 285 KB - [mhealth\\_v7i11e12442\\_app3.pdf](#) ]

---

### Multimedia Appendix 4

FLZ module on CF life satisfaction.

[PDF File (Adobe PDF File), 285 KB - [mhealth\\_v7i11e12442\\_app4.pdf](#) ]

---

### Multimedia Appendix 5

Self-developed questionnaire III concerning user behavior used in telephone interviews 4 and 8 weeks after introduction to the mHealth app.

[PDF File (Adobe PDF File), 407 KB - [mhealth\\_v7i11e12442\\_app5.pdf](#) ]

---

### Multimedia Appendix 6

Questionnaire I (SUS) concerning user behavior queried after 12 weeks of mobile health app use.

[PDF File (Adobe PDF File), 443 KB - [mhealth\\_v7i11e12442\\_app6.pdf](#) ]

---

### Multimedia Appendix 7

Self-developed questionnaire II concerning user behavior queried after 12 weeks of mHealth app use.

[PDF File (Adobe PDF File), 239 KB - [mhealth\\_v7i11e12442\\_app7.pdf](#) ]

---

## References

1. Naehrlich L, Burkhard M, Wosniok J, Working Group. Mukoviszidose eV - Federal Association Cystic Fibrosis. 2018. German Cystic Fibrosis Registry - Annual Report 2017 URL: [https://www.muko.info/fileadmin/user\\_upload/angebote/qualitaetsmanagement/register/berichtsbaende/annual\\_report\\_2017.pdf](https://www.muko.info/fileadmin/user_upload/angebote/qualitaetsmanagement/register/berichtsbaende/annual_report_2017.pdf) [accessed 2019-09-27]
2. Kerem E, Conway S, Elborn S, Heijerman H, Consensus Committee. Standards of care for patients with cystic fibrosis: a European consensus. *J Cyst Fibros* 2005 Mar;4(1):7-26 [FREE Full text] [doi: [10.1016/j.jcf.2004.12.002](https://doi.org/10.1016/j.jcf.2004.12.002)] [Medline: [15752677](https://pubmed.ncbi.nlm.nih.gov/15752677/)]
3. Jennings MT, Riekert KA, Boyle MP. Update on key emerging challenges in cystic fibrosis. *Med Princ Pract* 2014;23(5):393-402 [FREE Full text] [doi: [10.1159/000357646](https://doi.org/10.1159/000357646)] [Medline: [24434297](https://pubmed.ncbi.nlm.nih.gov/24434297/)]
4. Sawyer S, Afifi RA, Bearinger LH, Blakemore S, Dick B, Ezech AC, et al. Adolescence: a foundation for future health. *Lancet* 2012 Apr 28;379(9826):1630-1640. [doi: [10.1016/S0140-6736\(12\)60072-5](https://doi.org/10.1016/S0140-6736(12)60072-5)] [Medline: [22538178](https://pubmed.ncbi.nlm.nih.gov/22538178/)]
5. Nasr SZ. Cystic fibrosis in adolescents and young adults. *Adolesc Med* 2000 Oct;11(3):589-603. [Medline: [11060556](https://pubmed.ncbi.nlm.nih.gov/11060556/)]
6. Gudas L, Koocher GP, Wypij D. Perceptions of medical compliance in children and adolescents with cystic fibrosis. *J Dev Behav Pediatr* 1991 Aug;12(4):236-242. [doi: [10.1097/00004703-199108000-00004](https://doi.org/10.1097/00004703-199108000-00004)] [Medline: [1939678](https://pubmed.ncbi.nlm.nih.gov/1939678/)]

7. Bishay L, Sawicki GS. Strategies to optimize treatment adherence in adolescent patients with cystic fibrosis. *Adolesc Health Med Ther* 2016;7:117-124 [FREE Full text] [doi: [10.2147/AHMT.S95637](https://doi.org/10.2147/AHMT.S95637)] [Medline: [27799838](https://pubmed.ncbi.nlm.nih.gov/27799838/)]
8. Feierabend S, Rathgeb TP. mpfs. 2015. JIM 2015 Jugend, Information, (Multi-) Media URL: [https://www.mpfs.de/fileadmin/files/Studien/JIM/2015/JIM\\_Studie\\_2015.pdf](https://www.mpfs.de/fileadmin/files/Studien/JIM/2015/JIM_Studie_2015.pdf) [accessed 2019-09-27]
9. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. *J Med Internet Res* 2015 Dec 23;17(12):e287 [FREE Full text] [doi: [10.2196/jmir.5043](https://doi.org/10.2196/jmir.5043)] [Medline: [26701961](https://pubmed.ncbi.nlm.nih.gov/26701961/)]
10. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. *J Med Internet Res* 2016 May 16;18(5):e97 [FREE Full text] [doi: [10.2196/jmir.4883](https://doi.org/10.2196/jmir.4883)] [Medline: [27185295](https://pubmed.ncbi.nlm.nih.gov/27185295/)]
11. Hilliard M, Hahn A, Ridge AK, Eakin MN, Riekert KA. User preferences and design recommendations for an mHealth app to promote cystic fibrosis self-management. *JMIR Mhealth Uhealth* 2014 Oct 24;2(4):e44 [FREE Full text] [doi: [10.2196/mhealth.3599](https://doi.org/10.2196/mhealth.3599)] [Medline: [25344616](https://pubmed.ncbi.nlm.nih.gov/25344616/)]
12. MyCyFapp. 2018. A Project for Patients URL: <http://www.mycyfapp.eu/index.php/de/> [accessed 2018-08-26] [WebCite Cache ID 71xgVRR8C]
13. Google Play. 2018. CF MedCare Reminder App URL: <https://play.google.com/store/apps/details?id=nl.cone.cfmedcare&hl=en> [accessed 2018-08-26] [WebCite Cache ID 71xgOCsFw]
14. App Store - Apple. My Fight Against Cystic Fibrosis URL: <https://itunes.apple.com/de/app/my-fight-against-cystic-fibrosis/id918730808?mt=8> [accessed 2018-08-26] [WebCite Cache ID 71xhH79Nu]
15. Clinical Trials. 2018. CFfone: A Cell Phone Support Program for Adolescents With Cystic Fibrosis URL: <https://clinicaltrials.gov/ct2/show/NCT01183286> [accessed 2018-08-26] [WebCite Cache ID 71xgbZuRE]
16. Al-Saleh H, Kaufman M, Ratjen F. Beta Testing Of The My CF Smartphone/Tablet App. *American Journal of Respiratory and Critical Care Medicine* 2014;189:A6673.
17. Becker S, Brandl C, Meister S, Nagel E, Miron-Shatz T, Mitchell A, et al. Demographic and health related data of users of a mobile application to support drug adherence is associated with usage duration and intensity. *PLoS One* 2015;10(1):e0116980 [FREE Full text] [doi: [10.1371/journal.pone.0116980](https://doi.org/10.1371/journal.pone.0116980)] [Medline: [25629939](https://pubmed.ncbi.nlm.nih.gov/25629939/)]
18. Kinderhilfe Organtransplantation. URL: <http://app.kiohilfe.de/> [accessed 2018-08-26] [WebCite Cache ID 71xg6mtCJ]
19. Kinderhilfe Organtransplantation. 2014. Press Release URL: <https://www.kiohilfe.de/presse/artikel/85> [accessed 2019-06-06] [WebCite Cache ID 78vhtjwBK]
20. Rosenfeld M. An overview of endpoints for cystic fibrosis clinical trials: one size does not fit all. *Proc Am Thorac Soc* 2007 Aug 1;4(4):299-301. [doi: [10.1513/pats.200611-178HT](https://doi.org/10.1513/pats.200611-178HT)] [Medline: [17652489](https://pubmed.ncbi.nlm.nih.gov/17652489/)]
21. Brooke J. SUS-A quickdirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. *Usability Evaluation in Industry*. London: Taylor and Francis; 1996.
22. AttrakDiff Questionnaire. URL: <http://attrakdiff.de> [accessed 2018-08-26] [WebCite Cache ID 71xgE8tYu]
23. Hassenzahl M, Koller F, Burmester M. Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. In: *Mensch & Computer*. Germany: Vieweg & Teubner; 2003:187-196.
24. Hassenzahl MB, Koller F, Burmester M. Attrakdiff. 2008. Der User Experience (UX) auf der Spur: Zum Einsatz von [www.attrakdiff.de](http://www.attrakdiff.de) URL: [http://attrakdiff.de/files/up08\\_ux\\_auf\\_der\\_spur.pdf](http://attrakdiff.de/files/up08_ux_auf_der_spur.pdf) [accessed 2019-09-27]
25. Hassenzahl M. The interplay of beauty, goodness, and usability in interactive products. *Hum Comput Interact* 2004;19(4):319-349 [FREE Full text] [doi: [10.1207/s15327051hci1904\\_2](https://doi.org/10.1207/s15327051hci1904_2)]
26. Goldbeck L, Schmitz TG, Henrich G, Herschbach P. Questions on life satisfaction for adolescents and adults with cystic fibrosis: development of a disease-specific questionnaire. *Chest* 2003 Jan;123(1):42-48. [doi: [10.1378/chest.123.1.42](https://doi.org/10.1378/chest.123.1.42)] [Medline: [12527601](https://pubmed.ncbi.nlm.nih.gov/12527601/)]
27. Sherrill D, Lebowitz MD, Knudson RJ, Burrows B. Continuous longitudinal regression equations for pulmonary function measures. *Eur Respir J* 1992 Apr;5(4):452-462. [Medline: [1563504](https://pubmed.ncbi.nlm.nih.gov/1563504/)]
28. Deterding S, Dixon D, Khaled R, Nacke LE. From Game Design Elements to Gamefulness: Defining "Gamification". In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. 2011 Presented at: MindTrek'11; September 28-30, 2011; Tampere, Finland.
29. Roehrer E, Cummings E, Beggs S, Turner P, Hauser J, Micallef N, et al. Pilot evaluation of web enabled symptom monitoring in cystic fibrosis. *Inform Health Soc Care* 2013 Dec;38(4):354-365. [doi: [10.3109/17538157.2013.812646](https://doi.org/10.3109/17538157.2013.812646)] [Medline: [23957685](https://pubmed.ncbi.nlm.nih.gov/23957685/)]
30. Floch J, Zettl A, Fricke L, Weisser T, Grut L, Vilarinho T, et al. User needs in the development of a health app ecosystem for self-management of cystic fibrosis: user-centered development approach. *JMIR Mhealth Uhealth* 2018 May 8;6(5):e113 [FREE Full text] [doi: [10.2196/mhealth.8236](https://doi.org/10.2196/mhealth.8236)] [Medline: [29739742](https://pubmed.ncbi.nlm.nih.gov/29739742/)]
31. Calvo-Lerma J, Martinez-Jimenez CP, Lázaro-Ramos JP, Andrés A, Crespo-Escobar P, Stav E, MyCyFapp. Innovative approach for self-management and social welfare of children with cystic fibrosis in Europe: development, validation and implementation of an mHealth tool (MyCyFAPP). *BMJ Open* 2017 Mar 16;7(3):e014931 [FREE Full text] [doi: [10.1136/bmjopen-2016-014931](https://doi.org/10.1136/bmjopen-2016-014931)] [Medline: [28302638](https://pubmed.ncbi.nlm.nih.gov/28302638/)]

32. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. *Inform Health Soc Care* 2008 Sep;33(3):158-169. [doi: [10.1080/17538150802457562](https://doi.org/10.1080/17538150802457562)] [Medline: [18850399](https://pubmed.ncbi.nlm.nih.gov/18850399/)]
33. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. *Biomed Instrum Technol* 2012;Suppl:49-56. [doi: [10.2345/0899-8205-46.s2.49](https://doi.org/10.2345/0899-8205-46.s2.49)] [Medline: [23039777](https://pubmed.ncbi.nlm.nih.gov/23039777/)]
34. Diefenbach S, Kolb N, Hassenzahl M. The 'Hedonic' in Human-Computer Interaction. In: Proceedings of the 2014 conference on Designing interactive systems. 2014 Presented at: DIS'14; June 21-25, 2014; Vancouver, BC, Canada.
35. Gee L, Abbott J, Hart A, Conway SP, Etherington C, Webb AK. Associations between clinical variables and quality of life in adults with cystic fibrosis. *J Cyst Fibros* 2005 Mar;4(1):59-66 [FREE Full text] [doi: [10.1016/j.jcf.2004.12.005](https://doi.org/10.1016/j.jcf.2004.12.005)] [Medline: [15752683](https://pubmed.ncbi.nlm.nih.gov/15752683/)]

## Abbreviations

**BMI:** body mass index

**CF:** cystic fibrosis

**FEV1:** forced expiratory volume in 1 second

**FLZ:** questionnaire on life satisfaction (Fragebogen zur Lebenszufriedenheit)

**mHealth:** mobile health

*Edited by G Eysenbach; submitted 07.10.18; peer-reviewed by J Floch, X Ji, M Eakin; comments to author 03.02.19; revised version received 25.03.19; accepted 21.07.19; published 21.11.19.*

*Please cite as:*

*Rudolf I, Pieper K, Nolte H, Junge S, Dopfer C, Sauer-Heilborn A, Ringshausen FC, Tümmler B, von Jan U, Albrecht UV, Fuge J, Hansen G, Dittrich AM*

*Assessment of a Mobile App by Adolescents and Young Adults With Cystic Fibrosis: Pilot Evaluation*

*JMIR Mhealth Uhealth* 2019;7(11):e12442

URL: <https://mhealth.jmir.org/2019/11/e12442>

doi: [10.2196/12442](https://doi.org/10.2196/12442)

PMID: [31750841](https://pubmed.ncbi.nlm.nih.gov/31750841/)

©Isa Rudolf, Katharina Pieper, Helga Nolte, Sibylle Junge, Christian Dopfer, Annette Sauer-Heilborn, Felix C Ringshausen, Burkhard Tümmler, Ute von Jan, Urs-Vito Albrecht, Jan Fuge, Gesine Hansen, Anna-Maria Dittrich. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 21.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Mood Monitoring Over One Year for People With Chronic Obstructive Pulmonary Disease Using a Mobile Health System: Retrospective Analysis of a Randomized Controlled Trial

Maxine E Whelan<sup>1</sup>, BSc, PhD; Carmelo Velardo<sup>2</sup>, MSc, PhD; Heather Rutter<sup>1</sup>, BSc; Lionel Tarassenko<sup>2</sup>, MA, DPhil, FREng, FMedSci; Andrew J Farmer<sup>1</sup>, FRCGP, DM

<sup>1</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

<sup>2</sup>Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom

**Corresponding Author:**

Maxine E Whelan, BSc, PhD  
Nuffield Department of Primary Care Health Sciences  
University of Oxford  
Radcliffe Observatory Quarter  
Woodstock Road  
Oxford, OX2 6GG  
United Kingdom  
Phone: 44 7435330992  
Email: [maxine.whelan@phc.ox.ac.uk](mailto:maxine.whelan@phc.ox.ac.uk)

## Abstract

**Background:** Comorbid anxiety and depression can add to the complexity of managing treatment for people living with chronic obstructive pulmonary disease (COPD). Monitoring mood has the potential to identify individuals who might benefit from additional support and treatment.

**Objective:** We used data from the sElf-management and support proGrammE (EDGE) trial to examine: (1) the extent to which the mood-monitoring components of a mobile health system for patients with COPD were used by participants; (2) the levels of anxiety and depression symptoms among study participants; (3) the extent to which videos providing advice about coping with low mood were viewed; and (4) the characteristics of participants with differing levels of mood and utilization of mood monitoring.

**Methods:** A total of 107 men and women with a clinical diagnosis of COPD, aged  $\geq 40$  years old, were recruited to the intervention arm of the EDGE trial. Participants were invited to complete the Patient Health Questionnaire-8 and the Generalized Anxiety Disorder-7 test every four weeks using a tablet computer. Mood disturbance based on these measures was defined as a score  $\geq 5$  on either scale. Participants reporting a mood disturbance were automatically directed (signposted) to a stress or mood management video. Study outcomes included measures of health status, respiratory quality of life, and symptoms of anxiety and depression.

**Results:** Overall, 94 (87.9%) participants completed the 12-month study. A total of 80 participants entered at least one response each month for at least ten months. On average, 16 participants (range 8-38 participants) entered  $\geq 2$  responses each month. Of all the participants, 47 (50%) gave responses indicating a mood disturbance. Participants with a mood disturbance score for both scales ( $n=47$ ) compared with those without ( $n=20$ ) had lower health status ( $P=.008$ ), lower quality of life ( $P=.009$ ), and greater anxiety ( $P<.001$ ) and increased depression symptoms ( $P<.001$ ). Videos were viewed by 64 (68%) people over 12 months. Of the 220 viewing visualizations, 70 (34.7%) began after being signposted. Participants signposted to the stress management video (100%; IQR 23.3-100%) watched a greater proportion of it compared to those not signposted (38.4%; IQR 16.0-68.1%;  $P=.03$ ), whereas duration of viewing was not significantly different for the mood management video.

**Conclusions:** Monitoring of anxiety and depression symptoms for people with COPD is feasible. More than half of trial participants reported scores indicating a mood disturbance during the study. Signposting participants to an advisory video when reporting increased symptoms of a mood disturbance resulted in a longer view-time for the stress management video. The opportunity to elicit measures of mood regularly as part of a health monitoring system could contribute to better care for people with COPD.

(JMIR Mhealth Uhealth 2019;7(11):e14946) doi:[10.2196/14946](https://doi.org/10.2196/14946)

**KEYWORDS**

pulmonary disease, chronic obstructive; self-management; telemedicine; computers; handheld; anxiety; depression

## Introduction

Comorbid anxiety and depression affect as many as 50% of people with chronic obstructive pulmonary disease (COPD) [1,2] and can add to the complexity of managing treatment [3]. Depression and anxiety have been linked with mortality and hospital readmission, respectively [4]. Digital health studies to date involving people with COPD have focused on monitoring physiological variables, including oxygen saturation and pulse rate, symptoms, and identifying a practical daily measurement regimen to support self-management of their condition [5]. There is scope, however, for people with COPD to monitor not only vital signs and symptoms but also mood. Self-reporting mood can result in a greater awareness of one's own symptoms, which could result in several outcomes such as preventing deterioration in feelings of anxiety [6]. The use of digital technologies to support remote patient monitoring has become increasingly feasible for a wide range of health conditions due to the use of tablet computers and other readily available devices [7].

Monitoring mood is most often an active process that involves the manual entry of information, including symptoms of anxiety and depression [8]. Many mood-tracking mobile applications are available and could improve mental health and well-being but lack research evidence [9]. Technologies to self-report mood have provided clinicians and patients with a rich understanding of its variation among people with other conditions, and mental health applications are particularly considered a promising tool to extend mental health care beyond the clinic setting [10]. It is anticipated that the act of self-monitoring mood may reduce symptoms of anxiety and depression by increasing emotional self-awareness [11]. Bipolar disorder offers an exemplar where self-monitoring mood is prevalent [12], with self-monitoring mood on a daily basis being described as being of little or no inconvenience in people with bipolar disorder [13]. Digital technologies can be used to report measures of mood responses with ease. To date, there have been a number of such studies with bipolar patients, and findings suggest the time taken to complete mood questions is short [14], with rates of completion ranging from 73.4-100% [14-17].

We are not aware of any published studies involving digital health approaches to self-monitoring of mood in people with COPD. Current nonpharmacological guidelines recommend cognitive behavioral therapy, counselling, and self-help approaches [18]. It is known that untreated comorbid anxiety and depression can have overwhelming consequences on people with COPD [19], such that it impacts their physical, psychological, and social resilience [20]. For instance, depression in patients hospitalized for an acute exacerbation (a severe worsening of COPD symptoms) can adversely affect chances of survival [21,22] and has been associated with longer hospital length of stay, increased symptom burden, and reduced quality of life [23], as well as a reduction in daily step counts [24]. Consequently, monitoring of mood symptoms could be a logical addition to self-management interventions involving

people with COPD; however, it is currently unclear how well people with COPD might adhere to self-monitoring mood and how their symptoms of anxiety and depression may vary over time.

The aim of this analysis, carried out on data obtained as part of a randomized trial which involved mood monitoring within a wider mobile health (mHealth) self-management application [25], was to examine: (1) the extent to which the mood monitoring components of a COPD digital health system were used by participants over 12 months; (2) the levels of anxiety and depression symptoms among study participants; (3) the extent to which videos providing advice about coping with low mood were viewed; and (4) the characteristics of participants with differing levels of mood and utilization of mood self-monitoring.

## Methods

### Overview

A digital health system called self-management and support proGrammE (EDGE) was developed and customized for COPD patients [25]. The system supports self-management for patients by using a COPD symptom questionnaire, a Bluetooth-enabled pulse oximeter, and multimedia content. A component of the system included the collection of self-reported mood data by patients.

### Study Population

The data for this study was drawn from 110 participants allocated to the intervention arm of a 12-month randomized controlled trial of the EDGE COPD system. The trial recruited patients with a confirmed diagnosis of COPD, aged  $\geq 40$  years old, from primary care and community settings in the Thames Valley (United Kingdom), between 2013 and 2015 [25]. Trial eligibility criteria are reported in the protocol [26]. All participants gave their informed consent to take part. Ethics approval was received from the South Central, Berkshire Research Ethics Committee of the UK National Research Ethics Service (Ethics Ref: 12/SC/0437).

### Intervention

Each participant received a low-cost, internet-linked Android tablet computer (Samsung Galaxy Tab II 10.1) with the EDGE mHealth application installed, and an accompanying pulse oximeter (Nonin, Onyx II, 9560BT). All participants had unrestricted access to the tablet computer and pulse-monitoring device for 12 months. Participants were asked to complete a series of standard questions relating to their COPD symptoms daily and a mood questionnaire monthly. The mood questionnaire was completed monthly to reflect clinical practice and even though there are measures to capture daily variations in mood, intensive self-monitoring of mood was not considered necessary in our patient population. After completing the COPD symptom diary, participants were encouraged to wear the pulse oximeter for 30 seconds to record oxygen saturation levels and heart rate.



Participants were asked to use a mood monitoring tool comprising the following mood questions: a 4-item scale which included the Patient Health Questionnaire 2-item measure (PHQ-2) and the General Anxiety Disorder 2-item measure (GAD-2) [27]. These questions were first shown to participants two weeks after they started using the system and then at intervals every four weeks. Participants were prompted to complete the questionnaire via the EDGE system and were offered the opportunity to complete the questionnaire immediately or later. The prompt would remain on the tablet computer until the questionnaire was completed. If the anxiety (GAD-2) or the depression (PHQ-2) measures were scored as  $>1$  point then the relevant full questionnaire (the PHQ-8 or GAD-7) was shown, in line with national guidance [28-30]. PHQ-8 scores of 0-4, 5-9, 10-14, 15-19 and  $\geq 20$  were defined as none, mild, moderate, moderate-severe, and severe, respectively. GAD-7 scores of 0-4, 5-9, 10-14 and  $\geq 15$  were defined as none, mild, moderate, and severe, respectively. Participants were signposted to videos in the event scores of  $\geq 5$  points on either scale were reported by suggesting they might find it helpful to view either the stress or mood management video (or both). Participants were not able to track their mood over time, but if they scored above the  $\geq 5$  threshold at any point (or on multiple occasions) they would be signposted each time to the videos. The stress and mood management videos (with durations of 3 minutes 39 seconds and 1 minute 51 seconds, respectively) were also accessible to participants on the tablet computer at any time. A threshold of  $\geq 90\%$  of the duration of the video was applied to define a video as watched to distinguish them from brief video visualizations ( $<90\%$ ). The tablet computer recorded the day and time of entering data, and the day, time, and duration of viewing videos.

### Study Procedures

Research nurses provided participants with brief details on use of the EDGE system and gave out an information booklet outlining how to charge the devices. Participants were informed that the EDGE system was not a replacement for usual care, and that in the event of any deterioration they should contact their general practitioner (GP) or community respiratory nurse. Data was reviewed by a clinician at no less than 4-day intervals, and if either the PHQ-8 or GAD-7 scores were  $\geq 10$  then the GP was informed by letter.

### Data Collection

PHQ-8 and GAD-7 scores were recorded via the tablet computer. Age, sex, body mass index (BMI), COPD severity using the Global Initiative for Chronic Obstructive Lung Disease (GOLD) staging [31], Forced Expiratory Volume in 1 second ( $FEV_1\%$ ), health status using the EuroQol 5-Dimension Questionnaire (EQ-5D) [32], symptoms of anxiety and depression using the Symptom Checklist (SCL10 and SCL20a, respectively) [33], adherence to taking medications using the Medication Adherence Report Schedule (MARS) [34], Beliefs about Medications Questionnaire (BMQ) [35], respiratory quality of life using the St George's Respiratory Questionnaire for COPD (SGRQ-C) [36], smoking status (current or ex-smoker of  $\geq 2$  years or ex-smoker of  $<2$  years), and number of pack-years were recorded at baseline.

### Data Management

Self-reported mood data was collected via the tablet computer and linked to a study ID. All tablet computers were linked to the Internet by a subscriber identity module (SIM) card. Data was transferred from the EDGE mHealth application to a database held on a secure web server. Anonymized data were extracted for analysis. Participants who withdrew from the study were excluded from analysis.

### Statistical Analysis

Descriptive statistics (frequency, percentage) or mean (SD) or median (IQR) are reported. Comparisons of continuous and categorical variables used a two-tailed independent *t* test and a Chi-square analyses, respectively. The Mann-Whitney U test was completed for continuous variables when data were not normally distributed. The threshold for statistical significance was  $P < 0.05$  for all comparisons in this exploratory analysis.

Responses to the mood questions were assigned to the 4-week period in which they were completed. Some participants completed the mood questions on more than one occasion during each period of 4 weeks; up to three responses for each period were included, and the response with the greatest score was used for that month. If a participant did not respond to the mood questions for a given 4-week period, then this was recorded as a nonresponse.

Characteristics of participants who provided  $\geq 10$  months of responses and participants who provided  $<10$  responses were compared. A second comparison involved participants who recorded a score of  $\geq 5$  for both the GAD-7 and PHQ-8 questionnaires on at least one occasion during the 12 months compared to participants who recorded  $<5$  for the tablet-based mood measures throughout the trial.

A video session was defined as either: (1) happening after being signposted; or (2) watched freely with no prompt to watch. Signposted participants were immediately directed to the relevant video(s) where a score of  $\geq 5$  was recorded. Video visualizations that happened after a participant was signposted were defined as occurring within 30 minutes of entering the mood responses (within the range of observed response); other video visualizations were defined as being watched without being prompted. Participants signposted to the videos who watched at least 90% of its duration were profiled for their subsequent mood response to describe the short-term impact of this intervention.

## Results

### Summary

Only one of the 107 participants allocated to use the EDGE COPD system did not complete any mood questions. During the 12-month study period, 12 (11.3%) participants withdrew from the trial, of whom five gave the response, "too many things were going on" as a reason for stopping, five died, and two did not give a reason. Of these 12 participants, three withdrew before three months, two withdrew before six months, and seven withdrew before 12 months. In total, 94 participants completed

the study. The characteristics of these participants are reported in [Table 1](#).

**Table 1.** Participant characteristics.

Characteristics	Completing study (n=94)
Age (years), mean (SD)	69.7 (9.3)
Male, n (%)	57 (60.6)
<b>BMI<sup>a</sup>, n (%)</b>	
Underweight	3 (5.9)
Normal weight	14 (27.5)
Overweight	16 (31.4)
Obese	18 (35.3)
FEV <sub>1</sub> % <sup>b</sup> , mean (SD)	47.7 (15.9)
<b>GOLD<sup>c</sup> staging, n (%)</b>	
Moderate	37 (39.4)
Severe	40 (42.6)
Very severe	17 (18)
<b>Smoking status, n (%)</b>	
Current smoker	18 (19.2)
Exsmoker (<2 years)	13 (13.8)
Exsmoker (≥2 years)	63 (67)
Years smoking, median (IQR)	40 (29-55)
<b>MRC<sup>d</sup> dyspnea score, n (%)</b>	
2	15 (16)
3	64 (68.1)
4	15 (16)
EQ-5D <sup>e</sup> index, median (IQR)	0.6 (0.5-0.7)
SCL10 <sup>f</sup> , median (IQR)	3 (1-9)
SCL20a <sup>f</sup> , median (IQR)	10.5 (7-23)
SGRQ-C <sup>g</sup> , mean (SD)	56.5 (18.8)
BMQ <sup>h</sup> , mean (SD)	32.6 (6.2)
MARS <sup>i</sup> , median (IQR)	24 (23-25)

<sup>a</sup>BMI: body mass index.

<sup>b</sup>FEV<sub>1</sub> %: Forced Expiratory Volume in 1 second.

<sup>c</sup>GOLD: Global Initiative for Chronic Obstructive Pulmonary Disease.

<sup>d</sup>MRC: Medical Research Council.

<sup>e</sup>EQ-5D: EuroQol 5-Dimension Questionnaire.

<sup>f</sup>SCL: Symptom Checklist.

<sup>g</sup>SGRQ-C: St George's Respiratory Questionnaire for Chronic Obstructive Lung Disease.

<sup>h</sup>BMQ: Beliefs about Medications Questionnaire.

<sup>i</sup>MARS: Medication Adherence Report Schedule.

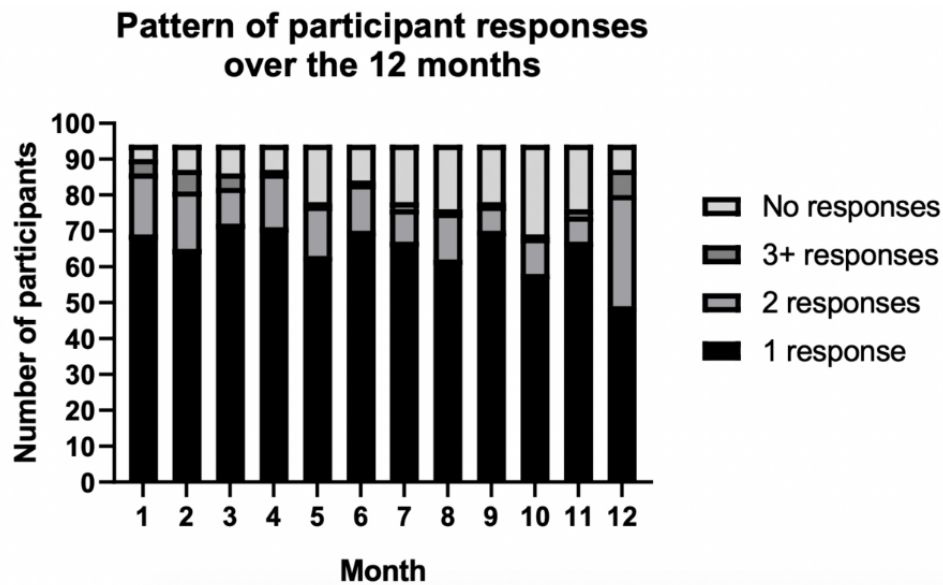
### Use of the Mood-Monitoring Tool

In total, there were 1200 responses to the mood questions. The distribution of responses is shown in [Figure 1](#). The total amount of missing data was 13.4%. There was a decline in response

over time, but 80 (85.1%) participants responded at least once to the mood questionnaire in ≥10 of the 12 months. Smoking status differed between those who used the system in each of the ≥10 months (n=80) compared to those who used the system

in <10 months (n=14) ( $P<.001$ ), with no other differences observed (see [Multimedia Appendix 1](#)).

**Figure 1.** Monthly frequency of responses to the mood questionnaires completed every four weeks.



**Levels of Mood Disturbance Over the Duration of the Study**

In the first month of using the system, 41 (43.6%) participants recorded a PHQ-8 score of  $\geq 5$ , and the total number of participants recording at least one elevated score increased to 68 (72.3%) by 12 months (Figure 2). In comparison, 33 (35.1%)

participants reported a GAD-7 score  $\geq 5$  in the first month, increasing to 53 (56.4%) participants by 12 months (Figure 2). A detailed breakdown of participant responses for both the PHQ-8 and GAD-7 scores, and their respective thresholds of severity for each of the 12 months, is reported in Figure 3 and Figure 4.

**Figure 2.** Cumulative frequency of participants with levels of anxiety and depression above their respective clinically defined thresholds. PHQ-8: Patient Health Questionnaire 8-item measure; GAD-7: General Anxiety Disorder 7-item measure.

**Cumulative frequency of participants reporting mild+ levels of anxiety (GAD-7) and depression (PHQ-8) over 12 months**

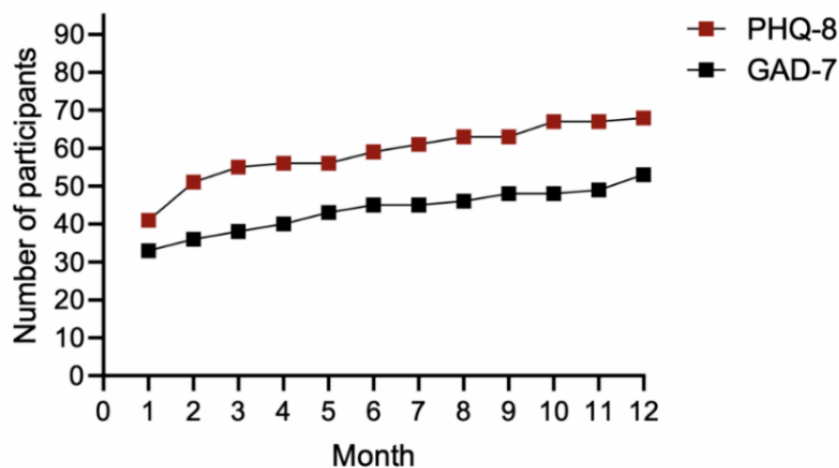


Figure 3. Proportion of, and degree of elevation for, symptoms of depression over the 12 months. PHQ-8: Patient Health Questionnaire 8-item measure.

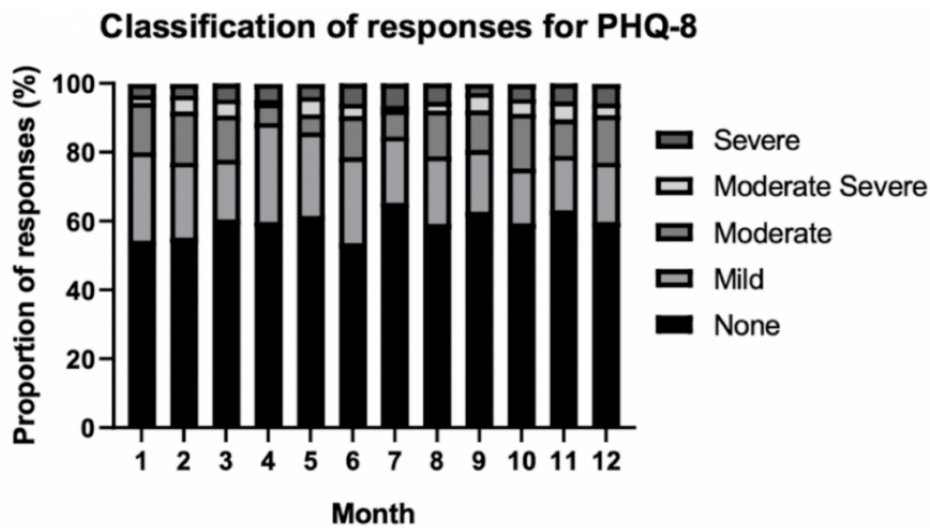
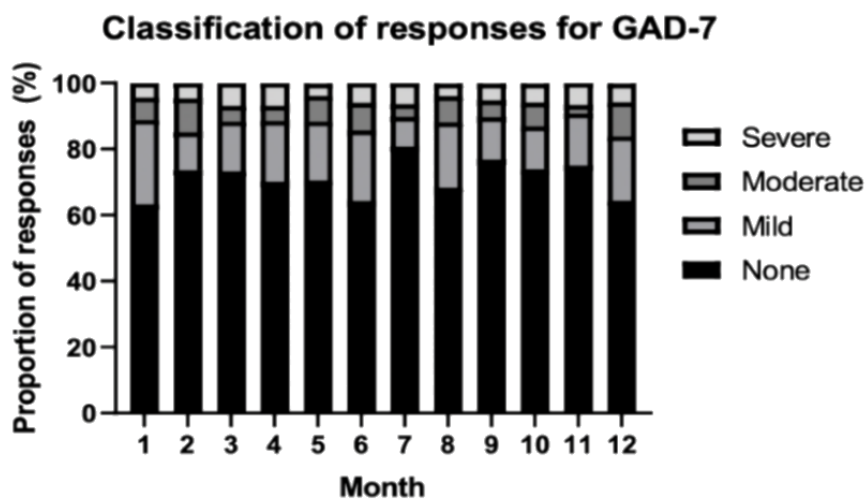


Figure 4. Proportion of, and degree of elevation for, symptoms of anxiety over the 12 months. GAD-7: General Anxiety Disorder 7-item measure.



**Comparison of People with Mood Disturbance to Those Without**

Overall, 47 (50%) participants had elevated scores for both the PHQ-8 and GAD-7 on at least one occasion and 20 (21.3%) participants recorded no elevated scores on either measure. The remaining participants only reported an elevated score for either PHQ-8 (n=21) or GAD-7 (n=6) during the 12-month monitoring period. Compared to those who reported no changes in the scores, those who responded with an elevated score for both scales during the 12 months (n=47) reported a lower health status (median 0.7 [IQR 0.6-0.8]; versus median 0.5 [IQR 0.4-0.07];  $P=.008$ ), reduced respiratory quality of life (mean 50.7 [SD 16.8] versus mean 63.3 [SD 17.8];  $P=.009$ ), greater symptoms of anxiety (median 1 [IQR 0-4] versus median 6 [IQR 2-13];  $P<.001$ ) and greater symptoms of depression (median 6 [IQR 3-9] versus median 22 [IQR 10-43];  $P<.001$ ), and no other differences were observed (see [Multimedia Appendix 2](#)).

**Accessing and Watching the Mood Management and Stress Videos**

In total, 51 (54.3%) participants accessed the mood management video at least once while 41 (43.6%) participants accessed the stress management video. A total of 64 (68.1%) participants watched one of the two videos and 28 (29.8%) participants watched both the mood and the stress management videos during the 12 months. The mood and stress management videos were watched an average of 0.8 (SD 1.0) and 0.9 (SD 1.1) times, respectively. Of the 202 video visualizations logged during the 12 months (an average of 3.2 video visualizations per participant), participants accessed the videos without being prompted a total of 132 times (78 for the mood management video and 54 for the stress video). The median proportion of the duration of the video watched was 49.5% (IQR 12.2-81.1%) for the mood video and 38.4% (IQR 16.0-68.1%) for the stress management videos.

## Proportion Watched by Those Signposted to the Mood Management and Stress Videos

There were a total of 471/1200 responses that used the mood-monitoring tool with a score of  $\geq 5$  points for the PHQ-8. In total, 39 video visualizations were watched after being signposted during the 12 months (recorded by 23 participants); 20 (51.3%) of these visualizations involved watching  $\geq 90\%$  of the video duration. The median proportion of the duration of the mood management video watched was 83.7% (IQR 23.2-100.0%). Similarly, 317/1200 responses recorded a score of  $\geq 5$  points for GAD-7 and were signposted to the stress management video. In total, 23 participants viewed 31 stress management video visualizations after being signposted during the 12 months and 20 (64.5%) visualizations were viewed for  $\geq 90\%$  of their duration. The median proportion of the stress management video watched was 100% (IQR 23.3-100%). Overall, 8 (14.3%) participants who reported an elevated score for both anxiety and depression during the 12 months watched  $\geq 90\%$  of both videos, while 21 (37.5%) participants did not watch either. In comparison, 4 (17.4%) participants who responded with no elevated scores for symptoms of either anxiety or depression watched at least one of these videos.

## Effect of Signposting Participants Versus Watching the Videos Freely

The proportion of the duration of the stress management video watched was significantly greater for the video visualizations that were signposted (median 100% [IQR 23.3-100.0%];  $P=.03$ ) compared to those freely watched (median 38.4% [IQR 16.0-68.1%];  $P=.03$ ) but was not significant for the mood management video. In total, 11 participants accessed the mood and stress management videos upon being signposted for both components at some point during the 12 months.

## The Impact of Signposting on Subsequent Mood Disturbance

Of the participants who were signposted to the mood management or stress video and watched  $\geq 90\%$  of its duration, 12 (54.5%) entered a lower score in the subsequent month, 4 (18.2%) entered a higher score, and the remaining 6 (27.3%) participants entered an identical score. In comparison, of the participants who were signposted but did not watch  $\geq 90\%$  of either video, 9 (52.9%) participants entered a lower score, 4 (23.5%) entered a higher score, and the remaining 2 participants (11.8%) entered an identical score.

## Discussion

### Principal Findings

This study confirms that it is both feasible and acceptable for people with COPD to complete monthly mood questions using a tablet computer-based digital health system. Many participants reported symptoms of anxiety and low mood at least once during the 12 months. We report how this group of participants reported a lower health status and quality of life compared to participants who did not report any symptoms of anxiety or depression. Signposting participants at the point of entering elevated mood responses did result in longer viewing times for the stress-management video but it remains unclear how to optimize

the process to make support available to people with COPD. In principle, mood self-monitoring could form an important component of future studies to consider the presence (and severity) of mood disturbance in people with COPD.

To our knowledge, this is the first study to evaluate incorporating longitudinal mood monitoring in digital health systems for people with COPD. Participants in this study were asked to answer a set of mood questions every 4 weeks over 12 months. Electronic self-monitoring of mood has been used extensively in people with bipolar disorder previously [37], with studies lasting 2 weeks [38], to 3 months [39], to 18 months [40]. Participants in the 2-week study were asked to respond to questions at 4 random times per day with no less than 2 hours between each survey [38]. This intense data collection strategy captured mood states and identified triggers in 10 people with bipolar disorder, and the median percentage of completed possible surveys was 78%. In comparison, Scharer et al asked participants to complete daily questions via a mobile application over 18 months and demonstrated its validity in recognizing both manic and depressive episodes [40]. The amount of reported missing data to date has ranged from 6.1-57.9% [37], while in the present analysis 13.4% was missing. This suggests that people with COPD were very compliant in self-monitoring mood over several months. However, the intensity of required reporting in this study was much lower compared with that in the bipolar literature and it is unclear to what extent the amount of missing data would increase if the required reporting was more frequent in people with COPD.

Despite instructions asking participants to complete the questions every four weeks in the current study, many participants responded more frequently. The mode of accessing the mood questions may have in part contributed to its ease of use. Previous studies have used web-based interfaces, personal digital assistants, and smartphones for data collection [37], but the current study used a tablet computer and the mood monitoring element was a part of a multicomponent digital intervention. Participants were also reminded to complete the questions after the 4-week period if they had not, and this reminder persisted on the tablet computer interface until the mood questions were answered. The high adherence rates may also be in part because the mood self-monitoring component formed part of a broader mHealth self-management system which was comprised of daily symptom monitoring and access to other information videos, including inhaler techniques and physical exercises. With a high degree of compliance to use of this element of the EDGE system, the findings support the use of self-monitoring mood in people with COPD even though systematic reviews [41,42] have highlighted that telemonitoring studies more often monitor other symptoms, including dyspnea and sputum [43-45], lung function [46], and physical activity in this patient group [47,48].

Monitoring mood in people with COPD may not yet be common, but the impact of mood disturbance is substantial. A meta-analysis of nearly 40,000 people with COPD highlighted that one quarter experienced clinically significant depressive symptoms (compared with one eighth in controls) [49]. We identified that half of people with COPD experienced clinically significant symptoms of anxiety and depression on at least one

occasion over the 12-month period. Not only is the prevalence important but also its influence on self-managing the condition.

Elevated symptoms for anxiety and depression have been associated with several markers of ill health. This includes fatigue and shortness of breath [50,51], poorer health-related quality of life [52-54], frequent hospitalization [55], comorbidities [56], functional limitations, and reduced exercise capacity [57-59]. People with a long-term condition, like COPD, also living with depression or anxiety have a worsened health status compared to people living with depression or anxiety alone [60]. People who reported elevated scores for the PHQ-8 and GAD-7 on at least one occasion over the 12 months had a significantly lower health status, reduced respiratory-related quality of life and elevated symptoms for both anxiety and depression. Symptoms of anxiety and depression are important determinants of health outcomes and health care utilization [61].

Most participants who were signposted to the videos did not watch the video. However, when participants did go on to watch the stress video after being signposted, they watched it for significantly longer. Signposting at the point of someone receiving an elevated score is an important opportunity to deliver relevant information. For instance, in the True Colours longitudinal mood-monitoring study, people with affective disorder were signposted to communicate with their relevant health care team to highlight any concerns with their mood symptoms [62]. In this analysis we focused on how people engaged with the stress management and mood videos, but this could extend to other domains of COPD, including self-reported shortness of breath. In that example, individuals could be directed to breathing technique videos, such as pursed lip breathing techniques, or it could extend even wider to other long-term conditions if people have comorbidities.

### Implications for Future Research and Clinical practice

This analysis demonstrates the current potential of monitoring mood in people with COPD. It appears a feasible and acceptable parameter to monitor over 12 months with a high rate of adherence. Previous work has suggested value in monitoring

several factors in people with COPD, including factors that affect activities of daily living, like symptoms [63]. Monitoring mood together with existing parameters of interest (including oxygen saturation and COPD symptoms) may provide a better insight into individual well-being. The findings also suggest that signposting people at the point of reporting an elevated score can lead to longer video viewing times and is a timely, resource-light feature. The integration of signposting to multimedia resources in digital health interventions is therefore encouraged, using health care system-endorsed material where possible, as is the presentation of feedback to allow participants to track their mood scores over time.

### Strengths and Limitations

The delivery of the mood monitoring component within the EDGE system was easy to use and nonobtrusive for participants. It formed part of a wider self-management system which included daily questions around COPD symptoms, demonstrating that the mood questions could be integrated with other components of self-management. As all participants received a reminder if they had an overdue response, we could not comment on the adherence rates if no reminder was sent. We were only able to determine if signposting resulted in an immediate viewing of the video (within 30 minutes) but could not comment on delayed watching habits. Participant insights would have been valuable to decipher why those participants did not follow the signposting.

### Conclusion

Mood monitoring in people with COPD is feasible, with most patients answering the mood questions over 12 months and with some participants entering multiple responses each month. With more than half of participants reporting elevated scores, plus the association this has with health status and quality of life, self-monitoring of mood could be an important, additional component to COPD self-management strategies. In addition, the incorporation of signposting for people with COPD to videos when entering an elevated score demonstrated promise, but further work is needed to optimize this approach.

---

### Acknowledgments

The authors would like to acknowledge the Nuffield Department of Primary Care Health Sciences' Clinical Trials Unit for supporting our request to access data from the original EDGE trial. AF and LT receive funding from the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. AF is an NIHR Senior Investigator.

---

### Authors' Contributions

AF and LT obtained funding for the original EDGE trial. MW, CV, and AF conceived and designed the analysis. CV and HR were involved in collecting the data. MW performed the analysis with support from CV and AF. MW and AF initially wrote the manuscript, with subsequent drafts viewed by CV, HR, and LT for edits.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Overview of statistical comparisons for adherent participants compared with those nonadherent with using the system. [[DOCX File, 18 KB - mhealth\\_v7i11e14946\\_app1.docx](#)]

## Multimedia Appendix 2

Overview of statistical comparisons for participants with elevated PHQ-8 and GAD-7 scores compared with those with nonelevated scores.

[[DOCX File , 18 KB - mhealth\\_v7i11e14946\\_app2.docx](#) ]

## Multimedia Appendix 3

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 2160 KB - mhealth\\_v7i11e14946\\_app3.pdf](#) ]

**References**

1. Dowson C, Laing R, Barraclough R, Town I, Mulder R, Norris K, et al. The use of the Hospital Anxiety and Depression Scale (HADS) in patients with chronic obstructive pulmonary disease: a pilot study. *N Z Med J* 2001 Oct 12;114(1141):447-449. [Medline: [11700772](#)]
2. Light RW, Merrill EJ, Despars JA, Gordon GH, Mutalipassi LR. Prevalence of depression and anxiety in patients with COPD. Relationship to functional capacity. *Chest* 1985 Jan;87(1):35-38. [doi: [10.1378/chest.87.1.35](#)] [Medline: [3965263](#)]
3. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS, GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001 Apr;163(5):1256-1276. [doi: [10.1164/ajrccm.163.5.2101039](#)] [Medline: [11316667](#)]
4. Gudmundsson G, Gislason T, Janson C, Lindberg E, Hallin R, Ulrik CS, et al. Risk factors for rehospitalisation in COPD: role of health status, anxiety and depression. *Eur Respir J* 2005 Sep;26(3):414-419 [FREE Full text] [doi: [10.1183/09031936.05.00078504](#)] [Medline: [16135721](#)]
5. Buekers J, De Boever P, Vaes A, Aerts J, Wouters E, Spruit M, et al. Oxygen saturation measurements in telemonitoring of patients with COPD: a systematic review. *Expert Rev Respir Med* 2018 Feb;12(2):113-123. [doi: [10.1080/17476348.2018.1417842](#)] [Medline: [29241369](#)]
6. Swinkels A, Giuliano TA. The Measurement and Conceptualization of Mood Awareness: Monitoring and Labeling One's Mood States. *Pers Soc Psychol Bull* 2016 Jul 02;21(9):934-949. [doi: [10.1177/0146167295219008](#)]
7. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. *Telemed J E Health* 2017 Jan;23(1):3-17 [FREE Full text] [doi: [10.1089/tmj.2016.0051](#)] [Medline: [27116181](#)]
8. Torous J, Powell A. Current research and trends in the use of smartphone applications for mood disorders. *Internet Interventions* 2015 May;2(2):169-173 [FREE Full text] [doi: [10.1016/j.invent.2015.03.002](#)]
9. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J Affect Disord* 2018 Feb;227:432-442. [doi: [10.1016/j.jad.2017.11.016](#)] [Medline: [29154165](#)]
10. 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](#)]
11. O'Toole MS, Jensen MB, Fentz HN, Zachariae R, Hougaard E. Emotion Differentiation and Emotion Regulation in High and Low Socially Anxious Individuals: An Experience-Sampling Study. *Cogn Ther Res* 2014 Mar 13;38(4):428-438. [doi: [10.1007/s10608-014-9611-2](#)]
12. Malhi G, Hamilton A, Morris G, Mannie Z, Das P, Outhred T. The promise of digital mood tracking technologies: are we heading on the right track? *Evid Based Ment Health* 2017 Nov;20(4):102-107. [doi: [10.1136/eb-2017-102757](#)] [Medline: [28855245](#)]
13. Saunders KEA, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: A qualitative study. *Eur Psychiatry* 2017 Mar;41:115-121 [FREE Full text] [doi: [10.1016/j.eurpsy.2016.11.005](#)] [Medline: [28135594](#)]
14. Juengst S, Graham K, Pulantara I, McCue M, Whyte E, Dicianno B, et al. Pilot feasibility of an mHealth system for conducting ecological momentary assessment of mood-related symptoms following traumatic brain injury. *Brain Inj* 2015;29(11):1351-1361. [doi: [10.3109/02699052.2015.1045031](#)] [Medline: [26287756](#)]
15. Tsanas A, Saunders KEA, Bilderbeck AC, Palmius N, Osipov M, Clifford GD, et al. Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord* 2016 Nov 15;205:225-233 [FREE Full text] [doi: [10.1016/j.jad.2016.06.065](#)] [Medline: [27449555](#)]
16. Schwartz S, Schultz S, Reider A, Saunders EFH. Daily mood monitoring of symptoms using smartphones in bipolar disorder: A pilot study assessing the feasibility of ecological momentary assessment. *J Affect Disord* 2016 Feb;191:88-93 [FREE Full text] [doi: [10.1016/j.jad.2015.11.013](#)] [Medline: [26655117](#)]
17. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005 Nov;43(11):1130-1139. [doi: [10.1097/01.mlr.0000182534.19832.83](#)] [Medline: [16224307](#)]

18. Heslop-Marshall K, Baker C, Carrick-Sen D, Newton J, Echevarria C, Stenton C, et al. Randomised controlled trial of cognitive behavioural therapy in COPD. *ERJ Open Res* 2018 Oct;4(4) [[FREE Full text](#)] [doi: [10.1183/23120541.00094-2018](https://doi.org/10.1183/23120541.00094-2018)] [Medline: [30479999](#)]
19. Yohannes A, Alexopoulos G. Depression and anxiety in patients with COPD. *Eur Respir Rev* 2014 Sep;23(133):345-349 [[FREE Full text](#)] [doi: [10.1183/09059180.00007813](https://doi.org/10.1183/09059180.00007813)] [Medline: [25176970](#)]
20. Spruit M, Franssen F, Rutten E, Wopereis S, Wouters E, Vanfleteren LEGW. A new perspective on COPD exacerbations: monitoring impact by measuring physical, psychological and social resilience. *Eur Respir J* 2016 Apr;47(4):1024-1027 [[FREE Full text](#)] [doi: [10.1183/13993003.01645-2015](https://doi.org/10.1183/13993003.01645-2015)] [Medline: [27037307](#)]
21. de Voogd JN, Wempe JB, Koëter GH, Postema K, van Sonderen E, Ranchor AV, et al. Depressive symptoms as predictors of mortality in patients with COPD. *Chest* 2009 Mar;135(3):619-625. [doi: [10.1378/chest.08-0078](https://doi.org/10.1378/chest.08-0078)] [Medline: [19029432](#)]
22. Atlantis E, Fahey P, Cochrane B, Smith S. Bidirectional associations between clinically relevant depression or anxiety and COPD: a systematic review and meta-analysis. *Chest* 2013 Sep;144(3):766-777. [doi: [10.1378/chest.12-1911](https://doi.org/10.1378/chest.12-1911)] [Medline: [23429910](#)]
23. Ng T, Niti M, Tan W, Cao Z, Ong K, Eng P. Depressive symptoms and chronic obstructive pulmonary disease: effect on mortality, hospital readmission, symptom burden, functional status, and quality of life. *Arch Intern Med* 2007 Jan 08;167(1):60-67. [doi: [10.1001/archinte.167.1.60](https://doi.org/10.1001/archinte.167.1.60)] [Medline: [17210879](#)]
24. Dueñas-Espín I, Demeyer H, Gimeno-Santos E, Polkey MI, Hopkinson NS, Rabinovich RA, et al. Depression symptoms reduce physical activity in COPD patients: a prospective multicenter study. *Int J Chron Obstruct Pulmon Dis* 2016;11:1287-1295 [[FREE Full text](#)] [doi: [10.2147/COPD.S101459](https://doi.org/10.2147/COPD.S101459)] [Medline: [27354787](#)]
25. Farmer A, Williams V, Velardo C, Shah SA, Yu L, Rutter H, et al. Self-Management Support Using a Digital Health System Compared With Usual Care for Chronic Obstructive Pulmonary Disease: Randomized Controlled Trial. *J Med Internet Res* 2017 May 03;19(5):e144 [[FREE Full text](#)] [doi: [10.2196/jmir.7116](https://doi.org/10.2196/jmir.7116)] [Medline: [28468749](#)]
26. Farmer A, Toms C, Hardinge M, Williams V, Rutter H, Tarassenko L. Self-management support using an Internet-linked tablet computer (the EDGE platform)-based intervention in chronic obstructive pulmonary disease: protocol for the EDGE-COPD randomised controlled trial. *BMJ Open* 2014 Jan 08;4(1):e004437 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2013-004437](https://doi.org/10.1136/bmjopen-2013-004437)] [Medline: [24401729](#)]
27. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord* 2010 Apr;122(1-2):86-95. [doi: [10.1016/j.jad.2009.06.019](https://doi.org/10.1016/j.jad.2009.06.019)] [Medline: [19616305](#)]
28. Kroenke K, Strine TW, Spitzer RL, Williams JBW, 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](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](#)]
29. Spitzer R, Kroenke K, Williams J, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](#)]
30. National IFHACE. Depression in adults with a chronic physical health problem: recognition and management. 2009 Oct. URL: <https://www.nice.org.uk/guidance/cG91> [accessed 2018-06-15]
31. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of chronic pulmonary disease. 2018. URL: [https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov\\_WMS.pdf](https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf) [accessed 2019-09-27]
32. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001 Jul;33(5):337-343. [doi: [10.3109/07853890109002087](https://doi.org/10.3109/07853890109002087)] [Medline: [11491192](#)]
33. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacol Bull* 1973 Jan;9(1):13-28. [Medline: [4682398](#)]
34. Morisky D, Green L, Levine D. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986 Jan;24(1):67-74. [doi: [10.1097/00005650-198601000-00007](https://doi.org/10.1097/00005650-198601000-00007)] [Medline: [3945130](#)]
35. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology & Health* 1999 Jan;14(1):1-24 [[FREE Full text](#)] [doi: [10.1080/08870449908407311](https://doi.org/10.1080/08870449908407311)]
36. Meguro M, Barley EA, Spencer S, Jones PW. Development and Validation of an Improved, COPD-Specific Version of the St. George Respiratory Questionnaire. *Chest* 2007 Aug;132(2):456-463. [doi: [10.1378/chest.06-0702](https://doi.org/10.1378/chest.06-0702)] [Medline: [17646240](#)]
37. Faurholt-Jepsen M, Munkholm K, Frost M, Bardram JE, Kessing LV. Electronic self-monitoring of mood using IT platforms in adult patients with bipolar disorder: A systematic review of the validity and evidence. *BMC Psychiatry* 2016 Jan 15;16:7 [[FREE Full text](#)] [doi: [10.1186/s12888-016-0713-0](https://doi.org/10.1186/s12888-016-0713-0)] [Medline: [26769120](#)]
38. Depp CA, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson TL, et al. Mobile interventions for severe mental illness: design and preliminary data from three approaches. *J Nerv Ment Dis* 2010 Oct;198(10):715-721 [[FREE Full text](#)] [doi: [10.1097/NMD.0b013e3181f49ea3](https://doi.org/10.1097/NMD.0b013e3181f49ea3)] [Medline: [20921861](#)]
39. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Res* 2014 Jun 30;217(1-2):124-127. [doi: [10.1016/j.psychres.2014.03.009](https://doi.org/10.1016/j.psychres.2014.03.009)] [Medline: [24679993](#)]



40. Schärer LO, Krienke UJ, Graf S, Meltzer K, Langosch JM. Validation of life-charts documented with the personal life-chart app - a self-monitoring tool for bipolar disorder. *BMC Psychiatry* 2015 Mar 14;15:49 [FREE Full text] [doi: [10.1186/s12888-015-0414-0](https://doi.org/10.1186/s12888-015-0414-0)] [Medline: [25885225](https://pubmed.ncbi.nlm.nih.gov/25885225/)]
41. Alwashmi M, Hawboldt J, Davis E, Marra C, Gamble J, Abu Ashour W. The Effect of Smartphone Interventions on Patients With Chronic Obstructive Pulmonary Disease Exacerbations: A Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth* 2016 Sep 01;4(3):e105 [FREE Full text] [doi: [10.2196/mhealth.5921](https://doi.org/10.2196/mhealth.5921)] [Medline: [27589898](https://pubmed.ncbi.nlm.nih.gov/27589898/)]
42. Lundell S, Holmner A, Rehn B, Nyberg A, Wadell K. Telehealthcare in COPD: a systematic review and meta-analysis on physical outcomes and dyspnea. *Respir Med* 2015 Jan;109(1):11-26 [FREE Full text] [doi: [10.1016/j.rmed.2014.10.008](https://doi.org/10.1016/j.rmed.2014.10.008)] [Medline: [25464906](https://pubmed.ncbi.nlm.nih.gov/25464906/)]
43. Halpin DMG, Laing-Morton T, Spedding S, Levy ML, Coyle P, Lewis J, et al. A randomised controlled trial of the effect of automated interactive calling combined with a health risk forecast on frequency and severity of exacerbations of COPD assessed clinically and using EXACT PRO. *Prim Care Respir J* 2011 Sep;20(3):324-31, 2 p following 331 [FREE Full text] [doi: [10.4104/pcrj.2011.00057](https://doi.org/10.4104/pcrj.2011.00057)] [Medline: [21687919](https://pubmed.ncbi.nlm.nih.gov/21687919/)]
44. Liu W, Wang C, Lin H, Lin S, Lee K, Lo Y, et al. Efficacy of a cell phone-based exercise programme for COPD. *Eur Respir J* 2008 Sep;32(3):651-659 [FREE Full text] [doi: [10.1183/09031936.00104407](https://doi.org/10.1183/09031936.00104407)] [Medline: [18508824](https://pubmed.ncbi.nlm.nih.gov/18508824/)]
45. Nguyen HQ, Donesky D, Reinke LF, Wolpin S, Chyall L, Benditt JO, et al. Internet-based dyspnea self-management support for patients with chronic obstructive pulmonary disease. *J Pain Symptom Manage* 2013 Jul;46(1):43-55 [FREE Full text] [doi: [10.1016/j.jpainsymman.2012.06.015](https://doi.org/10.1016/j.jpainsymman.2012.06.015)] [Medline: [23073395](https://pubmed.ncbi.nlm.nih.gov/23073395/)]
46. Jehn M, Donaldson G, Kiran B, Liebers U, Mueller K, Scherer D, et al. Tele-monitoring reduces exacerbation of COPD in the context of climate change--a randomized controlled trial. *Environ Health* 2013 Nov 21;12:99 [FREE Full text] [doi: [10.1186/1476-069X-12-99](https://doi.org/10.1186/1476-069X-12-99)] [Medline: [24261700](https://pubmed.ncbi.nlm.nih.gov/24261700/)]
47. Tabak M, Brusse-Keizer M, van der Valk P, Hermens H, Vollenbroek-Hutten M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. *Int J Chron Obstruct Pulmon Dis* 2014;9:935-944 [FREE Full text] [doi: [10.2147/COPD.S60179](https://doi.org/10.2147/COPD.S60179)] [Medline: [25246781](https://pubmed.ncbi.nlm.nih.gov/25246781/)]
48. Pedone C, Chiurco D, Scarlata S, Incalzi RA. Efficacy of multiparametric telemonitoring on respiratory outcomes in elderly people with COPD: a randomized controlled trial. *BMC Health Serv Res* 2013 Mar 06;13:82 [FREE Full text] [doi: [10.1186/1472-6963-13-82](https://doi.org/10.1186/1472-6963-13-82)] [Medline: [23497109](https://pubmed.ncbi.nlm.nih.gov/23497109/)]
49. Zhang MWB, Ho RCM, Cheung MWL, Fu E, Mak A. Prevalence of depressive symptoms in patients with chronic obstructive pulmonary disease: a systematic review, meta-analysis and meta-regression. *Gen Hosp Psychiatry* 2011;33(3):217-223. [doi: [10.1016/j.genhosppsy.2011.03.009](https://doi.org/10.1016/j.genhosppsy.2011.03.009)] [Medline: [21601717](https://pubmed.ncbi.nlm.nih.gov/21601717/)]
50. Jose A, Chelangara DS, S. S. Factors associated with anxiety and depression in chronic obstructive pulmonary disease. *Int J Res Med Sci* 2016;1074-1079 [FREE Full text] [doi: [10.18203/2320-6012.ijrms20160786](https://doi.org/10.18203/2320-6012.ijrms20160786)]
51. Doyle T, Palmer S, Johnson J, Babyak MA, Smith P, Mabe S, et al. Association of anxiety and depression with pulmonary-specific symptoms in chronic obstructive pulmonary disease. *Int J Psychiatry Med* 2013;45(2):189-202 [FREE Full text] [doi: [10.2190/PM.45.2.g](https://doi.org/10.2190/PM.45.2.g)] [Medline: [23977821](https://pubmed.ncbi.nlm.nih.gov/23977821/)]
52. Martinez Rivera C, Costan Galicia J, Alcázar Navarrete B, Garcia-Polo C, Ruiz Iturriaga LA, Herrejón A, et al. Factors Associated with Depression in COPD: A Multicenter Study. *Lung* 2016 Jun;194(3):335-343. [doi: [10.1007/s00408-016-9862-7](https://doi.org/10.1007/s00408-016-9862-7)] [Medline: [26932809](https://pubmed.ncbi.nlm.nih.gov/26932809/)]
53. Balcells E, Gea J, Ferrer J, Serra I, Orozco-Levi M, de Batlle J, PAC-COPD Study Group. Factors affecting the relationship between psychological status and quality of life in COPD patients. *Health Qual Life Outcomes* 2010 Sep 27;8:108 [FREE Full text] [doi: [10.1186/1477-7525-8-108](https://doi.org/10.1186/1477-7525-8-108)] [Medline: [20875100](https://pubmed.ncbi.nlm.nih.gov/20875100/)]
54. Tsiligianni IG, van der Molen T. A systematic review of the role of vitamin insufficiencies and supplementation in COPD. *Respir Res* 2010 Dec 06;11:171 [FREE Full text] [doi: [10.1186/1465-9921-11-171](https://doi.org/10.1186/1465-9921-11-171)] [Medline: [21134250](https://pubmed.ncbi.nlm.nih.gov/21134250/)]
55. Eisner MD, Blanc PD, Yelin EH, Katz PP, Sanchez G, Iribarren C, et al. Influence of anxiety on health outcomes in COPD. *Thorax* 2010 Mar;65(3):229-234 [FREE Full text] [doi: [10.1136/thx.2009.126201](https://doi.org/10.1136/thx.2009.126201)] [Medline: [20335292](https://pubmed.ncbi.nlm.nih.gov/20335292/)]
56. Tsai T, Livneh H, Lu M, Tsai P, Chen P, Sung F. Increased risk and related factors of depression among patients with COPD: a population-based cohort study. *BMC Public Health* 2013 Oct 19;13:976 [FREE Full text] [doi: [10.1186/1471-2458-13-976](https://doi.org/10.1186/1471-2458-13-976)] [Medline: [24138872](https://pubmed.ncbi.nlm.nih.gov/24138872/)]
57. Yohannes AM, Müllerová H, Hanania NA, Lavoie K, Tal-Singer R, Vestbo J, et al. Long-term Course of Depression Trajectories in Patients With COPD: A 3-Year Follow-up Analysis of the Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints Cohort. *Chest* 2016 Apr;149(4):916-926. [doi: [10.1016/j.chest.2015.10.081](https://doi.org/10.1016/j.chest.2015.10.081)] [Medline: [26836938](https://pubmed.ncbi.nlm.nih.gov/26836938/)]
58. Ng T, Niti M, Tan W, Cao Z, Ong K, Eng P. Depressive symptoms and chronic obstructive pulmonary disease: effect on mortality, hospital readmission, symptom burden, functional status, and quality of life. *Arch Intern Med* 2007 Jan 08;167(1):60-67. [doi: [10.1001/archinte.167.1.60](https://doi.org/10.1001/archinte.167.1.60)] [Medline: [17210879](https://pubmed.ncbi.nlm.nih.gov/17210879/)]
59. Blakemore A, Dickens C, Guthrie E, Bower P, Kontopantelis E, Afzal C, et al. Depression and anxiety predict health-related quality of life in chronic obstructive pulmonary disease: systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis* 2014;9:501-512 [FREE Full text] [doi: [10.2147/COPD.S58136](https://doi.org/10.2147/COPD.S58136)] [Medline: [24876770](https://pubmed.ncbi.nlm.nih.gov/24876770/)]

60. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* 2007 Sep 08;370(9590):851-858. [doi: [10.1016/S0140-6736\(07\)61415-9](https://doi.org/10.1016/S0140-6736(07)61415-9)] [Medline: [17826170](https://pubmed.ncbi.nlm.nih.gov/17826170/)]
61. Panagioti M, Scott C, Blakemore A, Coventry PA. Overview of the prevalence, impact, and management of depression and anxiety in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2014;9:1289-1306 [FREE Full text] [doi: [10.2147/COPD.S72073](https://doi.org/10.2147/COPD.S72073)] [Medline: [25419126](https://pubmed.ncbi.nlm.nih.gov/25419126/)]
62. Gordon-Smith K, Saunders K, Geddes J, Harrison P, Hinds C, Craddock N, et al. Large-scale roll out of electronic longitudinal mood-monitoring for research in affective disorders: Report from the UK bipolar disorder research network. *J Affect Disord* 2019 Mar 01;246:789-793. [doi: [10.1016/j.jad.2018.12.099](https://doi.org/10.1016/j.jad.2018.12.099)] [Medline: [30623825](https://pubmed.ncbi.nlm.nih.gov/30623825/)]
63. Sandstrom G, Lathia N, Mascolo C, Rentfrow P. Opportunities for Smartphones in Clinical Care: The Future of Mobile Mood Monitoring. *J. Clin. Psychiatry* 2016 Feb 24;77(02):e135-e137 [FREE Full text] [doi: [10.4088/jcp.15com10054](https://doi.org/10.4088/jcp.15com10054)]

## Abbreviations

**BMI:** body mass index

**BMQ:** beliefs about medications questionnaire

**COPD:** chronic obstructive pulmonary disorder

**FEV1%:** forced expiratory volume in 1 second

**EDGE:** sElf-management anD support proGrammE

**EQ-5D:** EuroQol 5-Dimension Questionnaire

**GAD:** General Anxiety Disorder measure

**GOLD:** Global Initiative for Chronic Obstructive Lung Disease

**GP:** general practitioner

**MARS:** Medication Adherence Report Schedule

**mHealth:** mobile health

**NIHR:** National Institute for Health Research

**PHQ:** Patient Health Questionnaire

**SCK:** Symptom Checklist

**SGRQ-C:** St George's Respiratory Questionnaire for Chronic Obstructive Lung Disease questionnaire

**SIM:** subscriber identity module

*Edited by L Buis; submitted 05.06.19; peer-reviewed by M Nahum, D Bakker; comments to author 30.07.19; revised version received 11.08.19; accepted 31.08.19; published 22.11.19.*

*Please cite as:*

Whelan ME, Velardo C, Rutter H, Tarassenko L, Farmer AJ

Mood Monitoring Over One Year for People With Chronic Obstructive Pulmonary Disease Using a Mobile Health System: Retrospective Analysis of a Randomized Controlled Trial

*JMIR Mhealth Uhealth* 2019;7(11):e14946

URL: <http://mhealth.jmir.org/2019/11/e14946/>

doi: [10.2196/14946](https://doi.org/10.2196/14946)

PMID: [31755872](https://pubmed.ncbi.nlm.nih.gov/31755872/)

©Maxine E Whelan, Carmelo Velardo, Heather Rutter, Lionel Tarassenko, Andrew J Farmer. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 22.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

---

Review

# Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review

---

Aurélie Najm<sup>1,2</sup>, MSc, MD; Laure Gossec<sup>3,4</sup>, MD, PhD; Catherine Weill<sup>5</sup>, PhD; David Benoist<sup>5</sup>, PhD; Francis Berenbaum<sup>6</sup>, MD, PhD; Elena Nikiphorou<sup>7</sup>, MD, PhD

<sup>1</sup>Department of Rheumatology, Nantes University Hospital, Nantes, France

<sup>2</sup>INSERM UMR 1238, Nantes University of Medicine, Nantes, France

<sup>3</sup>INSERM UMR S1136, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, Paris, France

<sup>4</sup>Rheumatology department, Pitié Salpêtrière Hospital, AP-HP, Paris, France

<sup>5</sup>Bibliothèque interuniversitaire de Santé, Paris Descartes University, Paris, France

<sup>6</sup>Department of rheumatology, Sorbonne Université, INSERM CRSA Saint-Antoine, AP-HP, Saint Antoine Hospital, Paris, France

<sup>7</sup>Department of Inflammation Biology, School of Immunology and Microbial Sciences, Faculty of Life Sciences & Medicine, King's College London, Paris, France

**Corresponding Author:**

Aurélie Najm, MSc, MD

Department of Rheumatology

Nantes University Hospital

1, place Alexis Ricordeau

Nantes, 44000

France

Phone: 33 633444793

Email: [aurelie.najm@gmail.com](mailto:aurelie.najm@gmail.com)

---

## Abstract

**Background:** Although the increasing availability of mobile health (mHealth) apps may enable people with rheumatic and musculoskeletal diseases (RMDs) to better self-manage their health, there is a general lack of evidence on ways to ensure appropriate development and evaluation of apps.

**Objective:** This study aimed to obtain an overview on existing mHealth apps for self-management in patients with RMDs, focusing on content and development methods.

**Methods:** A search was performed up to December 2017 across 5 databases. For each publication relevant to an app for RMDs, information on the disease, purpose, content, and development strategies was extracted and qualitatively assessed.

**Results:** Of 562 abstracts, 32 were included in the analysis. Of these 32 abstracts, 11 (34%) referred to an app linked to a connected device. Most of the apps targeted rheumatoid arthritis (11/32, 34%). The top three aspects addressed by the apps were pain (23/32, 71%), fatigue (15/32, 47%), and physical activity (15/32, 47%). The development process of the apps was described in 84% (27/32) of the articles and was of low to moderate quality in most of the cases. Despite most of the articles having been published within the past two years, only 5 apps were still commercially available at the time of our search. Moreover, only very few studies showed improvement of RMD outcome measures.

**Conclusions:** The development process of most apps was of low or moderate quality in many studies. Owing to the increasing RMD patients' willingness to use mHealth apps for self-management, optimal standards and quality assurance of new apps are mandatory.

(*JMIR Mhealth Uhealth* 2019;7(11):e14730) doi:[10.2196/14730](https://doi.org/10.2196/14730)

---

**KEYWORDS**

mobile health; self-management; arthritis; telemedicine; musculoskeletal diseases

## Introduction

### Background

Mobile health (mHealth) connects patients, their families, and health care professionals by creating a network with mobile and specialized devices with wearable sensors, recording health parameters, and gathering health data. Health information can be subsequently converted and transferred to physicians and other health care professionals involved in the care of patients via medical application interfaces. By enabling patients to access and share their health information, mHealth empowers patients to become more engaged and to take initiative in self-management and shared management of their health.

Since the first description of the concept of mHealth [1], its popularity has exponentially increased. This is primarily because of the fast expanding technological advances including the development of smartphones and fourth generation mobile communication system networks. The impressive popularity of mHealth apps in the last decade is reflected by the number of downloads in recent years, exceeding 200 million in 2010 [2].

On the wider mHealth market, various apps have been developed for different purposes. The latter include apps for disease prevention among healthy users [3] and apps for people with existing chronic health conditions [4]. A recent study, for example, demonstrated that apps can contribute to improve disease control in people with diabetes [5], hypertension [6], or asthma [7] and can help for the monitoring and self-management of obesity [8], mental health diseases [9], and multimorbidities [10]. A total of 17.3 million people report having rheumatic and musculoskeletal diseases (RMDs), the most frequent being low back pain and osteoarthritis (OA) [11]. As RMDs have multidimensional consequences on health, the increasing availability of apps has an important role to play in enabling people with RMDs to better self-manage their health [12]. Moreover, in a recent study of people with rheumatoid arthritis (RA), 86% agreed that an app to support self-management would be useful and welcomed [13]. This being said, the dramatic increase and adoption of mHealth apps and the business it generates raise some fundamental questions, such as (1) How is the scientific content controlled? and (2) How can we make sure that apps are appropriate for patients?

### Objectives

Despite the growing enthusiasm for this topic among physicians and researchers, there is a lack of evidence describing the development and evaluation of apps for people with RMDs that fulfill quality requirements for their implementation as part of routine care. This formed the rationale for this systematic review of literature as part of a larger project to inform points to consider for the development, evaluation, and implementation of mHealth apps for self-management of RMDs. The overarching aim of this systematic review was to obtain a clear view on existing mHealth apps for patients with RMD. Specific objectives were to better characterize (1) the target population of available apps, (2) their purpose and content, and (3) strategies of mHealth app development.

## Methods

### Search Strategy

A systematic literature review (SLR) was performed following the preferred reporting items for systematic reviews and meta-analyses methodology [14]. The search was performed using the Cochrane Library, EMBASE, MEDLINE, PsychINFO, Web of Science, and gray literature (internet and international rheumatology societies' websites) up to December 2017. Relevant keywords and Medical Subject Headings terms relative to 3 key domains were used: RMDs, self-management, and mHealth (see [Multimedia Appendices 1 and 2](#)). The search strategy was developed with support from 2 experienced librarians (DB and CW). The clinical questions and inclusion criteria were predefined according to the population, intervention, control, and outcomes (PICO) statement [15]. PICO is a framework that allows to facilitate literature search and to formulate the scientific questions. The target population was patients with any RMDs. The intervention was the description or use of any apps for self-management, irrespective of whether these apps were connected or not to a device.

Any articles describing the development, evaluation, usability, accessibility, effectiveness, and assessment of patient-reported outcomes (PROs) collected through the internet or through electronic apps, and satisfaction over the use of an app for disease self-management, were included. English language was applied as a limit for the articles. Double screening by 2 independent reviewers (AN and EN) was performed for all abstracts against inclusion/exclusion criteria with agreement of 99% in selected papers. Disagreements were resolved by discussion.

### Data Analysis

Data regarding type of app, target population, country of development, themes and objectives, development process, funding sources, and functionality of the app were collected. The development process was classified into 3 main categories: (1) patients or health care providers involved in both design and evaluation phases, (2) patients or health care providers involved in the evaluation but not in the design process, and (3) neither patients nor health care providers involved in the design or evaluation phases.

The commercial availability of the app was also checked on Google and Apple Store. Owing to the vast heterogeneity of the included studies, a meta-analysis was not considered appropriate. Descriptive statistics were performed using GraphPad Prism (GraphPad Software, San Diego)

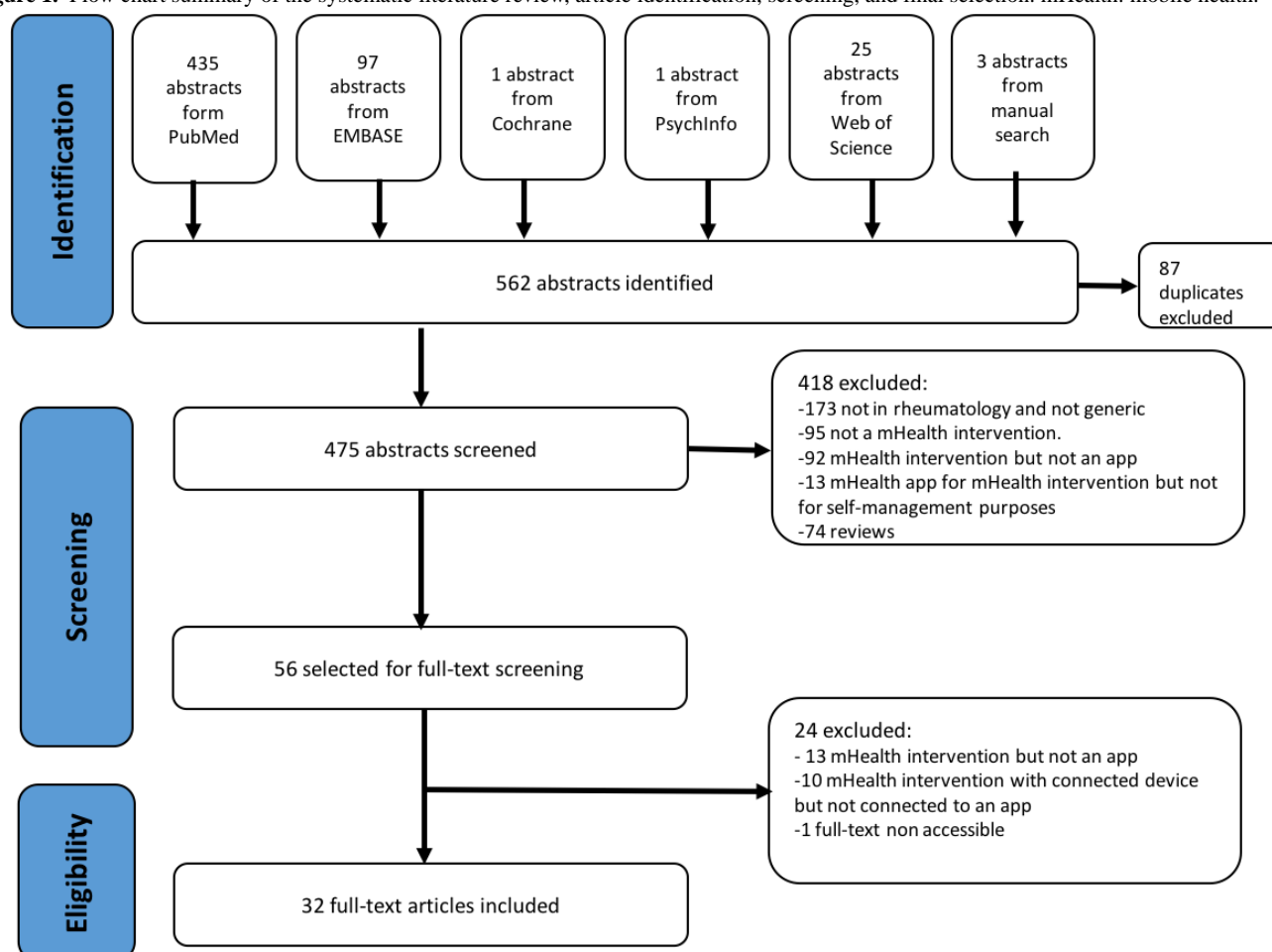
## Results

### Systematic Literature Search Output

The search identified 562 abstracts. Manual search through screening of national societies and patient associations' home pages yielded 1 additional reference. After duplicate exclusion, 475 abstracts were screened based on title and abstract. From these, 56 articles were identified as potentially relevant and selected for full-text assessment. After full-text assessment, 32

articles were considered suitable for inclusion in the analysis (Figure 1).

**Figure 1.** Flow chart summary of the systematic literature review, article identification, screening, and final selection. mHealth: mobile health.



### General Study Characteristics

Out of the 32 included studies, 28 were observational studies and 4 were randomized controlled trials (RCTs). Description of the objective for each study, the target disease, the country of origin, name of the app, and the type of data collected by the app of each RCT [16-21] and observational studies [22-47] is detailed in [Multimedia Appendix 3](#) [48].

### Target Disease of Mobile Health Apps, Country of Origin, and Funding Sources

Out of the 32 included articles, 13 (40%) referred to an app linked to a connected device. Most of the apps (26/32, 81%) were designed for the use of patients living with a specific rheumatic disease, distributed as follows: RA (11/32, 34%), fibromyalgia (5/32, 15%), juvenile idiopathic arthritis (4/32, 12%), OA (3/32, 9%), psoriatic arthritis and ankylosing spondylitis (AS; 1/32, 3%), spine disease (1/32, 3%), and ankle sprain (1/32, 3%). The other apps (6/32, 18%) were either designed for multiple diseases (eg, AS, RA, and systemic lupus erythematosus) or for the general population but used by patients living with RMDs in some studies. The great majority of the apps were developed in the United States (11/32, 34%), Japan (4/32, 12%), Canada (3/32, 9%), and Norway (3/32, 9%).

The funding sources were cited in 93% (30/32) of the articles and reported to be private in 34% (11/32) of the cases. The funding source was reported to be academic in 62% (20/32) of the articles.

### Purposes and Data Collected by Mobile Health Apps

Most of the apps were designed for self-monitoring and collection of specific outcome measures (22/32, 68%), the latter including patient-reported outcome measures (pain, fatigue, sleep, mood, and global well-being) and disease activity scores. Furthermore, many allowed self-visualization of the health data as a trend (17/22, 77%), such as disease activity scores and physical activity (measured by the number of steps). A few apps (7/32, 21%) aimed to promote physical activity through daily reminders and education on physical activity programs. For instance, 2 apps were designed to support coping mechanisms around pain management with relaxation therapy. Finally, 2 apps were designed primarily to help medication adherence through a tick-box option on the app when the medication is taken or through sending daily reminders with the possibility for the patient to edit the frequency of the reminders. None of the apps were reported as having the status of medical device.

Most of the apps addressed multiple disease features (detailed in [Table 1](#)).

**Table 1.** Features addressed by the different apps for rheumatoid arthritis and other rheumatic and musculoskeletal diseases.

Features addressed by the apps	Rheumatoid arthritis apps	Apps designed for other rheumatic and musculoskeletal diseases	
	Value, n (%)	Value, n (%)	Details on other diseases
Pain	7 (22)	13 (41)	JIA <sup>a</sup> , OA <sup>b</sup> , fibromyalgia, PsA/AS <sup>c</sup> , and spine disease
Fatigue	4 (13)	9 (28)	JIA, OA, and fibromyalgia
Physical activity	2 (6)	9 (28)	JIA and OA
Sleep	1 (3)	8 (25)	JIA, OA, and fibromyalgia
Disease Activity Score	8 (25)	1 (3)	PsA/AS
Health Assessment Questionnaire	6 (19)	2 (6)	JIA and PsA/AS
Mood	0 (0)	6 (19)	JIA, OA, and fibromyalgia
Global well-being (Short Form 36)	1 (3)	6 (19)	OA
Morning stiffness	1 (3)	4 (13)	JIA and PsA/AS
Depression/anxiety	1 (3)	2 (6)	Fibromyalgia
Medication/adherence	1 (3)	3 (9)	JIA, OA, and fibromyalgia
Tender joint count	1 (3)	3 (9)	JIA
Gait	4 (13)	0 (0)	— <sup>d</sup>
Social support	0 (0)	2 (6)	JIA
Work	0 (0)	2 (6)	JIA
Grip	1 (3)	0 (0)	—

<sup>a</sup>JIA: juvenile idiopathic arthritis.

<sup>b</sup>OA: osteoarthritis.

<sup>c</sup>PsA/AS: psoriatic arthritis/ankylosing spondyloarthritis.

<sup>d</sup>Not applicable.

### Development Process of the Apps

The development process of the app was not described at all in 9% (3/32) of the studies (Table 2). Only 15% (5/32) of articles stated that patients were included in the development of the apps. A qualitative phase occurred in only 18% (6/32) of the cases [19,27,30,33,36,37]. This qualitative phase consisted of

individual interviews (4 different studies), patient focus group (1 study), or patients focus groups and individual interviews (1 study). A mixed method approach was undertaken in 2 of those studies, with the addition of a patient survey or a Delphi procedure.

Health professionals and/or physicians were involved in the development or evaluation phase in 40% (13/32) of the studies.

**Table 2.** Description of the development phase and funding sources of the apps.

Reference	Quality of design and evaluation phase	Health care provider involvement	Patient involvement	Qualitative phase	Funding/development
Li et al, 2017 [16]	N/A <sup>a</sup>	N/A	N/A	N/A	Private
Skrepnik et al, 2017 [17]	B <sup>b</sup>	- <sup>c</sup>	+ <sup>d</sup>	-	Private
Kristjansdottir et al, 2013 [18]	A <sup>e</sup>	-	+	+	Public
Kristjansdottir et al, 2013 [19]	A	-	+	+	Public
Paxton et al, 2018 [20]	B	-	+	-	Private
Harmelink et al, 2017 [21]	B	-	+	-	Private
Bults et al, 2010 [22]	B	-	+	-	Public
Kvien et al, 2005 [23]	B	-	+	-	Public
Heiberg et al, 2007 [24]	B	-	+	-	Not reported
Richter et al, 2008 [25]	B	+	+	-	Public
Stinson et al, 2008 [26]	A	+	+	+	Public
Stinson et al, 2006 [27]	A	+	+	+	Public
Garcia-Palacios et al, 2014 [28]	B	+	+	-	Public
Nishiguchi et al, 2016 [29]	B	-	+	-	Public
de la Vega et al, 2018 [30]	A	+	+	+	Public
Salaffi et al, 2013 [31]	B	-	+	-	Private
Yen et al, 2016 [32]	B	+	+	-	Public
Khurana et al, 2016 [33]	A	-	+	+	Private
Kim et al, 2016 [34]	B	-	+	-	Public
Cai et al, 2017 [35]	A	+	+	+	Public
Revenas et al, 2015 [36]	A	+	+	+	Public and private
Synnott et al, 2015 [37]	B	+	+	-	Public
Bromberg et al, 2016 [38]	B	-	+	-	Public
Bromberg et al, 2014 [39]	B	-	+	-	Public
Okifuji et al, 2011 [40]	B	-	+	-	Private
Nishiguchi et al, 2014 [41]	B	+	+	-	Public
Shinohara et al, 2013 [42]	B	-	+	-	Not reported
Walker et al, 2017 [43]	B	+	+	-	Private
Yamada et al, 2012 [44]	B	-	+	-	Public
Espinoza et al, 2016 [45]	N/A	N/A	N/A	N/A	Private
Kim et al, 2016 [46]	B	+	+	-	Public
Twiggs et al, 2018 [47]	N/A	N/A	N/A	N/A	Private

<sup>a</sup>Not applicable.

<sup>b</sup>Patients or health care providers involved in the evaluation but not in the design process.

<sup>c</sup>Absent.

<sup>d</sup>Present.

<sup>e</sup>Patients or health care providers involved in both design and evaluation phases.

### Evaluation of the Apps

Physicians were rarely involved in app evaluation (5/32, 15%). Patients were more frequently involved in app evaluation but mostly indirectly through their adherence to the app (12/32,

70%). A total of 17 apps proposed a direct evaluation through the use of a satisfaction scale (12/32, 70%) and/or an open or closed questionnaire (9/32, 52%). Satisfaction scores and comments were generally positive.

## Commercial Availability of the Apps

Moreover, 16 (50%, 16/32) articles included in this review were published between 2016 and 2017. Only a few apps described in the publications were commercially available (4/32, 12%) at the time of our web-based search; all were free of cost.

## Quality of the Available Apps

We performed quality score on the apps, which were existing in the different online stores (iTunes and Google Play) using a validated app quality score, the Mobile Application Rating Scale (MARS) [48]. The MARS includes different quality subscale scores, which rate engagement, functionality, aesthetics, and information. The score was calculated for the available apps. MARS for the ReaApp (Google Play) was 3.01 out of 5 and 3.8 out of 5 for the iGetBetter app (iTunes). The other apps were either not available online (28/32) or not in English language (1/32) or not accessible for free (1/32).

## Effectiveness of the Apps

A total of 2 RCTs were included in the SLR, 1 study in OA [17] and 1 in fibromyalgia [18,19]. In the trial by Skrepnik et al [17], patients were randomized in 2 groups: a mobile OA app along with a wearable activity monitor with (intervention group) or without feedback (control group). A significant increase in the number of steps per day (1199 vs 467,  $P=.03$ ) as well as a reduction in pain from baseline during the 6-min walk test was shown in the intervention group. The trial by Kristjandottir et al [19] included a smartphone-delivered intervention with diaries and personalized feedback to patient living with fibromyalgia. The primary endpoint was met with a reduction of catastrophizing score in the treated group (mean 9.20, SD 5.85) compared with the control group (mean 15.71, SD 9.11;  $P<.001$ ).

## Discussion

### Principal Findings

To our knowledge, this is the first piece of work to identify literature published on self-management mHealth apps for patients living with RMDs. Our search yielded heterogeneous studies, referring to heterogeneous apps designed either for a specific rheumatic disease or for multiple diseases and also for the general public/healthy population. The large quantity and variety of information that was collected by the app and the relevance of collecting this information as part of a self-management initiative are questionable and not always clearly outlined.

The development process of most apps has been insufficient or not described in the screened existing literature, which raises questions around their credibility. Importantly, most of the apps, despite being designed for patient use, involved neither patients nor health care providers in their development phase. These findings are in line with recently published work, showing that health care professionals were involved in only 35% ( $n=7$ ) of the apps designed for RA patients [49]. This is a major concern as the absence of involvement of the relevant stakeholders might lead to inappropriate development tailored for the eventual user, including lack of assurance of content approval by specialists.

Our results highlight the unmet need for a standardization process to facilitate the convoluted and demanding processes required to develop mHealth apps. This matter goes beyond the scope of self-management apps for patients. For instance, Buijink et al [50] showed in a previous study that most mHealth apps designed for health care professionals were lacking authenticity details; authors, manufacturers, and distributors were not listed; and references were unavailable or out of date. Indeed, as mHealth apps are considered as medical devices by the US Food and Drug Administration, they should be subject to rigorous regulation [50]. These findings are consistent with our own findings and observations; indeed, the provenance and design process of the app as well as developers and funding sources are lacking details in many papers. Indeed, funding sources were not cited in more than half of the studies, which makes it difficult to identify clearly the nature of beneficiaries of such apps.

Moreover, despite more than half of the studies included in this review being published in the past 2 years, only a handful of apps were commercially available at the time of our search. This highlights the high turnover of apps developed for this purpose. One speculation could be that some of the apps were used in only 1 center and therefore have never been available for public use.

### Strengths and Limitations

It should be noted that our systematic review did not specifically address the question of reviewing apps existing in the Apple store. We focused on published literature on self-management mHealth apps, regardless of the type of RMDs. The very heterogeneous nature of the literature published on this topic and the relatively low number of relevant publications on the subject constitute the main limitations of this study. To ensure a most informative literature review, aside from systematically exploring the literature and other possible sources of information, we extracted data from the apps' content and development procedure when it was provided. The latter revealed that many apps were usually focusing on selected aspects of disease self-management. Most were giving the patients the opportunity to enter selected PROs, especially fatigue, pain, and sleep, and also collected information on physical activity. Disease activity scores were more rarely (28.2%) collected by the apps. No relevant data were available on the quantitative use of apps.

Our work is in line with a recent study by Grainger et al who assessed the quality of RA apps specifically, highlighting the fact that of the 19 apps analyzed, only 1 had functionality to allow both the calculation of a validated composite disease activity measure and the ability to track calculated patient data [51]. Another recent work showed that most of the apps designed for patients living with RA did not offer a comprehensive experience. Comparably with what we found in the literature, not all apps (75%) offered a symptom tracking experience, and when it was the case, only a few apps allowed collecting PROs, joint counts, and laboratory results [49].

Thinking forward, a development process under specific guidelines or recommendations is mandatory to improve physicians' as well as patients' confidence in future apps.



Having a meticulous development process in place can enhance appropriateness for the specific purpose they have been designed for, their scientific content and accuracy. By regulating the development process, the health care providers will also validate reliability of the scientific content and regulatory rules to ensure data protection and patient safety.

### Conclusions

In conclusion, despite patient willingness to use mHealth apps for self-management of their RMDs, better endeavors are needed

to provide an optimal standard and ensure the quality and safety of new apps. This work will be used to further inform European League Against Rheumatism (EULAR) points to consider for development, evaluation, and implementation of mHealth apps for self-management of RMDs by patients. We hope through this work to stimulate some careful considerations around mHealth app development and evaluation, which will lead to a general effort to improve their value.

### Acknowledgments

The authors would like to thank their task force members: Loreto Carmona, Gerd Burmester, Axel Finckh, Annamaria Iagnocco, Rik Lories, Francisca Sivera, Maxime Dougados, Zoltan Szekanez, Marie Kostine, John Pauling, Christophe Richez, Sofia Ramiro, Dieter Wiek, Petra Balázová, Simon Stones, Valentin Ritschl, and Yeliz Prior. EULAR funded this project (project number CL1102).

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

Search strategy for the hierarchical systematic literature review on mhealth Apps for disease self-management in patients with rheumatic and musculoskeletal diseases.

[[PDF File \(Adobe PDF File\), 81 KB - mhealth\\_v7i11e14730\\_app1.pdf](#)]

#### Multimedia Appendix 2

Venn Diagram demonstrating an example of article retrieval numbers using the largest database (Pubmed). N=435 represents the overlapping papers across three key domains searched.

[[PDF File \(Adobe PDF File\), 29 KB - mhealth\\_v7i11e14730\\_app2.pdf](#)]

#### Multimedia Appendix 3

Summary table of all 32 articles included in the review and type and purposes of studies included.

[[PDF File \(Adobe PDF File\), 103 KB - mhealth\\_v7i11e14730\\_app3.pdf](#)]

### References

1. Cisco C. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022 White Paper URL: [http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white\\_paper\\_c11-520862.html](http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html) [accessed 2019-09-10]
2. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. *J Med Syst* 2016 Mar;40(3):72. [doi: [10.1007/s10916-015-0425-x](https://doi.org/10.1007/s10916-015-0425-x)] [Medline: [26748792](https://pubmed.ncbi.nlm.nih.gov/26748792/)]
3. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. *J Med Internet Res* 2016 May 16;18(5):e97 [FREE Full text] [doi: [10.2196/jmir.4883](https://doi.org/10.2196/jmir.4883)] [Medline: [27185295](https://pubmed.ncbi.nlm.nih.gov/27185295/)]
4. Hartz J, Yingling L, Powell-Wiley TM. Use of mobile health technology in the prevention and management of diabetes mellitus. *Curr Cardiol Rep* 2016 Dec;18(12):130. [doi: [10.1007/s11886-016-0796-8](https://doi.org/10.1007/s11886-016-0796-8)] [Medline: [27826901](https://pubmed.ncbi.nlm.nih.gov/27826901/)]
5. Bashshur R, Shannon G, Krupinski E, Grigsby J. The taxonomy of telemedicine. *Telemed J E Health* 2011;17(6):484-494. [doi: [10.1089/tmj.2011.0103](https://doi.org/10.1089/tmj.2011.0103)] [Medline: [21718114](https://pubmed.ncbi.nlm.nih.gov/21718114/)]
6. Alessa T, Abdi S, Hawley MS, de Witte L. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, usability, and user satisfaction. *JMIR Mhealth Uhealth* 2018 Jul 23;6(7):e10723 [FREE Full text] [doi: [10.2196/10723](https://doi.org/10.2196/10723)] [Medline: [30037787](https://pubmed.ncbi.nlm.nih.gov/30037787/)]
7. Farzandipour M, Nabovati E, Sharif R, Arani MH, Anvari S. Patient self-management of asthma using mobile health applications: a systematic review of the functionalities and effects. *Appl Clin Inform* 2017 Oct;8(4):1068-1081 [FREE Full text] [doi: [10.4338/ACI-2017-07-R-0116](https://doi.org/10.4338/ACI-2017-07-R-0116)] [Medline: [29241254](https://pubmed.ncbi.nlm.nih.gov/29241254/)]
8. Oh B, Yi GH, Han MK, Kim JS, Lee CH, Cho B, et al. Importance of active participation in obesity management through mobile health care programs: substudy of a randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Jan 3;6(1):e2 [FREE Full text] [doi: [10.2196/mhealth.8719](https://doi.org/10.2196/mhealth.8719)] [Medline: [29298749](https://pubmed.ncbi.nlm.nih.gov/29298749/)]

9. Niendam TA, Tully LM, Iosif AM, Kumar D, Nye KE, Denton JC, et al. Enhancing early psychosis treatment using smartphone technology: a longitudinal feasibility and validity study. *J Psychiatr Res* 2018 Jan;96:239-246. [doi: [10.1016/j.jpsychires.2017.10.017](https://doi.org/10.1016/j.jpsychires.2017.10.017)] [Medline: [29126059](https://pubmed.ncbi.nlm.nih.gov/29126059/)]
10. Gray CS, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The electronic patient reported outcome tool: testing usability and feasibility of a mobile app and portal to support care for patients with complex chronic disease and disability in primary care settings. *JMIR Mhealth Uhealth* 2016 Jun 2;4(2):e58 [FREE Full text] [doi: [10.2196/mhealth.5331](https://doi.org/10.2196/mhealth.5331)] [Medline: [27256035](https://pubmed.ncbi.nlm.nih.gov/27256035/)]
11. Palazzo C, Ravaud JF, Papelard A, Ravaud P, Poiraudau S. The burden of musculoskeletal conditions. *PLoS One* 2014;9(3):e90633 [FREE Full text] [doi: [10.1371/journal.pone.0090633](https://doi.org/10.1371/journal.pone.0090633)] [Medline: [24595187](https://pubmed.ncbi.nlm.nih.gov/24595187/)]
12. Dixon WG, Michaud K. Using technology to support clinical care and research in rheumatoid arthritis. *Curr Opin Rheumatol* 2018 May;30(3):276-281 [FREE Full text] [doi: [10.1097/BOR.0000000000000485](https://doi.org/10.1097/BOR.0000000000000485)] [Medline: [29369089](https://pubmed.ncbi.nlm.nih.gov/29369089/)]
13. Azevedo R, Bernardes M, Fonseca J, Lima A. Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. *Rheumatol Int* 2015 Oct;35(10):1675-1685. [doi: [10.1007/s00296-015-3270-9](https://doi.org/10.1007/s00296-015-3270-9)] [Medline: [25903352](https://pubmed.ncbi.nlm.nih.gov/25903352/)]
14. 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](https://doi.org/10.1371/journal.pmed.1000097)] [Medline: [19621072](https://pubmed.ncbi.nlm.nih.gov/19621072/)]
15. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 1995;123(3):A12-A13. [doi: [10.7326/ACPJC-1995-123-3-A12](https://doi.org/10.7326/ACPJC-1995-123-3-A12)] [Medline: [7582737](https://pubmed.ncbi.nlm.nih.gov/7582737/)]
16. Li LC, Sayre EC, Xie H, Clayton C, Feehan LM. A community-based physical activity counselling program for people with knee osteoarthritis: feasibility and preliminary efficacy of the Track-OA study. *JMIR Mhealth Uhealth* 2017 Jun 26;5(6):e86 [FREE Full text] [doi: [10.2196/mhealth.7863](https://doi.org/10.2196/mhealth.7863)] [Medline: [28652228](https://pubmed.ncbi.nlm.nih.gov/28652228/)]
17. Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R. Assessing the impact of a novel smartphone application compared with standard follow-up on mobility of patients with knee osteoarthritis following treatment with Hylan G-F 20: a randomized controlled trial. *JMIR Mhealth Uhealth* 2017 May 9;5(5):e64 [FREE Full text] [doi: [10.2196/mhealth.7179](https://doi.org/10.2196/mhealth.7179)] [Medline: [28487266](https://pubmed.ncbi.nlm.nih.gov/28487266/)]
18. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain. Part 2: 11-month follow-up results of a randomized trial. *J Med Internet Res* 2013 Mar 28;15(3):e72 [FREE Full text] [doi: [10.2196/jmir.2442](https://doi.org/10.2196/jmir.2442)] [Medline: [23538392](https://pubmed.ncbi.nlm.nih.gov/23538392/)]
19. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. *J Med Internet Res* 2013 Jan 7;15(1):e5 [FREE Full text] [doi: [10.2196/jmir.2249](https://doi.org/10.2196/jmir.2249)] [Medline: [23291270](https://pubmed.ncbi.nlm.nih.gov/23291270/)]
20. Paxton RJ, Forster JE, Miller MJ, Geron KL, Stevens-Lapsley JE, Christiansen CL. A feasibility study for improved physical activity after total knee arthroplasty. *J Aging Phys Act* 2018 Jan 1;26(1):7-13 [FREE Full text] [doi: [10.1123/japa.2016-0268](https://doi.org/10.1123/japa.2016-0268)] [Medline: [28338406](https://pubmed.ncbi.nlm.nih.gov/28338406/)]
21. Harmelink KE, Zeegers AV, Tönis TM, Hullegie W, Nijhuis-van der Sanden MW, Staal JB. The effectiveness of the use of a digital activity coaching system in addition to a two-week home-based exercise program in patients after total knee arthroplasty: study protocol for a randomized controlled trial. *BMC Musculoskelet Disord* 2017 Jul 5;18(1):290 [FREE Full text] [doi: [10.1186/s12891-017-1647-5](https://doi.org/10.1186/s12891-017-1647-5)] [Medline: [28679400](https://pubmed.ncbi.nlm.nih.gov/28679400/)]
22. Bults RG, Knoppel DF, Widya IA, Schaake L, Hermens HJ. The myofeedback-based teletreatment system and its evaluation. *J Telemed Telecare* 2010;16(6):308-315. [doi: [10.1258/jtt.2010.006004](https://doi.org/10.1258/jtt.2010.006004)] [Medline: [20798424](https://pubmed.ncbi.nlm.nih.gov/20798424/)]
23. Kvien TK, Mowinckel P, Heiberg T, Dammann KL, Dale Ø, Aanerud GJ, et al. Performance of health status measures with a pen based personal digital assistant. *Ann Rheum Dis* 2005 Oct;64(10):1480-1484 [FREE Full text] [doi: [10.1136/ard.2004.030437](https://doi.org/10.1136/ard.2004.030437)] [Medline: [15843456](https://pubmed.ncbi.nlm.nih.gov/15843456/)]
24. Heiberg T, Kvien TK, Dale Ø, Mowinckel P, Aanerud GJ, Songe-Møller AB, et al. Daily health status registration (patient diary) in patients with rheumatoid arthritis: a comparison between personal digital assistant and paper-pencil format. *Arthritis Rheum* 2007 Apr 15;57(3):454-460 [FREE Full text] [doi: [10.1002/art.22613](https://doi.org/10.1002/art.22613)] [Medline: [17394232](https://pubmed.ncbi.nlm.nih.gov/17394232/)]
25. Richter JG, Becker A, Koch T, Nixdorf M, Willers R, Monser R, et al. Self-assessments of patients via Tablet PC in routine patient care: comparison with standardised paper questionnaires. *Ann Rheum Dis* 2008 Dec;67(12):1739-1741. [doi: [10.1136/ard.2008.090209](https://doi.org/10.1136/ard.2008.090209)] [Medline: [18647853](https://pubmed.ncbi.nlm.nih.gov/18647853/)]
26. Stinson JN, Petroz GC, Stevens BJ, Feldman BM, Streiner D, McGrath PJ, et al. Working out the kinks: testing the feasibility of an electronic pain diary for adolescents with arthritis. *Pain Res Manag* 2008;13(5):375-382 [FREE Full text] [doi: [10.1155/2008/326389](https://doi.org/10.1155/2008/326389)] [Medline: [18958308](https://pubmed.ncbi.nlm.nih.gov/18958308/)]
27. Stinson JN, Petroz GC, Tait G, Feldman BM, Streiner D, McGrath PJ, et al. e-Ouch: usability testing of an electronic chronic pain diary for adolescents with arthritis. *Clin J Pain* 2006;22(3):295-305. [doi: [10.1097/01.ajp.0000173371.54579.31](https://doi.org/10.1097/01.ajp.0000173371.54579.31)] [Medline: [16514331](https://pubmed.ncbi.nlm.nih.gov/16514331/)]

28. Garcia-Palacios A, Herrero R, Belmonte MA, Castilla D, Guixeres J, Molinari G, et al. Ecological momentary assessment for chronic pain in fibromyalgia using a smartphone: a randomized crossover study. *Eur J Pain* 2014 Jul;18(6):862-872. [doi: [10.1002/j.1532-2149.2013.00425.x](https://doi.org/10.1002/j.1532-2149.2013.00425.x)] [Medline: [24921074](https://pubmed.ncbi.nlm.nih.gov/24921074/)]
29. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment of rheumatoid arthritis disease activity using a smartphone application. Development and 3-month feasibility study. *Methods Inf Med* 2016;55(1):65-69. [doi: [10.3414/ME14-01-0106](https://doi.org/10.3414/ME14-01-0106)] [Medline: [26391694](https://pubmed.ncbi.nlm.nih.gov/26391694/)]
30. de la Vega R, Roset R, Galán S, Miró J. Fibroline: a mobile app for improving the quality of life of young people with fibromyalgia. *J Health Psychol* 2018 Jan;23(1):67-78. [doi: [10.1177/1359105316650509](https://doi.org/10.1177/1359105316650509)] [Medline: [27225285](https://pubmed.ncbi.nlm.nih.gov/27225285/)]
31. Salaffi F, Gasparini S, Ciapetti A, Gutierrez M, Grassi W. Usability of an innovative and interactive electronic system for collection of patient-reported data in axial spondyloarthritis: comparison with the traditional paper-administered format. *Rheumatology (Oxford)* 2013 Nov;52(11):2062-2070. [doi: [10.1093/rheumatology/ket276](https://doi.org/10.1093/rheumatology/ket276)] [Medline: [23955646](https://pubmed.ncbi.nlm.nih.gov/23955646/)]
32. Yen PY, Lara B, Lopetegui M, Bharat A, Ardoin S, Johnson B, et al. Usability and workflow evaluation of 'rheumatic disease activity' (READY). A mobile application for rheumatology patients and providers. *Appl Clin Inform* 2016 Nov 2;7(4):1007-1024 [FREE Full text] [doi: [10.4338/ACI-2016-03-RA-0036](https://doi.org/10.4338/ACI-2016-03-RA-0036)] [Medline: [27803949](https://pubmed.ncbi.nlm.nih.gov/27803949/)]
33. Khurana L, Durand EM, Gary ST, Otero AV, Gerzon MC, Beck J, et al. Subjects with osteoarthritis can easily use a handheld touch screen electronic device to report medication use: qualitative results from a usability study. *Patient Prefer Adherence* 2016;10:2171-2179 [FREE Full text] [doi: [10.2147/PPA.S94247](https://doi.org/10.2147/PPA.S94247)] [Medline: [27822018](https://pubmed.ncbi.nlm.nih.gov/27822018/)]
34. Kim CH, Chung CK, Choi Y, Shin H, Woo JW, Kim S, et al. The usefulness of a mobile device-based system for patient-reported outcomes in a spine outpatient clinic. *Spine J* 2016 Jul;16(7):843-850. [doi: [10.1016/j.spinee.2016.02.048](https://doi.org/10.1016/j.spinee.2016.02.048)] [Medline: [26961200](https://pubmed.ncbi.nlm.nih.gov/26961200/)]
35. Cai RA, Beste D, Chaplin H, Varakliotis S, Suffield L, Josephs F, et al. Developing and evaluating JIApp: acceptability and usability of a smartphone app system to improve self-management in young people with juvenile idiopathic arthritis. *JMIR Mhealth Uhealth* 2017 Aug 15;5(8):e121 [FREE Full text] [doi: [10.2196/mhealth.7229](https://doi.org/10.2196/mhealth.7229)] [Medline: [28811270](https://pubmed.ncbi.nlm.nih.gov/28811270/)]
36. 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 9;4(1):e22 [FREE Full text] [doi: [10.2196/resprot.3795](https://doi.org/10.2196/resprot.3795)] [Medline: [25665589](https://pubmed.ncbi.nlm.nih.gov/25665589/)]
37. Synnott J, Pedlow K, Bleakley C, Davies R, Nugent C, Moral-Munoz J, et al. ReApp – A Mobile App for the Rehabilitation of Ankle Sprains. In: Proceedings of the Conference on Ambient Intelligence for Health. 2015 Presented at: AmIHEALTH'15; December 1-4, 2015; Puerto Varas, Chile p. 61-67. [doi: [10.1007/978-3-319-26508-7\\_6](https://doi.org/10.1007/978-3-319-26508-7_6)]
38. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Prospective mediation models of sleep, pain, and daily function in children with arthritis using ecological momentary assessment. *Clin J Pain* 2016 Jun;32(6):471-477 [FREE Full text] [doi: [10.1097/AJP.0000000000000298](https://doi.org/10.1097/AJP.0000000000000298)] [Medline: [26340651](https://pubmed.ncbi.nlm.nih.gov/26340651/)]
39. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Self-reported pain and disease symptoms persist in juvenile idiopathic arthritis despite treatment advances: an electronic diary study. *Arthritis Rheumatol* 2014 Feb;66(2):462-469 [FREE Full text] [doi: [10.1002/art.38223](https://doi.org/10.1002/art.38223)] [Medline: [24504820](https://pubmed.ncbi.nlm.nih.gov/24504820/)]
40. Okifuji A, Bradshaw DH, Donaldson GW, Turk DC. Sequential analyses of daily symptoms in women with fibromyalgia syndrome. *J Pain* 2011 Jan;12(1):84-93 [FREE Full text] [doi: [10.1016/j.jpain.2010.05.003](https://doi.org/10.1016/j.jpain.2010.05.003)] [Medline: [20591745](https://pubmed.ncbi.nlm.nih.gov/20591745/)]
41. Nishiguchi S, Ito H, Yamada M, Yoshitomi H, Furu M, Ito T, et al. Self-assessment tool of disease activity of rheumatoid arthritis by using a smartphone application. *Telemed J E Health* 2014 Mar;20(3):235-240. [doi: [10.1089/tmj.2013.0162](https://doi.org/10.1089/tmj.2013.0162)] [Medline: [24404820](https://pubmed.ncbi.nlm.nih.gov/24404820/)]
42. Shinohara A, Ito T, Ura T, Nishiguchi S, Ito H, Yamada M, et al. Development of Lifelog Sharing System for Rheumatoid Arthritis Patients Using Smartphone. In: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2013 Presented at: EMBC'13; July 3-7, 2013; Osaka, Japan p. 7266-7269. [doi: [10.1109/EMBC.2013.6611235](https://doi.org/10.1109/EMBC.2013.6611235)]
43. Walker UA, Mueller RB, Jaeger VK, Theiler R, Forster A, Dufner P, et al. Disease activity dynamics in rheumatoid arthritis: patients' self-assessment of disease activity via WebApp. *Rheumatology (Oxford)* 2017 Oct 1;56(10):1707-1712 [FREE Full text] [doi: [10.1093/rheumatology/kex229](https://doi.org/10.1093/rheumatology/kex229)] [Medline: [28957553](https://pubmed.ncbi.nlm.nih.gov/28957553/)]
44. Yamada M, Aoyama T, Mori S, Nishiguchi S, Okamoto K, Ito T, et al. Objective assessment of abnormal gait in patients with rheumatoid arthritis using a smartphone. *Rheumatol Int* 2012 Dec;32(12):3869-3874. [doi: [10.1007/s00296-011-2283-2](https://doi.org/10.1007/s00296-011-2283-2)] [Medline: [22193221](https://pubmed.ncbi.nlm.nih.gov/22193221/)]
45. Espinoza F, le Blay P, Coulon D, Lieu S, Munro J, Jorgensen C, et al. Handgrip strength measured by a dynamometer connected to a smartphone: a new applied health technology solution for the self-assessment of rheumatoid arthritis disease activity. *Rheumatology (Oxford)* 2016 May;55(5):897-901. [doi: [10.1093/rheumatology/kew006](https://doi.org/10.1093/rheumatology/kew006)] [Medline: [26867731](https://pubmed.ncbi.nlm.nih.gov/26867731/)]
46. Kim K, Pham D, Schwarzkopf R. Mobile application use in monitoring patient adherence to perioperative total knee arthroplasty protocols. *Surg Technol Int* 2016 Apr;28:253-260. [Medline: [27042789](https://pubmed.ncbi.nlm.nih.gov/27042789/)]
47. Twiggs J, Salmon L, Kolos E, Bogue E, Miles B, Roe J. Measurement of physical activity in the pre- and early post-operative period after total knee arthroplasty for osteoarthritis using a Fitbit Flex device. *Med Eng Phys* 2018 Jan;51:31-40. [doi: [10.1016/j.medengphy.2017.10.007](https://doi.org/10.1016/j.medengphy.2017.10.007)] [Medline: [29117912](https://pubmed.ncbi.nlm.nih.gov/29117912/)]

48. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
49. Luo D, Wang P, Lu F, Elias J, Sparks JA, Lee YC. Mobile apps for individuals with rheumatoid arthritis: a systematic review. *J Clin Rheumatol* 2019 Apr;25(3):133-141. [doi: [10.1097/RHU.0000000000000800](https://doi.org/10.1097/RHU.0000000000000800)] [Medline: [29933327](https://pubmed.ncbi.nlm.nih.gov/29933327/)]
50. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med* 2013 Jun;18(3):90-92. [doi: [10.1136/eb-2012-100885](https://doi.org/10.1136/eb-2012-100885)] [Medline: [22923708](https://pubmed.ncbi.nlm.nih.gov/22923708/)]
51. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for people with rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. *JMIR Mhealth Uhealth* 2017 Feb 21;5(2):e7 [FREE Full text] [doi: [10.2196/mhealth.6956](https://doi.org/10.2196/mhealth.6956)] [Medline: [28223263](https://pubmed.ncbi.nlm.nih.gov/28223263/)]

## Abbreviations

**AS:** ankylosing spondylitis  
**EULAR:** European League Against Rheumatism  
**MARS:** Mobile Application Rating Scale  
**mHealth:** mobile health  
**OA:** osteoarthritis  
**PICO:** population, intervention, control, and outcomes  
**PROs:** patient-reported outcomes  
**RA:** rheumatoid arthritis  
**RCT:** randomized controlled trial  
**RMDs:** rheumatic and musculoskeletal diseases  
**SLR:** systematic literature review

*Edited by G Eysenbach; submitted 16.05.19; peer-reviewed by J Sellam, A Alunno, DR Wahl; comments to author 02.07.19; revised version received 26.07.19; accepted 30.07.19; published 26.11.19.*

*Please cite as:*

*Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E  
Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review  
JMIR Mhealth Uhealth 2019;7(11):e14730  
URL: <https://mhealth.jmir.org/2019/11/e14730>  
doi: [10.2196/14730](https://doi.org/10.2196/14730)  
PMID: [31769758](https://pubmed.ncbi.nlm.nih.gov/31769758/)*

©Aurélien Najm, Laure Gossec, Catherine Weill, David Benoist, Francis Berenbaum, Elena Nikiphorou. Originally published in *JMIR mHealth and uHealth* (<http://mhealth.jmir.org>), 26.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mHealth and uHealth*, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Effects of a Smartphone-Based Approach-Avoidance Intervention on Chocolate Craving and Consumption: Randomized Controlled Trial

Adrian Meule<sup>1,2</sup>, Dipl-Psych, PhD; Anna Richard<sup>1</sup>, BSc, MSc, PhD; Radomir Dinic<sup>3</sup>, BSc, MSc; Jens Blechert<sup>4</sup>, Dipl-Psych, PhD

<sup>1</sup>Schoen Clinic Roseneck, Prien am Chiemsee, Germany

<sup>2</sup>Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany

<sup>3</sup>Department of MultiMediaTechnology, Salzburg University of Applied Sciences, Puch, Austria

<sup>4</sup>Department of Psychology, University of Salzburg, Salzburg, Austria

**Corresponding Author:**

Adrian Meule, Dipl-Psych, PhD  
Schoen Clinic Roseneck  
Am Roseneck 6  
Priens am Chiemsee, 83209  
Germany  
Phone: 49 8051 68 100 174  
Email: [ameule@med.lmu.de](mailto:ameule@med.lmu.de)

## Abstract

**Background:** Repeatedly pushing high-calorie food stimuli away based on joystick movements has been found to reduce approach biases toward these stimuli. Some studies also found that such avoidance training reduced consumption of high-calorie foods.

**Objective:** This study aimed to test effects of a smartphone-based approach-avoidance intervention on chocolate craving and consumption, to make such interventions suitable for daily use.

**Methods:** Within a 10-day period, regular chocolate eaters (n=105, 86% female) performed five sessions during which they continuously avoided (ie, swiped upward) chocolate stimuli (experimental group, n=35), performed five sessions during which they approached and avoided chocolate stimuli equally often (placebo control group, n=35), or did not perform any training sessions (inactive control group, n=35). Training effects were measured during laboratory sessions before and after the intervention period and further continuously through daily ecological momentary assessment.

**Results:** Self-reported chocolate craving and consumption as well as body fat mass significantly decreased from pre- to postmeasurement across all groups. Ecological momentary assessment reports evidenced no differences in chocolate craving and consumption between intervention days and rest days as a function of the group.

**Conclusions:** A smartphone-based approach-avoidance training did not affect eating-related and anthropometric measures over and above measurement-based changes in this study. Future controlled studies need to examine whether other techniques of modifying food approach tendencies show an add-on benefit over conventional, monitoring-based intervention effects.

**Trial Registration:** AsPredicted 8203; <https://aspredicted.org/pt9df.pdf>.

(*JMIR Mhealth Uhealth* 2019;7(11):e12298) doi:[10.2196/12298](https://doi.org/10.2196/12298)

**KEYWORDS**

food; chocolate; craving; smartphone; mobile phone; mhealth; digital health; eating behavior

## Introduction

Training individuals to avoid appetitive stimuli has been found to reduce automatic approach tendencies toward these stimuli. For example, repeatedly pushing pictures of alcoholic beverages

away on a screen based on joystick movements has been found to reduce approach biases toward alcohol in heavy drinkers [1] and patients with alcohol use disorder [2,3]. Similar results have been obtained using pictures of high-calorie foods in samples of high trait food cravers [4] or individuals with obesity [5-7].

Although effects on actual consumption behaviors is less consistent [8-10], several studies point toward a decrease in craving for and consumption of appetitive substances through approach-avoidance training [11].

While traditional approach-avoidance tasks (AATs) and training are usually performed with joystick movements in front of a computer monitor, methods that make these techniques suitable for daily use are needed. One possibility for this is to implement AATs or training on smartphones. For example, 2 recent studies used a smartphone-based training during which participants were required to swipe pictures away or toward themselves to reduce body dissatisfaction [12] or procrastination [13]. Although these studies reported promising results (ie, changes in behavior because of the approach-avoidance intervention), interpretation was limited by the use of inactive (waitlist) control groups and by combining the training with conventional face-to-face treatment elements.

The aim of this study was, therefore, to evaluate a smartphone-based approach-avoidance training for reducing food craving and consumption in a randomized, fully controlled trial (ie, by comparing active training effects to placebo and no training groups). As chocolate is the most frequently craved food in Western societies [14,15], we restricted our study to chocolate-containing foods, similar to previous studies on approach-avoidance modification [16-18]. Specifically, participants were randomly assigned to 1 of 3 groups: during a 10-day period, they either performed five training sessions during which they continuously avoided pictures of chocolate-containing foods (upward swipes) and approached pictures of neutral objects (downward swipes; *experimental group*), performed 5 training sessions during which they approached and avoided food and neutral stimuli equally often (*placebo control group*), or did not perform an approach-avoidance training (*inactive control group*). All participants completed an AAT and reported their craving for and consumption of chocolate-containing foods before and after the 10-day period. Furthermore, previous studies found short-term effects of approach-avoidance training on food consumption (eg, reduced chocolate muffin consumption in a taste test immediately after an avoidance training session [18]). To capture such short-lived effects, participants reported their craving for and consumption of chocolate-containing foods on each evening during the 10-day period. This allowed us to examine both short-term training effects by comparing chocolate craving and consumption on intervention versus rest days during the 10-day period and long-term training effects by comparing pre- versus posttest values before and after the 10-day period.

We tested the following preregistered hypotheses ([Multimedia Appendix 1](#)):

- (1) Similar to findings showing that an approach bias modification training decreased approach bias toward high-calorie foods [4], we expected that approach bias toward chocolate-containing foods would decrease from pre- to posttest only in the experimental group but not in the 2 control groups.
- (2) Similar to findings showing that self-monitoring of snacking decreases snack food consumption [19], we expected that self-reported chocolate craving and consumption in the past 10

days would decrease from pre- to posttest in all 3 groups, as all participants were confronted with their chocolate consumption behavior during the study. However, because of craving- and consumption-reducing effects of approach-avoidance training found in previous studies [4,18], we expected that these decreases would be larger in the experimental group than in the placebo control group and the inactive control group.

- (3) Performing reaction time tasks involving palatable food pictures usually increases food craving from immediately before to immediately after the task [20,21]. Therefore, we expected that performing a chocolate-related AAT would induce chocolate craving, that is, current chocolate craving would be increased immediately after having performed the task compared with before. At pretest, we expected that these chocolate craving increases during the task would be similar in all 3 groups. As previous findings indicate that approach-avoidance training can decrease such food cue-induced craving [4], we expected that task-induced chocolate craving would be attenuated at posttest in the experimental group but not in the inactive control group. As participants in the placebo control group were confronted with the chocolate pictures more often than participants in the inactive control group, we expected that the placebo group would show an attenuation of task-induced chocolate craving at posttest as well because of habituation. Finally, we hypothesized that current hunger would be unaffected by the intervention, that is, would be similar across groups and measurements.

- (4) Given that short-term effects on food consumption have been reported in approach bias modification studies (ie, reduced consumption after a training session [18]), we expected that chocolate craving and consumption would be reduced on intervention days compared with rest days in the experimental group, and this difference would be larger than in the placebo control group.

In addition to these preregistered analyses, we also explored changes in body mass index and body fat mass as a function of group, examined whether any effects were moderated by baseline levels of trait chocolate craving and restrained eating, and tested whether groups differed in awareness of the study's aims.

## Methods

### Participants

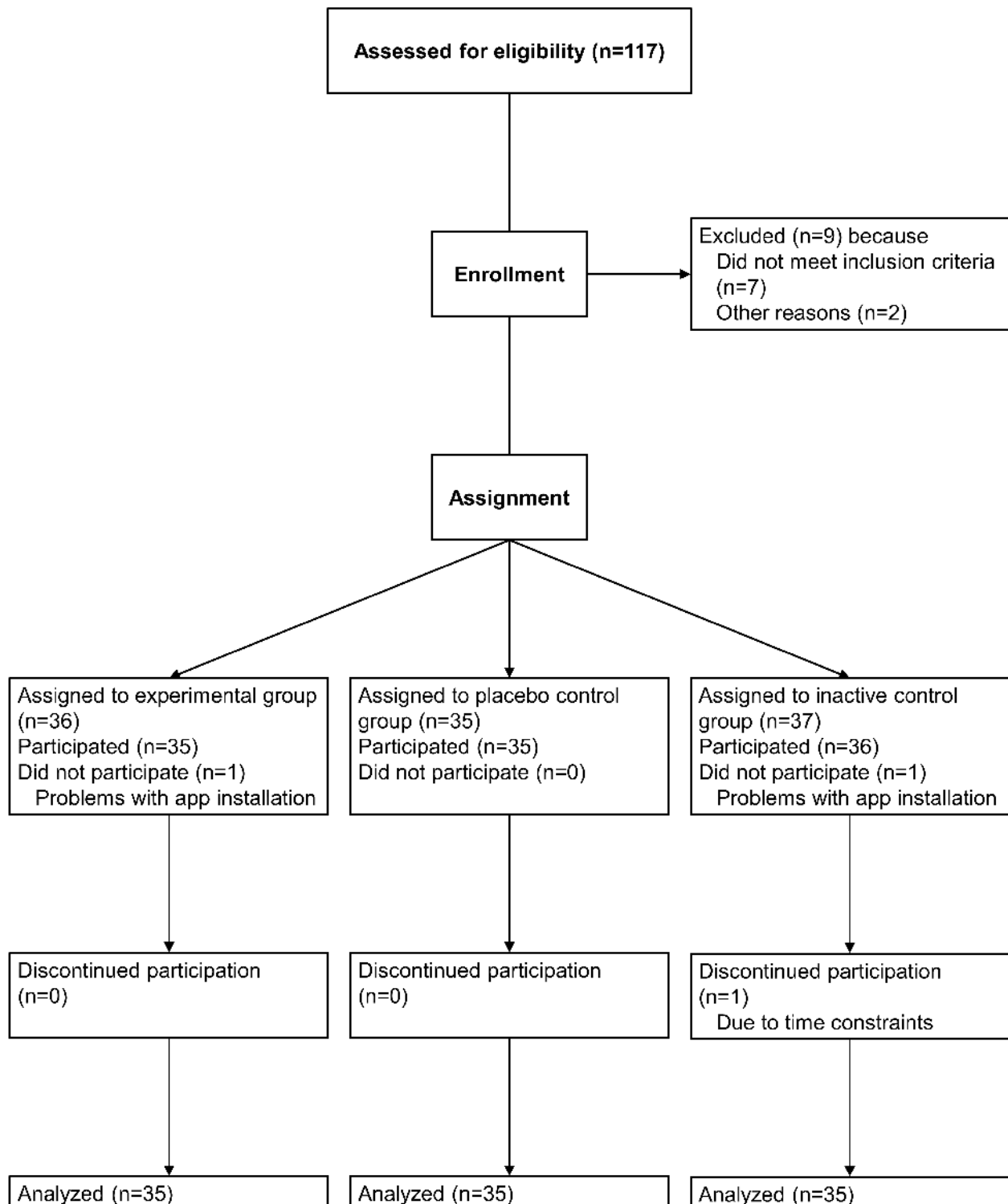
A power analysis was conducted with G\*Power version 3.1.9.2 [22] for repeated measures analysis of variance with a within-between interaction. This revealed that a sample size of 102 (ie, n=34 participants per group) would be sufficient to detect a small effect ( $f=0.1$ ), given an alpha level of .05, power of .80, 3 groups, 2 measurements, and a correlation of  $r=.80$  between repeated measures.

Participants were recruited at the University of Salzburg and through a local job advertisements website. Inclusion criteria were speaking fluent German, aged between 18 and 50 years, not being pregnant, and not having participated in similar studies in our laboratory. Recruitment advertisements also indicated that participants should be regular chocolate eaters (ie, several

times per week) and should not be underweight or currently dieting. A total of 117 individuals responded to the advertisements. A total of 9 participants were excluded before enrollment: 7 participants did not meet inclusion criteria (current pregnancy: n=1, non-German-speaking: n=2, already participated in similar studies in our laboratory: n=4), and 2 participants indicated that they recently decided to refrain from eating chocolate because of lactose intolerance and health

reasons (n=2; Figure 1). Of the remaining 108 individuals, 2 did not participate because of technical problems, and 1 discontinued participation (Figure 1). The final sample comprised 105 participants (85.7% female, 90/105) with a mean age of 23.4 years (SD 5.07) and a mean body mass index of 23.3 kg/m<sup>2</sup> (SD 4.14). The majority of participants had German (52.4%, 55/105) or Austrian (40.0%, 42/105) citizenship and were university students (94.3%, 99/105).

**Figure 1.** Flow of participants throughout the study. Note that while sample size was n=105 for the majority of analyses, sample size was n=104 for analyses involving body mass index at posttest and n=102 for analyses involving body fat mass at posttest because of missing data.



## Materials

### *Approach-Avoidance Task*

An AAT was employed to examine whether approach bias toward chocolate-containing foods changed from pre- to posttest as a function of group. The task was programmed in unity (Unity Technologies) and run on a 5-inch SAMSUNG Galaxy J3 smartphone (Samsung Electronics Austria GmbH). A total of 16 pictures of chocolate-containing foods and 16 pictures of nonedible objects were taken from the food pics database [23] (Multimedia Appendix 2). Pictures were matched regarding color, size, brightness, contrast, complexity, recognizability, and familiarity and have previously been used in a joystick-based AAT with which an approach bias toward food was found [24]. The task consisted of 2 blocks: participants were instructed to swipe pictures of food upward (=“away from yourself”) and swipe pictures of objects downward (=“toward yourself”) with the thumb of their dominant hand in 1 block and vice versa in the other block (block order was counterbalanced across participants). Within each block, each picture was presented twice in randomized order. Thus, participants pulled food, pushed food, pulled objects, and pushed objects in 32 trials each, totaling 128 trials. In each trial, 1 picture appeared in the center of the smartphone screen. Similar to joystick-based AATs [25], a zoom effect was employed: picture size increased when the picture was swiped downwards and decreased when the picture was swiped upwards. The picture then disappeared when reaching the border of the screen, and the next trial started (Multimedia Appendix 3).

### *Sociodemographic and Anthropometric Data*

Participants indicated their age, sex, handedness, education, and nationality. Body height (in cm) was measured with a wall-mounted stadiometer. Body weight (in kg) and fat mass (in %) were measured with the OMRON Body Composition Monitor BF511 (OMRON Healthcare Europe BV).

### *Chocolate Consumption*

To examine whether chocolate consumption changed from pre- to posttest as a function of group, participants responded to the question “How often did you consume chocolate-containing foods in the past ten days?” Responses were recorded on a rating slider anchored 0=*not at all* and 100=*very often*.

### *Food Cravings Questionnaire-Trait-Reduced*

The German, chocolate-adapted version of the Food Cravings Questionnaire-Trait-reduced (FCQ-T-r) [26] was used to examine whether groups differed at pretest, whether pretest scores moderated any intervention effects, and whether scores changed from pre- to posttest as a function of group. Participants are usually instructed to indicate how frequently each statement is true for them in general. However, to fit the purpose of this study, participants were instructed to indicate how frequently each statement was true for them in the past 10 days. The scale has 15 items (eg, “If I am craving chocolate, thoughts of eating it consume me” and “It is hard for me to resist the temptation to eat chocolate that is in my reach”), which are scored from 1=*never* to 6=*always*. Internal reliability was  $\alpha=.894$  at pretest and  $\alpha=.921$  at posttest.

### *Food Cravings Questionnaire-State*

The German, chocolate-adapted version of the Food Cravings Questionnaire-State (FCQ-S) [26] was used to measure current chocolate craving and hunger before and after the AAT. The scale has 15 items (12 items for the chocolate craving subscale and 3 items for the hunger subscale), which are scored from 1=*strongly disagree* to 5=*strongly agree*. Internal reliabilities of the chocolate craving subscale ranged between  $\alpha=.873$  and  $\alpha=.930$ , and internal reliabilities of the hunger subscale ranged between  $\alpha=.835$  and  $\alpha=.917$  in this study.

### *Restraint Scale*

The German version of the Restraint Scale [27] was used to examine whether groups differed in dietary restraint and whether dietary restraint moderated any intervention effects. The scale has 10 items, which are scored from 0 to 4 (items 1-4 and 10) and 0 to 3 (items 5-9) with different response options. Internal reliability was  $\alpha=.715$  in this study.

### *Dutch Eating Behavior Questionnaire*

The German version of the Dutch Behavior Questionnaire’s (DEBQ’s) restrained eating subscale [28] was used to examine whether groups differed in dietary restraint and whether dietary restraint moderated any intervention effects. The scale has 10 items, which are scored from 1=*never* to 5=*very often*. Internal reliability was  $\alpha=.879$  in this study.

### *Eating Disorder Examination-Questionnaire 8*

The German version of the Eating Disorder Examination-Questionnaire 8 (EDE-Q8) [29] was used to examine whether groups differed in eating disorder symptomatology. The scale has 8 items that are scored from 0=*no days/never/not at all* to 6=*every day/every time/very much*. Internal reliability was  $\alpha=.883$  in this study.

### *End-of-Day Questions*

On each evening during the 10-day period between pre- and posttest, participants answered questions on their smartphone using the application PsyDiary (MultimediaTechnology). Chocolate craving intensity was assessed with the question “How strong was your desire for chocolate-containing foods today (on average)?” Answers were recorded on a rating slider anchored 0=*very weak* and 100=*very strong*. Chocolate craving frequency was assessed with the question “How often did you have a desire for chocolate-containing foods today?” Answers were recorded on a rating slider anchored 0=*not at all* and 100=*very often*. Chocolate consumption quantity was assessed with the question “How many chocolate-containing foods did you consume today?” Answers were recorded on a rating slider anchored 0=*none* and 100=*a great many*. Chocolate consumption frequency was assessed with the question “How often did you consume chocolate-containing foods today?” Answers were recorded on a rating slider anchored 0=*not at all* and 100=*very often*.

### *Debriefing Questions*

Awareness of the study’s aims was assessed with the questions “Do you think that the aim of this study was to assess your behavior in relation to chocolate?” and “Do you think that the



aim of this study was to change your behavior in relation to chocolate?" Response options for both questions were *yes*, *no*, and *I don't know*.

## Procedure

The study was approved by the ethical review board of the University of Salzburg, and study design and hypotheses were preregistered at [aspredicted.org](https://aspredicted.org). The study was advertised as a study on "automatic reactions to chocolate-containing foods in daily life." That is, participants were not informed that the aim of the study was to change chocolate craving and consumption. Participants were randomly assigned to 1 of the 3 groups and were tested in the laboratory individually.

## Pretest

At pretest, participants signed informed consent and completed the FCQ-T-r, the question on chocolate consumption in the past 10 days, and the FCQ-S. Afterward, participants practiced the swipe movements in 2 blocks with 10 trials each (which included pictures of animals and household items that were not used in the main task) and then completed the AAT. Then, they completed the FCQ-S again, responded to the sociodemographic questions, and completed the Restraint Scale, the DEBQ, and the EDE-Q8. Subsequently, body height, weight, and fat mass were measured. Finally, participants installed the apps, and the experimenter explained their use, the remaining study procedures, and discussed any open questions. At the end of the day of the pretest, participants received the first prompt (ie, end-of-day questions) to familiarize them with the app (these data were discarded from analyses).

## Intervention Period

During the 10-day period between the pre- and posttest, all participants received the end-of-day questions on each evening at 9 pm and could respond to the questions until 10 pm. The experimental group additionally performed 5 training sessions (1 session on 5 days each). Training sessions were similar to the AAT used at pre- and posttest, except that pictures of food were always swiped upwards and pictures of objects were always swiped downwards (ie, there was no reversal of instructions between blocks). The placebo control group also performed 5 training sessions (1 session on 5 days each). Here, training sessions were equal to the AAT used at pre- and posttest, that is, pictures of food and objects were swiped upward or downward equally often. In both the experimental and placebo control group, intervention and rest days were pseudorandomized with a maximum of 3 consecutive intervention or rest days. On intervention days, the training session was available between 12 noon and 8 pm. (reminders were sent every 2 hours). The inactive control group did not perform any training sessions.

## Posttest

At posttest, participants again completed the FCQ-T-r, the question on chocolate consumption in the past 10 days, and the FCQ-S; performed the AAT; and then completed the FCQ-S again in the laboratory. Finally, they completed the debriefing questions, and body weight and fat mass were measured. Participation was reimbursed with course credits or €40. The amount of course credits or money was reduced when

participants did not complete all signals (ie, training sessions or end-of-day questions).

## Data Analyses

### Randomization Check and Compliance

We compared groups regarding baseline characteristics with analyses of variance (age, body mass index, body fat mass, chocolate consumption, FCQ-T-r scores, Restraint Scale scores, DEBQ scores, and EDE-Q8 scores) and Fisher exact tests (sex, handedness, education, and nationality). Furthermore, we compared groups regarding the number of completed training sessions (in %) and completed end-of-day questions (in %) with Kruskal-Wallis tests.

### Hypothesis 1

Erroneous trials (eg, swipes in the wrong direction) were excluded from analyses. These accounted for 7.27% of all trials at pretest and 10.4% of all trials at posttest. The number of valid trials did not differ between groups (Kruskal-Wallis tests: pretest  $P=.25$ , posttest  $P=.23$ ). Owing to the task setup, we were able to differentiate between 2 different reaction times: the time between picture appearance and participants' first touch on the screen (*touching time*) and the time between participants' first touch on the screen and picture disappearance (*dragging time*). Bootstrapped split-half reliability estimates for each condition (pull food, push food, pull objects, and push objects) were obtained using the R package *splithalf* [30] performing 5000 random splits. Reliability estimates for touching time ranged between  $r=.70$  and  $.77$  (Spearman-Brown-corrected  $r_{sb}=.82-.87$ ) at pretest and between  $r=.79$  and  $.81$  (Spearman-Brown-corrected  $r_{sb}=.88-.90$ ) at posttest. Reliability estimates for dragging time ranged between  $r=.69$  and  $.82$  (Spearman-Brown-corrected  $r_{sb}=.82-.90$ ) at pretest and between  $r=.63$  and  $.83$  (Spearman-Brown-corrected  $r_{sb}=.77-.90$ ) at posttest.

In line with joystick-based AAT studies [25], median reaction times were calculated. As outlined in the preregistration,  $3 \times 2 \times 2 \times 2$  analyses of variance for repeated measures were calculated with median reaction time data as dependent variables, *group* (experimental vs placebo control vs inactive control) as between-subjects factor and *measurement* (pre- vs posttest), *stimulus* (food vs objects), and *direction* (pull vs push) as within-subjects factors. This was done separately for touching time and for dragging time (which was not explicitly specified in the preregistration).

### Hypothesis 2

As outlined in the preregistration,  $3 \times 2$  analyses of variance for repeated measures were calculated with self-reported chocolate consumption and FCQ-T-r scores as dependent variables, *group* (experimental vs placebo control vs inactive control) as between-subjects factor, and *measurement* (pre- vs posttest) as within-subjects factor.

### Hypothesis 3

As outlined in the preregistration,  $3 \times 2 \times 2$  analyses of variance for repeated measures were calculated with FCQ-S scores (current chocolate craving and hunger) as dependent variables,

group (experimental vs placebo control vs inactive control) as between-subjects factor, and *measurement* (pre- vs posttest) and *task* (before vs after the task) as within-subjects factors.

#### Hypothesis 4

Responses to the end-of-day questions on intervention days on which participants did not complete the training session were excluded from analyses. These accounted for 47 signals (6.71%) of the possible 700 signals (10 days×70 participants [experimental + placebo control group]). As outlined in the preregistration, we applied linear mixed models using the R package *lme4* [31] to analyze the nested, longitudinal structure of the data. *Days* (0=rest day, 1=intervention day; Level 1) and *group* (0=experimental group, 1=placebo control group; Level 2) and their cross-level interaction *group* × *days* were used as predictors for chocolate craving intensity/frequency and for chocolate consumption quantity/frequency. We further explored whether pretest scores of the FCQ-T-r, Restraint Scale, and DEBQ at level 2 would modulate any effects. The level 1 predictor *days* was entered uncentered to the models, and the level 2 predictors *group*, *FCQ-T-r*, *Restraint Scale*, and *DEBQ* were grand-mean centered. The intercepts of all models were allowed to vary randomly. The data files and R-script for these analyses can be found in [Multimedia Appendix 4](#).

#### Exploratory Analyses

Analyses of variance for repeated measures with *group* (experimental vs placebo control vs inactive control) as between-subjects factor and *measurement* (pre- vs posttest) as within-subjects factor were calculated to examine changes in body mass index and body fat mass as a function of group. Moderation analyses were calculated with PROCESS [32] to examine whether FCQ-T-r scores at pretest, Restraint Scale scores, and DEBQ scores moderated any effects of group on chocolate consumption, body mass index, and body fat mass at posttest while controlling for pretest values. Restraint Scale scores and DEBQ scores were also tested as moderators of effects of group on FCQ-T-r scores at posttest while controlling for FCQ-T-r scores at pretest. Fisher exact tests were calculated to compare groups regarding the 2 debriefing questions. These analyses were not included in the preregistration protocol.

## Results

### Randomization Check and Compliance

Groups did not differ in any baseline characteristics (Table 1). Compliance was high for both completion of the training sessions (86.6%) and completion of the end-of-day questions (85.8%) and did not differ between groups (Table 1).

**Table 1.** Means and frequencies of study variables at pretest and compliance rates during the intervention phase as a function of the group (N=105).

Study variables	Experimental group (n=35)	Placebo control group (n=35)	Inactive control group (n=35)	Test statistics	P value
Age (years), mean (SD)	22.7 (3.36)	24.1 (6.13)	23.5 (5.37)	$F_{2,102}=0.64$ , $\eta_p^2=.012$	.53
Sex (female), n (%)	30 (85.7)	32 (91.4)	28 (80.0)	$\chi^2=1.8$ , $\Phi=.133$	.45
Handedness (right-handed), n (%)	28 (80.0)	32 (91.4)	33 (94.3)	$\chi^2=3.5$ , $\Phi=.194$	.23
Education (students), n (%)	33 (94.3)	32 (91.4)	34 (97.1)	$\chi^2=1.1$ , $\Phi=.101$	.87
Nationality (German), n (%)	16 (45.7)	21 (60.0)	18 (51.4)	$\chi^2=3.5$ , $\Phi=.184$	.47
Body mass index (kg/m <sup>2</sup> ), mean (SD)	23.5 (4.90)	23.3 (3.62)	23.0 (3.89)	$F_{2,102}=0.16$ , $\eta_p^2=.003$	.86
Body fat mass (%), mean (SD)	31.6 (9.91)	32.8 (7.52)	29.5 (9.00)	$F_{2,102}=1.29$ , $\eta_p^2=.025$	.28
Chocolate consumption (self-report), mean (SD)	55.6 (20.9)	61.5 (21.3)	58.1 (22.8)	$F_{2,102}=0.65$ , $\eta_p^2=.013$	.52
Food Cravings Questionnaire-Trait-reduced (chocolate version), mean (SD)	41.4 (8.54)	41.2 (10.5)	44.2 (12.2)	$F_{2,102}=0.90$ , $\eta_p^2=.017$	.41
Restraint Scale, mean (SD)	11.5 (5.05)	12.0 (4.53)	11.7 (4.69)	$F_{2,102}=0.11$ , $\eta_p^2=.002$	.89
Dutch Eating Behavior Questionnaire (restrained eating subscale), mean (SD)	2.04 (0.80)	2.03 (0.60)	2.16 (0.65)	$F_{2,102}=0.38$ , $\eta_p^2=.007$	.68
Eating Disorder Examination-Questionnaire 8, mean (SD)	0.97 (1.11)	1.00 (0.72)	1.23 (1.08)	$F_{2,102}=0.75$ , $\eta_p^2=.014$	.48
Training sessions compliance (%), mean (SD)	89.1 (17.7)	84.0 (19.3)	— <sup>a</sup>	—	.22
End-of-day questions compliance (%), mean (SD)	88.0 (17.5)	88.6 (16.1)	80.6 (25.6)	—	.22

<sup>a</sup>Not applicable.

## Hypothesis 1

### Touching Time

A main effect of *direction* ( $F_{1,102}=13.3$ ;  $P<.001$ ;  $\eta_p^2=.115$ ) indicated that participants touched the target stimuli faster in pull trials (mean 599 ms, SD 55.3) than in push trials (mean 606 ms, SD 56.0). There were significant main effects of *measurement* and *stimulus* and interaction effects *measurement*  $\times$  *stimulus* and *group*  $\times$  *measurement* (all  $P$ s $<.001$ ), which were qualified by a significant interaction *group*  $\times$  *measurement*  $\times$  *stimulus* ( $F_{2,102}=4.82$ ;  $P=.01$ ;  $\eta_p^2=.086$ ). However, as this interaction effect was small, did not include any direction effects, and post-hoc comparisons were inconclusive, it was not further interpreted. More information and a graphical depiction can be found in the supplementary material (Figure S1 in [Multimedia Appendix 5](#)). There was no significant main effect of *group* ( $F_{2,102}=2.17$ ;  $P=.12$ ;  $\eta_p^2=.041$ ) and no other significant interaction effects (all  $P$ s $>.16$ ).

### Dragging Time

A main effect of *stimulus* ( $F_{1,102}=9.46$ ;  $P=.003$ ;  $\eta_p^2=.085$ ) indicated that participants swiped food pictures (mean 248 ms, SD 44.5) faster than object pictures (mean 252 ms, SD 53.3). There were no other significant main or interaction effects (all  $P$  values $>.05$ ).

## Hypothesis 2

### Chocolate Craving

A main effect of *measurement* ( $F_{1,102}=11.7$ ;  $P=.001$ ;  $\eta_p^2=.103$ ) indicated that FCQ-T-r scores decreased from pretest (mean 42.3, SD 10.5) to posttest (mean 40.1, SD 11.7). There was no significant main effect of *group* ( $F_{2,102}=0.48$ ;  $P=.62$ ;  $\eta_p^2=.009$ ) and no significant interaction *group*  $\times$  *measurement* ( $F_{2,102}=0.79$ ;  $P=.46$ ;  $\eta_p^2=.015$ ).

### Chocolate Consumption

A main effect of *measurement* ( $F_{1,102}=10.3$ ;  $P=.002$ ;  $\eta_p^2=.092$ ) indicated that self-reported chocolate consumption decreased from pretest (mean 58.4, SD 21.6) to posttest (mean 51.8, SD 20.2). There was no significant main effect of *group* ( $F_{2,102}=0.35$ ;  $P=.71$ ;  $\eta_p^2=.007$ ) and no significant interaction *group*  $\times$  *measurement* ( $F_{2,102}=0.84$ ;  $P=.44$ ;  $\eta_p^2=.016$ ).

## Hypothesis 3

### Current Chocolate Craving

A main effect of *task* ( $F_{1,102}=20.7$ ;  $P<.001$ ;  $\eta_p^2=.169$ ) indicated that FCQ-S craving scores increased from before (mean 28.0, SD 7.55) to after the task (mean 29.5, SD 8.65). A main effect of *measurement* ( $F_{1,102}=17.6$ ;  $P<.001$ ;  $\eta_p^2=.147$ ) indicated the FCQ-S craving scores decreased from pretest (mean 30.1, SD 8.01) to posttest (mean 27.4, SD 9.20). There was no significant main effect of *group* ( $F_{2,102}=1.06$ ;  $P=.35$ ;  $\eta_p^2=.020$ ) and no significant interaction effects (all  $P$  values $>.46$ ).

### Hunger

A main effect of *task* ( $F_{1,102}=11.0$ ;  $P=.001$ ;  $\eta_p^2=.098$ ) indicated that FCQ-S hunger scores increased from before (mean 7.91,

SD 2.74) to after the task (mean 8.20, SD 3.01). There was no significant main effect of *measurement* ( $F_{1,102}=1.36$ ;  $P=.25$ ;  $\eta_p^2=.013$ ), no significant main effect of *group* ( $F_{2,102}=2.46$ ;  $P=.09$ ;  $\eta_p^2=.046$ ), and no significant interaction effects (all  $P$  values $>.43$ ).

## Hypothesis 4

### Chocolate Craving Intensity and Frequency

There was no significant effect of intervention versus rest days as a function of group (see Table S1 in [Multimedia Appendix 5](#)). Higher FCQ-T-r scores at pretest related to higher chocolate craving intensity and frequency, independent of days and group (Table S2). Restrained eating did not relate to chocolate craving intensity or frequency and did not interact with days or group (Table S3, Table S4).

### Chocolate Consumption Quantity and Frequency

There was no significant effect of intervention versus rest days as a function of group (Table S5). Higher FCQ-T-r scores at pretest related to higher chocolate consumption quantity and frequency, independent of days and group (Table S6). In addition, a significant *days*  $\times$  *FCQ-T-r* interaction indicated that participants with high trait chocolate craving scores consumed chocolate-containing foods more frequently on intervention than on rest days, irrespective of group (Table S6). Restrained eating did not relate to chocolate craving quantity or frequency and did not interact with days or group (Table S7, Table S8).

## Exploratory Analyses

### Body Mass Index

There were no significant main effects and no interaction effect *group*  $\times$  *measurement* (all  $P$  values $>.56$ ).

### Body Fat Mass

A main effect of *measurement* ( $F_{1,99}=4.43$ ;  $P=.04$ ;  $\eta_p^2=.043$ ) indicated that body fat mass decreased from pretest (mean 31.4, SD 8.49) to posttest (mean 31.1, SD 8.73). There was no significant main effect of *group* ( $F_{2,99}=0.80$ ;  $P=.45$ ;  $\eta_p^2=.016$ ) and no significant interaction *group*  $\times$  *measurement* ( $F_{2,99}=0.30$ ;  $P=.74$ ;  $\eta_p^2=.006$ ).

### Moderation Analyses

There were no significant interaction effects between *group* and *FCQ-T-r*, *Restraint Scale*, and *DEBQ* scores at pretest (all  $P$  values $>.24$ ).

### Debriefing Questions

A total of 93 participants (88.6%, 93/105) indicated that they thought the aim of the study was to assess their behavior in relation to chocolate, 4 participants (3.8%, 4/105) did not think so, and 8 participants (7.6%, 8/105) indicated that they did not know. There were no significant differences between groups ( $\chi^2=4.6$ ;  $P=.30$ ;  $\Phi=.224$ ). A total of 29 participants (27.6%, 29/105) indicated that they thought the aim of the study was to change their behavior in relation to chocolate, 61 participants (58.1%, 61/105) did not think so, and 15 participants (14.3%, 15/105) indicated that they did not know. Here, responses did

significantly differ between groups ( $\chi^2=9.63$ ;  $P=.04$ ;  $\Phi=.317$ ): more participants in the inactive control group ( $n=26$ ) did not think that the study's aim was to change their behavior than participants in both the experimental group ( $n=18$ ) and the placebo control group ( $n=17$ ), whereas the latter 2 groups did not differ from each other (based on follow-up  $z$  tests using  $\alpha=.05$ ).

## Discussion

### Summary of Results

This study examined effects of a smartphone-based approach-avoidance intervention on approach bias toward chocolate-containing foods and chocolate craving/consumption relative to placebo and no training conditions. The 3 groups were well matched at baseline, treatment adherence was high (87% completed training sessions), and study attrition was low. All dependent measures evidenced good-to-excellent reliability. However, a smartphone-based AAT did neither reveal an approach bias toward chocolate-containing foods at baseline nor a modulation through training. In fact, chocolate craving and consumption decreased throughout the study period in all 3 groups. This self-report finding was corroborated in that participants in all groups lost body fat. Crucially, only a minority of participants thought that this study's aim was to change their behavior, suggesting that these effects were not because of demand characteristics. Comparing chocolate craving and consumption on intervention versus rest days did not reveal any short-term effects of the training.

### Measuring and Modifying Approach-Avoidance Tendencies With Swipe Movements

To the best of our knowledge, this is the first study that aimed at measuring and changing an approach bias toward food stimuli based on swipe movements on smartphones. Although there are similar studies that examined effects of smartphone-based approach-avoidance training with swipe movements on procrastination and body dissatisfaction [12,13], these studies did not measure effects of the training on approach-avoidance tendencies. Thus, the lack of finding and modifying an approach bias toward chocolate-containing foods may be related to an insensitivity of our newly developed task to detect such effects. However, several arguments speak against such an interpretation. First, we used the same stimuli with which an approach bias toward food was detected in a comparable sample with a joystick-based AAT [24]. Second, the AAT in this study had moderate-to-good internal reliability [33] and, thus, unreliability of the task is unlikely to account for the current lack of findings. Third, increasing evidence indicates that the type of arm movements (flexion and extension; [34]) or distance change [35] is not essential for measuring or modifying approach-avoidance inclinations. For example, it has been found that upward and downward movements or framing actions as approach and avoidance suffice to modify stimulus evaluations [36]. Nevertheless, future research needs to determine whether other techniques such as moving the smartphone toward and away with arm movements [37,38] or using tilt movements [39] are better suited for detecting and changing approach-avoidance tendencies with smartphones. In addition, it has recently been

found that combining approach-avoidance actions with affective feedback produced stronger changes in food choices than conventional approach-avoidance training [40]. Thus, using such consequence-based approach-avoidance training may similarly enhance training effects with smartphone-based implementations.

### Effects of Monitoring Food Intake

Another consideration is that—even if the approach-avoidance training had an effect—it may have been masked by the general decreases in outcome variables across the study period that were observed regardless of group assignment. Specifically, we included daily end-of-day-questions in the study design to be able to examine short-term effects (ie, on the same day) of the single training sessions. However, these questions may have acted as a type of ecological momentary intervention [41]. For example, it has been shown that keeping a daily snack diary reduced snacking frequency, suggesting that cue monitoring suffices to decrease unhealthy food intake (irrespective of additional intervention modules; [42]), potentially through increased awareness for one's eating behavior. In fact, it has been found that self-monitoring in terms of completing a record of snacking once per day in the evening decreased snack food consumption even in samples that are not particularly motivated to change their behavior [19]. Thus, we cannot fully exclude the possibility that the intervention may have effects—albeit small—on eating behavior that were masked by effects of monitoring food intake.

### Limitations

Interpretation of results needs to consider the sample investigated in this study. Although we included both men and women with a body mass index ranging from underweight to obese, the majority of the sample were normal-weight women. It has been previously suggested that successful retraining of appetitive reactions and consumption behaviors may primarily be found in clinical samples [9]. Although we investigated a nonclinical sample, it is worth noting that our participants had above-average mean scores ( $>40$ ; Table 1) on the FCQ-T-r (mean scores were 35 in study 1 and 34 in study 2 in the validation studies; [26]), and their eating behavior was clearly impacted throughout the study period (ie, measures were sensitive to detect training-induced changes). This renders insufficient levels of trait chocolate craving as an explanation for these findings unlikely.

Several other methodological considerations might account for these results. For example, although we selected food stimuli with which we have previously detected an approach bias in a comparable sample using a joystick-based task [24], it may be that approach-avoidance training work better when using personalized stimuli, that is, pictures of foods that participants actually crave and consume regularly in their daily life. In related research on attentional bias, for example, it has been found that internal reliability of reaction time tasks can be increased when personalized stimuli are used [43]. Furthermore, we used relatively few training sessions (5), which may have been insufficient to produce meaningful changes in approach bias and eating behavior. However, evidence from joystick-based approach-avoidance training suggest that few

sessions suffice to detect such effects in relation to alcohol [44]. Yet, other smartphone-based studies did indeed use more frequent training sessions [12,13]. Thus, the number of training sessions required in smartphone-based approach-avoidance training need further examination. Finally, although we instructed participants regarding the meaning of upward and downward swipe movements, we did not assess whether they actually perceived the movements as pushing or pulling the pictures away from or toward themselves. Therefore, we cannot rule out the possibility that participants did not perceive the movements as intended, which could explain the lack of finding an approach bias and training effects.

## Conclusions

Repeatedly avoiding chocolate-containing foods in terms of (zoom out) upward swipe movements on smartphones did not change behavior related to these foods in this study. Owing to several methodological considerations, there is an urgent need for future research that determines the most effective way of measuring and changing approach-avoidance tendencies in daily life. General decreases in chocolate craving and consumption as well as body fat mass in this study may be because of the generally raised awareness of chocolate consumption throughout the study period. Thus, receiving daily prompts for monitoring food intake may be a cheap and efficient way to normalize food intake in individuals with eating disorders and facilitate weight loss in individuals with obesity.

---

## Acknowledgments

The authors would like to thank Thérèse Hamm, Maike Burkholder, Lina Lahmer, Lisa-Lucia Ernst, and Veronika Kainz for collecting the data.

This study was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC-StG-2014 639445 NewEat). The funder had no involvement in designing this study, data collection, analysis and interpretation of data, writing of this report, or in the decision to submit the article for publication.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Preregistration.

[[PDF File \(Adobe PDF File\), 131 KB - mhealth\\_v7i11e12298\\_app1.pdf](#) ]

---

### Multimedia Appendix 2

Stimuli.

[[RAR File , 2916 KB - mhealth\\_v7i11e12298\\_app2.rar](#) ]

---

### Multimedia Appendix 3

Example trials video.

[[MP4 File \(MP4 Video\), 2593 KB - mhealth\\_v7i11e12298\\_app3.mp4](#) ]

---

### Multimedia Appendix 4

Study data.

[[RAR File , 36 KB - mhealth\\_v7i11e12298\\_app4.rar](#) ]

---

### Multimedia Appendix 5

Supplementary material.

[[PDF File \(Adobe PDF File\), 912 KB - mhealth\\_v7i11e12298\\_app5.pdf](#) ]

---

### Multimedia Appendix 6

CONSORT-EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 98 KB - mhealth\\_v7i11e12298\\_app6.pdf](#) ]

---

## References

1. Wiers RW, Rinck M, Kordts R, Houben K, Strack F. Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 2010 Feb;105(2):279-287. [doi: [10.1111/j.1360-0443.2009.02775.x](https://doi.org/10.1111/j.1360-0443.2009.02775.x)] [Medline: [20078486](https://pubmed.ncbi.nlm.nih.gov/20078486/)]

2. Eberl C, Wiers RW, Pawelczack S, Rinck M, Becker ES, Lindenmeyer J. Approach bias modification in alcohol dependence: do clinical effects replicate and for whom does it work best? *Dev Cogn Neurosci* 2013 Apr;4:38-51 [FREE Full text] [doi: [10.1016/j.dcn.2012.11.002](https://doi.org/10.1016/j.dcn.2012.11.002)] [Medline: [23218805](https://pubmed.ncbi.nlm.nih.gov/23218805/)]
3. Wiers RW, Eberl C, Rinck M, Becker ES, Lindenmeyer J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol Sci* 2011 Apr;22(4):490-497. [doi: [10.1177/0956797611400615](https://doi.org/10.1177/0956797611400615)] [Medline: [21389338](https://pubmed.ncbi.nlm.nih.gov/21389338/)]
4. Brockmeyer T, Hahn C, Reetz C, Schmidt U, Friederich H. Approach bias modification in food craving—a proof-of-concept study. *Eur Eat Disord Rev* 2015 Sep;23(5):352-360. [doi: [10.1002/erv.2382](https://doi.org/10.1002/erv.2382)] [Medline: [26136357](https://pubmed.ncbi.nlm.nih.gov/26136357/)]
5. Mehl N, Mueller-Wieland L, Mathar D, Horstmann A. Retraining automatic action tendencies in obesity. *Physiol Behav* 2018 Aug 1;192:50-58. [doi: [10.1016/j.physbeh.2018.03.031](https://doi.org/10.1016/j.physbeh.2018.03.031)] [Medline: [29608998](https://pubmed.ncbi.nlm.nih.gov/29608998/)]
6. Warschburger P, Gmeiner M, Morawietz M, Rinck M. Battle of plates: a pilot study of an approach-avoidance training for overweight children and adolescents. *Public Health Nutr* 2018 Feb;21(2):426-434. [doi: [10.1017/S1368980017002701](https://doi.org/10.1017/S1368980017002701)] [Medline: [29108530](https://pubmed.ncbi.nlm.nih.gov/29108530/)]
7. Ferentzi H, Scheibner H, Wiers R, Becker ES, Lindenmeyer J, Beisel S, et al. Retraining of automatic action tendencies in individuals with obesity: a randomized controlled trial. *Appetite* 2018 Jul 1;126:66-72. [doi: [10.1016/j.appet.2018.03.016](https://doi.org/10.1016/j.appet.2018.03.016)] [Medline: [29596869](https://pubmed.ncbi.nlm.nih.gov/29596869/)]
8. Leeman RF, Nogueira C, Wiers RW, Cousijn J, Serafini K, DeMartini KS, et al. A test of multisession automatic action tendency retraining to reduce alcohol consumption among young adults in the context of a human laboratory paradigm. *Alcohol Clin Exp Res* 2018 Apr;42(4):803-814 [FREE Full text] [doi: [10.1111/acer.13613](https://doi.org/10.1111/acer.13613)] [Medline: [29450895](https://pubmed.ncbi.nlm.nih.gov/29450895/)]
9. Krishna A, Eder AB. No effects of explicit approach-avoidance training on immediate consumption of soft drinks. *Appetite* 2018 Nov 1;130:209-218. [doi: [10.1016/j.appet.2018.08.023](https://doi.org/10.1016/j.appet.2018.08.023)] [Medline: [30121309](https://pubmed.ncbi.nlm.nih.gov/30121309/)]
10. Becker D, Jostmann NB, Holland RW. Does approach bias modification really work in the eating domain? A commentary on Kakoschke et al (2017). *Addict Behav* 2018 Feb;77:293-294. [doi: [10.1016/j.addbeh.2017.02.025](https://doi.org/10.1016/j.addbeh.2017.02.025)] [Medline: [28237541](https://pubmed.ncbi.nlm.nih.gov/28237541/)]
11. Kakoschke N, Kemps E, Tiggemann M. Approach bias modification training and consumption: a review of the literature. *Addict Behav* 2017 Jan;64:21-28. [doi: [10.1016/j.addbeh.2016.08.007](https://doi.org/10.1016/j.addbeh.2016.08.007)] [Medline: [27538198](https://pubmed.ncbi.nlm.nih.gov/27538198/)]
12. Kollei I, Lukas CA, Loeber S, Berking M. An app-based blended intervention to reduce body dissatisfaction: a randomized controlled pilot study. *J Consult Clin Psychol* 2017 Nov;85(11):1104-1108. [doi: [10.1037/ccp0000246](https://doi.org/10.1037/ccp0000246)] [Medline: [28857576](https://pubmed.ncbi.nlm.nih.gov/28857576/)]
13. Lukas CA, Berking M. Reducing procrastination using a smartphone-based treatment program: a randomized controlled pilot study. *Internet Interv* 2018 Jun;12:83-90 [FREE Full text] [doi: [10.1016/j.invent.2017.07.002](https://doi.org/10.1016/j.invent.2017.07.002)] [Medline: [30135772](https://pubmed.ncbi.nlm.nih.gov/30135772/)]
14. Weingarten HP, Elston D. Food cravings in a college population. *Appetite* 1991 Dec;17(3):167-175. [doi: [10.1016/0195-6663\(91\)90019-O](https://doi.org/10.1016/0195-6663(91)90019-O)] [Medline: [1799279](https://pubmed.ncbi.nlm.nih.gov/1799279/)]
15. Richard A, Meule A, Reichenberger J, Blechert J. Food cravings in everyday life: an EMA study on snack-related thoughts, cravings, and consumption. *Appetite* 2017 Jun 1;113:215-223. [doi: [10.1016/j.appet.2017.02.037](https://doi.org/10.1016/j.appet.2017.02.037)] [Medline: [28249745](https://pubmed.ncbi.nlm.nih.gov/28249745/)]
16. Becker D, Jostmann NB, Wiers RW, Holland RW. Approach avoidance training in the eating domain: testing the effectiveness across three single session studies. *Appetite* 2015 Feb;85:58-65. [doi: [10.1016/j.appet.2014.11.017](https://doi.org/10.1016/j.appet.2014.11.017)] [Medline: [25447011](https://pubmed.ncbi.nlm.nih.gov/25447011/)]
17. Dickson H, Kavanagh DJ, MacLeod C. The pulling power of chocolate: effects of approach-avoidance training on approach bias and consumption. *Appetite* 2016 Apr 1;99:46-51. [doi: [10.1016/j.appet.2015.12.026](https://doi.org/10.1016/j.appet.2015.12.026)] [Medline: [26725150](https://pubmed.ncbi.nlm.nih.gov/26725150/)]
18. Schumacher SE, Kemps E, Tiggemann M. Bias modification training can alter approach bias and chocolate consumption. *Appetite* 2016 Jan 1;96:219-224. [doi: [10.1016/j.appet.2015.09.014](https://doi.org/10.1016/j.appet.2015.09.014)] [Medline: [26375357](https://pubmed.ncbi.nlm.nih.gov/26375357/)]
19. Maas J, Hietbrink L, Rinck M, Keijsers GP. Changing automatic behavior through self-monitoring: does overt change also imply implicit change? *J Behav Ther Exp Psychiatry* 2013 Sep;44(3):279-284. [doi: [10.1016/j.jbtep.2012.12.002](https://doi.org/10.1016/j.jbtep.2012.12.002)] [Medline: [23388468](https://pubmed.ncbi.nlm.nih.gov/23388468/)]
20. Meule A, Lutz AP, Krawietz V, Stützer J, Vögele C, Kübler A. Food-cue affected motor response inhibition and self-reported dieting success: a pictorial affective shifting task. *Front Psychol* 2014;5:216 [FREE Full text] [doi: [10.3389/fpsyg.2014.00216](https://doi.org/10.3389/fpsyg.2014.00216)] [Medline: [24659978](https://pubmed.ncbi.nlm.nih.gov/24659978/)]
21. Meule A, Skirde AK, Freund R, Vögele C, Kübler A. High-calorie food-cues impair working memory performance in high and low food cravers. *Appetite* 2012 Oct;59(2):264-269. [doi: [10.1016/j.appet.2012.05.010](https://doi.org/10.1016/j.appet.2012.05.010)] [Medline: [22613059](https://pubmed.ncbi.nlm.nih.gov/22613059/)]
22. Faul F, Erdfelder E, Lang A, Buchner A. G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/BF03193146](https://doi.org/10.3758/BF03193146)]
23. Blechert J, Meule A, Busch NA, Ohla K. Food-pics: an image database for experimental research on eating and appetite. *Front Psychol* 2014;5:617 [FREE Full text] [doi: [10.3389/fpsyg.2014.00617](https://doi.org/10.3389/fpsyg.2014.00617)] [Medline: [25009514](https://pubmed.ncbi.nlm.nih.gov/25009514/)]
24. Lender A, Meule A, Rinck M, Brockmeyer T, Blechert J. Measurement of food-related approach-avoidance biases: larger biases when food stimuli are task relevant. *Appetite* 2018 Jun 1;125:42-47. [doi: [10.1016/j.appet.2018.01.032](https://doi.org/10.1016/j.appet.2018.01.032)] [Medline: [29407526](https://pubmed.ncbi.nlm.nih.gov/29407526/)]
25. Rinck M, Becker ES. Approach and avoidance in fear of spiders. *J Behav Ther Exp Psychiatry* 2007 Jun;38(2):105-120. [doi: [10.1016/j.jbtep.2006.10.001](https://doi.org/10.1016/j.jbtep.2006.10.001)] [Medline: [17126289](https://pubmed.ncbi.nlm.nih.gov/17126289/)]
26. Meule A, Hormes JM. Chocolate versions of the Food Cravings Questionnaires. Associations with chocolate exposure-induced salivary flow and ad libitum chocolate consumption. *Appetite* 2015 Aug;91:256-265. [doi: [10.1016/j.appet.2015.04.054](https://doi.org/10.1016/j.appet.2015.04.054)] [Medline: [25913686](https://pubmed.ncbi.nlm.nih.gov/25913686/)]

27. Dinkel A, Berth H, Exner C, Rief W, Balck F. German version of the Restraint Scale for the assessment of restrained eating. *Diagnostica* 2005 Apr;51(2):67-74. [doi: [10.1026/0012-1924.51.2.67](https://doi.org/10.1026/0012-1924.51.2.67)]
28. Grunert S. Ein Inventar zur Erfassung von Selbstaussagen zum Ernährungsverhalten. *Diagnostica* 1989;35:167-179.
29. Kliem S, Möhle T, Zenger M, Strauß B, Brähler E, Hilbert A. The eating disorder examination-questionnaire 8: a brief measure of eating disorder psychopathology (EDE-Q8). *Int J Eat Disord* 2016 Jun;49(6):613-616. [doi: [10.1002/eat.22487](https://doi.org/10.1002/eat.22487)] [Medline: [26711183](https://pubmed.ncbi.nlm.nih.gov/26711183/)]
30. Parsons S. figshare. 2018. Splithalf: robust estimates of split half reliability (Version 4). URL: <https://figshare.com/articles/splithalf/5559175> [accessed 2019-09-25]
31. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):1-48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
32. Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis*. Second Edition. New York, NY: The Guilford Press; 2018.
33. Parsons S, Kruijt A, Fox E. Psychological science needs a standard practice of reporting the reliability of cognitive behavioural measurements. *PsyArXiv* 2018. [doi: [10.31234/osf.io/6ka9z](https://doi.org/10.31234/osf.io/6ka9z)]
34. Krieglmeyer R, de Houwer J, Deutsch R. On the nature of automatically triggered approach-avoidance behavior. *Emotion Rev* 2013 Jun 11;5(3):280-284. [doi: [10.1177/1754073913477501](https://doi.org/10.1177/1754073913477501)]
35. Eder AB, Rothermund K. When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *J Exp Psychol Gen* 2008 May;137(2):262-281. [doi: [10.1037/0096-3445.137.2.262](https://doi.org/10.1037/0096-3445.137.2.262)] [Medline: [18473659](https://pubmed.ncbi.nlm.nih.gov/18473659/)]
36. van Dessel P, Eder AB, Hughes S. Mechanisms underlying effects of approach-avoidance training on stimulus evaluation. *J Exp Psychol Learn Mem Cogn* 2018 Aug;44(8):1224-1241. [doi: [10.1037/xlm0000514](https://doi.org/10.1037/xlm0000514)] [Medline: [29648864](https://pubmed.ncbi.nlm.nih.gov/29648864/)]
37. Zech HG. ResearchGate. 2015. The Mobile Approach-Avoidance Task. URL: [https://www.researchgate.net/publication/282943650\\_The\\_mobile\\_approach-avoidance\\_task](https://www.researchgate.net/publication/282943650_The_mobile_approach-avoidance_task) [accessed 2019-09-25]
38. Cring C. Leiden Repository. 2017. Approach and Avoidance Tendencies: Using Smartphones to Study Behavioral Responses and Food Desires. URL: <https://openaccess.leidenuniv.nl/handle/1887/53247> [accessed 2019-09-25]
39. Kakoschke N, Hawker C, Castine B, de Courten B, Verdejo-Garcia A. Smartphone-based cognitive bias modification training improves healthy food choice in obesity: a pilot study. *Eur Eat Disord Rev* 2018 Sep;26(5):526-532. [doi: [10.1002/erv.2622](https://doi.org/10.1002/erv.2622)] [Medline: [30003634](https://pubmed.ncbi.nlm.nih.gov/30003634/)]
40. van Dessel P, Hughes S, de Houwer J. Consequence-based approach-avoidance training: a new and improved method for changing behavior. *Psychol Sci* 2018 Dec;29(12):1899-1910. [doi: [10.1177/0956797618796478](https://doi.org/10.1177/0956797618796478)] [Medline: [30312146](https://pubmed.ncbi.nlm.nih.gov/30312146/)]
41. 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](https://doi.org/10.1348/135910709X466063)] [Medline: [19646331](https://pubmed.ncbi.nlm.nih.gov/19646331/)]
42. Verhoeven AA, Adriaanse MA, de Vet E, Fennis BM, de Ridder DT. Identifying the 'if' for 'if-then' plans: combining implementation intentions with cue-monitoring targeting unhealthy snacking behaviour. *Psychol Health* 2014;29(12):1476-1492. [doi: [10.1080/08870446.2014.950658](https://doi.org/10.1080/08870446.2014.950658)] [Medline: [25099386](https://pubmed.ncbi.nlm.nih.gov/25099386/)]
43. Christiansen P, Mansfield R, Duckworth J, Field M, Jones A. Internal reliability of the alcohol-related visual probe task is increased by utilising personalised stimuli and eye-tracking. *Drug Alcohol Depend* 2015 Oct 1;155:170-174. [doi: [10.1016/j.drugalcdep.2015.07.672](https://doi.org/10.1016/j.drugalcdep.2015.07.672)] [Medline: [26239377](https://pubmed.ncbi.nlm.nih.gov/26239377/)]
44. Eberl C, Wiers RW, Pawelczack S, Rinck M, Becker ES, Lindenmeyer J. Implementation of approach bias re-training in alcoholism-how many sessions are needed? *Alcohol Clin Exp Res* 2014 Feb;38(2):587-594. [doi: [10.1111/acer.12281](https://doi.org/10.1111/acer.12281)] [Medline: [24164417](https://pubmed.ncbi.nlm.nih.gov/24164417/)]

## Abbreviations

- AAT:** approach-avoidance task
- DEBQ:** Dutch Behavior Questionnaire
- EDE-Q8:** Eating Disorder Examination-Questionnaire 8
- EMA:** ecological momentary assessment
- FCQ-S:** Food Cravings Questionnaire-State
- FCQ-T-r:** Food Cravings Questionnaire-Trait-reduced

*Edited by G Eysenbach; submitted 24.09.18; peer-reviewed by R Leeman, D Becker; comments to author 01.11.18; revised version received 03.02.19; accepted 24.03.19; published 01.11.19.*

*Please cite as:*

*Meule A, Richard A, Dinic R, Blechert J*

*Effects of a Smartphone-Based Approach-Avoidance Intervention on Chocolate Craving and Consumption: Randomized Controlled Trial*

*JMIR Mhealth Uhealth 2019;7(11):e12298*

URL: <https://mhealth.jmir.org/2019/11/e12298>

doi: [10.2196/12298](https://doi.org/10.2196/12298)

PMID: [31682584](https://pubmed.ncbi.nlm.nih.gov/31682584/)

©Adrian Meule, Anna Richard, Radomir Dinic, Jens Blechert. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 01.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# A Mobile-Based Comprehensive Weight Reduction Program for the Workplace (Health-On): Development and Pilot Study

Min Kyu Han<sup>1</sup>, MPH, MD; Belong Cho<sup>2,3</sup>, PhD, MD; Hyuktae Kwon<sup>3</sup>, MPH, PhD, MD; Ki Young Son<sup>4</sup>, MPH, PhD, MD; Hyejin Lee<sup>5</sup>, MPH, MD; Joo Kyung Lee<sup>6</sup>, MD; Jinho Park<sup>2,3</sup>, MPH, PhD, MD

<sup>1</sup>SK hynix International Medical Center, Wuxi, Jiangsu, China

<sup>2</sup>Department of Family Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

<sup>3</sup>Department of Family Medicine, Seoul National University Hospital, Seoul, Republic of Korea

<sup>4</sup>Department of Family Medicine, Asan Medical Center, Seoul, Republic of Korea

<sup>5</sup>Department of Family Medicine, Seoul National University Bundang Hospital, Seongnam City, Republic of Korea

<sup>6</sup>Department of Internal Medicine, Korea University Guro Hospital, Seoul, Republic of Korea

**Corresponding Author:**

Jinho Park, MPH, PhD, MD

Department of Family Medicine

Seoul National University Hospital

101 Daehak-ro Jongno-gu

Seoul, 03080

Republic of Korea

Phone: 82 2 2072 0865

Email: [kkolzzi0@gmail.com](mailto:kkolzzi0@gmail.com)

## Abstract

**Background:** There is a growing interest in mobile technology for obesity management. Despite the known effectiveness of workplace-based weight loss programs, there are few studies on mobile phone-delivered interventions.

**Objective:** This study aimed to develop and verify an integrated and personalized mobile technology-based weight control program, named Health-On, optimized for workplaces.

**Methods:** A weight reduction algorithm was developed for calorie prescription, continuous monitoring, periodic feedback and reevaluation, goal resetting, and offline intervention with behavior-changing strategies. A total of 30 obese volunteers (body mass index  $\geq 25$  kg/m<sup>2</sup>) participated in the 12-week Health-On pilot program. The primary outcome was weight reduction, and secondary outcomes were improved anthropometric measures, metabolic profiles, and fat computed tomography measures, all assessed pre- and postintervention.

**Results:** Health-On incorporated proprietary algorithms and several strategies intended to maximize adherence, using compatible online and offline interventions. The mean weight of 30 participants decreased by 5.8%, and median weight also decreased from 81.3 kg (interquartile range [IQR] 77.1-87.8) before intervention to 76.6 kg (IQR 70.8-79.5) after the 12-week intervention period ( $P < .001$ ). The metabolic profiles and fat measures (blood pressure, glycosylated hemoglobin, total cholesterol, triglyceride, high-density lipoprotein, low-density lipoprotein, alanine aminotransferase, and visceral and subcutaneous adipose tissue;  $P < .05$ ) also improved significantly.

**Conclusions:** In this single-group evaluation of 30 participants before and after the Health-On program, body weight decreased and metabolic profiles and fat measures improved. Follow-up studies are needed to assess effectiveness and long-term adherence.

(*JMIR Mhealth Uhealth* 2019;7(11):e11158) doi:[10.2196/11158](https://doi.org/10.2196/11158)

**KEYWORDS**

weight loss programs; smartphone; mobile phone; workplaces; obesity; obesity management

## Introduction

### Background

Obesity is a major global health problem [1], the cause of increased morbidity and mortality, and significant health care resources are expended on managing and preventing obesity and associated complications [2-5]. Of the options for the treatment of obesity, lifestyle interventions are foundational regardless of augmentation by drug therapy or bariatric surgery. [6,7]. Two principles are central to obesity lifestyle interventions: negative calorie balance and maintenance of it. For physicians, a clinical setting of fragmentary visits and short consultations makes individualized feedback and instruction problematic. For patients, a daily food intake or physical activity diary is intrusive and recollections are unreliable [8].

There is growing interest in lifestyle interventions using smart devices [9], and they can be effective tools to combat obesity, helping to overcome the limitations of current interventions [10,11]. They can also be helpful to achieve *comprehensive lifestyle modification* that is proved to be strongly effective to manage obesity [7,12]. Comprehensive strategy includes group intervention, online or offline interventions on diet and physical activity, frequent monitoring and feedback, and education.

Considering these things, workplace is one of the most optimal places for mobile-based comprehensive lifestyle interventions. Employee numbers will be fairly constant over the intervention period, and employer-provided resources (cafeteria and fitness center), messengers, and periodic health checkup services can

be incorporated into the interventions. The features of mobile health apps (eg, real-time lifestyle monitoring and participant interactivity) can be easily integrated in workplace interventions [13]. In addition, workplace lifestyle interventions are already known to be effective for weight reduction and increased productivity and in reducing financial burden on employees [14-16]. Despite these advantages, mobile-based interventions in practice are rare.

### Objectives

Against this background, we developed a comprehensive mobile app (Health-On) and embarked on a pilot study to verify its workplace feasibility and assess the clinical outcomes before and after the implementation of this program.

## Methods

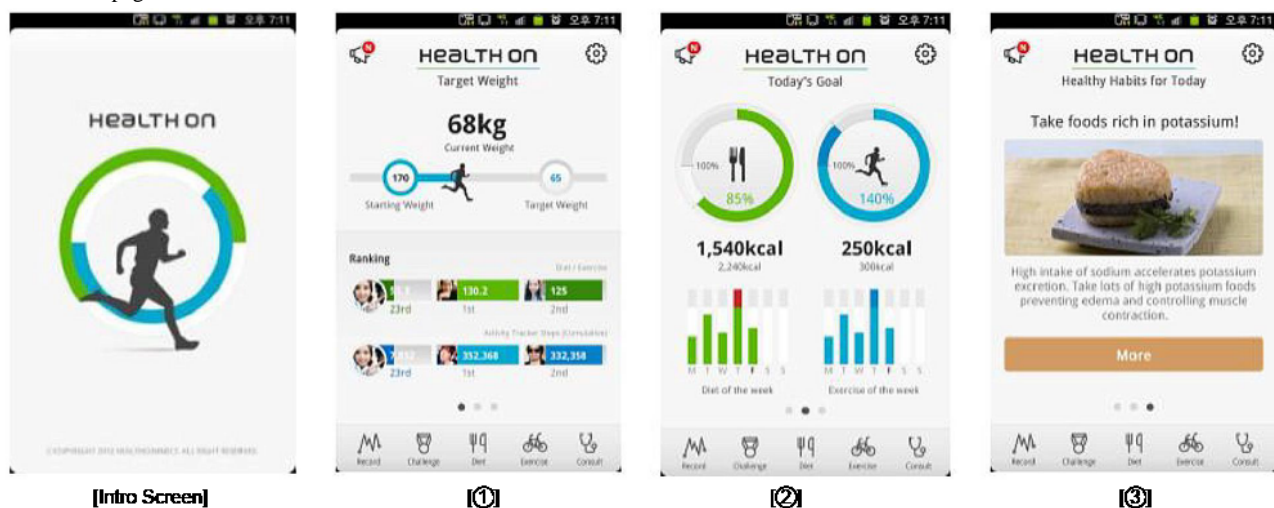
### Development of the Health-On App

Health-On is a program that combines online devices (mobile app and smart watch) and offline interventional resources (cafeteria, fitness center, and peer support; Figure 1).

Researchers first developed the Health-On mobile phone app using the Android software development kit (SDK r20.0.3) [17].

To accommodate the first principle, negative calorie balance, we developed an equation to calculate total energy expenditure (TEE) and an algorithm to prescribe diet and physical activities. For the second principle, maintenance of negative calorie balance, we created a daily diet or physical activity tracker and applied behavior change strategies.

Figure 1. Main pages.



### Equation for Calculating an Estimation of Total Energy Expenditure

To identify the calorie consumption or expenditure negative balance, we must calculate TEE and maintain a TEE greater than calorie intake. TEE is obtained by summing resting metabolic rate (RMR), the thermic effect of activity (TEA), and the thermogenic effect of food (TEF). We chose the predictive Cunningham equation for measuring RMR as it is an accurate formula for estimating RMR using fat-free mass and is close to measured values in Korea [18,19]. For measuring TEA, we used

the International Physical Activity Questionnaire Short Form (IPAQ-SF) [20,21]. However, IPAQ-SF does not include calories expended in daily activities, for example, showering or speaking. Thus, we estimated calorie expenditure from these to be 10% of RMR, on the premise that obese participants are sedentary [22]. TEF was assumed to be 10% of RMR + TEA [23].

TEE can be expressed as follows:  $TEE = RMR + TEA (1) + (2) + TEF$ , where (1)=kcal consumption from IPAQ-SF and (2)=kcal consumption not included in the IPAQ-SF (eg, walking

<10 min and other daily activities). This equation was used to maintain the negative calorie balance between dietary intake and physical activities.

### **Diet and Physical Activity Prescription Algorithm**

Following accepted guidelines [7], we set a recommended goal of 7% baseline body weight loss over 12 weeks, divided into 3.0%, 2.5%, and 1.5% targets for each 4-week period. We assumed that a 7000 kcal expenditure is needed to lose 1 kg of body weight [24]. Subjects themselves determined how many negative calories they needed from their control variables, for example, how much they should cut down on eating or how much they should increase physical activity to achieve their weight reduction goals. Then, the algorithm suggested the goals of daily calories from dietary intake necessary to achieve weight loss. These goals and variables were adjusted periodically to reflect weight change during the process. To achieve safe weight reduction, we set quantitative parameters of a minimum of 1200 kcal per day for women and 1500 kcal for men [7].

### **Convenient Method of Tracking Daily Diet and Physical Activity**

#### ***Diet***

Dietary intake could be entered in the app either by inputs on a pre-entered menu or by inputting using the search option.

#### ***Pre-Entered Inputs***

We applied various methods to minimize user effort in keying in calories. First, the workplace cafeteria daily menu nutritional data were entered into the app. In addition, menus of restaurants nearby the workplace were entered. Menus that were frequently entered by app users were also pre-entered, thereby automatically appearing in the app.

#### ***Input by Search***

We embedded the Canpro4.0 and FanTasy (Food and Nutrition Database System) databases [25,26]. When users consumed foods not on the pre-entered menu, they could search in the app. The typical Korean diet was provided in an easy-to-search select-and-enter list. Meals not on the list could be entered directly by the user.

#### ***Physical Activity***

Physical activity could be entered in the app either by automatic input or manually.

#### ***Automatic Input***

**Health-On enables connectivity with an activity tracker, for example, a smart watch,** to measure usual physical activity. The collected data (expended calories and step count) were automatically sent every hour to the app via Bluetooth and made available to users.

#### ***Manual Input***

The app enables entry of type, strength, and frequency and duration of aerobic exercise, strength exercise, and common sports, as activity is difficult to measure while using fitness equipment or doing aquatic exercise. This type of user-provided

information was calculated and automatically converted into expended calories based on user weight [27].

### **Behavior Change Strategies**

This program contains several strategies for encouraging behavior change that are effective for weight reduction [28]. We modified these strategies for mobile technology applicability (eg, individualized tailored feedback based on personal life log and ranking). We used resources available in the workplace, for example, the cafeteria and fitness center, to optimize advantages and maximize effectiveness [29].

#### ***Health Age***

Health-On can calculate health age from basic health information. We had previously found that a Web-based health risk appraisal (factoring in health age) can be effective for ascertaining health risks and motivating lifestyle modifications [30], and we adjusted this tool for our purposes.

#### ***Health Information***

A team of nurses, nutritionists, and exercise trainers devised educational material with diet and physical activity tips. This information was provided in the app daily to improve exercise and dietary habits and ensure effective weight loss. In addition, counseling with a nutritionist and an exercise trainer was made available through social networks.

#### ***Feedback on the Life Log***

We developed feedback for self-monitoring and self-reflection on periodic results. Each day, calorie intake was compared with the goal, and goal achievement provided an incentive for further progress. A nutritionist provided feedback on participants' dietary intake records. An exercise supervisor gave feedback based on comparisons between burned calories and exercise targets. Feedback was delivered via a pop-up.

#### ***History Query***

With the history query function, users could monitor their health examination results pre- and postprogram, weekly changes in body composition and measurements, and weekly dietary and exercise performance.

#### ***Competition***

Good-natured competition, a useful weight loss motivation method, was introduced for promoting Health-On continuous usage. Health-On automatically adds users' friends from contact lists, enabling competition for achievement scores or step counts.

#### ***Ranking***

The achievement score was based on the 7.0% weight loss success or failure compared with the starting weight plus proximity to monthly goals plus dietary input frequency or activity tracker usage.

#### ***Step Counts***

The cumulative number of steps over a period, recorded on the activity tracker's pedometer, was used to calculate rankings.

## Health-On Program Pilot Study

### Study Population

The primary purpose of this study was to verify Health-On's workplace feasibility and assess the clinical outcomes.

In April 2012, an announcement targeting SK telecom workers in a workplace about the recruitment of people with body mass index of more than 25 kg/m<sup>2</sup> [31] who were willing to control diet and exercise for weight loss had been made on the intranet. The Institutional Review Board of Seoul National University Hospital approved this study (IRB No.: H-1204-041-405).

### Exclusion Criteria

We excluded anyone from the recruited members (1) who had answered at least one *yes* to the questionnaires of Physical Activity Readiness Questionnaire [32], and their eligibility to this study was judged ineligible after counseling, (2) who had answered *yes* on the eating disorder survey and were judged ineligible for this study after counseling with a doctor, (3) who had systolic blood pressure (SBP)  $\geq 160$  mmHg or diastolic blood pressure (DBP)  $\geq 100$  mmHg, fasting blood sugar (FBS)  $\geq 160$  mg/dl, triglyceride  $\geq 500$  mg/dl, low-density lipoprotein-cholesterol (LDL-C)  $\geq 190$  mg/dl and judged ineligible after counseling with a doctor owing to identification of more than one abnormality, (4) who were obese and had received any pharmacological, procedural, or surgical treatment within a month, (5) who had undergone a drastic weight change (more than 10% of body weight) within a month, (6) who had suffered from or received a procedure because of severe illness such as myocardial infarction, stroke, cancer-related disorders, and hip surgery, (7) who had been suffering from thyroid disease, and (8) who were judged by the researcher as ineligible to participate. A total of 30 people were recruited for this study.

### Measurements

Anthropometric data (height, weight, and percentage of body fat were measured by Inbody 720), fasting blood samples, and computed tomography (CT) scans were collected via questionnaires. Data were collected twice, before and after the 12-week intervention. Participants provided written informed consent, and the Institutional Review Board approved the protocol.

The laboratory tests included alanine aminotransferase (ALT), aspartate aminotransferase, creatinine, lipid profile (total cholesterol, high-density lipoprotein-cholesterol (HDL-C), and triglyceride [TG]), FBS, and baseline glycosylated hemoglobin (HbA<sub>1c</sub>). Lipid profiles and blood glucose tests were performed after fasting.

Abdominal adipose tissue mass was estimated using cross-sectional images obtained by a standardized and validated CT technique [33]. Participants were examined in the supine position with a 16-detector row CT scanner (Somatom Sensation 16; Siemens Medical Solutions). A single umbilicus level 5-mm

slice image was obtained. The total abdominal adipose tissue area (subcutaneous adipose tissue area plus visceral adipose tissue (VAT) area) was calculated using specialized software (Rapidia 2.8; Infinitt) with the attenuation values within a range of  $-250$  to  $-50$  Hounsfield Units. We used VAT  $\geq 100$  cm<sup>2</sup> as the criterion for visceral obesity [34].

### Intervention

On the basis of these health indicators, weight reduction goals and target net kilocalories were individually determined through the Health-On algorithm.

For the first 4 weeks, a 600 to 700 kcal breakfast was provided to create a habit of not skipping breakfast. Participants were encouraged to join Group eXercise once a week at the company gym. Participants were able to check goal achievement by taking Inbody measurements regularly.

A specialist sent review data (LifeStyle Guide) every fortnight, and personal contact was scheduled after the initial health checkup, when checking monthly goals and during a poststudy follow-up.

### Outcomes

The primary outcome of the study was the change in weight of participants before and after the program. We also measured the change in percentage of body fat, lean body mass (kg), waist circumference (cm), serum TG, HDL-C, LDL-C, non-HDL-C, SBP and DBP (mmHg), fasting plasma glucose (mg/dL), and visceral fat as the secondary outcomes.

### Statistics

The continuous demographic variable and the baseline variable were summarized using descriptive statistics (means with standard deviations and medians with ranges). The categorical demography characteristics were summarized by frequency distribution and percentages. Comparison of the differences between pre- and poststudy outcomes was done with a Wilcoxon signed-rank test.

All analyses were conducted using STATA version 12.1 for Windows (StatCorp), with a *P* value  $< .05$  used to indicate statistically significant differences.

## Results

The interventions discussed in the Methods section are incorporated into the app.

### Health-On App (Online Intervention)

Health-On has 4 theme pages: main, diet, physical activity, and challenge and ranking. Each page allows users to easily see their achievements and to maximize user convenience and app effectiveness with a simple user interface (UI). The theme icons are on the tab menu, making movements between pages smooth and convenient (Figures 2-5).

Figure 2. Diet pages.

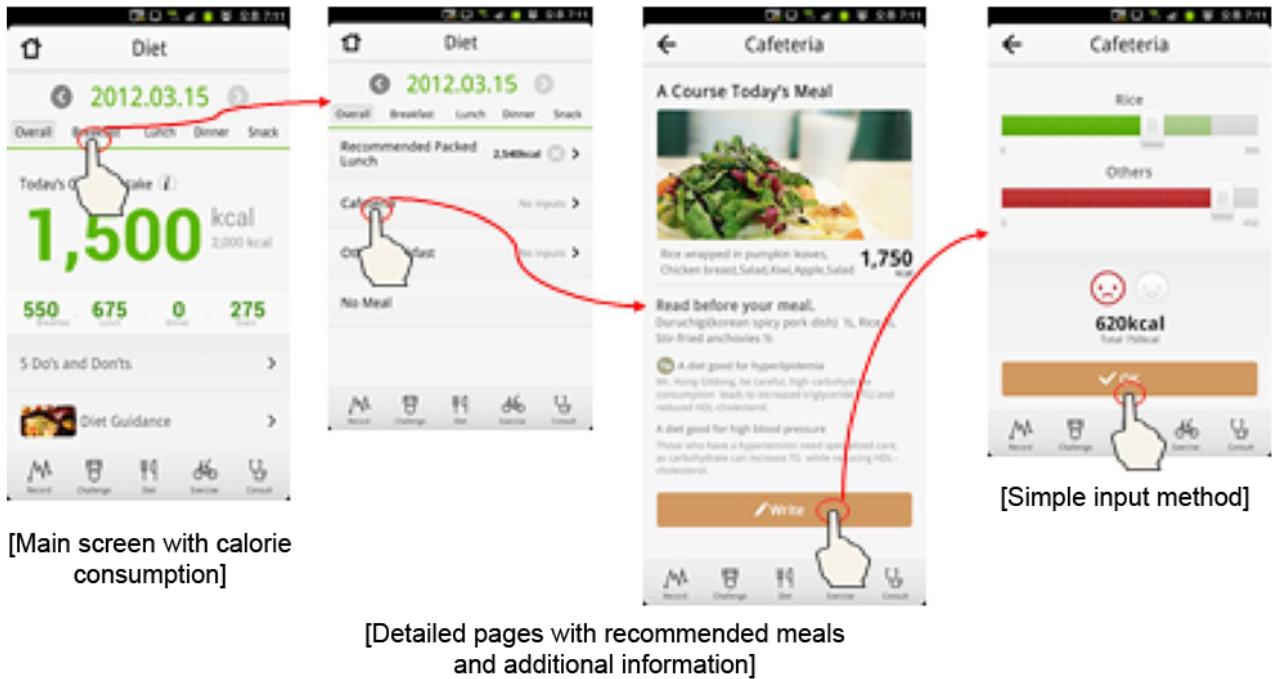


Figure 3. Physical activity pages: activity tracker.

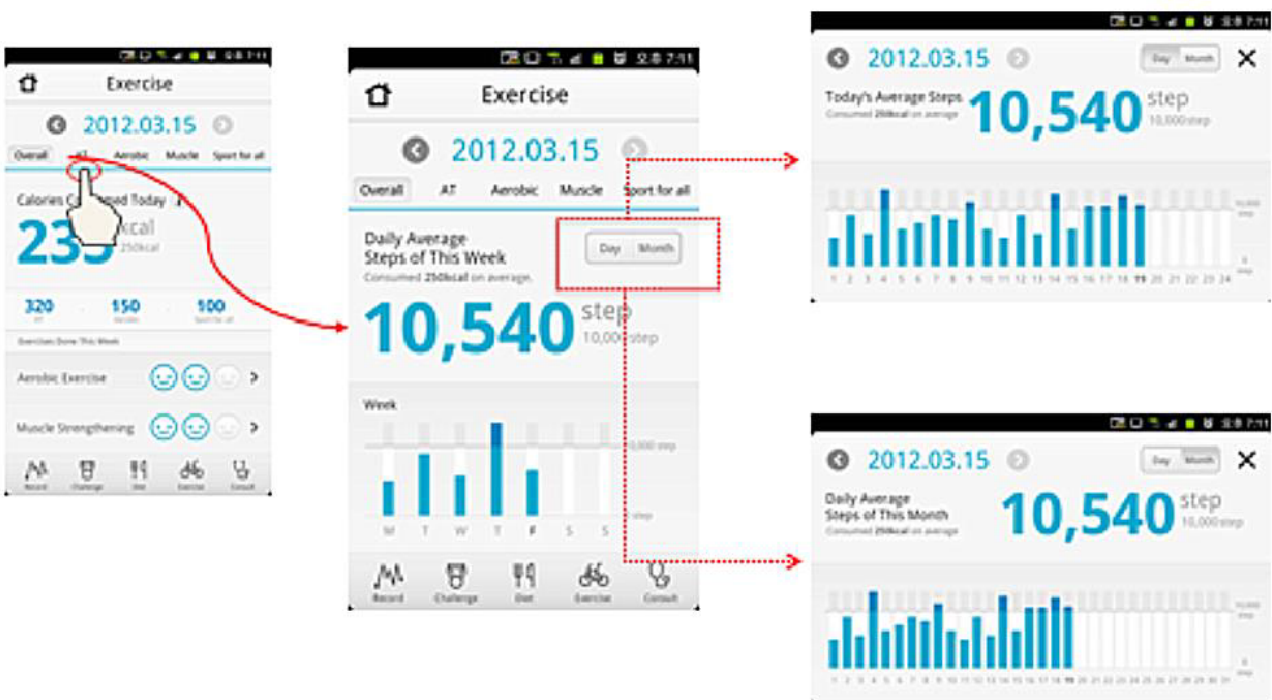


Figure 4. Physical activity pages: aerobic exercise.

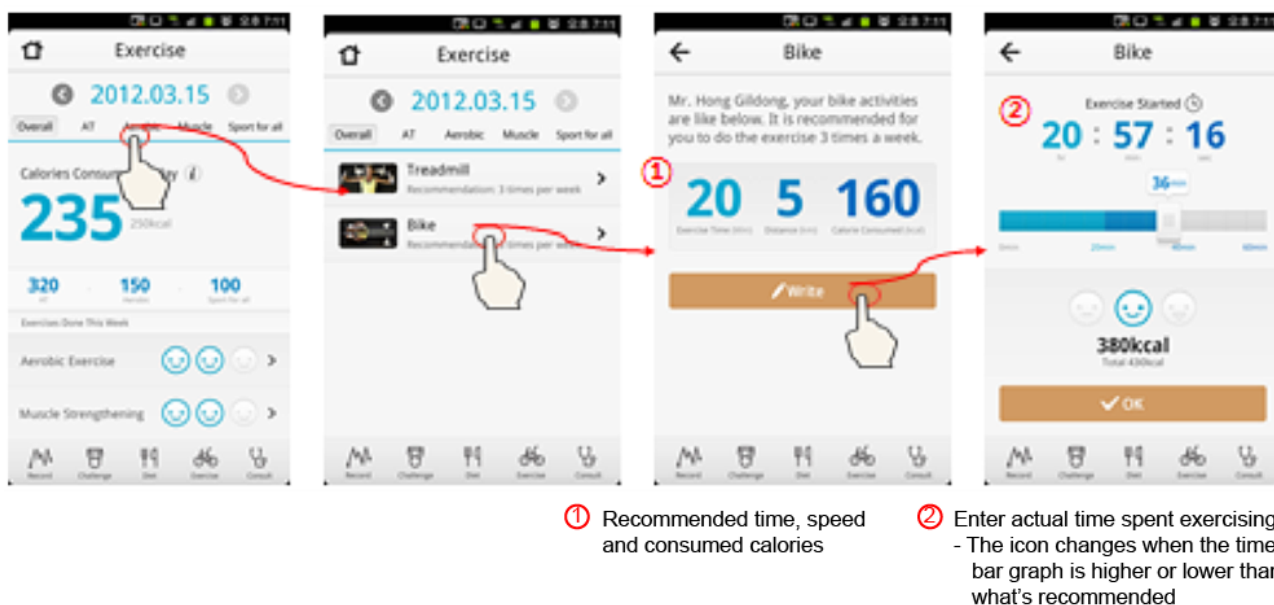
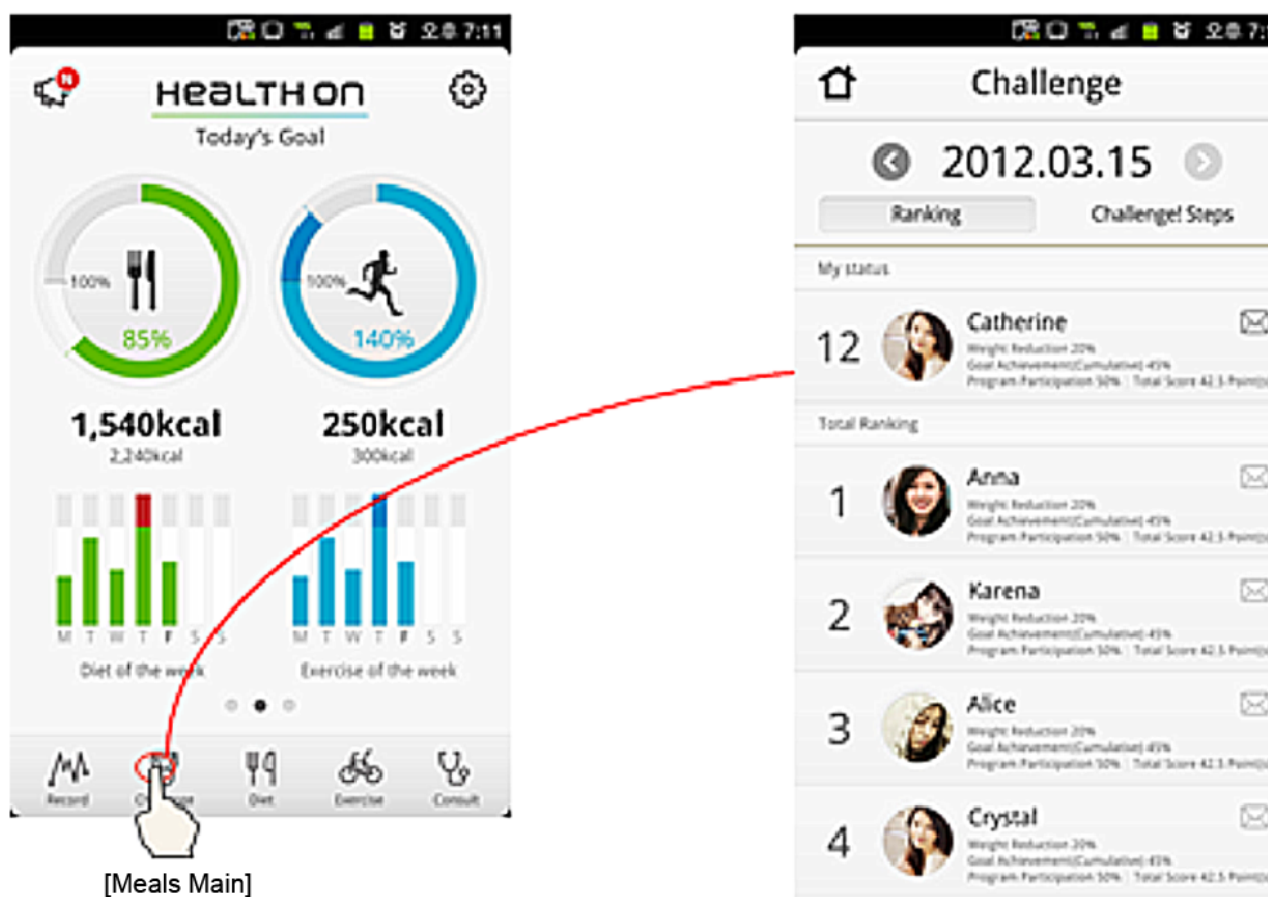


Figure 5. Challenge and ranking pages.



**Offline Intervention**

Recommended face-to-face components in internet-delivered weight reduction interventions are included; this increases intervention use and effectiveness [35]. Therefore, the offline intervention for Health-On enables nurses, nutritionists, and exercise trainers to inspect goal achievements and provide periodic feedback.

**Pilot Study**

Table 1 shows the baseline characteristics of the participants. Median age was 39 (IQR 35-42) years, and 28 (93.3%) were male. A total of 9 people (30%) had an education level above Master's degree, and the rest were college graduates, suggesting a relatively high level of education.

**Table 1.** Baseline characteristics of study participants (n=30).

Variable	Value
Age (years), median (IQR)	39 (35-42)
Male sex, n (%)	28.0 (93.3)
<b>Academic background, n (%)</b>	
Bachelor's degree	21 (70.0)
Graduate degree or higher	9 (30.0)
Current smoker, n (%)	11 (36.7)
Drinking frequency per week (days), median (IQR)	1.0 (1.0-2.0)
<b>Anthropometry, median (IQR)</b>	
Height (cm)	170.4 (166.5-173.0)
Weight (kg)	81.3 (77.1-87.8)
Waist circumference (cm)	96.8 (93.0-102.5)
BMI <sup>a</sup> (kg/m <sup>2</sup> )	28.0 (27.2-30.3)
<b>Bioimpedance measurement, median (IQR)</b>	
Lean body mass (kg)	55.4 (51.6-58.3)
Body fat (%)	28.2 (25.5-30.8)
<b>Metabolic profile, median (IQR)</b>	
Systolic BP <sup>b</sup> (mmHg)	128 (118-132)
Diastolic BP (mmHg)	75 (70-81)
Fasting plasma glucose (mg/dl)	92 (87-98)
HbA <sub>1c</sub> <sup>c</sup> (mg/dl)	5.6 (5.5-5.8)
Total cholesterol (mg/dl)	205 (174-228)
Triglyceride (mg/dl)	159 (108-214)
HDL <sup>d</sup> cholesterol (mg/dl)	45.5 (37.0-53.0)
LDL <sup>e</sup> cholesterol (mg/dl)	131.5 (97.0-155.0)
Non-HDL cholesterol (mg/dl)	154 (129-195)
AST <sup>f</sup> (IU/L)	24 (19-31)
ALT <sup>g</sup> (IU/L)	30.5 (18-59)
<b>Fat CT<sup>h</sup>, median (IQR)</b>	
CT fat ratio	0.64 (0.50-0.94)
Visceral fat (mm <sup>2</sup> )	140.5 (110.4-192.9)
Subcutaneous fat (mm <sup>2</sup> )	224.3 (184.0-292.8)

<sup>a</sup>BMI: body mass index (calculated as weight in kilograms divided by height in meters squared).

<sup>b</sup>BP: blood pressure.

<sup>c</sup>HbA<sub>1c</sub>: glycosylated hemoglobin.

<sup>d</sup>HDL: high-density lipoprotein.

<sup>e</sup>LDL: low-density lipoprotein.

<sup>f</sup>AST: aspartate aminotransferase.

<sup>g</sup>ALT: alanine aminotransferase.

<sup>h</sup>CT: computed tomography.

The median of variables was weight 81.3 kg (IQR 77.1-87.7), abdominal circumference 96.3 cm (IQR 93.0-102.5), and BMI 28.0 (IQR 27.2-30.3).

Visceral fat and subcutaneous fat were measured by the fat CT, with the median being 140.5 cm<sup>2</sup> (IQR 110.4-192.9) and 224.3 cm<sup>2</sup> (IQR 184.0-292.8), respectively.

Changes in anthropometric and metabolic profiles between pre- and postintervention are shown in Table 2. The mean body weight was decreased by 5.8%. The median of weight, waist circumference, BMI, lean body mass, and body fat percentage reduced significantly, as did most of the metabolic profiles, especially HbA<sub>1c</sub> and non-HDL-C. The changes in visceral fat and subcutaneous fat were statistically significant.

**Table 2.** Comparison of outcomes before and after Health-On program (n=30).

Outcomes	Baseline, median (IQR)	Final, median (IQR)	Difference, median (IQR)	P value <sup>a</sup>
<b>Anthropometric measurement</b>				
Weight (kg)	81.3 (77.1 to 87.8)	76.6 (70.8 to 79.5)	-6.2 (-8.4 to 3.9)	<.001
Waist circumference (cm)	96.8 (93.0 to 102.5)	88 (84.5 to 95.0)	-9.2 (-11 to 5.5)	<.001
BMI <sup>b</sup> (kg/m <sup>2</sup> )	28.0 (27.2 to 30.3)	25.7 (24.6 to 28.0)	-2.2 (-3.4 to 1.5)	<.001
<b>Bioimpedance measurement</b>				
LBM <sup>c</sup> (kg)	55.4 (51.6 to 58.3)	54.0 (51.2 to 57.1)	-1.0 (-2.1 to 0.3)	.003
Body fat (%)	28.2 (25.5 to 30.8)	24.1 (20.9 to 27.6)	-4.65 (-6.5 to 1.8)	<.001
<b>Metabolic profile</b>				
SBP <sup>d</sup> (mmHg)	128 (118 to 132)	121 (107 to 127)	-5.5 (-14 to 2)	.002
DBP <sup>e</sup> (mmHg)	75 (70 to 81)	70 (63 to 82)	-4 (-10 to 2)	.08
Fasting plasma glucose (mg/dl)	92 (87 to 98)	90 (86 to 99)	-1.5 (-8 to 3)	.38
HbA <sub>1c</sub> <sup>f</sup> (mg/dl)	5.6 (5.5 to 5.8)	5.4 (5.2 to 5.6)	-0.2 (-0.4 to 0.1)	<.001
Total cholesterol (mg/dl)	205 (174 to 228)	185 (158 to 198)	-15.5 (-31 to 7)	.001
Triglyceride (mg/dl)	159 (108 to 214)	89 (57 to 124)	-63 (-113 to 25)	<.001
HDL <sup>g</sup> cholesterol (mg/dl)	45.5 (37.0 to 53.0)	51 (43 to 62)	6.5 (1 to 10)	.007
LDL <sup>h</sup> cholesterol (mg/dl)	131.5 (97.0 to 155.0)	107.5 (88 to 125)	-17.5 (-31 to 3)	.001
Non-HDL cholesterol (mg/dl)	154 (129 to 195)	126.5 (107 to 145)	-21.5 (-36 to 11)	<.001
AST <sup>i</sup> (IU/L)	24 (19 to 31)	22.5 (17 to 32)	-1.5 (-9 to 4)	.23
ALT <sup>j</sup> (IU/L)	30.5 (18 to 59)	20.5 (16 to 30)	-5 (-28 to 1)	.006
<b>Fat CT<sup>k</sup></b>				
CT fat ratio	0.64 (0.50 to 0.94)	0.63 (0.4 to 0.92)	-0.05 (-0.19 to 0.06)	.09
Visceral fat (cm <sup>2</sup> )	140.5 (110.4 to 192.9)	95.6 (732.4 to 133.0)	-39.4 (-60.3 to 16.3)	<.001
Subcutaneous fat (cm <sup>2</sup> )	224.3 (184.0 to 292.8)	165.0 (117.3 to 237.9)	-56.4 (-77.0 to 20.6)	<.001

<sup>a</sup>Wilcoxon signed-rank test.

<sup>b</sup>BMI: Body Mass Index (calculated as weight in kilograms divided by height in meters squared).

<sup>c</sup>LBM: lean body mass.

<sup>d</sup>SBP: systolic blood pressure.

<sup>e</sup>DBP: diastolic blood pressure.

<sup>f</sup>HbA<sub>1c</sub>: glycosylated hemoglobin.

<sup>g</sup>HDL: high-density lipoprotein.

<sup>h</sup>LDL: low-density lipoprotein.

<sup>i</sup>AST: aspartate aminotransferase.

<sup>j</sup>ALT: alanine aminotransferase.

<sup>k</sup>CT: computed tomography.



## Discussion

### Principal Findings

The purpose of this study was to develop and verify a comprehensive mobile-based workplace weight reduction program.

Health-On has several strengths compared with other weight reduction apps. Although there are more than 10,000 such apps, a recent review reported that many have insufficient evidence-based content vis-a-vis US government diet and exercise recommendations [35,36]. Health-On incorporates a scientifically evidenced algorithm for estimating negative calorie balance and behavioral strategies established as effective for weight management. In addition, requiring people to complete preprogram health questionnaires, enabled avoidance of participants with health risks. Therefore, Health-On can minimize possible adverse effects, often neglected by other apps.

We created a more convenient method of keying in data. As noted in previous self-monitoring studies, the less intrusive the tool, the higher the rate of adherence [37]. We devised several convenient ways to improve user adherence by collecting data on dietary intake and physical activities. The accuracy and sustainability of keeping a food diary is important; therefore, we attempted to create a user-friendly and nonintrusive UI. Instead of the food diary method, we installed a volume bar for pre-entered diet fields so people can more conveniently record food consumption. This significantly relieves the onerous task of name searching and certainly increases adherence to and sustainability of the program compared with a paper food diary. As for monitoring physical activities, the entry method is more convenient than the paper diary technique that most other apps require. A better approach, such as activity tracking via wearable devices and auto-synchronizing data via Bluetooth, can improve program adherence and compensate for participant self-reporting limitations. Furthermore, physical activities can be quantified and evaluated more accurately.

This program incorporated clinically proven effective behavior change strategies for facilitating weight loss, changes in diet and exercise, and preventing relapse [24,28,37,38]. It is known that if, when deciding the target weight loss, users' diet-exercise preferences are reflected on the target, it will result in greater success [39]. Furthermore, Health-On enables patients to monitor their lifestyles through an everyday life log and allows professionals to provide feedback and educational information. Behavioral changes are promoted through self-monitoring, education, feedback, and competition.

This comprehensive approach combines interventional components that have the strongest effect on obesity (eg, healthy meals) and workplace wellness [15,40]. Most apps do not integrate diet, physical activity, and behavior change strategies [15,41].

Workplace weight management is a highly effective approach to intervention [3,7,16], and Health-On is one of the mobile-based workplace interventions that maximize the advantages and suitability of workplace lifestyle interventions.

Effective population-based weight management is feasible in workplaces, from which supplementary benefits flow, for example, increased productivity, lower absenteeism, and reduced medical costs [3,7,15,16].

In a meta-analytic review, Verweij et al [42] analyzed the effectiveness of workplace interventions targeting physical activity and/or dietary behavior on outcomes. Their study delivered moderate quality evidence that workplace physical activity and dietary behavior interventions significantly reduce body weight (9 studies; mean difference -1.19 kg [95% CI -1.64 to -0.74]) [43]. These studies did not use mobile interventions. Stephans et al [43] performed a 3-month randomized controlled trial using a behavior-based mobile phone app. They reported that the difference in weight change between groups was statistically significant (mobile phone group -1.8 kg vs Control group +0.3kg;  $P=.03$ ). Direct comparison is difficult owing to research methodological differences, but the weight loss of Health-On is considerable. Changes in anthropometry, bioimpedance measures, metabolic profile, and visceral fat were also significant [42].

During calorie restriction, muscle wasting prevention is important because muscles play an important role in improving metabolic profiles and reducing insulin resistance [44]. Nevertheless, without the simultaneous modification of diet and physical activity, losing weight with a low-calorie diet alone might reduce fat and also fat-free mass [45]. It may superficially seem like a significant reduction in weight but resisting metabolic rate and insulin sensitivity would become lower because of a reduction in muscle mass, thus resulting in the tendency of weight regain [45]. This study was characterized with a body fat reduction -4.65% median difference (MD; IQR -6.5 to -1.8), and a decrease in muscle mass 1.0 kg MD (IQR -2.1 to -0.3). By preventing muscle loss that could generally occur with a lower calorie diet, it could be concluded that there had been an ideal weight loss achievement with decreased tendency to long-term weight regain.

Interestingly, there was a significant reduction in visceral fat levels despite the short intervention period. Accumulation of VAT is a clinically important marker as it increases insulin resistance to induce metabolic syndrome and heightens cardiovascular risks; therefore, the results of this study demonstrate that Health-On could be effective in managing obesity and lowering the cardiovascular disease risks [46,47].

Through the reduction in obesity rates in workers, as many studies' results previously showed, improved productivity, and reduced absent rates, a reduction in expenses could be expected from the employer [48-50].

### Limitations

This study had several limitations:

1. This study was designed as a pilot test so that direct weight reduction effect of Health-On could be measured. This design is limited in that it does not allow direct comparisons with other forms of treatment. Prospective, randomized trials with appropriate controls are needed.
2. Health-On was developed on an energy balance equation, which was based on scientific evidence such as

Cunningham's equation, which estimates RMR indirectly [51]. Metabolic rate and calorie expenditure are assumed, not determined through individual differences such as genetic susceptibility to obesity.

3. IPAQ-SF is the desired instrument to measure physical activities, but it considers those within the previous 7 days. There can be differences between daily calorie expenditure calculated from IPAQ-SF and the actual level of physical activity. Although we attempted to compensate for these shortcomings with activity trackers, discrepancies could occur [52].

Despite the short period of this Health-On study, weight reduction was substantial. However, the persistence of weight reduction was unpredictable. Thus, further research for postprogram body weight changes and management is indicated.

### Conclusions

Health-On is a promising workplace intervention tool that can be used in similar environments, for example, universities and the military, with minimal modifications. The results of this study could form a base for designing randomized clinical trials for comparison with conventional weight loss programs. Henceforth, future research should focus on the additional benefits and longitudinal effects of this program.

### Acknowledgments

This study was supported by the Health Connect Co that was created as a joint venture by Seoul National University Hospital and SK Telecom (HC-12-RD-HM-001).

### Conflicts of Interest

None declared.

### References

1. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser 2000;894:i-xii, 1. [Medline: [11234459](#)]
2. 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 [FREE Full text] [doi: [10.1016/S0140-6736\(14\)60460-8](#)] [Medline: [24880830](#)]
3. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet* 2011 Aug 27;378(9793):815-825. [doi: [10.1016/S0140-6736\(11\)60814-3](#)] [Medline: [21872750](#)]
4. Wilson PW, 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 9;162(16):1867-1872. [doi: [10.1001/archinte.162.16.1867](#)] [Medline: [12196085](#)]
5. Specchia ML, Veneziano MA, Cadeddu C, Ferriero AM, Mancuso A, Ianuale C, et al. Economic impact of adult obesity on health systems: a systematic review. *Eur J Public Health* 2015 Apr;25(2):255-262. [doi: [10.1093/eurpub/cku170](#)] [Medline: [25320051](#)]
6. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: [10.1056/NEJMoa012512](#)] [Medline: [11832527](#)]
7. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J Am Coll Cardiol* 2014 Jul 1;63(25 Pt B):2985-3023 [FREE Full text] [doi: [10.1016/j.jacc.2013.11.004](#)] [Medline: [24239920](#)]
8. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res* 2013 Apr 15;15(4):e32 [FREE Full text] [doi: [10.2196/jmir.2283](#)] [Medline: [23587561](#)]
9. Gilmore LA, Duhé AF, Frost EA, Redman LM. The technology boom: a new era in obesity management. *J Diabetes Sci Technol* 2014 May;8(3):596-608 [FREE Full text] [doi: [10.1177/1932296814525189](#)] [Medline: [24876625](#)]
10. Balk-Møller NC, Poulsen SK, Larsen TM. Effect of a nine-month web- and app-based workplace intervention to promote healthy lifestyle and weight loss for employees in the social welfare and health care sector: a randomized controlled trial. *J Med Internet Res* 2017 Apr 10;19(4):e108 [FREE Full text] [doi: [10.2196/jmir.6196](#)] [Medline: [28396303](#)]
11. Oh B, Cho B, Han MK, Choi H, Lee MN, Kang H, et al. The effectiveness of mobile phone-based care for weight control in metabolic syndrome patients: randomized controlled trial. *JMIR Mhealth Uhealth* 2015 Aug 20;3(3):e83 [FREE Full text] [doi: [10.2196/mhealth.4222](#)] [Medline: [26293568](#)]

12. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MD, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. *J Med Internet Res* 2019 Jan 08;21(1):e298 [FREE Full text] [doi: [10.2196/jmir.9609](https://doi.org/10.2196/jmir.9609)] [Medline: [30622090](https://pubmed.ncbi.nlm.nih.gov/30622090/)]
13. Kim J, Oh S, Steinhubl S, Kim S, Bae WK, Han JS, et al. Effectiveness of 6 months of tailored text message reminders for obese male participants in a worksite weight loss program: randomized controlled trial. *JMIR Mhealth Uhealth* 2015 Feb 3;3(1):e14 [FREE Full text] [doi: [10.2196/mhealth.3949](https://doi.org/10.2196/mhealth.3949)] [Medline: [25648325](https://pubmed.ncbi.nlm.nih.gov/25648325/)]
14. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
15. Dobbs R, Sawers C, Thompson F, Manyika J, Woetzel J, Child P, et al. McKinsey & Company. 2014. How the World Could Better Fight Obesity URL: <https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/how-the-world-could-better-fight-obesity> [accessed 2019-10-01]
16. Kumanyika SK, Obarzanek E, Stettler N, Bell R, Field AE, Fortmann SP, American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention. Population-based prevention of obesity: the need for comprehensive promotion of healthful eating, physical activity, and energy balance: a scientific statement from American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention (formerly the expert panel on population and prevention science). *Circulation* 2008 Jul 22;118(4):428-464. [doi: [10.1161/CIRCULATIONAHA.108.189702](https://doi.org/10.1161/CIRCULATIONAHA.108.189702)] [Medline: [18591433](https://pubmed.ncbi.nlm.nih.gov/18591433/)]
17. Android Developers. Android Studio URL: <https://developer.android.com/studio/> [accessed 2018-05-10] [WebCite Cache ID 6zItoM9wI]
18. Lee SH, Kim EK. Accuracy of predictive equations for resting metabolic rates and daily energy expenditures of police officials doing shift work by type of work. *Clin Nutr Res* 2012 Jul;1(1):66-77 [FREE Full text] [doi: [10.7762/cnr.2012.1.1.66](https://doi.org/10.7762/cnr.2012.1.1.66)] [Medline: [23429979](https://pubmed.ncbi.nlm.nih.gov/23429979/)]
19. Chang UJ, Lee KR. Correlation between measured resting energy expenditure and predicted basal energy expenditure in female college students. *J Korean Soc Food Sci Nutr* 2005;34(2):196-201 [FREE Full text] [doi: [10.3746/jkfn.2005.34.2.196](https://doi.org/10.3746/jkfn.2005.34.2.196)]
20. Sjostrom M, Ekelund U, Poortvliet E, Hurtig-Wennlöf A, Yngve A. Assessment of physical activity using IPAQ (version 4) and activity monitors (CSA). *Meas Phys Educ Exerc Sci* 2000;4(4):263-264 [FREE Full text]
21. Oh JY, Yang YJ, Kim BS, Kang JH. Validity and reliability of Korean version of International Physical Activity Questionnaire (IPAQ) short form. *J Korean Med Sci* 2007 Jul;28(7):532-541 [FREE Full text]
22. Brooks GA, Butte NF, Rand WM, Flatt J, Caballero B. Chronicle of the Institute of Medicine physical activity recommendation: how a physical activity recommendation came to be among dietary recommendations. *Am J Clin Nutr* 2004 May;79(5):921S-930S. [doi: [10.1093/ajcn/79.5.921S](https://doi.org/10.1093/ajcn/79.5.921S)] [Medline: [15113740](https://pubmed.ncbi.nlm.nih.gov/15113740/)]
23. Willett W. *Nutritional Epidemiology*. New York, NY: Oxford University Press; 1998.
24. Cowburn G, Hillsdon M, Hankey CR. Obesity management by life-style strategies. *Br Med Bull* 1997;53(2):389-408. [doi: [10.1093/oxfordjournals.bmb.a011619](https://doi.org/10.1093/oxfordjournals.bmb.a011619)] [Medline: [9246842](https://pubmed.ncbi.nlm.nih.gov/9246842/)]
25. The Korean Nutrition Society. Computer aided nutritional analysis program (CANPRO) version 4 URL: [http://www.kns.or.kr/pagegenerater.asp?catalogid=kns2008&language=ko&pagecode=sub08\\_02a](http://www.kns.or.kr/pagegenerater.asp?catalogid=kns2008&language=ko&pagecode=sub08_02a) [accessed 2012-04-05]
26. Korea Food and Drug Administration. FANTASY (Food And Nutrient daTA SYstem). URL: <http://kissna.kfda.go.kr/> [accessed 2012-04-05]
27. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011 Aug;43(8):1575-1581. [doi: [10.1249/MSS.0b013e31821ece12](https://doi.org/10.1249/MSS.0b013e31821ece12)] [Medline: [21681120](https://pubmed.ncbi.nlm.nih.gov/21681120/)]
28. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, 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](https://doi.org/10.1016/j.amepre.2013.07.005)] [Medline: [24139771](https://pubmed.ncbi.nlm.nih.gov/24139771/)]
29. Ritterband LM, Tate DF. The science of internet interventions. Introduction. *Ann Behav Med* 2009 Aug;38(1):1-3. [doi: [10.1007/s12160-009-9132-5](https://doi.org/10.1007/s12160-009-9132-5)] [Medline: [19816750](https://pubmed.ncbi.nlm.nih.gov/19816750/)]
30. Park JH, Kwon H, Oh S, Lee CM, Cho B. Update and validation of a national health risk appraisal tool in Korea. *J Public Health (Oxf)* 2013 Mar;35(1):107-114. [doi: [10.1093/pubmed/fds061](https://doi.org/10.1093/pubmed/fds061)] [Medline: [22888146](https://pubmed.ncbi.nlm.nih.gov/22888146/)]
31. Oh SW, Shin S, Yun YH, Yoo T, Huh B. Cut-off point of BMI and obesity-related comorbidities and mortality in middle-aged Koreans. *Obes Res* 2004 Dec;12(12):2031-2040 [FREE Full text] [doi: [10.1038/oby.2004.254](https://doi.org/10.1038/oby.2004.254)] [Medline: [15687405](https://pubmed.ncbi.nlm.nih.gov/15687405/)]
32. Adams R. Revised physical activity readiness questionnaire. *Can Fam Physician* 1999 Apr;45:992, 995, 1004-992, 995, 1005 [FREE Full text] [Medline: [10216799](https://pubmed.ncbi.nlm.nih.gov/10216799/)]
33. Leahey TM, Kumar R, Weinberg BM, Wing RR. Teammates and social influence affect weight loss outcomes in a team-based weight loss competition. *Obesity (Silver Spring)* 2012 Jul;20(7):1413-1418 [FREE Full text] [doi: [10.1038/oby.2012.18](https://doi.org/10.1038/oby.2012.18)] [Medline: [22310234](https://pubmed.ncbi.nlm.nih.gov/22310234/)]
34. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J Med Internet Res* 2013 Jul 16;15(7):e146 [FREE Full text] [doi: [10.2196/jmir.2665](https://doi.org/10.2196/jmir.2665)] [Medline: [23859884](https://pubmed.ncbi.nlm.nih.gov/23859884/)]

35. The Wall Street Journal. 2011. Software for Shaping Up: Our Picks for Some of the Best Apps to Monitor and Improve Diet and Fitness URL: <https://www.wsj.com/articles/SB10001424052748703961104576148732585957902> [accessed 2018-05-10] [WebCite Cache ID 6zIwGxW4J]
36. Breton ER, Fuemmeler BF, Abrams 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](https://doi.org/10.1007/s13142-011-0076-5)] [Medline: [24073074](https://pubmed.ncbi.nlm.nih.gov/24073074/)]
37. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. *J Am Diet Assoc* 2011 Jan;111(1):92-102 [FREE Full text] [doi: [10.1016/j.jada.2010.10.008](https://doi.org/10.1016/j.jada.2010.10.008)] [Medline: [21185970](https://pubmed.ncbi.nlm.nih.gov/21185970/)]
38. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, van Horn L, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. *Circulation* 2010 Jul 27;122(4):406-441. [doi: [10.1161/CIR.0b013e3181e8edf1](https://doi.org/10.1161/CIR.0b013e3181e8edf1)] [Medline: [20625115](https://pubmed.ncbi.nlm.nih.gov/20625115/)]
39. Mühlbacher A, Bethge S. Preferences of overweight and obese patients for weight loss programmes: a discrete-choice experiment. *Int J Integr Care* 2013;13:e034 [FREE Full text] [doi: [10.5334/ijic.1113](https://doi.org/10.5334/ijic.1113)] [Medline: [24179457](https://pubmed.ncbi.nlm.nih.gov/24179457/)]
40. Wing RR, Hill JO. Successful weight loss maintenance. *Annu Rev Nutr* 2001;21:323-341. [doi: [10.1146/annurev.nutr.21.1.323](https://doi.org/10.1146/annurev.nutr.21.1.323)] [Medline: [11375440](https://pubmed.ncbi.nlm.nih.gov/11375440/)]
41. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res* 2014 Aug 15;16(8):e192 [FREE Full text] [doi: [10.2196/jmir.3469](https://doi.org/10.2196/jmir.3469)] [Medline: [25131661](https://pubmed.ncbi.nlm.nih.gov/25131661/)]
42. Verweij LM, Coffeng J, van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. *Obes Rev* 2011 Jun;12(6):406-429. [doi: [10.1111/j.1467-789X.2010.00765.x](https://doi.org/10.1111/j.1467-789X.2010.00765.x)] [Medline: [20546142](https://pubmed.ncbi.nlm.nih.gov/20546142/)]
43. Stephens JD, Yager AM, Allen J. Smartphone technology and text messaging for weight loss in young adults: a randomized controlled trial. *J Cardiovasc Nurs* 2017;32(1):39-46 [FREE Full text] [doi: [10.1097/JCN.0000000000000307](https://doi.org/10.1097/JCN.0000000000000307)] [Medline: [26646593](https://pubmed.ncbi.nlm.nih.gov/26646593/)]
44. Moon S. Low skeletal muscle mass is associated with insulin resistance, diabetes, and metabolic syndrome in the Korean population: the Korea National Health and Nutrition Examination Survey (KNHANES) 2009-2010. *Endocr J* 2014;61(1):61-70 [FREE Full text] [doi: [10.1507/endocrj.ej13-0244](https://doi.org/10.1507/endocrj.ej13-0244)] [Medline: [24088600](https://pubmed.ncbi.nlm.nih.gov/24088600/)]
45. Stiegler P, Cunliffe A. The role of diet and exercise for the maintenance of fat-free mass and resting metabolic rate during weight loss. *Sports Med* 2006;36(3):239-262. [doi: [10.2165/00007256-200636030-00005](https://doi.org/10.2165/00007256-200636030-00005)] [Medline: [16526835](https://pubmed.ncbi.nlm.nih.gov/16526835/)]
46. Lebovitz HE, Banerji MA. Point: visceral adiposity is causally related to insulin resistance. *Diabetes Care* 2005 Sep;28(9):2322-2325. [doi: [10.2337/diacare.28.9.2322](https://doi.org/10.2337/diacare.28.9.2322)] [Medline: [16123512](https://pubmed.ncbi.nlm.nih.gov/16123512/)]
47. Seo D, Niu J. Evaluation of internet-based interventions on waist circumference reduction: a meta-analysis. *J Med Internet Res* 2015 Jul 21;17(7):e181 [FREE Full text] [doi: [10.2196/jmir.3921](https://doi.org/10.2196/jmir.3921)] [Medline: [26199208](https://pubmed.ncbi.nlm.nih.gov/26199208/)]
48. Naito M, Nakayama T, Okamura T, Miura K, Yanagita M, Fujieda Y, HIPOP-OHP Research Group. Effect of a 4-year workplace-based physical activity intervention program on the blood lipid profiles of participating employees: the high-risk and population strategy for occupational health promotion (HIPOP-OHP) study. *Atherosclerosis* 2008 Apr;197(2):784-790. [doi: [10.1016/j.atherosclerosis.2007.07.026](https://doi.org/10.1016/j.atherosclerosis.2007.07.026)] [Medline: [17868680](https://pubmed.ncbi.nlm.nih.gov/17868680/)]
49. Muto T, Yamauchi K. Evaluation of a multicomponent workplace health promotion program conducted in Japan for improving employees' cardiovascular disease risk factors. *Prev Med* 2001 Dec;33(6):571-577. [doi: [10.1006/pmed.2001.0923](https://doi.org/10.1006/pmed.2001.0923)] [Medline: [11716652](https://pubmed.ncbi.nlm.nih.gov/11716652/)]
50. Aldana SG, Merrill RM, Price K, Hardy A, Hager R. Financial impact of a comprehensive multisite workplace health promotion program. *Prev Med* 2005 Feb;40(2):131-137. [doi: [10.1016/j.ypmed.2004.05.008](https://doi.org/10.1016/j.ypmed.2004.05.008)] [Medline: [15533521](https://pubmed.ncbi.nlm.nih.gov/15533521/)]
51. Horton ES. Introduction: an overview of the assessment and regulation of energy balance in humans. *Am J Clin Nutr* 1983 Dec;38(6):972-977. [doi: [10.1093/ajcn/38.6.972a](https://doi.org/10.1093/ajcn/38.6.972a)] [Medline: [6650453](https://pubmed.ncbi.nlm.nih.gov/6650453/)]
52. Chastin SFM, Culhane B, Dall PM. Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). *Physiol Meas* 2014 Nov;35(11):2319-2328. [doi: [10.1088/0967-3334/35/11/2319](https://doi.org/10.1088/0967-3334/35/11/2319)] [Medline: [25341050](https://pubmed.ncbi.nlm.nih.gov/25341050/)]

## Abbreviations

- ALT:** alanine aminotransferase
- CT:** computed tomography
- DBP:** diastolic blood pressure
- FBS:** fasting blood sugar
- HDL-C:** high-density lipoprotein-cholesterol
- IPAQ-SF:** International Physical Activity Questionnaire Short Form
- IQR:** interquartile range
- LDL-C:** low-density lipoprotein-cholesterol
- MD:** median difference

**RMR:** resting metabolic rate  
**SBP:** systolic blood pressure  
**TEA:** thermic effect of activity  
**TEE:** total energy expenditure  
**TG:** triglyceride  
**UI:** user interface  
**VAT:** visceral adipose tissue

*Edited by G Eysenbach; submitted 06.06.18; peer-reviewed by B Oh, J Apolzan; comments to author 29.09.18; revised version received 23.02.19; accepted 19.08.19; published 04.11.19.*

*Please cite as:*

*Han MK, Cho B, Kwon H, Son KY, Lee H, Lee JK, Park J*

*A Mobile-Based Comprehensive Weight Reduction Program for the Workplace (Health-On): Development and Pilot Study*

*JMIR Mhealth Uhealth 2019;7(11):e11158*

*URL: <https://mhealth.jmir.org/2019/11/e11158>*

*doi: [10.2196/11158](https://doi.org/10.2196/11158)*

*PMID: [31682576](https://pubmed.ncbi.nlm.nih.gov/31682576/)*

©Min Kyu Han, Belong Cho, Hyuktae Kwon, Ki Young Son, Hyejin Lee, Joo Kyung Lee, Jinho Park. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 04.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Development of a Mobile Phone App for Measuring Striking Response Time in Combat Sports: Cross-Sectional Validation Study

Victor Coswig<sup>1\*</sup>, BSc, MSc, PhD; Jader Sant' Ana<sup>2\*</sup>, BSc, MSc; Maicon Nascimento Coelho<sup>2\*</sup>, BSc; Antonio Renato Pereira Moro<sup>2\*</sup>, BSc, MSc, PhD; Fernando Diefenthaler<sup>2\*</sup>, BSc, MSc, PhD

<sup>1</sup>Universidade Federal do Pará, Departamento de Educação Física, Castanhal, Brazil

<sup>2</sup>Universidade Federal de Santa Catarina, Departamento de Educação Física, Florianópolis, Brazil

\* all authors contributed equally

**Corresponding Author:**

Fernando Diefenthaler, BSc, MSc, PhD

Universidade Federal de Santa Catarina

Departamento de Educação Física

Campus Universitário

Reitor João David Ferreira Lima

Florianópolis, 88040-900

Brazil

Phone: 55 4837219000 ext 4779

Email: [fernando.diefenthaler@ufsc.br](mailto:fernando.diefenthaler@ufsc.br)

## Abstract

**Background:** TReaction is a mobile app developed to determine strike response time at low cost and with easy application in combat sports. However, the validity and accuracy of the response time obtained by the TReaction app has not yet been evaluated.

**Objective:** This study aimed to test the validity and reliability of the TReaction app in measuring motor response time in combat sports.

**Methods:** A total of two athletes performed 59 strikes to assess the response time upon visual stimulus using the TReaction app simultaneously with a high-speed camera. Accuracy of the measure was verified using a computer simulator programmed to discharge visual stimuli and obtain the response time. Pearson correlation, Student *t* test for dependent samples, and the Bland-Altman analysis were performed. Accuracy was verified using the intraclass correlation coefficient. Effect size (*g*) and the typical error of measurement (TEM) were calculated. The significance level was set at  $P < .05$ .

**Results:** No significant difference ( $P = .56$ ) was found between both systems. The methods presented a very strong correlation ( $r = 0.993$ ). The magnitude of differences was trivial ( $g < 0.25$ ), and TEM was 1.4%. These findings indicate a high accuracy between the computer screen and the mobile app measures to determine the beginning of the task and the response time.

**Conclusions:** Our findings suggest that the TReaction app is a valid tool to evaluate the response time in combat sports athletes.

(JMIR Mhealth Uhealth 2019;7(11):e14641) doi:[10.2196/14641](https://doi.org/10.2196/14641)

## KEYWORDS

reaction time; martial arts; mobile apps; software validation

## Introduction

### Background

The success of combat sports is often determined by fast motor actions in response to a given stimulus [1-3]. Response time, defined as the time required to perform a voluntary movement after a stimulus [4,5], is an important variable that can be improved and monitored in combat sports conditioning [6-9].

A greater success during competition could be associated with faster response time value [10,11].

### Technology Systems to Measure Response Time

To measure response time from a given visual or audible stimulus, different systems have been proposed to determine the time to perform a specific task or motor action. Most systems use cameras with high sampling frequency [12,13] or

accelerometers [14,15] to obtain spatial and temporal information of a determined motor gesture or task. Moreover, some systems measure response time based on visual stimuli shown on a computer screen in time intervals as previously defined. In these cases, response time to visual stimulus is measured from the moment that the user presses the computer keyboard [16,17].

### Limitations of Systems to Measure Response Time

Despite the quick results provided by those systems in which computer keyboards are used, they lack specificity (specifically motor gestures) and ecological validity, as combat sports actions are not performed during the task. In addition, systems involving kinematics (motion analysis) or accelerometry to measure response time usually require more time to prepare, collect, and analyze the data. Moreover, these systems are expensive and often have low applicability to athlete training routines.

### Objective and Research Questions

To minimize the aforementioned limitations, the use of a specific mobile phone app (TReaction app, ETS4ME, São José, SC, Brazil) has been proposed to determine the response time of different motor actions (strikes) with a low-cost, highly applicable, and ecologically valid solution. The TReaction app measures the motor response time between the flash fire and sound waves produced from contact with the target through the mobile phone microphone. However, as far as we know, the response time measurement validity obtained by the TReaction app has not been tested. In fact, the use of mobile apps has been growing exponentially in fitness, health, and sports science, which raises the need for scientific evidence [18]. Therefore, this study aimed to test the validity and reliability of the TReaction app in measuring motor response time in combat sports.

## Methods

### Experimental Design

Data collection for this study was performed in 2 sessions. In the first session, a response time test was performed using the TReaction app simultaneously with a high-speed camera (visual system). In the second session, the TReaction app measurement reliability was validated with a repeated measures design using a simulator. This simulator was programmed to verify the accuracy of the visual stimuli and the response time with previously programmed fixed intervals.

### Subjects

The 2 subjects of this study were selected intentionally. These individuals were male black belt athletes with significant experience in regional and national taekwondo competitions. Furthermore, each selected subject had to have at least 10 years of experience in a modality of combat sport. Before participating in the study, all participants were informed and provided written consent about the specific aims and methods of the study along with possible study risks. The local research ethics committee approved this study (protocol number 145.882). All procedures were performed in accordance with the Declaration of Helsinki.

### Procedures

#### Motor Response Time

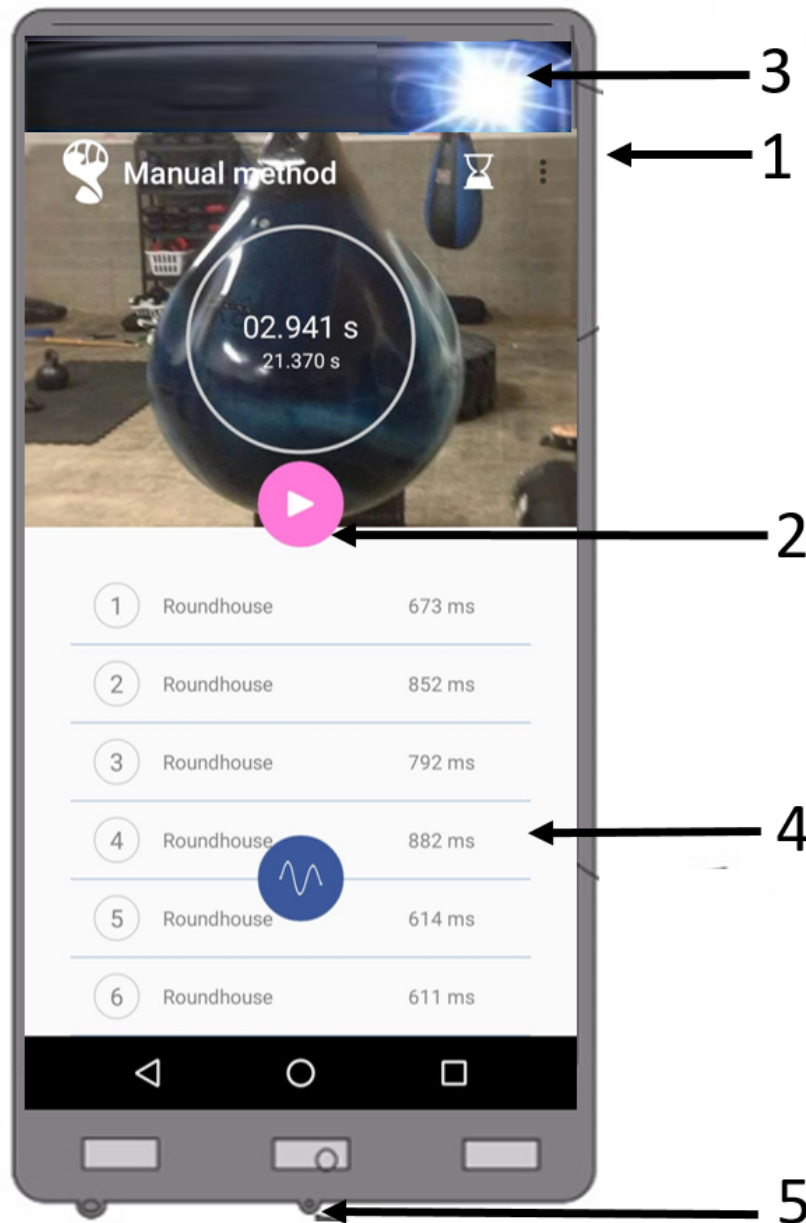
The TReaction app was used to determine the motor response time and the accuracy of the measurement was validated through a repeated measures design that examined motor gesture accuracy during contact with an impact pad. The measurements of values obtained by the mobile app (TReaction app) were compared with those obtained using a digital high-speed camera, which is considered the gold standard method.

#### Mobile App

The response time obtained from the TReaction app is provided as shown in Figure 1 and described below:

- Number 1 represents the mobile phone;
- Number 2 represents a play button via touchscreen technology present in the mobile app interface and that activates the mobile phone camera flash through an algorithm software developed in Java programming language using integrated development environment Android Studio (Apache Software Foundation, Maryland, EUA) and that allows to acquire and process data;
- Number 3 represents the mobile phone camera flash, emitting the visual stimulus and indicates the beginning of the response time to perform a gesture or task to reach a contact target (ie, reach an impact pad for kicks);
- Number 4 represents when contact with the target occurs and a sound wave is generated and processed through the software developed to accurately identify and point response time test results obtained;
- Number 5 represents the mobile microphone to audio capture.

**Figure 1.** Representation of the TReaction app system. See text above for explanation of numbered points.



### **High-Speed Camera**

To accurately identify the strike response time (defined as the time interval between the light signal and the moment of contact with the target), we performed a simultaneous motion analysis using a high-speed digital camera (Fastec TS4, Motion Engineering Company) with a sampling frequency of 1000 Hz and compared it with repeated measures. The response time was determined using a software analysis (Fastec FasMotion, Motion Engineering Company), capturing the first frame when the flashlight appeared and the moment of the material restitution caused by the contact of the foot with the contact pad.

### **Motor Response Time Measurement Protocol**

Both athletes performed a total of 59 strikes using the roundhouse kick technique at the height of the trunk. To

calculate the response time, 2 different mobile phone devices were used (30 strikes registered with a Samsung Galaxy J7 Neo with Android 7.0 system, and 29 with a Motorola G Moto G Play with Android 6.0.1 system). The tests were conducted as shown in Figure 2, with strikes being performed aiming to reach the impact pad in response to the visual stimulus (camera flash signal). To perform the strike, the athlete positioned himself in front of the impact pad that was held by an experienced researcher. The mobile device with the TReaction app was held by another researcher and positioned sideways to the impact pad so that the camera flash was directed to the athlete's vision field. The high-speed camera was positioned perpendicular to the task plane of motion and in such a way to include the athlete, the impact pad, and the mobile's camera flash. Both athletes had previous experience with this protocol.



**Figure 2.** Set up of the protocol test position to identify response time of the strikes.



### **Accuracy Measurement Protocol**

To verify the accuracy of the measurement over repeated measures, a software running with an algorithm developed in Java language, simulating the strike sound and simultaneous change in the computer monitor screen color (Notebook Acer aspire E 15, Acer), was used. The monitor screen was alternating between black and white with the following epochs previously fixed at 3014, 3004, 3014, and 3002 ms. This configuration was chosen to test the consistency of the changes in the monitor screen. First, the TReaction app was set up in the consecutive audio capture mode to verify response time reliability via audio acquisition. Then it was set up in the flash drive mode repeatedly for each auditory stimulus emitted by the computer, with following epochs previously fixed at 3014, 3004, 3014, and 3002 ms. At this moment, the high-speed camera was used to simultaneously record repeated responses of the mobile phone camera flash and the color change of the monitor screen.

### **Statistical Analysis**

Descriptive statistics (mean [SD]) were used to present the results, and normality was verified using the Shapiro-Wilk test. Pearson linear correlation was used to verify the relationship between response time obtained by the TReaction app and high-speed camera. A paired Student *t* test was used to test differences between both methods. The magnitude of the

differences was verified from the effect size (*g*), and it was then scaled for trivial (0.25), small (0.25 to 0.50), moderate (0.50 to 1.0), or large (1.0) effects [19]. The typical error of measurement (TEM) was estimated by dividing the SD of the difference score by  $\sqrt{2}$  [20]. The Bland and Altman [21] analyses were used to test agreement between the methods. The intraclass correlation coefficient (ICC) was established to verify the accuracy of repeated measures. Data processing and analysis were performed using Microsoft Office Excel 2007, SPSS 17.0 (IBM Corp) and GraphPad Prism 5.01. The statistical significance level was set at  $P < .05$ .

### **Results**

Table 1 shows the mean and SD values, level of significance of the differences, the magnitude of the effect size, and the TEM for the response time obtained by the high-speed camera and the TReaction app.

For mean response time values between the high-speed camera and the TReaction app, there was no difference, and the magnitude of the differences obtained was classified as trivial. The methods to determine the response time presented a strong correlation, and the TEM pointed to the confidence in the measurement obtained from the TReaction app in relation to the high-speed camera.

The Bland and Altman plots describing the agreement of response time measurements obtained with a high-speed camera and the TReaction app is presented in Figure 3. Response time presented a bias of 13.05 ms. The 95% limits of agreement for

differences of response time between the high-speed camera and TReaction app were 27.06 ms or -14.01 ms and 40.11 ms (below and above bias, respectively).

**Table 1.** Mean and standard deviation for the response time obtained using the high-speed camera and TReaction app (n=59).

Parameter	High-speed camera, mean (SD)	TReaction app, mean (SD)	r <sup>a</sup> value	P value	g (ES <sup>b</sup> ) <sup>c</sup>	TEM <sup>d</sup> (%)
Response time (ms)	673 (115)	686 (116)	0.993	.56	0.113	9.2 (1.4)

<sup>a</sup>Pearson linear correlation.

<sup>b</sup>ES: effect size.

<sup>c</sup>Magnitude of the differences.

<sup>d</sup>TEM: typical error of measurement (absolute and relative values).

**Figure 3.** Bland and Altman plots describing the agreement of measurements of response time obtained with high speed camera and TReaction app.

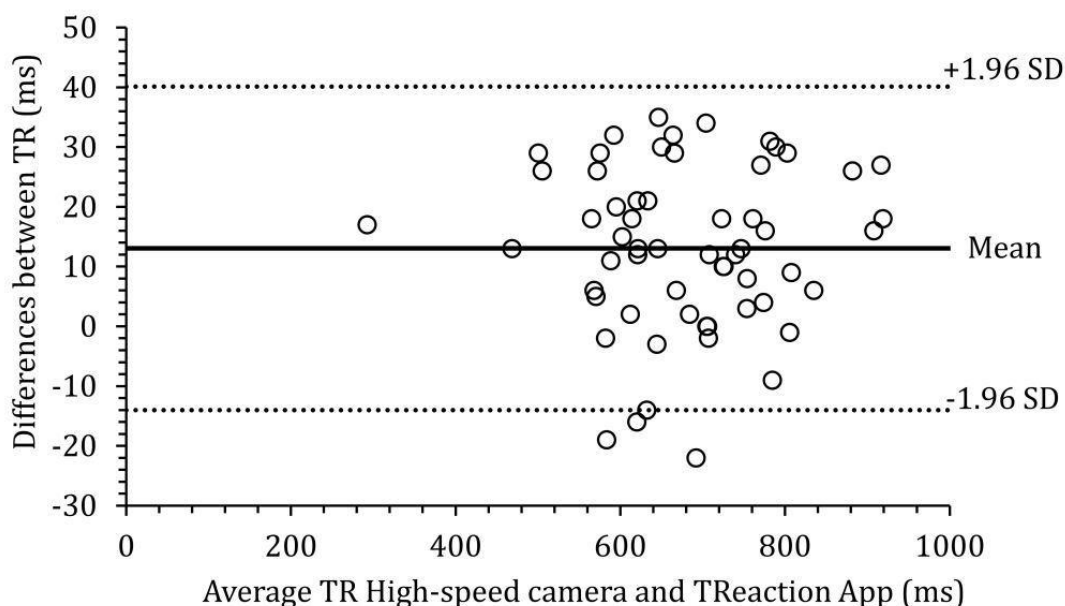


Table 2 shows the values of the repeated measures of audio emission time and computer screen color change and the capture of the audio system obtained with the TReaction app. There was no difference in mean time values in comparison with the values obtained by the repeated measures test for change in the

computer screen color and to the response time obtained by the audio capture using the TReaction app. The magnitude of the differences obtained was classified as trivial and the TEM pointed to reliability showed a strong correlation established by the ICC for the repeated measures.

**Table 2.** Mean and standard deviation for the comparison of the repeated measures of the audio emission time and the screen change of the simulator, as well as the capture of the audio system obtained with the TReaction app (n=52).

Parameter	Audio and screen (computer), mean (SD)	Audio TReaction, mean (SD)	P value	g (ES <sup>a</sup> ) <sup>b</sup>	TEM <sup>c</sup> (%)	ICC <sup>d</sup>
Time (ms)	3009 (5.5)	3008 (5.6)	.37	0.011	0.04 (0.001)	0.998

<sup>a</sup>ES: effect size.

<sup>b</sup>Magnitude of the differences.

<sup>c</sup>TEM: typical error of measurement (absolute and relative values).

<sup>d</sup>ICC: intraclass correlation coefficient.

Table 3 shows repeated measurement values of the mobile phone camera flash opening time to trigger the motor response onset and the computer screen change time. There was no difference between mean values of the flash opening time and values obtained by the repeated measurements test using the TReaction

app and the measures of the color change in the computer screen. The magnitude of the differences obtained was classified as trivial, and the TEM pointed to accuracy; a strong correlation was established between the repeated measurements to determine the visual stimulus to perform a motor response.

**Table 3.** Mean and standard deviation for the comparison between values of the flash opening time and values obtained by the repeated measurements test using the TReaction app and the measures of the color change in the computer screen (n=52).

Parameter	Flash (Fastec)	Screen (Fastec)	$r^a$ value	$P$ value	$g$ (ES) <sup>b,c</sup>	TEM <sup>d</sup> (%)
Time (ms)	3007 (16.5)	3009 (5.5)	0.846	.43	0.103	0.8 (0.03)

<sup>a</sup>Pearson linear correlation.

<sup>b</sup>ES: effect size.

<sup>b</sup>Magnitude of the differences.

<sup>c</sup>TEM: typical error of measurement (absolute and relative values).

## Discussion

### Principal Findings

This study aimed to verify the validity of the TReaction app, which was designed to evaluate the motor response time of combat sports athletes. In addition, this study aimed to verify if the setup used by the TReaction app to determine the beginning of the task (flashlight stimulus) and the time of contact with the target (time to perform the task) would reach acceptable levels of accuracy. To the best of our knowledge, this is the first study to investigate the validity of the TReaction app. The main results showed no differences between the mean response time of all strikes evaluated using the TReaction app and those obtained by the high-speed camera, which is supported by the effect size observed (trivial). The response time results showed a strong correlation between both TReaction app and high-speed camera (gold standard) methods. Besides, the TEM in our study (9.2 ms) is lower than the value (30 ms) observed previously with single-beam photocells across short-sprint studies [22].

When using the TReaction app, considerations must be given to the values of differences by the inference and the degrees of agreement of the Bland and Altman plot. Our results show an average difference of 13 ms for response time between TReaction app and high-speed camera. Response time presented a bias of 13.05 ms, and the 95% limit of agreement for the differences between the high-speed camera and TReaction app was 27.06 ms with 3 individual samples slightly outside the limit of agreement. Although the response time observed in our study presented bias higher than the value (0.35 ms) detected by [23], the variance of the magnitude of different values can be considered similar to the values observed by the authors (95% limits of agreement -11.78 to 12.39 ms). In the above-mentioned study, the authors also observed 3 individual samples slightly outside the limit of agreement, and the difference between systems to measure response time during the task was higher than 17 ms for an average time of 246 ms.

This study demonstrated that the difference between the TReaction app and high-speed camera observed is associated with the distinct method that the TReaction app uses to calculate the time between the visual stimulus (flash fire) to perform the task and the moment to determine the motor response time. The TReaction app calculates response time based on the sound wave generated by the impact with the target object. This measure presented a high ICC and no differences between the audio output and the previously programmed computer software.

The kinematics is among the most accurate systems used to measure motor gestures associated with performance such as vertical jump [24,25], especially when the goal is to accurately measure the performance of motor gestures in combat sports [13,26]. However, high cost and low applicability of these kinematics systems make it challenging to control essential parameters in combat sports. Therefore, the cost-effectiveness and high applicability of utilizing the TReaction app make it a viable alternative for accurate response time measurements in combat sports.

### Study Limitations

One limitation of the TReaction app may be associated with the small variability observed in the firing time of the flash over repeated measures, which increases the probability of incorrectly identifying the moment of the initial stimulus to perform the task. However, in this study, a high-speed camera (1000 Hz) was used to establish the response time with maximum accuracy, and no differences were observed when compared with the response time values obtained by TReaction app. Another aspect to consider when interpreting our findings is regarding the limited sample size, as a higher variability can be expected from international and national level athletes (coefficient of variation 2.9% and 6.1%, respectively) in response time when repeated measures of kicks are performed [27], which means that our perspectives are limited for this group. Although we believe that it did not reduce the strength of our findings, future studies should be encouraged to investigate TReaction feasibility in different modalities and subjects' characteristics, such as competitive levels.

### Final Considerations

Finally, the reliability observed for computer screen color change time, previously programmed to occur simultaneously to the auditory stimulus at predefined times, was also verified. Results observed allow the use of the screen color change as an alternative to be used as a visual stimulus to initiate a motor task. Therefore, we suggest that the TReaction app can accurately measure response time during different techniques used in combat sports. This is relevant because response time is an important variable for combat sports [8,10,15,26]. As the TReaction app is a viable alternative to obtain and control the response time for combat sports athletes, it can possibly be used by athletes in other sports, physical trainers, and coaches to obtain time interval measurements between stimulus onset and the response time required for contacting a target with a strike.

## Conclusions

The results of this study suggest that TReaction app is a valid and reliable tool for measuring response time in combat sports. However, the TEM and observed interference of agreement should be considered when comparing the response time values obtained by TReaction app with other kinematic measurements.

The TReaction app provides an easily executable test that provides relevant sport-specific data regarding motor response

time. Outcomes from the TReaction app suggest that athletes and coaches involved in combat sports can obtain reliable continuous motivational feedback leading to accurate diagnostic and performance criteria. The TReaction app can achieve and better control training-induced adaptations as well as help choose interventions that will help achieve maximum combat sports success. In addition, the TReaction app showed to be a valid tool that is both cost-effective and highly applicable to combat sports practitioners.

## Acknowledgments

The authors wish to thank the Pro-Rector for Research and Post-Graduation of Federal University of Pará for financial support and National Council for Scientific and Technological Development (CNPq).

## Authors' Contributions

VC, JS, MNC, ARPM, and FD established the research design, conducted data analysis, discussed the results, and wrote the final version of the manuscript.

## Conflicts of Interest

JS and MNC are the TReaction app developers.

## References

1. Bouhlel E, Jouini A, Gmada N, Nefzi A, Ben Abdallah K, Tabka Z. Heart rate and blood lactate responses during Taekwondo training and competition. *Sci Sport* 2006 Oct;21(5):285-290 [FREE Full text] [doi: [10.1016/j.scispo.2006.08.003](https://doi.org/10.1016/j.scispo.2006.08.003)]
2. Özdemir RA, Kirazcı S, Ugras A. Simple reaction time and decision making performance after different physical workloads: an examination with elite athletes. *Int J Hum Sci* 2010;7(2):655-670.
3. Loturco I, Franchini E, Abad CC. A comparative study of specific reaction time in elite boxers: differences between jabs and crosses. *J Athl Enhancement* 2015 Feb 8;04(03):307-310 [FREE Full text] [doi: [10.4172/2324-9080.1000199](https://doi.org/10.4172/2324-9080.1000199)]
4. Magill R, Anderson D. *Motor Learning Concepts and Applications*. Eighth Edition. Boston, USA: McGraw-Hill; 2007.
5. Botwinick J, Thompson LW. Premotor and motor components of reaction time. *J Exp Psychol* 1966 Jan;71(1):9-15. [doi: [10.1037/h0022634](https://doi.org/10.1037/h0022634)] [Medline: [5902149](https://pubmed.ncbi.nlm.nih.gov/5902149/)]
6. Falco C, Estevan I, Vieten M. Kinematical Analysis of Five Kicks in Taekwondo. *Portuguese J Sport Sci* 2011;11(2):219-222.
7. Kim JW, Kwon MS, Yenuga SS, Kwon YH. The effects of target distance on pivot hip, trunk, pelvis, and kicking leg kinematics in Taekwondo roundhouse kicks. *Sports Biomech* 2010 Jun;9(2):98-114. [doi: [10.1080/14763141003799459](https://doi.org/10.1080/14763141003799459)] [Medline: [20806845](https://pubmed.ncbi.nlm.nih.gov/20806845/)]
8. Cojocariu A. Measurement of reaction time in Qwan Ki Do. *Biol Sport* 2011;28(2):139-143. [doi: [10.5604/947454](https://doi.org/10.5604/947454)]
9. Macovei S, Lambu EA, Lambu IS. Study about the reaction time relation to sports performance in Karate Do. *Sci Mov Heal* 2013;8(Suppl 2):228-233.
10. Mori S, Ohtani Y, Imanaka K. Reaction times and anticipatory skills of karate athletes. *Hum Mov Sci* 2002 Jul;21(2):213-230. [doi: [10.1016/S0167-9457\(02\)00103-3](https://doi.org/10.1016/S0167-9457(02)00103-3)] [Medline: [12167300](https://pubmed.ncbi.nlm.nih.gov/12167300/)]
11. Darby D, Moriarity J, Pietrzak R, Kutcher J, McAward K, McCrory P. Prediction of winning amateur boxers using pretournament reaction times. *J Sports Med Phys Fitness* 2014 Jun;54(3):340-346. [Medline: [24739297](https://pubmed.ncbi.nlm.nih.gov/24739297/)]
12. Hermann G, Scholz M, Vieten M, Kohloeffel M. Reaction and Performance Time of Taekwondo Top Athletes Demonstrating the Baldung Chagi. In: *Proceedings of the 26th International Conference on Biomechanics in Sports*. 2008 Presented at: ISBS'08; July 14-18, 2008; Seoul, Korea.
13. Falco C, Molina-García J, Alvarez O, Estevan I. Effects of target distance on select biomechanical parameters in taekwondo roundhouse kick. *Sports Biomech* 2013 Nov;12(4):381-388. [doi: [10.1080/14763141.2013.776626](https://doi.org/10.1080/14763141.2013.776626)] [Medline: [24466650](https://pubmed.ncbi.nlm.nih.gov/24466650/)]
14. Sant'Ana JS, Diefenthaeler F, Pupo J, Detanico D, Guglielmo LG, Santos SG. Anaerobic evaluation of taekwondo athletes. *Int Sport Med J* 2014;15:492-499.
15. Sant'Ana J, Franchini E, da Silva V, Diefenthaeler F. Effect of fatigue on reaction time, response time, performance time, and kick impact in taekwondo roundhouse kick. *Sports Biomech* 2017 Jun;16(2):201-209. [doi: [10.1080/14763141.2016.1217347](https://doi.org/10.1080/14763141.2016.1217347)] [Medline: [27592682](https://pubmed.ncbi.nlm.nih.gov/27592682/)]
16. Quel EM, Bennett SJ, Lopez-Adan E, Zapico AG, Morales FS. Choice reaction time is not related to competition success in karate combat. *Eur J Hum Mov* 2015;35:41-50.
17. Sabina M, Elena-Andreea L, Ioana-Sorina L. Profile of some psychomotor capacities in top performance Karate Do athletes. *Procedia - Soc Behav Sci* 2014;117:262-268. [doi: [10.1016/j.sbspro.2014.02.211](https://doi.org/10.1016/j.sbspro.2014.02.211)]

18. Muntaner-Mas A, Martinez-Nicolas A, Lavie CJ, Blair SN, Ross R, Arena R, et al. A systematic review of fitness apps and their potential clinical and sports utility for objective and remote assessment of cardiorespiratory fitness. *Sports Med* 2019 Apr;49(4):587-600 [FREE Full text] [doi: [10.1007/s40279-019-01084-y](https://doi.org/10.1007/s40279-019-01084-y)] [Medline: [30825094](https://pubmed.ncbi.nlm.nih.gov/30825094/)]
19. Rhea MR. Determining the magnitude of treatment effects in strength training research through the use of the effect size. *J Strength Cond Res* 2004 Nov;18(4):918-920. [doi: [10.1519/14403.1](https://doi.org/10.1519/14403.1)] [Medline: [15574101](https://pubmed.ncbi.nlm.nih.gov/15574101/)]
20. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000 Jul;30(1):1-15. [doi: [10.2165/00007256-200030010-00001](https://doi.org/10.2165/00007256-200030010-00001)] [Medline: [10907753](https://pubmed.ncbi.nlm.nih.gov/10907753/)]
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 Feb 8;1(8476):307-310. [doi: [10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)] [Medline: [2868172](https://pubmed.ncbi.nlm.nih.gov/2868172/)]
22. Wong P, Chaouachi A, Chamari K, Dellal A, Wisloff U. Effect of preseason concurrent muscular strength and high-intensity interval training in professional soccer players. *J Strength Cond Res* 2010 Mar;24(3):653-660. [doi: [10.1519/JSC.0b013e3181aa36a2](https://doi.org/10.1519/JSC.0b013e3181aa36a2)] [Medline: [19816215](https://pubmed.ncbi.nlm.nih.gov/19816215/)]
23. Abbasi-Kesbi R, Memarzadeh-Tehran H, Deen MJ. Technique to estimate human reaction time based on visual perception. *Healthc Technol Lett* 2017 Apr;4(2):73-77 [FREE Full text] [doi: [10.1049/htl.2016.0106](https://doi.org/10.1049/htl.2016.0106)] [Medline: [28461901](https://pubmed.ncbi.nlm.nih.gov/28461901/)]
24. Balsalobre-Fernández C, Tejero-González CM, del Campo-Vecino J, Bavaresco N. The concurrent validity and reliability of a low-cost, high-speed camera-based method for measuring the flight time of vertical jumps. *J Strength Cond Res* 2014 Feb;28(2):528-533. [doi: [10.1519/JSC.0b013e318299a52e](https://doi.org/10.1519/JSC.0b013e318299a52e)] [Medline: [23689339](https://pubmed.ncbi.nlm.nih.gov/23689339/)]
25. García-López J, Peleteiro J, Rodríguez-Marroyo JA, Morante JC, Herrero JA, Villa JG. The validation of a new method that measures contact and flight times during vertical jump. *Int J Sports Med* 2005 May;26(4):294-302. [doi: [10.1055/s-2004-820962](https://doi.org/10.1055/s-2004-820962)] [Medline: [15795814](https://pubmed.ncbi.nlm.nih.gov/15795814/)]
26. Santos VG, Santos VR, Felipe LJ, Almeida JW, Bertuzzi R, Kiss MA, et al. Caffeine reduces reaction time and improves performance in simulated-contest of taekwondo. *Nutrients* 2014 Feb 10;6(2):637-649 [FREE Full text] [doi: [10.3390/nu6020637](https://doi.org/10.3390/nu6020637)] [Medline: [24518826](https://pubmed.ncbi.nlm.nih.gov/24518826/)]
27. Pozo J, Bastien G, Dierick F. Execution time, kinetics, and kinematics of the mae-geri kick: comparison of national and international standard karate athletes. *J Sports Sci* 2011 Nov;29(14):1553-1561. [doi: [10.1080/02640414.2011.605164](https://doi.org/10.1080/02640414.2011.605164)] [Medline: [21995835](https://pubmed.ncbi.nlm.nih.gov/21995835/)]

## Abbreviations

**ICC:** intraclass correlation coefficient

**TEM:** typical error of measurement

*Edited by G Eysenbach; submitted 07.05.19; peer-reviewed by J Koral, L Cheong; comments to author 13.07.19; revised version received 19.07.19; accepted 19.07.19; published 11.11.19.*

*Please cite as:*

*Coswig V, Sant' Ana J, Coelho MN, Pereira Moro AR, Diefenthaler F*

*Development of a Mobile Phone App for Measuring Striking Response Time in Combat Sports: Cross-Sectional Validation Study*  
*JMIR Mhealth Uhealth* 2019;7(11):e14641

URL: <http://mhealth.jmir.org/2019/11/e14641/>

doi: [10.2196/14641](https://doi.org/10.2196/14641)

PMID: [31710294](https://pubmed.ncbi.nlm.nih.gov/31710294/)

©Victor Coswig, Jader Sant' Ana, Maicon Nascimento Coelho, Antonio Renato Pereira Moro, Fernando Diefenthaler. Originally published in *JMIR Mhealth and Uhealth* (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mhealth and uhealth*, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Impact of a Mobile App–Based Health Coaching and Behavior Change Program on Participant Engagement and Weight Status of Overweight and Obese Children: Retrospective Cohort Study

Victor Cueto<sup>1,2</sup>, MD, MS; C Jason Wang<sup>1,3</sup>, MD, PhD; Lee Michael Sanders<sup>1</sup>, MD, MPH

<sup>1</sup>Division of General Pediatrics, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States

<sup>2</sup>Division of General Internal Medicine, Department of Medicine, Rutgers New Jersey Medical School, Newark, NJ, United States

<sup>3</sup>Center for Policy, Outcomes, and Prevention, Stanford University, Stanford, CA, United States

**Corresponding Author:**

Victor Cueto, MD, MS

Division of General Internal Medicine

Department of Medicine

Rutgers New Jersey Medical School

150 Bergen Street, H-251

Newark, NJ, 07101

United States

Phone: 1 973 972 5672

Email: [vcueto@alumni.stanford.edu](mailto:vcueto@alumni.stanford.edu)

**Related Article:**

Comment in: <https://mhealth.jmir.org/2021/2/e17492>

## Abstract

**Background:** Effective treatment of obesity in children and adolescents traditionally requires frequent in-person contact, and it is often limited by low participant engagement. Mobile health tools may offer alternative models that enhance participant engagement.

**Objective:** The aim of this study was to assess child engagement over time, with a mobile app–based health coaching and behavior change program for weight management, and to examine the association between engagement and change in weight status.

**Methods:** This was a retrospective cohort study of user data from *Kurbo*, a commercial program that provides weekly individual coaching via video chat and supports self-monitoring of health behaviors through a mobile app. Study participants included users of *Kurbo* between March 2015 and March 2017, who were 5 to 18 years old and who were overweight or obese (body mass index; BMI  $\geq$  85th percentile or  $\geq$  95th percentile) at baseline. The primary outcome, engagement, was defined as the total number of health coaching sessions received. The secondary outcome was change in weight status, defined as the change in BMI as a percentage of the 95th percentile (%BMI<sub>p95</sub>). Analyses of outcome measures were compared across three initial commitment period groups: 4 weeks, 12 to 16 weeks, or 24 weeks. Multivariable linear regression models were constructed to adjust outcomes for the independent variables of sex, age group (5-11 years, 12-14 years, and 15-18 years), and commitment period. A sensitivity analysis was conducted, excluding a subset of participants involuntarily assigned to the 12- to 16-week commitment period by an employer or health plan.

**Results:** A total of 1120 participants were included in analyses. At baseline, participants had a mean age of 12 years (SD 2.5), mean BMI percentile of 96.6 (SD 3.1), mean %BMI<sub>p95</sub> of 114.5 (SD 16.5), and they were predominantly female 68.04% (762/1120). Participant distribution across commitment periods was 26.07% (292/1120) for 4 weeks, 61.61% (690/1120) for 12-16 weeks, and 12.32% (138/1120) for 24 weeks. The median coaching sessions (interquartile range) received were 8 (3-16) for the 4-week group, 9 (5-12) for the 12- to 16-week group, and 19 (11-25) for the 24-week group ( $P<.001$ ). Adjusted for sex and age group, participants in the 4- and 12-week groups participated in  $-8.03$  (95% CI  $-10.19$  to  $-5.87$ ) and  $-9.34$  (95% CI  $-11.31$  to  $-7.39$ ) fewer coaching sessions, compared with those in the 24-week group ( $P<.001$ ). Adjusted for commitment period,

sex, and age group, the overall mean change in %BMIp95 was  $-0.21$  (95% CI  $-0.25$  to  $-0.17$ ) per additional coaching session ( $P < .001$ ).

**Conclusions:** Among overweight and obese children using a mobile app-based health coaching and behavior change program, increased engagement was associated with longer voluntary commitment periods, and increased number of coaching sessions was associated with decreased weight status.

(*JMIR Mhealth Uhealth* 2019;7(11):e14458) doi:[10.2196/14458](https://doi.org/10.2196/14458)

## KEYWORDS

child obesity; mHealth; mobile apps; health coaching; health behavior; self-monitoring; behavior change

## Introduction

### Background

A total of 1 in 3 children in the United States is either overweight or obese [1,2]. Obese children and adolescents are at risk for carrying their excess weight into adulthood and developing multiple comorbidities, including diabetes and coronary disease as adults [3,4]. Clinical management guidelines and treatment algorithms call for a staged approach to the treatment of overweight and obese children, which aims to promote healthy lifestyle changes through behavioral counseling [4,5]. Effective behavioral interventions for pediatric obesity involve multiple components focused on promoting healthy eating and exercise habits [6]. The most effective interventions involve supporting both children and parents to set goals, incorporate stimulus control, utilize problem solving, and participate in self-monitoring while working to achieve behavior changes [6]. The US Preventive Services Task Force (USPSTF) recommends primary providers to either provide or refer children with obesity to comprehensive intensive behavioral interventions aimed at decreasing excess weight and improving overall weight status [6]. However, child participation in comprehensive intensive behavioral interventions and clinical weight management programs is often low, with considerably high program attrition [7-9]. Common barriers to participation in weight management programs identified by families include concerns regarding affordability, inflexible scheduling, conflicts with other activities, time commitment, distance and transportation, and misalignment between expectations and program services [8,10,11]. Conversely, facilitators of participation for children and families may include tailored treatment plans and individualized health coaching [10-12]. Mobile health (mHealth) and telehealth technologies may provide a unique opportunity to overcome barriers to participation in obesity treatment by providing individualized interventions on a family's timeline and in their home environments [13-15].

Among obesity treatment trials for adults, mHealth tools appear to successfully assist patients in managing comorbidities, such as diabetes, improve physical activity and dietary behaviors, and achieve meaningful weight loss [16-22]. mHealth interventions in children and adolescents have been found to be effective at improving health behaviors and health outcomes across a wide reach of conditions [23]. In obesity treatment trials for children, the effect of mHealth tools is less clear, as they have been primarily studied as a component of larger, multifaceted interventions [13,15,24,25]. mHealth tools

incorporated into obesity prevention and treatment trials for children vary in their ability to improve weight outcomes [15,24,25]. However, mobile technology has been shown to be well accepted, feasible, and effective at supporting self-monitoring and promoting changes of physical activity and dietary behaviors [13,24-26]. However, overall, there is limited evidence for the efficacy of mHealth interventions as stand-alone treatment modalities for pediatric obesity and weight management [13]. Although there are a number of commercially available mobile apps targeting weight-related physical activity and dietary behaviors in children, reviews of commercial apps have found that most lack high-quality information, include only a few behavior change techniques (BCTs), and are not rooted in evidence-based behavior change theories [27,28]. As a result, some have suggested the need to rigorously evaluate commercial and stand-alone mHealth interventions aimed at promoting health behavior change among overweight or obese children [13,27,28].

### Objectives

In this study, we aimed to assess the engagement of overweight or obese children with a commercially available mHealth tool (*Kurbo*), which provides individualized health coaching and self-monitoring support designed to improve diet and physical activity behaviors. The primary aim was to describe and compare the engagement of participants with health coaching sessions, as a condition of their commitment period. The secondary aim was to examine the association between coaching sessions received and the change in a participant's weight status over time. We hypothesized that participants with longer commitment periods would engage in both more coaching sessions and have a trend toward greater weight loss.

## Methods

### Design

This was a retrospective cohort study of participants in a commercial, mobile app-based platform and program (*Kurbo*) designed to promote health behavior change and weight management through self-monitoring and health coaching support. The research team and investigators had no role in the development of the mobile app platform or the creation and delivery of program content.

### Program

The *Kurbo* mobile app platform (Figure 1) and program was designed to promote behavior change and encourage healthy lifestyle choices [29]. The program content and health coaching

incorporate multiple BCTs, consistent with established taxonomy for behavior change interventions [30]. The BCTs emphasized in the program are linked to multiple theoretical frameworks, including the theory of reasoned action, theory of planned behavior, social cognitive theory, and control theory, as well as operant conditioning and the information motivation behavioral skills model [31]. The program design was also

informed by a model of supportive accountability, which emphasizes the essential role of human support in mHealth interventions [32]. The mobile app and program include 2 primary components: (1) self-monitoring of eating and physical activity behaviors through a mobile app interface and (2) individualized coaching sessions through video chat.

**Figure 1.** Mobile app platform.



The self-monitoring component of the program employs the BCTs of self-monitoring of behavior and monitoring of outcomes of behavior [30]. Participants are encouraged to use the mobile app to log their daily food intake, using food categories adapted from the evidence-based traffic light system [33,34]. The traffic light system categorizes foods into 3 groups: unrestricted healthy green-light foods, less healthy yellow-light foods that should be eaten with caution, and unhealthy red-light foods that should be avoided [34]. The aim of this approach to eating behaviors is to encourage participants to gradually increase consumption of healthy foods (green lights) and decrease unhealthy foods (red lights) over time. This approach incorporates behavior substitution and habit formation [30]. Participants are also asked to self-monitor their physical activity behavior by logging the duration of activities in the mobile app, while working toward a goal of 60 min of moderate-to-vigorous physical activity each day.

The individualized coaching sessions component is provided by individuals who are hired and trained as coaches by *Kurbo*. Participants are paired with the same coach for the duration of their participation with the program, which aims to provide social support and accountability [30,32]. Coaches monitor the participant's dietary and physical activity behaviors through a Web-based dashboard, which serves to reinforce self-monitoring and allows for feedback of behavior [30]. Each coaching session lasts about 15 min, and it is made available on a weekly basis. Coaching sessions emphasize review of past behavior and outcome goals, as well as support future outcome and behavior goal setting [30]. Coaches encourage self-talk and identification of behavioral cues, as well as assist with tailored problem solving and action planning [30]. For example, coaches may discuss environmental cues and identification of triggers to eat red-light foods while supporting goal setting and action planning for choosing more green-light foods. Additional topics addressed during coaching sessions may include understanding food labels



and portion sizes. After each coaching session, participants receive an email from their coach, with praise for goals met and a tailored plan regarding goals set for the next week. Of note, parents are strongly encouraged to participate in coaching sessions; however, the program is not prescriptive regarding parental involvement. Similarly, the program allows for individually tailored coaching, but it is not specifically designed to address family dynamics or the age of participant.

In addition to coaching sessions, participants are also able to contact their coach between coaching sessions, via short message service text messages, email, or in-app messaging. Independent of the *Kurbo* mobile app platform, participants also have access to supplementary resources, including an emailed e-workbook, biweekly email newsletter, physical activity demonstration videos, blog posts, and downloadable healthy-eating cookbooks. These supplementary resources highlight BCTs by providing instructions on how to perform behaviors, support restructuring of the physical environment, and allow for social comparison.

Finally, the program also makes use of the BCT of a behavioral contract [30].

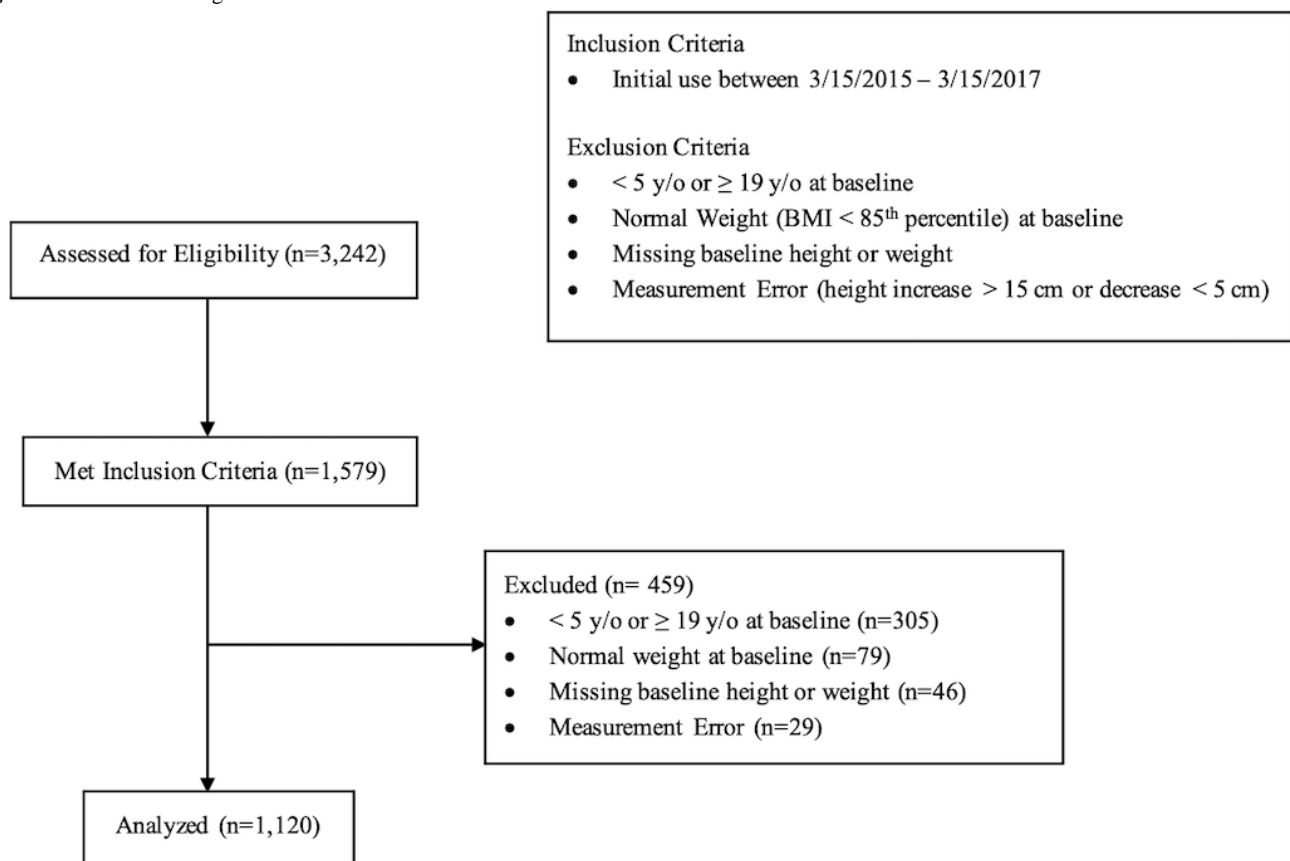
### Participants and Data Source

The study examined a retrospectively identified cohort of participants who initially utilized the *Kurbo* mobile app and program between March 15, 2015 and March 15, 2017. Participants were not recruited for the purpose of conducting the study. Deidentified participant data from the *Kurbo* data registry were provided to the investigators for the purpose of the study. The study received an exemption from the Stanford University School of Medicine Institutional Review Board. All data were independently reviewed and inspected by the research team to confirm that the predetermined inclusion and exclusion criteria were met before analysis.

### Inclusion Criteria

The inclusion criterion (Figure 2) applied was participation during a defined 2-year period from March 15, 2015, to March 15, 2017.

Figure 2. Cohort flow diagram.



### Exclusion Criteria

The exclusion criteria (Figure 2) included age less than 5 years or age greater than or equal to 19 years upon initial use of the program. Additional exclusion criteria included a normal weight status (body mass index; BMI < 85<sup>th</sup> percentile) at baseline, as well as any data measurement errors, including missing baseline height, missing baseline weight, or any height velocity measurements exceeding 15 cm increase or 5 cm decrease (on the basis of established height velocity reference values) [35].

### Commitment Periods

Study participants were voluntarily subscribed or assigned to 1 of 3 commitment periods: 4 weeks, 12 to 16 weeks, or 24 weeks. Each participant was supported by either an employer-benefited plan, a health insurer-benefited plan, or a self-paid plan. The cost of the program is covered by either the parent (self-pay), a parent's employer, or a family health insurance plan. Self-pay rates are dependent on the commitment period. Participants in self-paid plans voluntarily chose from 1 of 3 commitment periods: 4 weeks, 12 weeks, or 24 weeks.

Those in employer- or health insurer–benefited plans were contractually assigned to 1 of 2 commitment periods: 12 weeks or 16 weeks. Of note, there were only 16 participants in the 16-week commitment period, and for analytic purposes, these participants were combined with the 12-week period to form a 12- to 16-week commitment period group. All participants had the ability to renew or change their plan at the end of the initial commitment period; however, data regarding renewals or changes were not available for analysis.

## Measures

### Baseline Characteristics

Participant demographic characteristics were limited to self-reported age and sex, provided by either child or parent. Age in years at baseline was used to create 4 distinct age group categories. The age group categories were defined as 5 to 11, 12 to 14, and 15 to 18 years old. These age groups were informed by commonly reported groupings from population prevalence and large intervention studies [2,36]. The data registry objectively captured the payment source used for the program. Socioeconomic data, such as race, ethnicity, family income, or education were not reported. Self-reported baseline weight and height measurements were combined with age and sex, to calculate and derive the corresponding age and sex-specific BMI, BMI percentiles, and BMI expressed as a percentage of the 95<sup>th</sup> BMI percentile (%BMI<sub>p95</sub>) at baseline. This was accomplished using SAS code, developed by the Centers for Disease Control and Prevention (CDC) for this purpose. All participants' baseline weight was categorized according to CDC criteria, on the basis of age and sex-specific BMI percentile thresholds for children and adolescents. According to these criteria, weight status was defined as either overweight ( $\geq 85^{\text{th}}$  to  $< 95^{\text{th}}$  BMI percentile) or obese ( $\geq 95^{\text{th}}$  BMI percentile). Obese participants were also categorized according to %BMI<sub>p95</sub>, a measurement of relative BMI, which is recommended by the CDC for children and adolescents with severe obesity [37]. This recommendation is based on analyses that have shown BMI scores to be poorly reflective of adiposity in youth with very high BMI measures and severe obesity. The measure of %BMI<sub>p95</sub> is a more reliable measure of adiposity among obese youth and recommended for studies with a significant proportion of severely obese (%BMI<sub>p95</sub> $\geq 120$ ) children or adolescents [38]. Obese participants were additionally categorized according to 3 distinct classes of obesity, that is, Class I to III obesity, which are reflective of cardiometabolic risk and commonly used in obesity prevalence studies [2,4,39]. Class I is defined as  $\geq 95^{\text{th}}$  BMI percentile; Class II corresponds with %BMI<sub>p95</sub>  $\geq 120$  to  $< 140\%$  or BMI  $\geq 35$ , whichever is lower; and Class III applies to %BMI<sub>p95</sub>  $\geq 140$  or BMI  $\geq 40$ , whichever is lower.

### Primary Outcome: Participant Engagement

The primary outcome and measure of participant engagement was *coaching sessions*, defined as the total cumulative number of individual coaching sessions received by a participant during the participation period.

Other measures of participant engagement included participation period, program retention, coaching messages, dietary events,

and physical-activity events. *Participation period* was defined as the total number of weeks between when the participant signed up for the program and the last recorded interaction with the app. The last recorded interaction with the app included logging of dietary or physical activity, an in-app text message sent to a health coach, or a coaching session. *Program retention* was defined as having a total participation period that was greater or equal to the intended commitment period in weeks. *Coaching messages* was defined as the cumulative total number of individual in-app text messages sent by each participant to his/her assigned health coach during the participation period. *Dietary events* was defined as the cumulative total number of self-reported individual foods logged by a given participant. *Physical activity events* was defined as the cumulative total number of self-reported individual physical activities logged by each participant during the participation period. The data registry did not capture whether dietary or physical activity events were self-reported by the participant or a parent.

### Secondary Outcome: Change in Weight Status

A secondary outcome of the study was change in the participant's weight status, defined as the *change in %BMI<sub>p95</sub>* between the self-reported baseline and endpoint measurements recorded during the participation period. The baseline measurement of %BMI<sub>p95</sub> was derived from the participant's initial self-reported weight and height measurement entered into the mobile app. The endpoint measurement of %BMI<sub>p95</sub> was derived from the participant's last self-reported height and weight measurement. There was no predetermined time interval between baseline and endpoint measurements.

## Analysis

Descriptive statistics were used to compare baseline characteristics (age, age group, sex, BMI percentile, %BMI<sub>p95</sub>, weight category, obesity class, and payment source), primary outcome (participant engagement), and secondary outcome (change in weight status) across the 3 commitment periods. Categorical variables were expressed as absolute values and corresponding percentages. Normally distributed continuous variables are reported as a mean with standard deviation (SD). The continuous engagement measures analyzed had nonnormal distribution patterns, each of these are reported as a median with interquartile range (IQR). Differences of measures across commitment period groups were explored using Chi-square tests for categorical variables and analysis of variance for normally distributed continuous measures. Similarly, differences across commitment periods for nonparametric continuous measures were analyzed using Kruskal-Wallis tests. Significance of change in weight status within commitment periods was analyzed using paired two-tailed *t* tests.

Multivariable linear regression models were constructed to examine 2 sets of associations: (1) between the primary outcome (number of coaching sessions) and each commitment period (reference of 24-week period) and (2) between the primary outcome (number of coaching sessions) and the secondary outcome (change in %BMI<sub>p95</sub>). Each multivariable model included adjustment for significant baseline differences in age group and sex. A sensitivity analysis, excluding involuntary participants (ie, health plan or employer supported), was

performed to isolate differences associated with voluntariness of commitment period ([Multimedia Appendix 1](#)). All analyses were conducted using SAS Institute Inc software (SAS University Edition/SAS Studio 3.71).

## Results

### Participants

Of the 3242 participants assessed for eligibility, 1579 met the inclusion criteria. Of those, 305 participants were excluded for being outside the age range, normal weight at baseline (79), missing baseline data (46), or data measurement error. This yielded a final analytic sample of 1120 study participants, displayed in ([Figure 1](#)).

### Baseline Characteristics

The baseline characteristics for the study sample by commitment period are displayed in ([Table 1](#)). Overall, 292 participants were

in the 4-week commitment period, 690 participants were in the 12- to 16-week group, and 138 participants were in the 24-week commitment period. Mean age at baseline was 12 years (SD 2.5), and most participants 68.04% (762/1120) were female. There were no statistically significant differences across the 3 groups in age or sex. The majority of participants 76.61% (858/1120) were categorized as obese, with mean BMI percentile (SD) of 96.6 (3.1) and mean (SD) %BMI<sub>p95</sub> of 114.5 (16.5). Children in the 24-week group were more likely to be classified as obese, when compared with those in the 4- and 12-week groups (118/138, 85.5% vs 218/292, 74.7% and 522/690, 75.7%, respectively,  $P=.03$ ). The predominant payment source (743/1120, 66.34%) was self-pay. The distribution of baseline characteristics for the study sample by age group are displayed in [Multimedia Appendix 1](#). Except for payment source, all baseline characteristics differed significantly across age groups.

**Table 1.** Baseline participant characteristics by commitment period.

Baseline characteristics	All periods (N=1120)	4 weeks (n=292)	12-16 weeks <sup>a</sup> (n=690)	24 weeks (n=138)	P value
Age (years), mean (SD)	12.0 (2.5)	11.9 (2.2)	12.0 (2.7)	12.0 (2.4)	.89 <sup>b</sup>
<b>Age group, n (%)</b>					.61 <sup>c</sup>
5-11 years	573 (51.16)	150 (51.4)	350 (50.7)	73 (52.9)	
12-14 years	392 (35.00)	109 (37.3)	237 (34.4)	46 (33.3)	
15-18 years	155 (13.84)	33 (11.3)	103 (14.9)	19 (13.8)	
<b>Sex, n (%)</b>					.20 <sup>c</sup>
Male	358 (31.96)	103 (35.3)	218 (31.6)	37 (26.8)	
Female	762 (68.04)	189 (64.7)	472 (64.1)	101 (73.2)	
Body mass index percentile, mean (SD)	96.6 (3.1)	96.4 (3.3)	96.5 (3.1)	97.3 (2.5)	.009 <sup>b</sup>
%BMIp95 <sup>d</sup> , mean (SD)	114.5 (16.5)	113.4 (17.5)	114.1 (19.3)	118.8 (19.3)	.01 <sup>b</sup>
<b>Weight category<sup>e</sup>, n (%)</b>					.03 <sup>c</sup>
Overweight	262 (23.39)	74 (25.3)	168 (24.4)	20 (14.5)	
Obese	858 (76.61)	218 (74.7)	522 (75.7)	118 (85.5)	
<b>Obesity class<sup>f</sup>, n (%)</b>					.39 <sup>c</sup>
Class I	508 (59.21)	128 (58.7)	315 (60.3)	65 (55.1)	
Class II	236 (27.51)	65 (29.8)	140 (26.8)	31 (26.3)	
Class III	114 (13.29)	25 (11.5)	67 (12.8)	22 (18.6)	
<b>Payment source<sup>g</sup>, n (%)</b>					<.001 <sup>c</sup>
Self-pay	743 (66.34)	292 (100)	314 (45.5)	138 (100)	
Health plan	278 (24.82)	— <sup>h</sup>	278 (40.3)	—	
Employer	99 (8.84)	—	99 (14.3)	—	

<sup>a</sup>12 to 16 weeks includes n=674 participants with 12-week and n=16 participants with 16-week commitment periods.

<sup>b</sup>Analysis of variance.

<sup>c</sup>Chi-square test.

<sup>d</sup>%BMIp95: percentage of the 95th BMI percentile.

<sup>e</sup>Categories by Centers for Disease Control and Prevention body mass index percentile for Age and Sex. Overweight (BMI Percentile  $\geq 85$  and  $< 95$ th), Obese (BMI Percentile  $\geq 95$ th).

<sup>f</sup>Obesity Class I ( $\geq 95$ th to  $< 120$  %BMIp95), Class II ( $\geq 120$  to  $< 140$  %BMIp95, or BMI  $\geq 35$ ), Class III ( $\geq 140$  %BMIp95, or BMI  $\geq 40$ ), inclusive of N 858 participants categorized as obese.

<sup>g</sup>The 4 weeks and 24 weeks commitment periods consisted entirely of Self-pay participants, accordingly data for Health plan and Employer are not applicable.

<sup>h</sup>Not applicable.

## Participant Engagement

The engagement of participants with the mobile app program, compared across commitment periods, is displayed in (Table 2).

The primary outcome of median number of coaching sessions received was 8 (IQR 3-15) for the 4-week group, 9 (IQR 5-12) for the 12- to 16-week group, and 19 (IQR 11-25) for the 24-week group ( $P < .001$ ). Overall, the median number of coaching sessions per participant was 9, with an IQR of 5 to 15. The median (IQR) values for other engagement measures were as follows: participation period, 15 weeks (IQR 12-30); the number of coaching messages, 3 (IQR 0-10); the number

of logged dietary events, 174 (IQR 83-325); and the number of logged physical activity events, 42 (IQR 15-91). Median weeks of participation differed across commitment periods: 16 weeks (IQR 8-36) for the 4-week group, 14 weeks (IQR 12-22) for the 12- to 16-week group, and 30 weeks (IQR 22-51) for the 24-week group. Overall, program retention was high, with 79.91% (895/1120) of participants remaining engaged with the program for at least the duration of their commitment period. Program retention across commitment periods was 92.5% (270/292) for the 4-week group, 76.8% (530/690) for the 12- to 16-week group, and 68.8% (95/138) for the 24-week group ( $P < .001$ ). In addition, the engagement of all participants stratified by age group is shown in (Table 3). There were no statistically significant differences in engagement by age group.

**Table 2.** Engagement of participants with mobile app-based program by commitment period, among all participants (N=1120).

Engagement measures	All periods (N=1120)	4 weeks (n=292)	12-16 weeks <sup>a</sup> (n=690)	24 weeks (n=138)	P value
Coaching sessions <sup>b</sup> , median (IQR <sup>c</sup> )	9 (5-15)	8 (3-16)	9 (5-12)	19 (11-25)	<.001 <sup>d</sup>
Coaching messages <sup>e</sup> , median (IQR)	3 (0-10)	4 (0-11)	3 (0-9)	6 (1-14)	<.001 <sup>d</sup>
Dietary events <sup>f</sup> , median (IQR)	174 (83-325)	163 (80-321)	153 (76-278)	335 (188-596)	<.001 <sup>d</sup>
Physical activity events <sup>g</sup> , median (IQR)	42 (15-91)	42 (15-98)	36 (13-78)	76 (33-152)	<.001 <sup>d</sup>
Participation period <sup>h</sup> , median (IQR)	15 (12-30)	16 (8-36)	14 (12-22)	30 (22-51)	<.001 <sup>d</sup>
Program retention <sup>i</sup> , n (%)	895 (79.91)	270 (92.5)	530 (76.8)	95 (68.8)	<.001 <sup>j</sup>

<sup>a</sup>Includes 674 participants with 12-week and 16 participants with 16-week commitment periods.

<sup>b</sup>Median of total number of coaching sessions between participant and coach.

<sup>c</sup>IQR: interquartile range.

<sup>d</sup>Kruskal-Wallis Test.

<sup>e</sup>Median of total number of text messages from participant to coach.

<sup>f</sup>Median of total number of dietary event food logs recorded by participants (n=1100), otherwise missing.

<sup>g</sup>Median of total number of physical activity event logs recorded by participants (n=1078), otherwise missing.

<sup>h</sup>Median of total weeks between sign up and last recorded interaction with the app.

<sup>i</sup>Proportion of participants who completed equal or greater weeks than initial commitment period.

<sup>j</sup>Chi-square test.

**Table 3.** Engagement of participants with mobile app-based program by age group, among all participants (N=1120).

Engagement measures	All age groups, (N=1120)	5-11 years (n=573)	12-14 years (n=392)	15-18 years (n=155)	P value
Coaching sessions <sup>a</sup> , median (IQR <sup>b</sup> )	9 (5-15)	10 (5-15)	9 (5-14)	10 (6-15)	.77 <sup>c</sup>
Coaching messages <sup>d</sup> , median (IQR)	3 (0-10)	3 (0-10)	3 (0-12)	3 (0-10)	.94 <sup>c</sup>
Dietary events <sup>e</sup> , median (IQR)	174 (83-325)	171 (80-318)	175 (84-330)	177 (87-342)	.68 <sup>c</sup>
Physical activity events <sup>f</sup> , median (IQR)	42 (15-91)	44 (15-89)	41 (15-92)	36 (13-102)	.89 <sup>c</sup>
Participation period <sup>g</sup> , median (IQR)	15 (12-30)	15 (12-32)	15 (11-27)	14 (12-27)	.33 <sup>c</sup>
Program retention <sup>h</sup> , n (%)	895 (79.91)	493(79.1)	314 (80.1)	128 (82.6)	.62 <sup>i</sup>

<sup>a</sup>Median of total number of coaching sessions between participant and coach.

<sup>b</sup>IQR: interquartile range.

<sup>c</sup>Kruskal-Wallis Test.

<sup>d</sup>Median of total number of text messages from participant to coach.

<sup>e</sup>Median of total number of dietary event food logs recorded by participants (n=1100), otherwise missing.

<sup>f</sup>Median of total number of physical activity event logs recorded by participants (n=1078), otherwise missing.

<sup>g</sup>Median of total weeks between sign-up and last recorded interaction with the app.

<sup>h</sup>Proportion of participants who completed equal or greater weeks than initial commitment period.

<sup>i</sup>Chi-square test.

Results of unadjusted and adjusted models for the primary outcome (number of coaching sessions per participant) are displayed in [Table 4](#). After adjustment for child sex and age group (with the 24-week group as reference), the 12- to 16-week group was associated with the fewest number of coaching sessions per participant, with a beta-coefficient of  $-9.34$  (95% CI  $-11.30$  to  $-7.39$ ), in contrast to the 4-week group, with a beta-coefficient of  $-8.03$  (95% CI  $-10.19$  to  $-5.87$ ). The results

of sensitivity analyses restricted to only self-pay (voluntary) participants are included in [Multimedia Appendix 1](#). In this subpopulation, the 4-week group was associated with the fewest number of coaching sessions per participant, with a beta-coefficient of  $-8.06$  (95% CI  $-10.56$  to  $-5.56$ ), in contrast to the 12-to 16-week group, with a beta-coefficient of  $-6.02$  (95% CI  $-8.48$  to  $-3.56$ ).

**Table 4.** Factors associated with total number of coaching sessions, among all participants (N=1120).

Participant factors	Unadjusted <sup>a</sup> beta-coefficient (95% CI)	P value	Adjusted <sup>b</sup> beta-coefficient (95% CI)	P value
<b>Age group (years)</b>				
Age 5-11 years (reference: 15-18 years)	-0.17 (-2.14 to 1.79)	.86	-0.13 (-2.03 to 1.78)	.89
Age 12-14 years (reference: 15-18 years)	-0.59 (-2.68 to 1.44)	.56	-0.53 (-2.53 to 1.46)	.59
<b>Sex</b>				
Male (reference: female)	-1.38 (-2.77 to 0.01)	.05	-1.15 (-2.50 to 0.19)	.09
<b>Commitment period</b>				
4 weeks (reference: 24 weeks)	-8.15 (-10.31 to -5.98)	<.001	-8.03 (-10.19 to -5.87)	<.001
12-16 weeks (reference: 24 weeks)	-9.41 (-11.36 to -7.46)	<.001	-9.34 (-11.30 to -7.39)	<.001

<sup>a</sup>Unadjusted bivariate linear regression model of coaching sessions outcome as a function of age group, sex, or commitment period.

<sup>b</sup>Adjusted multivariable linear regression model of coaching sessions outcome adjusted as a function of age group, sex, and commitment period.

### Change in Weight Status

Within each commitment period, the mean change between baseline and endpoint for %BMI<sub>p95</sub> was -5.4 (95% CI -6.2 to -4.5) for 4 weeks ( $P<.001$ ), -4.8 (95% CI -5.3 to -4.3) for 12 to 16 weeks ( $P<.001$ ), and -6.9 (95% CI -8.3 to -5.6) for 24 weeks ( $P<.001$ ). Compared across age groups, the mean change of %BMI<sub>p95</sub> was -5.6 (SD 7.9) for 5 to 11 year olds, -4.7 (SD 5.9) for 12 to 14 year olds, and -5.2 (SD 5.6) for 15 to 18 year olds ( $P=.09$ ). Adjusting for age group and sex within each commitment period, the beta-coefficient per coaching session was -0.25 (95% CI -0.32 to -0.18) for the 4-week group, -0.16 (95% CI -0.21 to -0.11) for the 12-week group, and -0.26 (95% CI -0.34 to -0.18) for the 24-week group ( $P<.001$ ). The multivariable model of all participants, adjusted for initial commitment period, age group, and sex, demonstrated an overall beta-coefficient decrease of -0.21 (95% CI -0.25 to -0.17) in %BMI<sub>p95</sub> per each coaching session received ( $P<.001$ ).

## Discussion

### Principal Findings

This retrospective study described the engagement of a large cohort of children and adolescents, with a multicomponent mobile app-based comprehensive behavioral program aimed at promoting healthy dietary and exercise lifestyle behaviors. Unlike traditional behavioral interventions and clinical weight management programs, which largely rely on in-person visits and sessions [6,40,41], mHealth programs, such as the one studied, enable participants to self-monitor health behaviors and receive health coaching at their own pace. As such, the findings of this study generally add to the growing evidence base for mHealth tools, more specifically for mobile app-based comprehensive behavioral programs to support health behavior change for overweight or obese children and adolescents.

Our findings of overall engagement with a median of 9 (IQR 5-15) coaching sessions during the participation period is notable for an mHealth program. This level of engagement, although considered low intensity by USPSTF criteria, is comparable with contact levels of in-person weight management programs [40,41]. We also documented consistent levels of participant engagement with other measures, such as self-monitoring of

physical activity and dietary habits. The documented engagement of participants with both individualized coaching and self-monitoring is an important finding, given that these components represent behavioral change techniques that are known to be effective in managing pediatric obesity [42]. Furthermore, overall program retention with this mHealth program was high, and attrition was 20.9%, which is considerably lower than attrition rates of between 37% and 41%, reported for traditional in-person (non-mHealth) weight management programs [8].

Finally, in this observational study, we found a significant association between the number of coaching sessions and the change in self-reported weight status during participation in the program. This association suggests that greater exposure to coaching sessions in this study was correlated with increased weight reduction. USPSTF analyses of intensive in-person interventions suggest a dose-response relationship between intervention hours received and beneficial changes in weight, with effective programs requiring at least 26 contact hours [6]. Although this study design was unable to capture a participant's absolute contact hours with the app or coaching sessions, the observed association between weight change and coaching sessions suggests a possible lower dose or lower threshold effect for this mHealth intervention, when compared with traditional in-person interventions. However, further research is necessary to validate the nature and magnitude of this association, as well as to rule out other explanations, including selection bias or reporter bias.

### Strengths and Limitations

The study has limitations common to other studies of digital health and mHealth interventions [15]. As an observational retrospective study design, the study also lacked an independent control or comparison group. The study is subject to some reporting bias, as all measures, including anthropometrics, were self-reported by participants. The study is also subject to selection bias, both because enrollment and participation were self-directed and because full functionality required access to an internet-connected mobile device. Another limitation was unmeasured confounders, including race/ethnicity, parental educational attainment, and household income, which limits adjustment and generalization, particularly for low-income

populations [2]. However, the inclusion of employer-benefited and health plan-benefited participants, which may have included both commercial and government sponsored plans, likely increased the heterogeneity of the sample. Finally, the study as designed was unable to fully account for other factors that may influence intervention fidelity, including quality of coaching sessions, role of parents, clustering by coach, or use of other resources. Specifically, the study could not account for the degree of parental involvement and supervision among younger participants as compared with older participants. Nonetheless, the findings do not support a significant difference between age groups.

Still, rigorous and independent studies of digital health and mHealth interventions are limited, and this study meaningfully contributes to the literature on digital health and mHealth interventions focused on obesity treatment and prevention. First, the intervention was comprehensively designed with multiple components that have been shown to be effective strategies for weight management and behavior change in pediatric populations [6,30,33,34,42]. This is in contrast to the often limited BCTs employed in other mHealth tools [27,28]. Second, the intervention leveraged effective BCTs to support self-guided behavior change and utilized health coaches in a family context [24]. Third, the study population included a broad distribution of age groups, including preadolescent children, adolescents, and late adolescents. In fact, the age, sex, and weight characteristics of the study population are similar to that of traditional in-person medical weight management programs [40,41]. In addition to obese participants, the intervention also included overweight participants, which suggests generalizability of the findings beyond this study. Finally, it is notable that engagement with the program and change in weight

status was consistent across age groups that include children, preadolescents, and adolescents.

## Implications

Our findings have implications for clinical care, population health, and public policy. Clinicians providing obesity treatment may consider the incorporation of mHealth programs, such as the one studied here, as an adjunct to clinic visits and traditional medical management strategies. Health care systems aiming to improve population health management efforts might find these types of mHealth solutions more accessible for providing access to care for patients in rural areas where availability of providers may be limited or to patients in urban areas, who may be restricted by long commute times or have limited transportation options. Finally, public health leaders and policy makers may be encouraged by the role that emerging digital technologies could play in addressing obesity at the community level.

## Conclusions

This study of a mobile app-based health behavior change and health coaching program among a large cohort of overweight and obese participants demonstrated high participant engagement. Increased engagement with coaching sessions was associated with longer voluntary commitment periods. Overall program retention was higher than that reported for similar in-person intensive behavioral interventions and weight management programs. Participant engagement with coaching sessions was associated with decreases in weight status (%BMI<sub>p95</sub>). Taken together, these findings highlight the potential of mHealth platforms as a promising model for delivering behavioral interventions that support weight management and behavior change for overweight or obese children and adolescents.

---

## Acknowledgments

VC was affiliated with the Division of General Pediatrics, Department of Pediatrics at Stanford University School of Medicine at the time this study was conducted, and is currently affiliated with Rutgers New Jersey Medical School Division of General Internal Medicine, Department of Medicine. VC conducted this study with support from Stanford's Health Resources and Services Administration (HRSA) Hispanic Center of Excellence Grant (D34HP16047) as well as the Stanford Child Health Research Institute (CHRI) and Lucile Packard Children's Hospital (LPCH) Stanford. The contents of this publication represent the views of the authors and do not necessarily reflect the views of HRSA, CHRI, or LPCH. The authors thank *Kurbo* for providing data for the purpose of this study. *Kurbo* was not involved in the design of the study, the analysis and interpretation of the data, or the preparation and submission of this manuscript.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Supplementary tables – Baseline participant characteristics by age group, all participants (N=1120) and Factors associated with total number of coaching sessions, among self-pay (voluntary) participants (N=743).

[PDF File (Adobe PDF File), 86 KB - [mhealth\\_v7i11e14458\\_app1.pdf](https://mhealth.jmir.org/2019/11/e14458_app1.pdf) ]

---

## References

1. Ogden CL, Carroll MD, Lawman HG, Fryar CD, Kruszon-Moran D, Kit BK, et al. Trends in obesity prevalence among children and adolescents in the United States, 1988-1994 through 2013-2014. *J Am Med Assoc* 2016 Jun 7;315(21):2292-2299 [FREE Full text] [doi: [10.1001/jama.2016.6361](https://doi.org/10.1001/jama.2016.6361)] [Medline: [27272581](https://pubmed.ncbi.nlm.nih.gov/27272581/)]

2. Skinner AC, Ravanbakht SN, Skelton JA, Perrin EM, Armstrong SC. Prevalence of obesity and severe obesity in US children, 1999-2016. *Pediatrics* 2018 Sep;141(3):e20173459 [FREE Full text] [doi: [10.1542/peds.2018-1916](https://doi.org/10.1542/peds.2018-1916)] [Medline: [30177517](https://pubmed.ncbi.nlm.nih.gov/30177517/)]
3. Ajala O, Mold F, Boughton C, Cooke D, Whyte M. Childhood predictors of cardiovascular disease in adulthood. A systematic review and meta-analysis. *Obes Rev* 2017 Sep;18(9):1061-1070. [doi: [10.1111/obr.12561](https://doi.org/10.1111/obr.12561)] [Medline: [28545166](https://pubmed.ncbi.nlm.nih.gov/28545166/)]
4. Kelly AS, Barlow SE, Rao G, Inge TH, Hayman LL, Steinberger J, American Heart Association Atherosclerosis, Hypertension, Obesity in the Young Committee of the Council on Cardiovascular Disease in the Young, Council on Nutrition, Physical Activity and Metabolism, Council on Clinical Cardiology. Severe obesity in children and adolescents: identification, associated health risks, and treatment approaches: a scientific statement from the American Heart Association. *Circulation* 2013 Oct 8;128(15):1689-1712. [doi: [10.1161/CIR.0b013e3182a5cfb3](https://doi.org/10.1161/CIR.0b013e3182a5cfb3)] [Medline: [24016455](https://pubmed.ncbi.nlm.nih.gov/24016455/)]
5. American Academy of Pediatrics: Institute for Health Childhood Weight. 2015. Algorithm for the Assessment and Management of Childhood Obesity in Patients 2 Years and Older URL: [https://ihcw.aap.org/Documents/Assessment%20and%20Management%20of%20Childhood%20Obesity%20Algorithm\\_FINAL.pdf](https://ihcw.aap.org/Documents/Assessment%20and%20Management%20of%20Childhood%20Obesity%20Algorithm_FINAL.pdf) [accessed 2019-07-12]
6. US Preventive Services Task Force, Grossman DC, Bibbins-Domingo K, Curry SJ, Barry MJ, Davidson KW, et al. Screening for obesity in children and adolescents: US preventive services task force recommendation statement. *J Am Med Assoc* 2017 Jun 20;317(23):2417-2426. [doi: [10.1001/jama.2017.6803](https://doi.org/10.1001/jama.2017.6803)] [Medline: [28632874](https://pubmed.ncbi.nlm.nih.gov/28632874/)]
7. Skelton JA, Beech BM. Attrition in paediatric weight management: a review of the literature and new directions. *Obes Rev* 2011 May;12(5):e273-e281 [FREE Full text] [doi: [10.1111/j.1467-789X.2010.00803.x](https://doi.org/10.1111/j.1467-789X.2010.00803.x)] [Medline: [20880126](https://pubmed.ncbi.nlm.nih.gov/20880126/)]
8. Dhaliwal J, Nosworthy NM, Holt NL, Zwaigenbaum L, Avis JL, Rasquinha A, et al. Attrition and the management of pediatric obesity: an integrative review. *Child Obes* 2014 Dec;10(6):461-473. [doi: [10.1089/chi.2014.0060](https://doi.org/10.1089/chi.2014.0060)] [Medline: [25496035](https://pubmed.ncbi.nlm.nih.gov/25496035/)]
9. Skelton JA, Goff DC, Ip E, Beech BM. Attrition in a multidisciplinary pediatric weight management clinic. *Child Obes* 2011 Jun 20;7(3):185-193 [FREE Full text] [doi: [10.1089/chi.2011.0010](https://doi.org/10.1089/chi.2011.0010)] [Medline: [21966612](https://pubmed.ncbi.nlm.nih.gov/21966612/)]
10. Hampl S, Paves H, Laubscher K, Eneli I. Patient engagement and attrition in pediatric obesity clinics and programs: results and recommendations. *Pediatrics* 2011 Sep;128(Suppl 2):S59-S64 [FREE Full text] [doi: [10.1542/peds.2011-0480E](https://doi.org/10.1542/peds.2011-0480E)] [Medline: [21885646](https://pubmed.ncbi.nlm.nih.gov/21885646/)]
11. Hampl S, Demeule M, Eneli I, Frank M, Hawkins MJ, Kirk S, et al. Parent perspectives on attrition from tertiary care pediatric weight management programs. *Clin Pediatr (Phila)* 2013 Jun;52(6):513-519. [doi: [10.1177/0009922813482515](https://doi.org/10.1177/0009922813482515)] [Medline: [23539682](https://pubmed.ncbi.nlm.nih.gov/23539682/)]
12. Rice KG, Jumamil RB, Jabour SM, Cheng JK. Role of health coaches in pediatric weight management. *Clin Pediatr (Phila)* 2017 Feb;56(2):162-170. [doi: [10.1177/0009922816645515](https://doi.org/10.1177/0009922816645515)] [Medline: [27169713](https://pubmed.ncbi.nlm.nih.gov/27169713/)]
13. Turner T, Spruijt-Metz D, Wen CK, Hingle MD. Prevention and treatment of pediatric obesity using mobile and wireless technologies: a systematic review. *Pediatr Obes* 2015 Dec;10(6):403-409 [FREE Full text] [doi: [10.1111/jjpo.12002](https://doi.org/10.1111/jjpo.12002)] [Medline: [25641770](https://pubmed.ncbi.nlm.nih.gov/25641770/)]
14. Kozak AT, Buscemi J, Hawkins MA, Wang ML, Breland JY, Ross KM, et al. Technology-based interventions for weight management: current randomized controlled trial evidence and future directions. *J Behav Med* 2017 Feb;40(1):99-111 [FREE Full text] [doi: [10.1007/s10865-016-9805-z](https://doi.org/10.1007/s10865-016-9805-z)] [Medline: [27783259](https://pubmed.ncbi.nlm.nih.gov/27783259/)]
15. Chaplais E, Naughton G, Thivel D, Courteix D, Greene D. Smartphone interventions for weight treatment and behavioral change in pediatric obesity: a systematic review. *Telemed J E Health* 2015 Oct;21(10):822-830. [doi: [10.1089/tmj.2014.0197](https://doi.org/10.1089/tmj.2014.0197)] [Medline: [26290954](https://pubmed.ncbi.nlm.nih.gov/26290954/)]
16. Allen JK, Stephens J, Patel A. Technology-assisted weight management interventions: systematic review of clinical trials. *Telemed J E Health* 2014 Dec;20(12):1103-1120 [FREE Full text] [doi: [10.1089/tmj.2014.0030](https://doi.org/10.1089/tmj.2014.0030)] [Medline: [25409001](https://pubmed.ncbi.nlm.nih.gov/25409001/)]
17. Levine DM, Savarimuthu S, Squires A, Nicholson J, Jay M. Technology-assisted weight loss interventions in primary care: a systematic review. *J Gen Intern Med* 2015 Jan;30(1):107-117 [FREE Full text] [doi: [10.1007/s11606-014-2987-6](https://doi.org/10.1007/s11606-014-2987-6)] [Medline: [25134692](https://pubmed.ncbi.nlm.nih.gov/25134692/)]
18. Sepah SC, Jiang L, Ellis RJ, McDermott K, Peters AL. Engagement and outcomes in a digital diabetes prevention program: 3-year update. *BMJ Open Diabetes Res Care* 2017;5(1):e000422 [FREE Full text] [doi: [10.1136/bmjdr-2017-000422](https://doi.org/10.1136/bmjdr-2017-000422)] [Medline: [28948027](https://pubmed.ncbi.nlm.nih.gov/28948027/)]
19. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. *J Med Internet Res* 2015 Oct 23;17(10):e240 [FREE Full text] [doi: [10.2196/jmir.4897](https://doi.org/10.2196/jmir.4897)] [Medline: [26499966](https://pubmed.ncbi.nlm.nih.gov/26499966/)]
20. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. *J Cardiovasc Nurs* 2013;28(4):320-329 [FREE Full text] [doi: [10.1097/JCN.0b013e318250a3e7](https://doi.org/10.1097/JCN.0b013e318250a3e7)] [Medline: [22635061](https://pubmed.ncbi.nlm.nih.gov/22635061/)]
21. Mateo GF, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. *J Med Internet Res* 2015 Nov 10;17(11):e253 [FREE Full text] [doi: [10.2196/jmir.4836](https://doi.org/10.2196/jmir.4836)] [Medline: [26554314](https://pubmed.ncbi.nlm.nih.gov/26554314/)]
22. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. *PLoS One* 2016;11(11):e0166718 [FREE Full text] [doi: [10.1371/journal.pone.0166718](https://doi.org/10.1371/journal.pone.0166718)] [Medline: [27861583](https://pubmed.ncbi.nlm.nih.gov/27861583/)]



23. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. *JAMA Pediatr* 2017 May 1;171(5):461-469 [FREE Full text] [doi: [10.1001/jamapediatrics.2017.0042](https://doi.org/10.1001/jamapediatrics.2017.0042)] [Medline: [28319239](https://pubmed.ncbi.nlm.nih.gov/28319239/)]
24. Taveras EM, Marshall R, Kleinman KP, Gillman MW, Hacker K, Horan CM, et al. Comparative effectiveness of childhood obesity interventions in pediatric primary care: a cluster-randomized clinical trial. *JAMA Pediatr* 2015 Jun;169(6):535-542. [doi: [10.1001/jamapediatrics.2015.0182](https://doi.org/10.1001/jamapediatrics.2015.0182)] [Medline: [25895016](https://pubmed.ncbi.nlm.nih.gov/25895016/)]
25. Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. *Pediatrics* 2014 Sep;134(3):e723-e731. [doi: [10.1542/peds.2014-1012](https://doi.org/10.1542/peds.2014-1012)] [Medline: [25157000](https://pubmed.ncbi.nlm.nih.gov/25157000/)]
26. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. *Am J Clin Nutr* 2017 Jun;105(6):1327-1335. [doi: [10.3945/ajcn.116.150995](https://doi.org/10.3945/ajcn.116.150995)] [Medline: [28446496](https://pubmed.ncbi.nlm.nih.gov/28446496/)]
27. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. *Int J Behav Nutr Phys Act* 2017 Jun 24;14(1):83 [FREE Full text] [doi: [10.1186/s12966-017-0538-3](https://doi.org/10.1186/s12966-017-0538-3)] [Medline: [28646889](https://pubmed.ncbi.nlm.nih.gov/28646889/)]
28. Schoffman DE, Turner-McGrievy G, Jones SJ, Wilcox S. Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? *Transl Behav Med* 2013 Sep;3(3):320-325 [FREE Full text] [doi: [10.1007/s13142-013-0206-3](https://doi.org/10.1007/s13142-013-0206-3)] [Medline: [24073184](https://pubmed.ncbi.nlm.nih.gov/24073184/)]
29. Kurbo. Health Coaching for Kids, Teens, Families. URL: <https://kurbo.com/>
30. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
31. 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](https://doi.org/10.1037/0278-6133.27.3.379)] [Medline: [18624603](https://pubmed.ncbi.nlm.nih.gov/18624603/)]
32. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to ehealth interventions. *J Med Internet Res* 2011 Mar 10;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
33. Epstein LH, Gordy CC, Raynor HA, Beddome M, Kilanowski CK, Paluch R. Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity. *Obes Res* 2001 Mar;9(3):171-178 [FREE Full text] [doi: [10.1038/oby.2001.18](https://doi.org/10.1038/oby.2001.18)] [Medline: [11323442](https://pubmed.ncbi.nlm.nih.gov/11323442/)]
34. Epstein LH, Myers MD, Raynor HA, Saelens BE. Treatment of pediatric obesity. *Pediatrics* 1998 Mar;101(3 Pt 2):554-570. [Medline: [12224662](https://pubmed.ncbi.nlm.nih.gov/12224662/)]
35. Kelly A, Winer KK, Kalkwarf H, Oberfield SE, Lappe J, Gilsanz V, et al. Age-based reference ranges for annual height velocity in US children. *J Clin Endocrinol Metab* 2014 Jun;99(6):2104-2112 [FREE Full text] [doi: [10.1210/jc.2013-4455](https://doi.org/10.1210/jc.2013-4455)] [Medline: [24601728](https://pubmed.ncbi.nlm.nih.gov/24601728/)]
36. Jasik CB, King EC, Rhodes E, Sweeney B, Mietus-Snyder M, Grow HM, et al. Characteristics of youth presenting for weight management: retrospective national data from the power study group. *Child Obes* 2015 Oct;11(5):630-637. [doi: [10.1089/chi.2015.0053](https://doi.org/10.1089/chi.2015.0053)] [Medline: [26440387](https://pubmed.ncbi.nlm.nih.gov/26440387/)]
37. Freedman DS, Berenson GS. Tracking of BMI scores for severe obesity. *Pediatrics* 2017 Sep;140(3):e20171072 [FREE Full text] [doi: [10.1542/peds.2017-1072](https://doi.org/10.1542/peds.2017-1072)] [Medline: [28830920](https://pubmed.ncbi.nlm.nih.gov/28830920/)]
38. Freedman DS, Butte NF, Taveras EM, Lundeen EA, Blanck HM, Goodman AB, et al. BMI z-scores are a poor indicator of adiposity among 2- to 19-year-olds with very high BMIs, NHANES 1999-2000 to 2013-2014. *Obesity (Silver Spring)* 2017 Apr;25(4):739-746 [FREE Full text] [doi: [10.1002/oby.21782](https://doi.org/10.1002/oby.21782)] [Medline: [28245098](https://pubmed.ncbi.nlm.nih.gov/28245098/)]
39. Skinner AC, Perrin EM, Moss LA, Skelton JA. Cardiometabolic risks and severity of obesity in children and young adults. *N Engl J Med* 2015 Oct;373(14):1307-1317. [doi: [10.1056/nejmoa1502821](https://doi.org/10.1056/nejmoa1502821)]
40. Kirk S, Armstrong S, King E, Trapp C, Grow M, Tucker J, et al. Establishment of the pediatric obesity weight evaluation registry: a national research collaborative for identifying the optimal assessment and treatment of pediatric obesity. *Child Obes* 2017 Feb;13(1):9-17. [doi: [10.1089/chi.2016.0060](https://doi.org/10.1089/chi.2016.0060)] [Medline: [27732057](https://pubmed.ncbi.nlm.nih.gov/27732057/)]
41. Quattrin T, Roemmich JN, Paluch R, Yu J, Epstein LH, Ecker MA. Efficacy of family-based weight control program for preschool children in primary care. *Pediatrics* 2012 Oct;130(4):660-666 [FREE Full text] [doi: [10.1542/peds.2012-0701](https://doi.org/10.1542/peds.2012-0701)] [Medline: [22987879](https://pubmed.ncbi.nlm.nih.gov/22987879/)]
42. Martin J, Chater A, Lorencatto F. Effective behaviour change techniques in the prevention and management of childhood obesity. *Int J Obes (Lond)* 2013 Oct;37(10):1287-1294. [doi: [10.1038/ijo.2013.107](https://doi.org/10.1038/ijo.2013.107)] [Medline: [23756676](https://pubmed.ncbi.nlm.nih.gov/23756676/)]

## Abbreviations

- %BMI<sub>p95</sub>**: BMI as percentage of the 95th percentile
- BCT**: behavior change technique
- BMI**: body mass index

**CDC:** Centers for Disease Control and Prevention  
**CHRI:** Child Health Research Institute  
**HRSA:** Health Resources and Services Administration  
**IQR:** interquartile range  
**LPCH:** Lucile Packard Children's Hospital  
**mHealth:** mobile health  
**USPSTF:** US Preventive Services Task Force

*Edited by G Eysenbach; submitted 20.04.19; peer-reviewed by S Armstrong, J Wilroy; comments to author 15.06.19; revised version received 12.07.19; accepted 14.08.19; published 15.11.19.*

*Please cite as:*

*Cueto V, Wang CJ, Sanders LM*

*Impact of a Mobile App-Based Health Coaching and Behavior Change Program on Participant Engagement and Weight Status of Overweight and Obese Children: Retrospective Cohort Study*

*JMIR Mhealth Uhealth 2019;7(11):e14458*

*URL: <http://mhealth.jmir.org/2019/11/e14458/>*

*doi: [10.2196/14458](https://doi.org/10.2196/14458)*

*PMID: [31730041](https://pubmed.ncbi.nlm.nih.gov/31730041/)*

©Victor Cueto, C Jason Wang, Lee Michael Sanders. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 15.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# A Baby Formula Designed for Chinese Babies: Content Analysis of Milk Formula Advertisements on Chinese Parenting Apps

Jing Zhao<sup>1,2</sup>, MPH; Mu Li<sup>1,2</sup>, PhD; Becky Freeman<sup>1</sup>, PhD

<sup>1</sup>School of Public Health, Sydney Medical School, The University of Sydney, Sydney, Australia

<sup>2</sup>China Studies Centre, The University of Sydney, Sydney, Australia

**Corresponding Author:**

Jing Zhao, MPH

School of Public Health

Sydney Medical School

The University of Sydney

Edward Ford Building A27

Sydney, 2006

Australia

Phone: 61 04 0650 2156

Email: [jzha5010@uni.sydney.edu.au](mailto:jzha5010@uni.sydney.edu.au)

## Abstract

**Background:** China is the largest market for infant formula. With the increasing use of smartphones, apps have become the latest tool used to promote milk formula. Formula manufacturers and distributors both have seized the popularity of apps as an avenue for marketing.

**Objective:** This study aimed to identify and analyze milk formula ads featured on Chinese pregnancy and parenting apps, to build the first complete picture of app-based milk formula marketing techniques being used by milk formula brand variants on these apps, and to more fully understand the ad content that potentially undermines public health messaging about infant and young child feeding.

**Methods:** We searched for free-to-download Chinese parenting apps in the 360 App Store, the biggest Android app store in China. The final sample consisted of 353 unique formula ads from the 79 apps that met the inclusion criteria. We developed a content analysis coding tool for categorizing the marketing techniques used in ads, which included a total of 22 coding options developed across 4 categories: emotional imagery, marketing elements, claims, and advertising disclosure.

**Results:** The 353 milk formula ads were distributed across 31 companies, 44 brands, and 79 brand variants. Overall, 15 of 31 corporations were international with the remaining 16 being Chinese owned. An image of a natural pasture was the most commonly used emotional image among the brand variants (16/79). All variants included branding elements, and 75 variants linked directly to e-shops. Special price promotions were promoted by nearly half (n=39) of all variants. A total of 5 variants included a celebrity endorsement in their advertising. A total of 25 of the 79 variants made a product quality claim. Only 14 variants made a direct advertisement disclosure.

**Conclusions:** The purpose of marketing messages is to widen the use of formula and normalize formula as an appropriate food for all infants and young children, rather than as a specialized food for those unable to breastfeed. Policy makers should take steps to establish an appropriate regulatory framework and provide detailed monitoring and enforcement to ensure that milk formula marketing practices do not undermine breastfeeding norms and behaviors.

(*JMIR Mhealth Uhealth* 2019;7(11):e14219) doi:[10.2196/14219](https://doi.org/10.2196/14219)

**KEYWORDS**

infant formula; food policy; health promotion; marketing; mobile app; parenting; breastfeeding; advertisement

## Introduction

Despite the strong health message that *breast is best* for growing infants [1-4], and breastfeeding also benefits mothers [5], no

country in the world fully meets the recommended breastfeeding guidelines according to a 2017 report by the United Nations International Children's Emergency Fund and the World Health Organization (WHO) [6]. In the 194 nations evaluated in the

report, only 40% of children younger than 6 months are breastfed exclusively, and only 23 countries have exclusive breastfeeding rates above 60% [6]. One of the powerful environmental factors influencing breastfeeding is the ubiquitous presence of breast milk substitutes (BMS) marketing [7-9]. The global milk formula market, which is primarily composed of BMS, reached sales of US \$44.8 billion in 2014 and is set to reach US \$70.6 billion by 2019 [10]. China's, including Hong Kong, contribution to the global sales growth was over 50% during 2010 to 2015 [11]. China is the largest market for infant formula, valued at US \$17.8 million in 2016 and is projected to more than double in value by 2019 [10].

Omnipresent marketing of BMS negatively affects breastfeeding practices [10]. Examples of this harmful marketing include provision of free products in maternity facilities [12], promotion by health workers [13], and mass media [12] and Web-based advertising [14]. With the increasing use of smartphones, apps have become the latest tool used to promote milk formula. Women utilize pregnancy and parenting apps as primary sources of information and emotional support [15-17]. In China, the number of active users of the most popular parenting app, Babytree, has reached 20 million per month in 2018 [18]. BMS manufacturers and distributors have seized parenting app popularity as an avenue for marketing. In a previous study involving the content analysis of Chinese infant feeding apps, 20 out of 26 apps were found to host infant formula banner ads on their homepages, and 12 apps included e-commerce stores that both sold and advertised infant formula [19]. Although there is emerging research on how milk formulas are being marketed in digital media, little of this research has closely examined apps [9] or focused on China. Given the exponential growth in popularity of pregnancy and parenting apps, there is a need to understand the milk formula marketing techniques on these apps. Equally, although marketing case studies of specific formula companies or specific types of BMS help to highlight the issue in China [20], a more complete picture of milk formula marketing strategies on digital media is needed.

The International Code of Marketing of Breast Milk Substitutes (the Code), published by the WHO in 1981, is an international health policy framework to regulate the marketing of BMS [21]. The Code prohibits all advertising to the public of BMS, including in digital media [21]. The Code recognizes that BMS have a legitimate role to play, when these are necessary. What the Code aims to do is to protect parents from irresponsible and biased marketing of BMS, to make sure that their choices are made based on full, impartial information, rather than misleading, inaccurate, or biased marketing claims, with the goal of protecting breastfeeding and promoting a healthy diet. There should be no advertising or other promotion to the general public of products within the scope of the Code, that is, BMS (any milk formula up to 3 years of age) and pictures or text that idealize the use of BMS should not be used. However, the Code, on its own, is not legally enforceable. The Code states that governments should take action to enforce the principles and aims of the Code, including the implementation of legislation. The Code has been adopted by more than 70% of countries, including China [22]. The Chinese Rules Governing the Administration of Marketing of Breast Milk Substitutes were

adopted by 6 government sectors in 1995 [13]. However, driven by commercial interests, China repealed these legal measures in 2016, without any replacement, substantially weakening protection from harmful BMS promotion [21]. The industry uses its immense financial resources to impede country's efforts to adopt the Code [21] and the money spent on marketing BMS dwarfs that invested by governments promoting breastfeeding [22]. In the United States, for example, BMS companies spend an average of US \$30 per infant every year on product promotion, compared with the US \$0.21 per baby invested in breastfeeding promotion [23].

Given that this marketing heavily influences infant and young child feeding preferences and choices [10], understanding how all milk formula products are being promoted on pregnancy and parenting apps is essential. The primary aim of this study was to identify and analyze milk formula ads featured on Chinese parenting apps. To begin to build the first complete picture of the milk formula marketing techniques being used by milk formula brands on these apps, in this study, we assessed the range of milk formula brands being advertised on apps, advertising techniques employed by these brands, and the scope of health and nutrition claims embedded in app ads. Our aim was to more fully understand the ad content being deployed by formula brands that potentially undermines public health messaging about infant and young child feeding.

## Methods

### App Selection

Android is the most popular smartphone operating system in China, with an almost two-thirds share of the total market [24]. In February 2018, we accessed the monthly free-download ranking list in the *Pregnancy and Parenting* category of the 360 App Store, the biggest Android app store in China. In that list, apps were ranked by monthly number of times downloaded and all publicly available for free download. We included apps that targeted mothers or mothers-to-be, provided pregnancy and early parenting information, featured tools, like vaccine and pregnancy reminders, and provided support through online social forums. A total of 556 apps were initially included and any apps that had been downloaded below 1000 times per month shown in the ranking list were excluded. We also excluded any apps that were focused on children above 3 years of age; broken or contained a dead link; designed mainly for playing games, reading bedtime stories, singing lullabies, or predicting children's height; or assisted in choosing a baby name. JZ initially screened each app based on the description page and its associated image in the 360 App store. When JZ was unsure about including a particular app during screening, it was discussed and screened together with ML. The final 79 apps were selected for milk formula ad identification and collection. No ethics approval was required for this study as all data collected were publicly available.

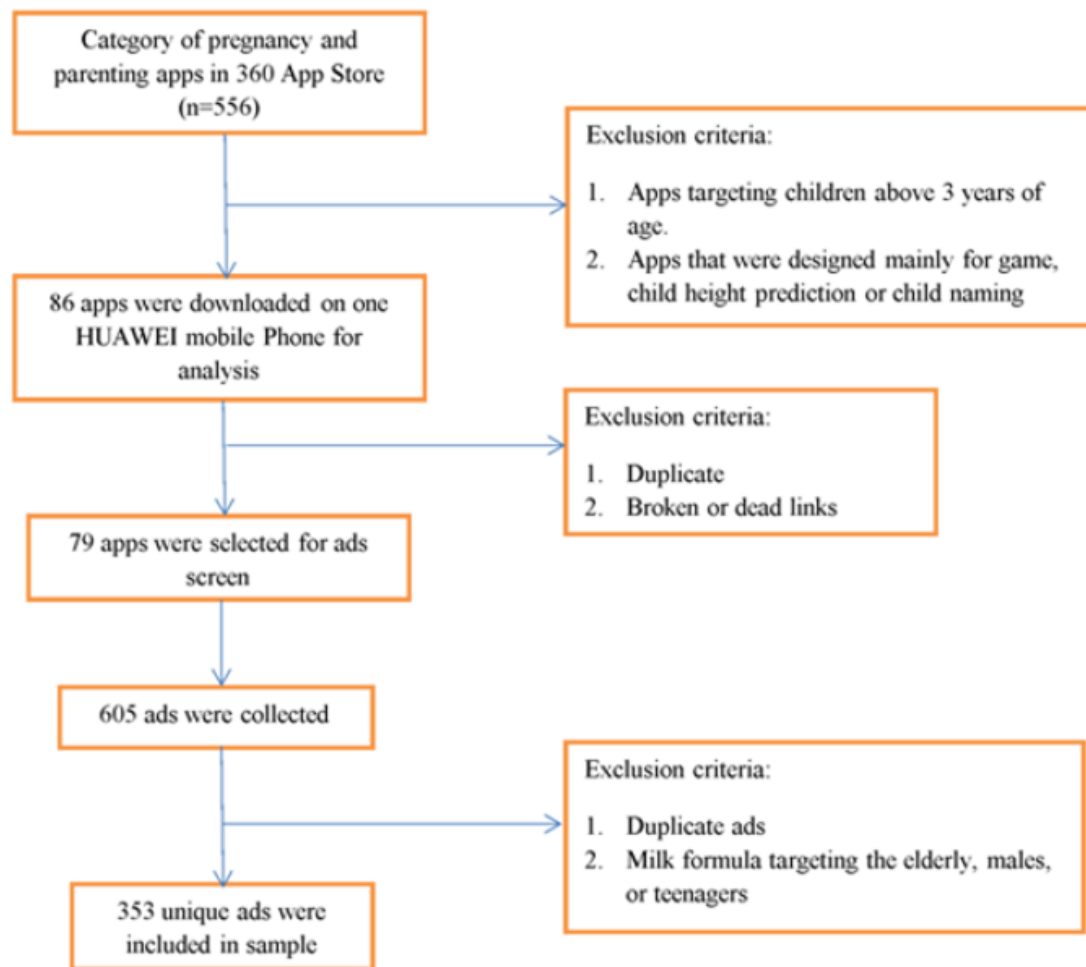
### Advertisement Identification

Ads promoting any type of milk formula were defined as formula ads for this study. Between April and May 2018, we scrolled through the entire app loading page and e-commerce page of each of the included 79 apps to identify and document

the formula ads. In addition, as banner ads have been shown to capture audience attention [25], we included any formula ads that appeared on each app opening home page. Initially, all of the formula ads from the selected app pages were captured by screenshot and kept as digital files for further screening analysis. All of the formula ads were then screened, and any duplicate

ads and ads not related to milk formula products, such as those for infant food products and milk bottles, were excluded. In addition, any ads for milk formula designed for the elderly, adult men, or teenagers were also excluded. The final sample comprised 353 unique formula ads (Figure 1).

**Figure 1.** Flow chart for milk formula ads collection.



## Advertisement Analysis

Manufacturers frequently offer different variations of the same branded products known as brand variants [26]. In our analysis, we observed that brand variants were very different in terms of name, color, design, options, style, features, and packaging portfolios. Brand variants are used to microsegment consumers with highly differentiated product and also help to ameliorate the effects of marketing restrictions [27]. Therefore, for this analysis, we were most interested in how these brand variants use a range of marketing elements and techniques.

The formula company, brand, and brand variants featured in each ad were identified. Descriptive statistics were used to assess the range of companies, brands, and variants of milk formula, and the frequency of different milk types and target ages. If a single ad promoted more than one type of formula milk or only promoted a company or a brand itself, then that ad was categorized as brand advertising.

On the basis of previous content analysis studies of infant and toddler food ads in magazines [28], food and beverages on Facebook [29], and Web-based advertising of infant formula [30], we developed a content analysis coding tool for categorizing the marketing techniques used in ads in our study. Coding was completed manually by JZ. Initially, this tool was pilot tested on 30 ads randomly chosen from the examined apps. Each of the 30 ads were assessed against the coding tool categories to evaluate appropriateness, completeness, and relevancy to public health communication and also to determine what data would be collected from the ads. After this pilot, we deleted 2 options, *fear emotional appeal* and *vouchers, offers, rebates*, and added 2 options, *online shop* and *delivery*, to the coding tool. Finally, a total of 22 coding options were developed across 4 categories, namely, emotional appeal, marketing elements, products claims, and advertisement disclosure (for details, see Table 1).

**Table 1.** Coding categories and options of milk formula ad content.

Coding category and coding options	Definition or examples
<b>Emotional imagery (based on ad imagery)</b>	
Happy child	Image of a smiling baby.
Parental love	Image of a parent hugging or kissing a baby.
Satisfied baby drinking from a milk bottle or holding a formula can	Image of a cute baby drinking milk from a bottle or holding a milk formula can.
Relieved mother, pregnant women	Image of a mother or pregnant woman relieved or at ease.
Beautiful natural environment, pasture	Image of a mountain pasture, clear sky, and natural landscape.
<b>Marketing elements</b>	
Branding elements	Any logos, colors, and trademarks.
Special price promotions	Limited-time offers, discount, or other reduced price.
Sampling	Consumers are given some quantity of a product for no charge (either new or established products).
Coupons	A document or e-voucher exchanged for financial discount with product purchase.
Bonus packs	Special packaging that provides consumers with extra quantities of products at no extra charge.
Competitions, prizes, and giveaways	Any contest involving a participant entry, including content about the product or the brand (such as photos and article) created by users.
Cartoon characters or spokes characters	Any nonhuman characters that are used to promote a product or a brand—giraffes, sheep, bees, bears, and other cartoon images.
Sponsorships and partnerships	Any events that the brand supports or other brands with which the brand partners.
Celebrity endorsement	People with an entertainment or media profile.
Online shop	The ad is directly linked to the products in e-shop, and app user could buy the product online.
Delivery	Products are delivered directly to consumer.
<b>Claims</b>	
Food safety claim	Purity, no contamination, imported milk source, good quality of milk source, whole package imported, and a long brand history.
Product quality claim	Organic, goat milk, premium or gold or super package, scientific based or medical evidence based or medical recommendations based or trusted by parents and health professionals.
Health claims	Nutritionally balances, and provides good nutrition to, the infant or children, supports healthy growth/better overall health, helps to support digestive system, good for the brain/better brain power, and specialized for allergic baby.
Nutrition claims	Contains DHA <sup>a</sup> or ARA <sup>b</sup> , Omega 3, or milk fat globule; contains fiber; contains probiotics; contains protein or amino acids; contains lactose; contains vitamin.
Claims idealizing the use of milk formula	Texts aim to idealize the use of milk formula, including wording <i>closest to, inspired by, and following the example of</i> human breast milk; <i>closer to breast milk</i> .
Advertisement disclosure	All advertising has to be identified with the word <i>Advertisement</i> .

<sup>a</sup>DHA: docosahexaenonic acid.

<sup>b</sup>ARA: arachidonic acid.

First, any emotional imagery included in the ads was coded. Emotional images were coded to one of the following 5 options: (1) happy child, (2) parental love, (3) satisfied baby drinking from a milk bottle or holding a formula can, (4) relieved mother, pregnant women, and (5) beautiful natural environment, pasture. Any duplicate images that appeared in multiple ads were not assessed. Next for each ad, the presence of any of the 11 marketing element categories was recorded. Marketing element

categories included branding, special price promotions, sampling, coupons, bonus pack, contests, prizes and giveaways, sponsorships, spokes characters and celebrity endorsement, and free delivery. The third coding category, claims, was made up of 5 coding options: food safety claims, product quality claims, health claims, nutrition claims, and claims idealizing the use of formula, with several types of claim under each option. Finally, for the last coding category, disclosure, it was noted whether

the ad included text identifying that it was advertising content. According to China's *Interim Measures for the Administration of Internet Advertising*, which was launched in 2016, all digital advertising must be clearly identified with the word *advertising*. Ads were initially coded using the exact Chinese characters that appeared in ads, and then the coding results are reported here in English.

All content coding was completed by JZ and then tested for accuracy by a research assistant independently coding 40 ads randomly selected from the final sample. Cohen kappa was used to measure intercoder reliability using SPSS version 22.0 statistical software (SPSS Inc, an IBM Company). Kappa value was 0.81. Definitions for infant formula and substitutes is provided in [Textbox 1](#).

**Textbox 1.** Definitions of infant formula, follow-on formula, and breast milk substitutes.

- Breast milk substitutes refer to any food for children (up to 3 years of age) being marketed or otherwise presented as a partial or total replacement for breast milk, whether suitable for that purpose or not [21]
- Milk formula refers to the wider range of milk powders for all ages available on the market [21]
- Infant formula refers to milk formula products intended for infants during the recommended exclusive breastfeeding phase (typically 0-6 months of age) [21]
- Follow-on formula refers to milk formula products intended for older infants, as they begin to receive complementary foods, and young children (typically 6-12 months of age) [21]
- Toddler milk: A fortified milk-based product only suitable for children older than 12 months (12-36 months) [21]
- Child milk: A fortified milk-based product only suitable for children older than 36 months [21]

## Results

### General Characteristics

In total, 353 unique ads from the 79 apps were collected for analysis. The 353 milk formula ads were distributed across 31 companies, 44 brands, and 79 brand variants. Just over half of the milk formula companies (15/31) were international corporations with the remaining 16 Chinese-owned. Many of the international companies marketed multiple brands and brand variants with the majority of brands (27/44, 61%) and brand variants (49/79, 62%) being produced by international companies. In all, 70 of the 353 ads were distributed across 10 variants from 4 brands manufactured by the Danone Company. This was followed by Nestle, with 52 ads across 11 variants under 4 brands and then Mead Johnson with 34 ads across 7 brand variants under 6 brands.

More than one-thirds of brand variants (n=28) advertised more than one type of milk formula. [Multimedia Appendix 1](#) shows

the full distribution of the types of milk formula across brands and brand variants that appeared in the ads. Infant formula was advertised by 24 brand variants, follow-on formula by 30 brand variants, and toddler milk formula by 40 brand variants. A further 11 brand variants promoted formula for children, and 1 variant included a formula for older children. A total of 10 brand variants were for milk formula that was promoted as being for pregnant and breastfeeding women. Another 3 variants claimed to have formulas that were specifically designed for babies born prematurely. There were 11 brands and 15 variants identified in general brand ads that were not for any specific type of milk formula.

### Emotional Imagery

An image of a natural pasture was the most commonly used emotional image among all 79 brand variants (n=16). Followed by an image of parental love (n=11), happy child (n=7), relieved mother/pregnant women (n=5), and a baby drinking/holding milk bottle (n=4; [Table 2](#)).

**Table 2.** Emotional imagery, marketing elements, and advertisement disclosure presented in 79 milk formula brand variant ads.

Coding category and option	Occurrence <sup>a</sup> , n (%)
<b>Emotional imagery</b>	
Beautiful natural environment, pasture	16 (20)
Parental love	11 (14)
Happy child	7 (9)
Relieved mother, pregnant women	5 (6)
Satisfied baby drinking milk bottle or holding formula can	4 (5)
<b>Marketing elements</b>	
Branding elements	79 (100)
Online shop	75 (95)
Special price promotions	39 (49)
Competitions, prizes, and giveaways	27 (34)
Delivery	23 (29)
Coupons	20 (25)
Bonus packs	17 (22)
Cartoon characters/spokes characters	16 (20)
Sampling	14 (18)
Celebrity endorsement	5 (6)
Sponsorships and partnerships	4 (5)
Advertisement disclosure	32 (41)

<sup>a</sup>Some brand variants contained more than one emotional imagery or marketing element.

## Marketing Elements

All variants included branding elements, such as logos, trademarks, and brand colors, and linked directly to e-shops. Special price promotions, such as discounts, sales, and limited-time offers, were promoted by nearly half (39/79, 49%) of all brand variants. Competitions, prizes, or giveaways on the condition of supplying contact details were used by 27 variants. A user-generated content (UGC) contest was held by one brand variant (Enfinitas) where consumers were asked to submit a photo of or a story about the milk formula to win a gift. A total of 23 variants offered free delivery service. Coupons, distributed either by the brand or online store, were available in the ads of 20 variants. A total of 17 variants offered a bonus pack and 16 variants included licensed characters to promote their products, such as cartoon giraffes, sheep, bees, and bears. Time-limited product sampling was identified in the ad of 14 variants of both established formulas and new products. A total of 5 variants included a celebrity endorsement in their advertising. These celebrities were all well known in China, included an actress who is a mother and is followed by nearly 30 million fans on the social media site Weibo [31]. In all, 4 brand variants engaged in sponsored activities, such as Wyeth Promama advertised its official mothers' club, Biostime Supreme sponsored an education activity when traveling with young kids, Nan HA promoted a Web-based parent personality test, and Aptamil German sponsored a TEDx program titled *Trust is visible* in China, where 5 well-known individuals shared their stories about trust and relationships. The brand variant name and logo

were all displayed prominently in association with the sponsored activities. Many brand variant ads contained more than one marketing element, and some marketing elements appeared in more than one brand variant advertising, but each brand variant was counted only once for each element (Table 2).

## Claims

Many brand variant ads contained more than one claim across the 5 coding options. Although some type of claims appeared in more than one ad for a single brand variant, each unique type of claim was counted only once (Table 3). A total of 34 brand variants made at least one type of food safety claim. A total of 27 domestic brand variants included at least one food safety claim, with 7 stating to be *100% imported source* and a further 20 highlighting *100% produced and packaged overseas*. Overall, 9 brand variants claimed to use good quality of milk source, and 8 brand variants claimed to have a *long brand history*. At least one type of product quality claim was made by 25 brand variants. The most identified product quality claim among brand variants was *super/gold/premium* (n=20), then *organic* (n=7) and *scientific/evidence based* (n=5), with the claim *goat milk* being used in 4 brand variants. A total of 24 brand variants made at least one health claim, with 8 claiming to improve digestion and absorption, 7 claiming to protect against allergies, 6 claiming to support growth and development/better overall health, and 3 claiming to improve brain development. For the 24 brand variants that made at least one nutrition claim, the phrase "Contains DHA/ARA, Omega 3, or Milk Fat Globule Membrane," was the most frequently identified claim (n=7),



whereas 5 variants highlighted the addition of sphingomyelin or choline; 4 mentioned added protein or amino acids, 4 were said to contain probiotics, 2 claimed to include lactose-containing milk formula, and 1 claim related to each of vitamins and fiber. In the claim category idealizing the use of

milk formula, 10 brand variants used words idealizing milk formula and bottle-feeding, including *thousands of mothers' choice*, *similar to breast milk*, or *love*. [Multimedia Appendices 2 and 3](#) are typical examples of milk formula ads examined in the study.

**Table 3.** Prevalence of 79 milk formula brand variants making claims by type of claim.

Claim type	Brand variant making this claim type <sup>a</sup> , n
<b>Food safety claims</b>	34
Whole package imported	20
Good quality of milk source	9
Long brand history	8
Imported milk source	7
Purity, no contamination	2
<b>Product quality claims</b>	25
Premium, gold package, super	20
Organic	7
Scientific based, medical evidence based/medical recommendations based/trusted by parents and health professionals.	5
Goat milk	4
<b>Health claims</b>	24
Helps to support digestive and absorption system	8
Specialized for allergic baby	7
Nutritionally balances and provides good nutrition to infant or children/supports healthy growth/better overall health	6
Good for the brain/better brain power	3
<b>Nutrition claims</b>	24
Contains DHA <sup>b</sup> /ARA <sup>c</sup> , Omega 3, or milk fat globule membrane	7
Contains sphingomyelin or choline	5
Contains protein or amino acids	4
Contains probiotics	4
Contains lactose	2
Contains fiber	1
Contains vitamin	1
<b>Claims idealizing the use of milk formula</b>	10
Including <i>thousands of mothers' choice/similar to breast milk/love</i>	10

<sup>a</sup>Some brand variants made multiple types of claims under each claim option.

<sup>b</sup>DHA: docosahexaenonic acid.

<sup>c</sup>ARA: arachidonic acid.

## Advertisement Disclosure

In total, only 14 variants made a direct advertisement disclosure. In addition, a further 18 variants included the Chinese character for “product/commodity” instead of advertisement. A total of 2 variants (NAN Pro and S-26 Promil Ultima) stated *breast milk is the best* at the bottom of their ads (see [Table 3](#)).

## Discussion

### The Influence of Milk Formula Branding on Breastfeeding

This study identified a total of 79 brand variants that included multiple types of milk formula, with 24 of these brand variants advertising infant milk formula, 30 follow-on formula, and 40 toddler formula. This reflects a full-scale breach of the Code, which applies advertising restrictions to any milk specifically

marketed for feeding infants and young children up to the age of 3 years. Since 1986, the WHO has maintained the position that follow-on formula is nutritionally unnecessary, but most countries only apply partial advertising restrictions to infant formula [21]. Toddler milk formula is reportedly set to outperform sales of infant and follow-on milk formula [20], which might explain in part our findings that 40 variants advertised toddler milk formula. Manufacturers are continuing to focus on the growth prospects in the toddler milk formula market to widen their consumer base to include older children. Strikingly, we identified 3 brand variants of milk formula marketed specifically for premature babies, which directly contradicts the WHO advice that “preterm birth infants who are able to breastfeed should be put to the breast as soon as possible after birth when they are clinically stable and should be exclusively breastfed until 6 months of age” [32].

In addition, 11 brand variants promoted child (aged >36 months) milk formula, with one brand variant, Enfakid, only promoting child milk formula; the remaining variants advertised more than one type of milk formula. As the Code is only designed to prevent companies from promoting milk formula and food for children up to 3 years of age, these products and promotions skirt advertising regulations. Moreover, China launched new infant milk formula registration measures in 2015 that came into effect on January 1, 2018. This legislation limits the number of brands per manufacturer to 3 for each of the development stages under 3 years of age [33]. This means child milk formula is not explicitly regulated by the new regulation, which may also help explain the 11 brand variants identified in our study that included a child milk formula.

Brand extension and brand-focused advertising help to establish brand awareness, preference, and loyalty to formula from fetus through to early childhood, this includes influencing pregnant women and mothers who are breastfeeding their infants. We found 10 variants of a special type of milk formula marketed as being for breastfeeding/pregnant women. Marketing an increasing range of milk formulas for different age groups weakens the impact of restrictions on infant formula (0-6 months). Sharing a brand that is identified with other life stages of formula is likely to influence infant feeding behavior to the same extent as direct advertising, as consumers are unable to differentiate between the two [34]. Similarly, ads for follow-on formula are perceived by pregnant women and mothers as promoting infant formula [35]. Research carried out in Australia [34], the United Kingdom [36], and Italy [35] showed how difficult it is for pregnant women and mothers to identify the difference between promotion of follow-on or toddler formula and promotion of infant formula. As products in these older age group categories are often branded, packaged, and labeled in ways that resemble infant formula, they can also be erroneously introduced in the first 6 months of life. Consumers then assume that the claims made in these brand variants are also true for the infant formula variants [35].

Manufacturers use persuasive marketing techniques to reinforce their brand identity, such as special price promotion, coupons, prizes, and giveaways. This is despite the fact that the Code prohibits any point of sale advertising or any promotional technique to induce sales [21]. Of particular note is the UGC

contest, consumers were asked to submit photos and content that expressed their *love* of the promoted milk formula. UGC is a form of peer endorsement that increases brand authenticity and enhances consumer trust. This distinctive social media marketing tactic builds on the effectiveness of competitions by allowing companies and consumers to collaborate and build brands together [35]. The explosion of e-commerce and the ubiquitous promotion and availability of milk formula in the Chinese online environment encourages and enables consumers to directly purchase products within pregnancy and parenting apps. Both domestic and overseas manufacturers have enjoyed strong sales through online shopping [37].

Some brand variants advertised with claims of *love* and *similar to breast milk* in Chinese characters, which could increase exposure to idealized bottle feeding and raise desire to use milk formula. Pregnant women and new mothers exposed to BMS ads containing the words *close* and *similar* believe infant formula is nutritionally equivalent to human milk [38]. This is another clear violation of the Code that states not to contain pictures or text that idealizes the use of BMS. In addition, the most commonly used (by 16 variants) image of *beautiful natural pasture and clear sky* implies the pure source and high quality of milk formula, which plays to Chinese mothers' food safety concerns around domestic brands [39]. This is due to the 2008 Chinese milk scandal where milk formula was found to contain high concentrations of melamine and was estimated to cause tens of thousands of infants and children to develop severe kidney problems [40].

### The Influence of Marketing Claims on Breastfeeding

The 2008 milk product scandal in China has also led to high demand for imported infant formula. Large international companies exploit these product fears to create a premium image of their brands. Meanwhile, domestic companies are attempting to rehabilitate their brand image while also making similar reassuring claims that they use *100% imported source* or *100% produced and packaged overseas* that are paired with emotional images of nature. For example, the brand variant-NAN pro claimed that its formula is imported from Switzerland and included an image of a snow-capped mountain. Premium brand variants featured heavily in our results. Marketing of so-called premium products with the associated *premium* pricing is leading to confusion among consumers and having a significant financial impact on families, especially in Asian countries [8,11]. A case study from Singapore found that parents perceived that more expensive or premium products are of higher quality, as they lack sufficient understanding of the nutritional content of formula milk [41]. A survey of Chinese bottle-feeding mothers found that 65% of mothers chose organic infant milk because they were willing to pay more for their baby's food [42].

Heath claims also included the claim of formula with added nutrients and associated benefits. For example, 3 brand variants claimed to contain added ingredients that improve the baby's intelligence, which may leave some mothers with the impression that their own breast milk is inferior to formula [9]. Again, this is despite the Code clearly stating not to permit the formula to be marketed as being equal to or comparable with breast milk [21]. This violation is not unique to the Chinese market. In the

United Kingdom, an analysis of 13 online parental chat rooms found that the single most repeated idea across the sites was that formula was closest to breast milk, a statement that was originated from the Aptamil (Danone) marketing [43]. A qualitative study of 4 focus groups reported that participants found ads confusing in terms of how formula-feeding is superior to, inferior to, or the same as breastfeeding [38]. We identified 7 brand variants that claimed to help preventing infants from developing allergies, and 8 brand variants claimed their products were good for the baby's digestive health. However, the evidence for these claimed benefits is not established [44]. A US study found that more than half of the 22 infant formula products reviewed were marketed with claims, whereas none of which were backed by publicly available scientific evidence [45].

### Policy Change and Actions Needed

The Chinese national government must take responsibility for ensuring the Code is implemented with adequate enforcement measures. Advertising regulations that restrict a broader range of marketing techniques are urgently needed. Monitoring advertisers' compliance with such restrictions forms a strong basis for regulation and is particularly important for online and digital media. Moreover, a significant monetary penalty should be applied if milk formula companies or other app platforms are found to be breaking laws. Concurrently, when compared with the innovative strategies used by the milk formula industry, a more active approach is needed to promote breastfeeding by public health authorities, rather than the simple message that breast milk is the best. Scaling up health professional advice to breastfeed alongside engaging media campaigns or social mobilization events such as the national breastfeeding day or world breastfeeding week is crucial. Brazil, for example, is widely recognized for implementing a successful National Breastfeeding Program that has made a substantial improvement in breastfeeding exclusivity and duration [46]. This society-wide program includes regulation of the commercialization of infant formula and foods, a strict enforcement of the Code, training for health workers and the development of mother-to-mother support groups, maternity leave extended to 6 months,

introduction of the Baby Friendly Hospitals Initiative, and investment in over 200 human milk banks [46]. Public health professionals should also learn from the milk formula industry in terms of constructing messages and applying proven marketing techniques in promoting breastfeeding. In China in particular, there are few health professional-endorsed parenting apps that advocate for breastfeeding and healthy infant feeding, and the space is primarily occupied by commercial entities [19].

### Strengths and Limitations

A key strength of the study is that we developed and established a comprehensive coding tool to manually analyze milk formula ad content, including text and images, in a non-English language context. This content analysis also has some limitations. First, this study is only a snapshot of a small number of available free-to-download parenting apps. Second, although the ad collection was not confined to one type of milk formula, we may have missed some milk formula brands or variants, as many products are being continually added and removed from the market. Although a second independent coder was used to test the accuracy of the ad coding and high agreement was achieved, it is possible that ad characteristics could be missed or miscoded.

### Conclusions

This is the first study to analyze milk formula ads found on the popular Chinese parenting apps. Products that function as BMS should not be freely advertised [10]. The purpose of these marketing messages is to widen the use of formulas and normalize formula as an appropriate food for all infants and young children rather than as a specialized food for those unable to breastfeed. The present analysis affirms a need for greater efforts in implementation, monitoring, and enforcement of the Code. Implementation of the Code is not possible without adequate funds and allocated budgets from national and local governments. As stated, China adopted the Code in 1995 but repealed these legal measures in 2016. The government should enshrine the Code in law to both ensure effective monitoring and comprehensive enforcement of unethical BMS marketing practices and to create an environment where breastfeeding is normal, accepted, and protected.

---

### Acknowledgments

The authors would like to thank Dr Nina Berry for assistance with establishing the ad content coding process. All data generated or analyzed during this study are available within this study and its supplementary information file.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

Distribution of the types of milk formula across brands and brand variants that appeared in the advertisements.

[DOCX File, 36 KB - [mhealth\\_v7i11e14219\\_app1.docx](#) ]

---

#### Multimedia Appendix 2

Example of a2 Platium formula advertisement featuring emotional imagery, marketing elements and claims.

[PNG File, 1455 KB - [mhealth\\_v7i11e14219\\_app2.png](#) ]

## Multimedia Appendix 3

Example of NAN Pro formula advertisement featuring emotional imagery, marketing elements and claims.

[PNG File, 513 KB - [mhealth\\_v7i11e14219\\_app3.png](#)]

**References**

1. Heinig MJ. Host defense benefits of breastfeeding for the infant. Effect of breastfeeding duration and exclusivity. *Pediatr Clin North Am* 2001 Feb;48(1):105-23, ix. [doi: [10.1016/s0031-3955\(05\)70288-1](#)] [Medline: [11236719](#)]
2. Jones G, Steketee RW, Black RE, Bhutta ZA, Morris SS, Bellagio Child Survival Study Group. How many child deaths can we prevent this year? *Lancet* 2003 Jul 5;362(9377):65-71. [doi: [10.1016/S0140-6736\(03\)13811-1](#)] [Medline: [12853204](#)]
3. Newton ER. Breastmilk: the gold standard. *Clin Obstet Gynecol* 2004 Sep;47(3):632-642. [doi: [10.1097/01.grf.0000136184.19927.98](#)] [Medline: [15326427](#)]
4. Quigley MA, Kelly YJ, Sacker A. Breastfeeding and hospitalization for diarrheal and respiratory infection in the United Kingdom Millennium Cohort Study. *Pediatrics* 2007 Apr;119(4):e837-e842. [doi: [10.1542/peds.2006-2256](#)] [Medline: [17403827](#)]
5. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality. *Lancet* 2000 Feb 5;355(9202):451-455. [Medline: [10841125](#)]
6. World Health Organization. Babies and mothers worldwide failed by lack of investment in breastfeeding. *Saudi Med J* 2017;38(9):974-975 [FREE Full text]
7. Mason F, Rawe K, Wright S. Save the Children. 2013. Superfood for Babies: How Overcoming Barriers to Breastfeeding Will Save Children's Lives URL: <https://www.savethechildren.org/content/dam/global/reports/health-and-nutrition/baby-superfood-asia.pdf?vanityurl=BF-Report> [accessed 2019-03-31] [WebCite Cache ID 77GoJ9M3D]
8. Baker P, Smith J, Salmon L, Friel S, Kent G, Iellamo A, et al. Global trends and patterns of commercial milk-based formula sales: is an unprecedented infant and young child feeding transition underway? *Public Health Nutr* 2016 Oct;19(14):2540-2550. [doi: [10.1017/S1368980016001117](#)] [Medline: [27211798](#)]
9. Piwoz EG, Huffman SL. The impact of marketing of breast-milk substitutes on WHO-recommended breastfeeding practices. *Food Nutr Bull* 2015 Dec;36(4):373-386. [doi: [10.1177/0379572115602174](#)] [Medline: [26314734](#)]
10. Rollins NC, Bhandari N, Hajeerhoy N, Horton S, Lutter CK, Martines JC, Lancet Breastfeeding Series Group. Why invest, and what it will take to improve breastfeeding practices? *Lancet* 2016 Jan 30;387(10017):491-504. [doi: [10.1016/S0140-6736\(15\)01044-2](#)] [Medline: [26869576](#)]
11. Changing Markets. 2017. Milking It: How Milk Formula Companies Are Putting Profits Before Science URL: <https://changingmarkets.org/wp-content/uploads/2017/10/Milking-it-Final-report-CM.pdf> [accessed 2019-03-29] [WebCite Cache ID 77EVvYSWg]
12. Xu F, Liu X, Binns CW, Xiao C, Wu J, Lee AH. A decade of change in breastfeeding in China's far north-west. *Int Breastfeed J* 2006 Nov 24;1:22 [FREE Full text] [doi: [10.1186/1746-4358-1-22](#)] [Medline: [17123450](#)]
13. Liu A, Dai Y, Xie X, Chen L. Implementation of international code of marketing breast-milk substitutes in China. *Breastfeed Med* 2014 Nov;9(9):467-472. [doi: [10.1089/bfm.2014.0053](#)] [Medline: [25026262](#)]
14. Zhang Y, Carlton E, Fein SB. The association of prenatal media marketing exposure recall with breastfeeding intentions, initiation, and duration. *J Hum Lact* 2013 Nov;29(4):500-509. [doi: [10.1177/0890334413487256](#)] [Medline: [23686404](#)]
15. Derbyshire E, Dancy D. Smartphone medical applications for women's health: what is the evidence-base and feedback? *Int J Telemed Appl* 2013;2013:782074 [FREE Full text] [doi: [10.1155/2013/782074](#)] [Medline: [24454354](#)]
16. Hearn L, Miller M, Fletcher A. Online healthy lifestyle support in the perinatal period: what do women want and do they use it? *Aust J Prim Health* 2013;19(4):313-318. [doi: [10.1071/PY13039](#)] [Medline: [23899373](#)]
17. Johnson SA. 'Maternal devices', social media and the self-management of pregnancy, mothering and child health. *Societies* 2014 Jun 13;4(2):330-350. [doi: [10.3390/soc4020330](#)]
18. Useit. 2018. Seven Wheat Research Institute: 2018 Maternal and Child App Industry Analysis Report URL: <https://www.useit.com.cn/thread-19156-1-1.html> [accessed 2019-03-30] [WebCite Cache ID 77EVIImNqs]
19. Zhao J, Freeman B, Li M. How do infant feeding apps in China measure up? A content quality assessment. *JMIR Mhealth Uhealth* 2017 Dec 6;5(12):e186 [FREE Full text] [doi: [10.2196/mhealth.8764](#)] [Medline: [29212627](#)]
20. SlideShare. 2016. Market Overview: Identifying New Trends and Opportunities in the Global Infant Formula Market URL: <https://www.slideshare.net/Euromonitor/market-oveview-identifying-new-trends-and-opportunities-in-the-global-infant-formula-market> [accessed 2019-03-29] [WebCite Cache ID 77EW8WAe4]
21. World Health Organization, UNICEF, IBFAN. Marketing of Breast-milk Substitutes: National Implementation of the International Code. Geneva: World Health Organization; 2018.
22. Lutter CK. The international code of marketing of breast-milk substitutes: lessons learned and implications for the regulation of marketing of foods and beverages to children. *Public Health Nutr* 2013 Oct;16(10):1879-1884. [doi: [10.1017/S1368980012004235](#)] [Medline: [23034164](#)]

23. International Baby Food Action Network. State of the Code by Country, 2006 URL: <http://ibfan.org/code-watch-reports> [accessed 2019-03-29] [WebCite Cache ID 77EWoAxJI]
24. Statista. 2018. Market Share of Mobile Operating Systems in China From January 2013 to August 2019 URL: <https://www.statista.com/statistics/262176/market-share-held-by-mobile-operating-systems-in-china/> [accessed 2019-03-29] [WebCite Cache ID 77EX4glwz]
25. Sajjacholapunt P, Ball LJ. The influence of banner advertisements on attention and memory: human faces with averted gaze can enhance advertising effectiveness. *Front Psychol* 2014;5:166 [FREE Full text] [doi: [10.3389/fpsyg.2014.00166](https://doi.org/10.3389/fpsyg.2014.00166)] [Medline: [24624104](https://pubmed.ncbi.nlm.nih.gov/24624104/)]
26. Bergen M, Dutta S, Shugan SM. Branded variants: a retail perspective. *J Mark Res* 1996;33(1):9-19. [doi: [10.2307/3152009](https://doi.org/10.2307/3152009)]
27. Greenland SJ. Cigarette brand variant portfolio strategy and the use of colour in a darkening market. *Tob Control* 2015 Mar;24(e1):e65-e71. [doi: [10.1136/tobaccocontrol-2013-051055](https://doi.org/10.1136/tobaccocontrol-2013-051055)] [Medline: [23988859](https://pubmed.ncbi.nlm.nih.gov/23988859/)]
28. Chen Y, Chang J, Gong Y. A content analysis of infant and toddler food advertisements in Taiwanese popular pregnancy and early parenting magazines. *J Hum Lact* 2015 Aug;31(3):458-466. [doi: [10.1177/0890334415576513](https://doi.org/10.1177/0890334415576513)] [Medline: [25766374](https://pubmed.ncbi.nlm.nih.gov/25766374/)]
29. Freeman B, Kelly B, Baur L, Chapman K, Chapman S, Gill T, et al. Digital junk: food and beverage marketing on Facebook. *Am J Public Health* 2014 Dec;104(12):e56-e64. [doi: [10.2105/AJPH.2014.302167](https://doi.org/10.2105/AJPH.2014.302167)] [Medline: [25322294](https://pubmed.ncbi.nlm.nih.gov/25322294/)]
30. Abrahams SW. Milk and social media: online communities and the International Code of Marketing of Breast-milk Substitutes. *J Hum Lact* 2012 Aug;28(3):400-406. [doi: [10.1177/0890334412447080](https://doi.org/10.1177/0890334412447080)] [Medline: [22674963](https://pubmed.ncbi.nlm.nih.gov/22674963/)]
31. Weibo. Zhang Ziyi URL: [https://www.weibo.com/u/1914100420?is\\_hot=1#\\_rnd1545009243734](https://www.weibo.com/u/1914100420?is_hot=1#_rnd1545009243734) [accessed 2018-11-02]
32. Edmond K, Bahl R. World Health Organization. 2007. Optimal Feeding of the Low Birth Weight Infants: Technical Review URL: [https://apps.who.int/iris/bitstream/handle/10665/43602/9789241595094\\_eng.pdf?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/43602/9789241595094_eng.pdf?sequence=1) [accessed 2018-11-20] [WebCite Cache ID 77EXwRpGN]
33. Chen E. mondaq. 2017. China: China CFDA Finalizes Infant Formula Registration Rules URL: <http://www.mondaq.com/china/x/659418/food+drugs+law/China+CFDA+Finalizes+Infant+Formula+Registration+Rules> [accessed 2019-03-30] [WebCite Cache ID 77EYEdUcY]
34. Berry NJ, Jones SC, Iverson D. Toddler milk advertising in Australia: infant formula advertising in disguise? *Australasian Mark J* 2012;20(1):24-27. [doi: [10.1016/j.ausmj.2011.10.011](https://doi.org/10.1016/j.ausmj.2011.10.011)]
35. Cattaneo A, Pani P, Carletti C, Guidetti M, Mutti V, Guidetti C, Follow-on Formula Research Group. Advertisements of follow-on formula and their perception by pregnant women and mothers in Italy. *Arch Dis Child* 2015 Apr;100(4):323-328. [doi: [10.1136/archdischild-2014-306996](https://doi.org/10.1136/archdischild-2014-306996)] [Medline: [25512963](https://pubmed.ncbi.nlm.nih.gov/25512963/)]
36. Family Health Service, The Department of Health Hong Kong SAR Government. Survey on Mothers' Views of Formula Milk Promotion and Information on Infant and Young Child Feeding. 2013. URL: [https://www.fhs.gov.hk/english/archive/files/reports/Survey\\_on\\_Mothers\\_views\\_on\\_FM\\_promotion\\_full\\_2016\\_final.pdf](https://www.fhs.gov.hk/english/archive/files/reports/Survey_on_Mothers_views_on_FM_promotion_full_2016_final.pdf)
37. Mason FG. Save the Children. 2018. Don't Push It: Why the Formula Milk Industry Must Clean Up Its Act URL: <https://resourcecentre.savethechildren.net/node/13218/pdf/dont-push-it.pdf> [accessed 2018-02-04]
38. Parry K, Taylor E, Hall-Dardess P, Walker M, Labbok M. Understanding women's interpretations of infant formula advertising. *Birth* 2013 Jun;40(2):115-124. [doi: [10.1111/birt.12044](https://doi.org/10.1111/birt.12044)] [Medline: [24635466](https://pubmed.ncbi.nlm.nih.gov/24635466/)]
39. Ji A, Wong YI, Cai T, Liu J. Infant formula safety concerns and consequences in China. *World J Pediatr* 2014 Feb;10(1):7-9 [FREE Full text] [doi: [10.1007/s12519-014-0447-3](https://doi.org/10.1007/s12519-014-0447-3)] [Medline: [24464657](https://pubmed.ncbi.nlm.nih.gov/24464657/)]
40. Parry J. China's tainted milk scandal spreads around world. *Br Med J* 2008 Oct 1;337:a1890. [doi: [10.1136/bmj.a1890](https://doi.org/10.1136/bmj.a1890)] [Medline: [18829644](https://pubmed.ncbi.nlm.nih.gov/18829644/)]
41. The Competition Commission of Singapore. 2017. Market Inquiry into the Supply of Formula Milk for Infants and Young Children in Singapore URL: <https://www.ccs.gov.sg/media-and-publications/publications/studies-research-papers/market-studies> [accessed 2018-08-08]
42. Cheryl N. Mintel. 75% of Chinese Mums Feed Their Baby Organic Infant Milk Formula URL: <https://www.mintel.com/press-centre/food-and-drink/three-in-four-chinese-mums-feed-their-baby-organic-infant-milk-formula> [accessed 2018-05-09]
43. The Caroline Walker Trust. I Hear its Closest to Breastmilk: A Review of the Discussions of Parents and Parents to be Around Formula and Formula Feeding on Web Discussion Sites URL: [https://www.cwt.org.uk/wp-content/uploads/2014/07/Formula\\_PTBRreport.pdf](https://www.cwt.org.uk/wp-content/uploads/2014/07/Formula_PTBRreport.pdf) [accessed 2018-07-04]
44. Abrams SA. Is it time to put a moratorium on new infant formulas that are not adequately investigated? *J Pediatr* 2015 Mar;166(3):756-760. [doi: [10.1016/j.jpeds.2014.11.003](https://doi.org/10.1016/j.jpeds.2014.11.003)] [Medline: [25491093](https://pubmed.ncbi.nlm.nih.gov/25491093/)]
45. Belamarich PF, Bochner RE, Racine AD. A critical review of the marketing claims of infant formula products in the United States. *Clin Pediatr (Phila)* 2016 May;55(5):437-442. [doi: [10.1177/0009922815589913](https://doi.org/10.1177/0009922815589913)] [Medline: [26054781](https://pubmed.ncbi.nlm.nih.gov/26054781/)]
46. Pérez-Escamilla R. Breastfeeding in Brazil: major progress, but still a long way to go. *J Pediatr (Rio J)* 2017;93(2):107-110 [FREE Full text] [doi: [10.1016/j.jped.2016.10.003](https://doi.org/10.1016/j.jped.2016.10.003)] [Medline: [27816509](https://pubmed.ncbi.nlm.nih.gov/27816509/)]

## Abbreviations

**BMS:** breast milk substitutes

**Code:** The International Code of Marketing of Breast Milk Substitutes

**UGC:** user-generated content

**WHO:** World Health Organization

*Edited by G Eysenbach; submitted 31.03.19; peer-reviewed by D Vollmer Dahlke, N Bartle, H Yang; comments to author 08.08.19; revised version received 06.09.19; accepted 24.09.19; published 29.11.19.*

*Please cite as:*

*Zhao J, Li M, Freeman B*

*A Baby Formula Designed for Chinese Babies: Content Analysis of Milk Formula Advertisements on Chinese Parenting Apps*

*JMIR Mhealth Uhealth 2019;7(11):e14219*

*URL: <http://mhealth.jmir.org/2019/11/e14219/>*

*doi: [10.2196/14219](https://doi.org/10.2196/14219)*

*PMID: [31782743](https://pubmed.ncbi.nlm.nih.gov/31782743/)*

©Jing Zhao, Mu Li, Becky Freeman. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 29.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# An Interactive Parent-Targeted Text Messaging Intervention to Improve Oral Health in Children Attending Urban Pediatric Clinics: Feasibility Randomized Controlled Trial

Belinda Borrelli<sup>1,2</sup>, PhD; Michelle Henshaw<sup>2,3</sup>, DDS; Romano Endrighi<sup>1</sup>, PhD; William G Adams<sup>4</sup>, MD; Timothy Heeren<sup>5</sup>, PhD; Rochelle K Rosen<sup>6,7</sup>, PhD; Beth Bock<sup>6,8</sup>, PhD; Scott Wertz<sup>9</sup>

<sup>1</sup>Center for Behavioral Science Research, Boston University Henry M Goldman School of Dental Medicine, Boston, MA, United States

<sup>2</sup>Center for Research to Evaluate & Eliminate Dental Disparities, Department of Health Policy & Health Services Research, Boston University Henry M Goldman School of Dental Medicine, Boston, MA, United States

<sup>3</sup>Office of Global & Population Health, Boston University Henry M Goldman School of Dental Medicine, Boston, MA, United States

<sup>4</sup>Department of Pediatrics, Boston Medical Center & Boston University School of Medicine, Boston, MA, United States

<sup>5</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States

<sup>6</sup>Centers for Behavioral and Preventive Medicine, The Miriam Hospital, Providence, RI, United States

<sup>7</sup>Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, RI, United States

<sup>8</sup>Department of Psychiatry and Human Behavior, Alpert Medical School at Brown University, Providence, RI, United States

<sup>9</sup>Agile Health, Inc, Lincolnshire, IL, United States

**Corresponding Author:**

Belinda Borrelli, PhD

Center for Behavioral Science Research

Boston University Henry M Goldman School of Dental Medicine

560 Harrison Ave, 3rd floor

Boston, MA, 02118

United States

Phone: 1 617 358 3358

Email: [belindab@bu.edu](mailto:belindab@bu.edu)

## Abstract

**Background:** Effective preventive treatments for dental decay exist, but caries experience among preschoolers has not changed, with marked disparities in untreated decay. Despite near-universal use of SMS text messaging, there are no studies using text messages to improve the oral health of vulnerable children.

**Objective:** This randomized controlled feasibility trial aimed to test the effects of oral health text messages (OHT) versus a control (child wellness text messages or CWT). OHT was hypothesized to outperform CWT on improving pediatric oral health behaviors and parent attitudes.

**Methods:** Parents with a child aged <7 years were recruited at urban clinics during pediatric appointments (79% [41/52] below poverty line; 66% [36/55] black) and randomized to OHT (text messages on brushing, dental visits, bottle and sippy cups, healthy eating and sugary beverages, and fluoride) or CWT (text messages on reading, safety, physical activity and development, secondhand smoke, and stress) groups. Automated text messages based on Social Cognitive Theory were sent twice each day for 8-weeks. Groups were equivalent on the basis of the number of text messages sent, personalization, interactivity, and opportunity to earn electronic badges and unlock animated characters. Assessments were conducted at baseline and 8 weeks later. Data were analyzed with linear mixed-effects models.

**Results:** A total of 55 participants were randomized (28 OHT and 27 CWT). Only one participant dropped out during the text message program and 47 (24 OHT and 23 CWT) completed follow up surveys. Response rates exceeded 68.78% (1040/1512) and overall program satisfaction was high (OHT mean 6.3; CWT mean 6.2; 1-7 scale range). Of the OHT group participants, 84% (21/25) would recommend the program to others. Overall program likeability scores were high (OHT mean 5.90; CWT mean 6.0; 1-7 scale range). Participants reported high perceived impact of the OHT program on brushing their child's teeth, motivation to address their child's oral health, and knowledge of their child's oral health needs (mean 4.7, 4.6, and 4.6, respectively; 1-5 scale range). At follow up, compared with CWT, OHT group participants were more likely to brush their children's teeth twice per day (odds ratio [OR] 1.37, 95% CI 0.28-6.50) and demonstrated improved attitudes regarding the use of fluoride (OR

3.82, 95% CI 0.9-16.8) and toward getting regular dental checkups for their child (OR 4.68, 95% CI 0.24-91.4). There were modest, but not significant, changes in motivation ( $F_{1,53}=0.60$ ;  $P=.45$ ) and self-efficacy ( $F_{1,53}=0.24$ ;  $P=.63$ ) to engage in oral health behaviors, favoring OHT ( $d=0.28$  and  $d=0.16$  for motivation and self-efficacy, respectively).

**Conclusions:** The OHT program demonstrated feasibility was well utilized and appealing to the target population and showed promise for efficacy.

(*JMIR Mhealth Uhealth* 2019;7(11):e14247) doi:[10.2196/14247](https://doi.org/10.2196/14247)

## KEYWORDS

oral health; mHealth; text message; dental caries; health behavior

## Introduction

### Background

Although there are effective preventive treatments for dental decay, caries experience among preschoolers has remained relatively unchanged for the past two decades [1]. However, not all populations share the burden of the disease equally, demonstrated by the persistent and marked disparities in caries experience, untreated decay, and the lack of dental care access by both race or ethnicity and income [1]. Finding innovative strategies to reduce the prevalence and severity of this disease in high-risk populations is essential to reducing disparities. For young children, the role of the primary caregiver is especially important in reducing caries risk [2].

Community health centers provide comprehensive and cost-effective primary health care for America's most underserved communities. Nationally, there are almost 1400 community health centers with over 11,000 locations that treat 28 million patients per year, of which 8.4 million are children [3]. Community health centers provide care primarily to low-income persons (91%), the under- or uninsured (49% Medicaid; government-funded medical and dental insurance for low-income individuals; and 23% uninsured), and racial and ethnic minority groups (63%) [3]—the same groups that are at highest risk for caries. Despite guidelines that the first dental visit should occur by age 1 year, less than 2% of children enrolled in Medicaid meet this recommendation, and only 9% of 1- and 2-year olds receive preventive dental visits [4]. In contrast, almost 90% of US children attend well-child primary care visits [5]. This high attendance rate, coupled with the fact that during the first 3 years of life, children have 12 well-child visits scheduled, provides an infrastructure with the potential to reach both children at high risk for caries and their parents.

Short message service (SMS) text messaging may be one way to reach busy parents in their everyday lives with evidenced-based information. Over 95% of adults in the United States regularly use SMS text messaging, with no disparities in race, ethnicity, and income [6,7]. The advantages of an SMS text messaging intervention are as follows: access anytime and place, ability to tailor to content timing and intensity, provision of real-time coping strategies to users in everyday settings, few barriers to participation, interactive functionality in real time, low participant burden, reduced cost burden on the health care system [8,9], and high potential for dissemination.

SMS text messaging interventions have been effective across a wide variety of behaviors, such as smoking cessation [10],

medication adherence, diabetes care [11], and weight management [12,13]. Although there are 3 studies that use SMS text messaging to improve pediatric oral health, they have small samples and short-term outcomes (1 week-3 months) and lack rigorous controls [14-16]. Only 1 was conducted in the United States and was limited in that it used text messages as *reminders* only, and text messages were sent for only 7 days [14].

We developed a text message program focused on motivating adherence to pediatric oral health behaviors. The program content and structure was based on clinical guidelines [17,18] and recommendations from a multidisciplinary scientific advisory board [19]. We also interviewed medical assistants, nurses, and pediatricians ( $n=9$ ) to assess their opinions about the program and on how to integrate it into the clinic flow. We conducted 11 focus groups with parents ( $n=63$ ) to develop text message content, match content to participants' literacy levels, design program preferences (ie, features, structure, length, and badges), incorporate cultural considerations, identify knowledge gaps, and map text message content onto a theoretical model and mediators (Social Cognitive Theory) [20-22]. To ensure functionality, we conducted an internal pilot (9 users) followed by a usability study during which participants from the target population ( $n=21$ ) used the text message program for one month and completed self-report questionnaires and qualitative interviews at the middle and end of the month regarding their opinions on program content, badges, and structure as well as satisfaction, comprehension, and perceived impact on hypothesized mediators and behavior.

### Objective

In this paper, we report the results from the text message program that we developed using the iterative process outlined above. Parents who attended the target pediatric clinics and have children under the age of 7 years were randomized into the pilot feasibility randomized controlled trial employing a parallel design and using a 1:1 allocation ratio to receive the oral health text messages (OHT) or child wellness text messages (CWT) for 2 months. Our comparison group, CWT, was developed by our team and scientific advisory board, using clinical guidelines [23,24]. The aim of the pilot was to test recruitment processes and assess participant satisfaction and the potential impact of OHT on putative Social Cognitive Theory mediators, oral health knowledge and attitudes, and pediatric oral health behaviors. We hypothesized that participant satisfaction would be equivalent between conditions (to preserve internal validity), and participants randomized to OHT group would experience positive changes in relevant Social Cognitive Theory constructs (eg, motivation, self-efficacy, and outcome



expectations) and report improved knowledge, attitudes, and oral health behaviors (tooth brushing).

## Methods

### Participants

Potential participants were parents (or caregivers) of children who were patients of pediatric clinics in 2 community health centers in an urban area of Boston, United States. The majority of patients in these clinics receive Medicaid (>88%). Participants were recruited to participate in our study, described as a *child wellness study*, by research assistants in pediatric waiting rooms and clinic staff referral. The research assistants administered informed consent to potential participants, and those who consented documented the consent in writing and were asked to complete baseline questionnaires. Participants were then randomized (using random number functions in a 1:1 allocation ratio using the statistical package SAS 9.4 (TS1M5) platform by SAS Institute Inc) to receive either OHT or CWT. Randomization triggered a text message asking the participant to *opt into* the program. A permuted randomized block design was used, stratified by clinic, child age, and history of caries. Research assistants were masked to treatment condition, as participants' first text messages were delivered 24 hours after enrollment.

Parents were considered eligible if they met the following criteria: aged  $\geq 18$  years and had a child aged 6 months to 7 years who received medical care at one of the target community health centers; lived in the greater Boston area and were not planning on moving for 8 weeks; spoke, understood, and read either English or Spanish fluently; had a mobile phone with unlimited SMS text messaging capability; texted at least one time in the past month; adequate ability to read health-related material [25]; were not enrolled in another mobile phone child health or wellness study; reported no abuse of alcohol or drugs [26]; and no previous serious mental illness. The study received ethical approval from our human subject institutional review board along with review and approval by the ethical committees at the community health centers. Participant recruitment and follow up assessment took place from March through May 2017.

### Procedure

#### Structure of the Programs

OHT and CWT were matched on program duration (8 weeks) and dose (2 text messages per day for 1 month followed by 1 text message per day for 1 month), engagement strategies (quizzes on fun facts, birthday text messages, ability to earn child-friendly animated badges for goal achievement, and ability to *unlock* higher levels of animated badges for engaging in the target behavior), and personalization and customization that allowed for the tailoring of message content. Text messages in both conditions were interactive, focusing on problem-solving barriers to behavior change. For OHT, the target goal was brushing every day, twice per day, and for CWT, the target goal was reading every day. Both programs provided feedback on progress toward goal attainment via ability to electronically view a *trophy case* of badges earned so far and motivational text messages. Participants could also participate in *challenge*

*weeks* during which they were given daily electronic badges for achieving the target behavior. The text messages were fully automated, but all incoming text messages could be monitored and responded to in real time via a dashboard interface. Responding to participants directly was necessary if the system did not recognize a text and, therefore, could not produce an automated response—for example, some participants sent a *smiley* emoji or asked a study-related question. The dashboard interface also allowed the study team to respond to participants who did not respond to assessments and allowed for rapid adjustment of personalization settings, such as changing the language of the text messages (ie, English to Spanish or Spanish to English). The dashboard was not used to communicate additional intervention content to participants. Text messages were delivered by Agile Health, Inc. Their system is Health Insurance Portability and Accountability Act compliant, and all data are encrypted in transit and at rest.

#### Oral Health Text Messages

OHT received core topic text messages (tooth brushing and cleaning gums and visiting the dentist) and choice topic text messages (bedtime routine, bottle and sippy cup use, sugar-sweetened beverages, healthy eating, getting fluoride, and fun facts). In month 1, participants received 1 text message from the core topics and 1 text message from the choice topics each day; in month 2, only 1 text message was sent per day, alternating between core and choice topics. Participants earned weekly badges depicting colorful dental-related images if they met the goal of brushing twice each day, working toward the goal of earning a *SuperTooth Hero* badge at the end of each of the 2 months (*Charlie Chew* and *Molly Molar*). Participants could also opt into a *challenge week* in which they were cued daily to brush their child's teeth to unlock *bonus SuperTooth Heroes* (*Faye Fluoride* and *Captain Chomp*) upon achieving brushing twice a day. SuperTooth Heroes were 4 different anthropomorphized teeth wearing capes and holding toothbrushes. The badges and heroes were accompanied by a description of the goal that was achieved. Other text message features specific to OHT included a *dentist finder* (geared to find pediatric dentists), photos and video of brushing technique, and Web links (eg, amount of sugar in popular food and drinks).

#### Child Wellness Text Messages

Participants who were randomized to CWT received core topic text messages focused on the promotion of reading and safety in the home. Choice topic text messages were as follows: healthy sleep and behavior, safety hazards, child development, physical activity, stress tips, and eliminating second hand smoke. CWT earned weekly badges depicting cartoon animals for reading to their child every day. If they earned 4 different weekly badges (each badge specifying the reading goal achieved), they were entitled to *book buddy* badges (anthropomorphized books) each month. They could also earn 2 *mystery book buddies* for the completion of challenge weeks. The badges and buddies were accompanied by a statement of the goal that was achieved. Other text message features specific to CWT included strategies for handling challenging child behaviors, Web links for community resources for parents, and photos or video of important concepts.

## Measures

Surveys were self-administered on the Web at baseline, before randomization, and at the 2-month follow up (after the end of daily text message programs). Participants were compensated a total of US \$25 for completing the baseline survey and US \$40 for completing the follow up survey. To prevent participant expectations from unduly influencing the results, assessments of CWT (reading and safety) and OHT (oral health behavior) outcomes were given to all participants. As the main purpose of the study was to report on OHT, we have presented the measures and results of only those outcomes.

## Sociodemographics

Sociodemographics including age, sex, education, income, race and ethnicity, marital and employment status, and child characteristics were obtained through self-report at baseline.

## Program Satisfaction

Program satisfaction measures were given at the conclusion of the text message programs (2 months after baseline). We used several indices of satisfaction because of the multidimensional nature of program satisfaction. First, we measured the *share-worthiness* of the text messages by asking whether participants showed the text messages to others and the extent to which they believed that the text messages would be helpful to family and friends (ranges from 1=*not at all helpful* to 7=*very helpful*). The perceived quality of the text messages were measured with 2 items from the Mobile Application Rating Scale (MARS) [27]; one assessing perceived quality through a star rating (1 star=*one of the worst text message programs*, 3 stars=*average*, and 5 stars=*one of the best text message programs*) and the other assessing how much longer they would have liked to receive the text messages (range from 1=*1 month* to 5=*5 months*). Satisfaction with each program component and overall program satisfaction were assessed with 11 items, each rated on a 1 to 7 scale (range from 1=*not satisfied at all* to 7=*very much satisfied*). The *likeability* of each program component was assessed with 10 items, each rated on a 1 to 7 scale (range from 1=*did not like it at all* to 7=*liked it very much*); we also computed an overall likeability score by averaging across the items. We assessed the perceived impact of the text messages with a 6-item scale from the MARS [27]. This scale assessed the extent to which participants randomized to the OHT group perceived that the program had an impact on their knowledge about oral health, their motivation to brush their child's teeth, their attitude toward changing their oral health practices, and likelihood of actual behavior change [27]. Each item was rated on a 5-point Likert scale (range from 1=*strongly disagree* to 5=*strongly agree*). We also assessed the perceived impact of the program on 6 key behaviors that corresponded to core and choice topics in the OHT. Participants rated the extent to which the program had an impact on each behavior using a 7-point Likert scale (range from 1=*not at all* to 7=*very much*) [27].

## Child Brushing and Fluoride Use

Child tooth brushing was assessed as *never*, *sometimes but not every day*, *once a day*, *twice a day*, and *more than twice a day* [28]. Responses were collapsed into 2 levels: yes=achieved

brushing recommendations (*twice a day or more than twice a day*) versus no (*never*, *sometimes but not every day*, or *once a day*). In the OHT group alone, brushing was assessed weekly through text messages in which participants were asked how many days in the last 7 days were their child's teeth brushed and how many times each day (1, 2, or more). We computed a brushing behavior variable by multiplying the number of days (0-7) by the number of times per day the child's teeth were brushed each week. Use of fluoride toothpaste was assessed with 1 item, "when your child's teeth are brushed, is fluoride toothpaste usually used (yes/no)?" [28].

## Attitudes Toward Oral Health

Attitudes toward oral health were assessed with 3 items from the Basic Risk Factors Questionnaire [29-32]: (1) "Children can get cavities as soon as their first tooth comes in," (2) "It is best to use toothpaste with fluoride when brushing a child's teeth," and (3) "Children's teeth should be brushed the last thing before bed." Participants rated each on a 4-point scale (1=*strongly disagree*, 2=*somewhat disagree*, 3=*somewhat agree*, and 4=*strongly agree*). For analyses, *strongly disagree*, *somewhat disagree*, and *somewhat agree* were collapsed into a *not strongly agree* category and compared against the *strongly agree* group.

## Social Cognitive Theory Constructs

Social Cognitive Theory constructs were assessed with measures from the Basic Risk Factors Questionnaire [32]. Outcome expectations or beliefs that engaging in a behavior will produce a desired outcome was measured with 3 items: (1) "Limiting my child's intake of sugary foods and drinks can help prevent cavities," (2) "Drinking tap water can help prevent cavities," and (3) "Regular dental checkups help keep children's teeth and mouth healthy," each rated on a 4 point scale (1=*strongly disagree*, 2=*somewhat disagree*, 3=*somewhat agree*, and 4=*strongly agree*). *Strongly disagree*, *somewhat disagree*, and *somewhat agree* were collapsed into *not strongly agree* and compared against *strongly agree*. Motivation was assessed with 4 items measuring participants' degree of desire to engage in recommended oral health behaviors, each rated on a scale ranging from 1 to 5 (higher scores reflect higher motivation). Self-efficacy was assessed with 4 items that assessed participants' perceived degree of confidence in their ability to engage in recommended oral health behaviors, each item rated on a scale ranging from 1 to 5 (higher scores reflect greater self-efficacy).

## Program Engagement

Engagement was collected automatically through program interaction. We computed *dose received* by dividing the number of text messages sent to participants each week by the number of participants and then averaging across all weeks. A *total response rate* was computed by computing the number of participant-submitted responses to texts in which a response was expected and dividing this by the number of possible responses. An *assessment response rate* was computed by dividing the number of participants that responded to assessment texts by the number of participants. The number of *unsolicited* user texts (texts sent by users where a response was not expected

such as *emojis* and *thank you*) was an additional index of user engagement. We also tracked the number of participants who opted into *challenge weeks* in which users were to set a daily behavioral goal (for the OHT group participants, the *challenge* was brushing their child's teeth twice per day, and for the CWT group participants, the *challenge* was reading to their child for 10 min each day). For OHT group, we also assessed the proportion of participants choosing each *choice* module across the 8-week program. This was computed by combining the number of OHT participants that selected any given *choice* module every week the module was available and dividing it by the number of potential module choices (product of number of weeks a module was made available and sample size).

### Analytic Plan

At baseline, study groups were compared on sociodemographic characteristics using independent sample *t* tests for continuous variables or chi-square tests for categorical variables. The baseline characteristics of participants who did not complete the follow up survey ( $n=8$ ) were also compared with the rest of the sample using *t* tests or chi-square tests as appropriate. User engagement and interaction with the program data were summarized and compared between study groups. Program satisfaction was compared between groups with *t* tests, and perceived program impact descriptives are presented for the OHT group.

Changes in oral health attitudes and behaviors from baseline to follow up in the OHT group compared with the CWT group were analyzed through models for longitudinal data with a group-by-time interaction representing the intervention effect. For binary outcomes, Generalized Estimating Equations (GEE) logistic regression for longitudinal data estimated the odds of achieving a behavior (eg, brushing recommendations or use of fluoride toothpaste) at follow up compared with

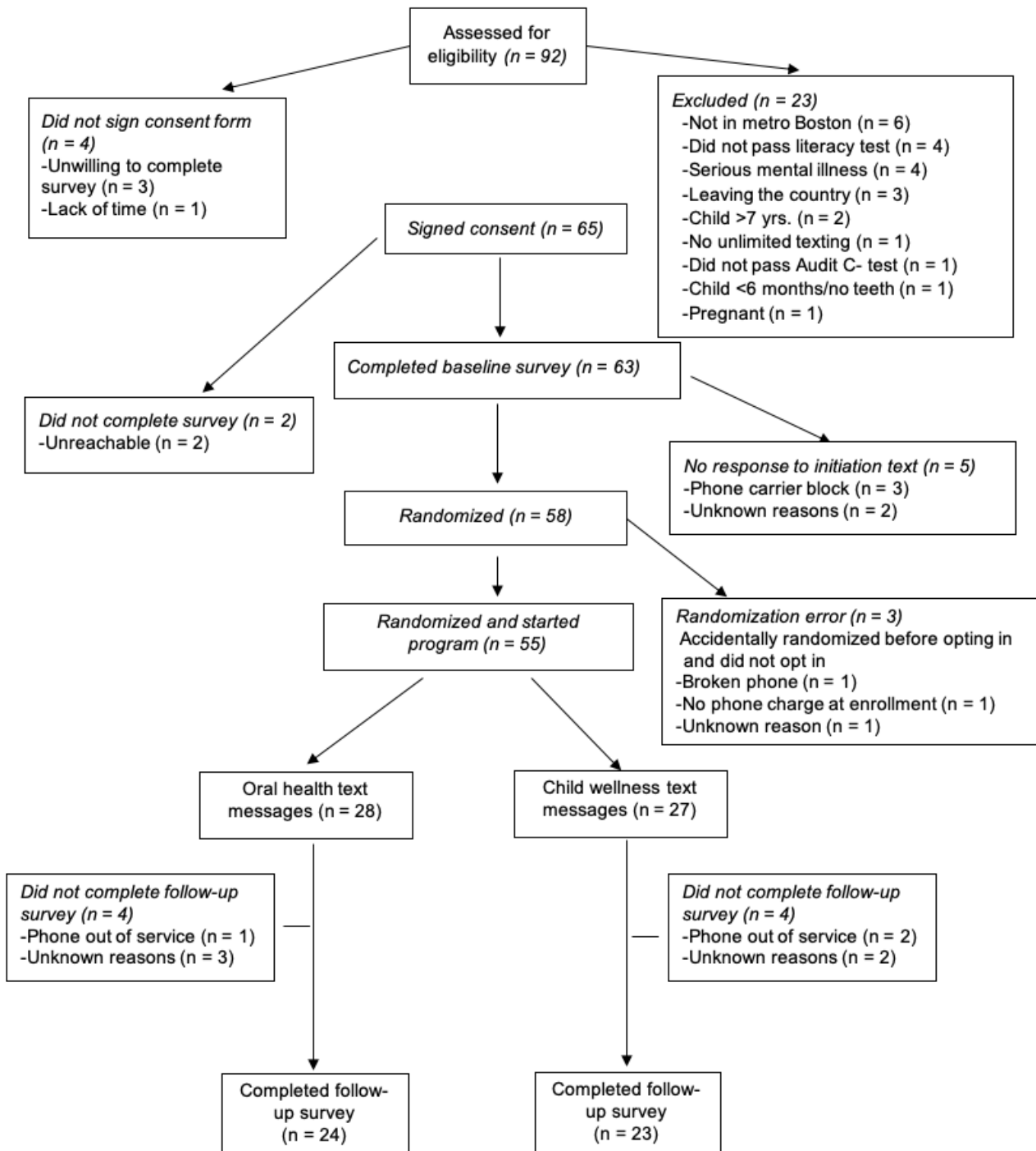
baseline, for those in the OHT versus CWT group. For outcome expectations and attitudes toward oral health, GEE logistic regression estimated the odds of strongly versus not strongly agreeing to each construct item at follow up versus baseline, for the OHT versus CWT group. For continuous outcomes, mixed-effects linear regression models were used to compare changes in group means from baseline to follow up, in the OHT versus CWT group. Effect sizes for continuous measures are presented as Cohen *d* calculated on the change score between baseline and follow up. Effect sizes for binary measures are presented as odds ratios (OR) with 95% confidence intervals. Analyses were conducted on participants who completed the follow up assessment ( $n=47$ ).

## Results

### Overview

As shown in [Figure 1](#), 55 out of the 65 individuals (55/65, 85%) who were eligible and signed the informed consent were randomized and initiated the text messages programs; 47 out of the 55 randomized participants (47/55, 85%) completed the end of treatment assessment with no significant differences in completion rates between groups. Participants who did not complete the follow up survey ( $n=8$ ) were significantly younger (mean age in years 26.8 vs 31.6;  $P=.02$ ). Satisfaction data were completed by 48 participants. Only 1 participant dropped out of the text message program. Owing to a technical problem, there were 3 CWT that were scheduled to be delivered but were not delivered. A total of 2 of these text messages applied only to participants who had toddlers, and 1 text message applied only to those who chose the stress management module. Therefore, this technical problem affected <2% of the text messages. There were no other unintended or harmful effect to participants.

Figure 1. Consolidated Standards of Reporting Trials diagram.



**Sociodemographics**

Compared with OHT, CWT group participants were more likely to be employed ( $\chi^2_4=14.7$ ;  $P=.005$ ) and have received information about children’s dental health at pediatric visits in the past year ( $\chi^2_1=4.1$ ;  $P=.04$ ). The mean age of the child was

2.7 years (SD 1.7) in the OHT group and 3.0 years (SD 1.9) in the CWT group. Out of the 28 children, 6 (6/28, 21%) in the OHT group, and out of the 27 children, 5 (5/27, 18%) in the CWT group had a history of cavities, reported via parent self-report. The baseline characteristics of the participants are shown in [Table 1](#).

**Table 1.** Baseline characteristics of participants by treatment group.

Baseline characteristics	Oral health text messages (n=28), n (%)	Child wellness text messages (n=27), n (%)	All (N=55), n (%)
Female	28 (100)	25 (93)	53 (96)
Parent/caregiver age (year), mean (SD)	31.0 (6.4)	31.0 (6.9)	31.0 (6.6)
Below poverty line	23 (88) <sup>a</sup>	18 (69) <sup>a</sup>	41 (79)
Less than high school education	6 (21)	2 (7)	8 (15)
Full/part time employment <sup>b</sup>	8 (29)	21 (78)	29 (53)
Married/engaged/live together	10 (36)	5 (19)	15 (27)
<b>Race/ethnicity</b>			
Black—African American	18 (64)	18 (67)	36 (66)
White—non-Hispanic	4 (14)	0 (0)	4 (7)
Asian	2 (7)	1 (4)	3 (5)
Hispanic (black/white)	3 (11)	6 (22)	9 (16)
Multiracial/other	1 (4)	2 (7)	3 (6)
Received dental health information <sup>c</sup>	9 (32)	16 (59)	25 (45)
Use fluoride toothpaste	12 (46) <sup>a</sup>	12 (48) <sup>d</sup>	24 (47)
<b>Mobile communication preference<sup>e</sup></b>			
Text message	21 (78)	21 (78)	42 (78)
Phone call	2 (7)	2 (7)	4 (7)
No preference	4 (15)	4 (15)	8 (15)
<b>Texting frequency<sup>e</sup></b>			
Every day	21 (78)	24 (89)	45 (83)
Most days	6 (22)	1 (4)	7 (13)
Occasionally	0 (0)	2 (7)	2 (4)
<b>Number of text messages sent per day</b>			
1-3	4 (14)	2 (7)	6 (11)
4-10	7 (25)	9 (33)	16 (29)
≥11	17 (61)	16 (59)	33 (60)

<sup>a</sup>n=26.<sup>b</sup>P=.005.<sup>c</sup>P=.04.<sup>d</sup>n=25.<sup>e</sup>Oral health text messages, n=27.

## Program Satisfaction

In the OHT group, 8 out of 25 participants (8/25, 32%) showed the text messages to others compared with 11 out of 23 (11/23, 47%) in the CWT group ( $\chi^2_1=1.0$ ;  $P=.32$ ). Participants in both groups believed that the text messages would be helpful to family/friends (1-7 scale; OHT mean 5.7, SD 1.6; CWT mean 6.0, SD 1.0;  $t_{47}=-0.9$ ;  $P=.35$ ), and 84% (21/25) of OHT group participants said that they would recommend the program to others. In the OHT group, 15 out of 22 (15/22, 68%) participants rated the program 4 stars or more compared with 16 out of 21

(16/21, 78%) in the CWT group ( $\chi^2_1=0.6$ ;  $P=.34$ ). Out of 19 participants, 12 (12/19, 63%) in the OHT group wanted the program to last up to 2 months longer, and 7 participants (7/19, 37%) indicated that they wanted the program to last  $\geq 3$  months longer, which was not significantly different from the CWT group ( $\chi^2_1=1.8$ ;  $P=.18$ ). Satisfaction ratings were generally high for program components (Table 2). The mean overall level of program satisfaction was also high with no significant group differences (1-7 scale range; OHT mean 6.3, SD 1.4; CWT mean 6.2, SD 0.9;  $t_{46}=0.4$ ;  $P=.70$ ).

**Table 2.** Participant's satisfaction with the text message program.

Program satisfaction scale items	Oral health text messages (n=25), mean (SD)	Child wellness text messages (n=23), mean (SD)
The ability to choose topics of interest to me	6.29 (1.4)	6.21 (1.0)
The ability to earn electronic badges	5.66 (1.8)	5.72 (1.2)
The ability to unlock levels of badges	5.63 (1.9)	5.66 (1.3)
The level of program customization to child	6.25 (1.4)	5.78 (1.2)
Receipt of support when needed	6.20 (1.1)	5.65 (1.0)
The amount of information in the text messages	6.24 (1.3)	5.86 (1.3)
The quality of the information in the text messages	6.08 (1.4)	5.86 (1.0)
Relevancy of program was for self and family	6.32 (1.4)	5.95 (1.1)
Age-appropriateness of the messages for child	6.00 (1.4)	6.21 (0.8)
The trustworthiness of the information	6.33 (0.8)	6.17 (0.9)
The degree to which the text messages apply to their family	5.96 (1.6)	6.00 (1.0)

With regard to the *likeability* of program components (Table 3), there were no significant differences between groups in overall *likeability* scores (1-7 scale range; OHT mean 5.90, SD 1.40; CWT mean 6.00, SD 0.80;  $t_{46}=-0.4$ ;  $P=.70$ ). Participants in both groups liked the frequency of the text messages (mean 5.75, SD 1.60), timing of the text messages (mean 5.74, SD 1.4) and text message features such as the ability to set goals (mean 6.21, SD 1.00), choose topics of interest (mean 6.40, SD 1.00), earn weekly badges (mean 5.68, SD 1.60), and participate in challenge weeks (mean 6.55, SD 0.70). With regard to the perceived impact of the program, OHT group participants indicated that taking part in the program increased their awareness of their child's oral health (1-5 scale range; mean 4.64, SD 1.10), increased their knowledge and understanding of their child's oral health needs (mean 4.64, SD 1.00), increased

their motivation to address their child's oral health (mean 4.64, SD 1.00), improved their attitude toward their child's oral health (mean 4.64, SD 0.80), encouraged them to bring their child to the dentist for regular checkups (mean 4.48, SD 1.10), and helped them ensure their child's teeth were brushed (mean 4.68, SD 1.0). OHT group participants also indicated that the program had a positive impact on brushing their child's teeth (1-7 scale range; mean 6.40, SD 1.50), increasing the amount of their child's tap water consumption (mean 5.48, SD 2.00), and decreasing their child's consumption of sugar-sweetened beverages (mean 6.32, SD 1.30) and sugary foods (mean 6.12, SD 1.50). OHT group participants also indicated that the program had a positive impact on their willingness to take their child to the dentist (mean 5.84, SD 2.0) and overall knowledge about their child's oral health (mean 6.20, SD 1.5).

**Table 3.** Likeability of the text message program.

Likeability scale items	Oral health text messages (n=25), mean (SD)	Child wellness text messages (n=23), mean (SD)
Responding to text message questions on a daily basis	5.52 (1.8)	5.82 (1.2)
Responding to text messages about the frequency of the target behavior (brushing or reading)	5.88 (1.6)	6.00 (1.0)
Setting goals	6.41 (0.9)	6.00 (1.2)
The ability to choose text message topics of interest to you	6.25 (1.4)	6.56 (0.6)
Receiving information about a particular topic	6.29 (1.4)	6.30 (0.8)
The frequency with which texts were delivered	5.64 (2.0)	5.86 (1.3)
The time of the day texts were received	5.83 (1.6)	5.65 (1.2)
Earning weekly badges	5.65 (2.0)	5.72 (1.4)
Earning monthly animated characters ( <i>SuperTooth Heroes</i> or <i>Book Buddies</i> )	5.60 (2.0)	5.47 (1.3)
Participating in a <i>challenge week</i>	6.79 (0.5)	6.30 (0.8)

### Child Brushing and Fluoride Use

The proportion of participants meeting pediatric brushing recommendations increased from baseline to follow up in both groups, with the effect size favoring OHT (OR 1.37, 95% CI 0.28-6.50). In the OHT group, brushing was also assessed

weekly through text messages, and increased rates of brushing was reported over the course of the 8-week program (time effect  $F_{1,94}=8.4$ ;  $P=.005$ ; see [Multimedia Appendix 1](#)). The proportions of participants using fluoride toothpaste to brush their child's teeth increased from baseline to follow up in both groups, with the effect size favoring CWT (OR 0.81, 95% CI 0.25-2.65).

## Attitudes Toward Oral Health

As shown in Table 4, the proportion of participants strongly agreeing that “children can get cavities in baby’s teeth” increased from baseline to follow up in both groups, with the effect size favoring the OHT group (OR 1.98, 95% CI 0.54-7.18). The proportion of participants strongly agreeing

that “it is best to use toothpaste with fluoride when brushing a child’s teeth” increased from baseline to follow up in the OHT group alone (OR 3.82, 95% CI 0.90-16.80), and the proportion strongly agreeing that “children’s teeth should be brushed the last thing before bed” increased in both groups, with the effect size favoring the OHT group (OR 2.07, 95% CI 0.10-41.50).

**Table 4.** Oral health behaviors and attitudes.

Oral health behavior and attitudes	Baseline		Follow up		Odds Ratio (95% CI)
	OHT <sup>a</sup> n/N, n (%)	CWT <sup>b</sup> n/N, n (%)	OHT n/N, n (%)	CWT n/N, n (%)	
<b>Oral health behavior</b>					
Achieved child brushing recommendations <sup>c</sup>	17/26 (65)	15/24 (62)	17/22 (77)	16/23 (70)	1.37 (0.28-6.5)
Used fluoride toothpaste to brush child’s teeth	12/26 (46)	12/25 (48)	14/22 (64)	14/21 (67)	0.81 (0.25-2.65)
<b>Attitudes toward oral health (strongly agree)</b>					
Children can get cavities in baby’s teeth	6/25 (24)	10/26 (38)	11/24 (46)	11/22 (50)	1.98 (0.54-7.18)
Best to use fluoride toothpaste for children	9/21 (43)	15/24 (62)	15/22 (68)	12/22 (54)	3.82 (0.90-16.8)
Should brush teeth last thing before bed	21/25 (84)	22/26 (85)	23/24 (96)	21/23 (91)	2.07 (0.10-41.50)
<b>Outcome expectations (strongly agree)</b>					
Limiting sugary foods helps prevent cavities	19/25 (76)	19/26 (73)	21/24 (87)	18/23 (78)	1.68 (0.23-11.80)
Drinking tap water helps prevent cavities	5/22 (23)	4/26 (15)	8/22 (36)	4/20 (20)	1.42 (0.24-8.60)
Dental checkups help keep teeth healthy	20/26 (77)	23/26 (88)	23/24 (96)	21/23 (91)	4.70 (0.24-91.40)

<sup>a</sup>OHT: oral health text message.

<sup>b</sup>CWT: child wellness text message.

<sup>c</sup>Brushing twice per day, 6 or more days per week.

## Social Cognitive Theory Constructs

With regard to outcome expectations, Table 4 shows that the proportion of participants strongly agreeing that “limiting children’s intake of sugary foods helps prevent cavities” increased from baseline to follow up in both groups, with the effect size favoring the OHT group (OR 1.68, 95% CI 0.23-11.80). There was a similar pattern for the belief that “drinking tap water can help prevent cavities,” with the effect size favoring the OHT group (OR 1.42, 95% CI 0.24-8.60), and for “regular dental check-ups help keep children’s teeth and mouth healthy,” with the effect size in favor of the OHT group (OR 4.68, 95% CI 0.24-91.40).

Participants’ motivation to engage in oral health promoting behaviors score increased between baseline (OHT mean 4.31 and CWT mean 4.42) and follow up (OHT mean 4.64 and CWT mean 4.52) in both groups, with the effect size favoring the OHT group ( $d=0.28$ ) and no significant between-group differences ( $F_{1,53}=0.60$ ;  $P=.45$ ). A similar pattern emerged for self-efficacy. Scores increased between baseline (OHT mean 4.50 and CWT mean 4.60) and follow up (OHT mean 4.76 and CWT mean 4.76) in both groups, with the effect size favoring the OHT group ( $d=0.16$ ) and no significant between-group differences ( $F_{1,53}=0.24$ ;  $P=.63$ ).

## Program Engagement

The mean weekly number of texts sent was comparable between groups (15.3 in OHT vs 15.4 in CWT), indicating dose equivalence. Aggregating over the 8-week study period, in the OHT group, there were 1040 responses out of the 1512 (1040/1512, 68.8% overall total response rate) possible responses, whereas in the CWT group, there were 1167 responses out of the 1463 (1167/1463, 79.8% overall total response rate) possible responses. Response to text messages did not taper off as the program progressed, remaining consistent across weeks for both groups (see Multimedia Appendix 2). Averaging across the study period, of the 28 OHT group participants, the mean number of participants responding to weekly assessment texts was 19.3 (19.3/28, 69% overall assessment response rate). Of the 27 CWT group participants, the mean number of participants responding to weekly assessment texts was 19 (19/27, 70% overall assessment response rate). Of the 28 OHT group participants, 16 (16/28, 57%) opted into at least one of the 2 possible *challenge weeks*. Of the 27 CWT group participants, 17 (17/27, 63%) opted into at least one of the 2 possible *challenge weeks*. Participants in the OHT group sent a total of 439 *unsolicited* text messages and participants in the CWT group sent a total of 475 *unsolicited* text messages during the 8 week program. With regard to choosing *choice* topics, aggregating over 8 weeks when *bedtime routine* was offered, 42% (35/84) selected the module; when *bottle/sippy cup* was offered, 21% (23/112) selected the module;

when *fun facts* was offered, 30% (25/84) selected the module; when *getting fluoride* was offered, 32% (27/84) selected the module; when *healthy eating* was offered, 51% (57/112) selected the module; and when *sugar-sweetened beverages* was offered, 26% (29/112) selected the module.

## Discussion

### Principal Findings

Sustainability of health behavior change is greatest if interventions are integrated into existing channels and woven into the fabric of people's lives. Our study incorporated both of these elements by partnering with pediatric clinics regularly visited by underserved families and using text messages, a preferred and near-universal form of communication. Text message interventions have the advantage of reaching large segments of previously unreachable populations with evidenced-based information, in real time and real-life settings. No previous studies have tested text messages to improve the oral health of at-risk children in a randomized controlled trial, matching for treatment dose and intensity. Our pilot study showed proof of concept of our OHT intervention with 4 principal findings: (1) OHT was perceived as highly acceptable and satisfactory, (2) participants in both conditions demonstrated a high level of engagement, (3) OHT had an impact on parent's attitudes toward oral health and social cognitive mediators, and (4) the program showed preliminary effectiveness at increasing brushing behaviors among those randomized to the OHT versus CWT group.

The high levels of acceptability and satisfaction reported by participants could be a function of the fact that we co-designed the program content and structure with the target population. Focus groups and interviews enabled us to ascertain participant preferences about the *surface structure* of the program (look and feel and images of the electronic badges) and the *deep structure* of the program (values and beliefs of the population) [33]. Collecting both quantitative and qualitative data during program development helped ensure that the words and images used were acceptable and incorporated cultural preferences and ensure that we could address any knowledge gaps and myths about oral health behaviors. We assessed satisfaction not only for the program as a whole but also for each program component, which has been rarely reported in the literature but is essential for the design of effective programs. Aside from self-report of satisfaction, another indicator of program satisfaction is whether or not participants report sharing the text messages with others. A large minority of our participants indicated that they shared the texts with family and friends. Thus, the program could also have a *contagion* effect, that is, unmeasured effects within each person's social network.

Our strategies for engagement not only included *static* strategies, such as personalization and customization, but also included *dynamic* strategies that required participant interaction, such as quizzes on fun facts, the ability to earn child-friendly electronic badges, and enticement to *unlock* access to other characters. Program engagement was high on all 3 indicators and did not significantly differ between groups. It is important to have several indicators of engagement to avoid masking differential

engagement rates, such as those that might exist between overall program responses and responses to research-related text messages. In addition, few, if any studies, have reported on *unsolicited* texts from participants, but this is also an important indicator of engagement because proactive responding could be an indicator of deeper processing of the text messages by participants, rather than simply reacting to text messages. Of the 3 studies that have used text messages for pediatric oral health, only 1 reported on program engagement, but that program was only 7 days in duration [14].

Few studies have examined the effect of a pediatric oral health intervention on parental attitudes, and previous studies have supported the association between attitudes and behavior [28]. Our findings indicated that, with the exception of using fluoridated toothpaste to brush their child's teeth, OHT consistently showed promise for changing caregiver attitudes and behaviors toward their child's oral health. No studies to date have used theory (Social Cognitive Theory) to develop a comprehensive oral health intervention to improve children's oral health through text messages. It may be that the development of text messages grounded in theory is an important factor in improving both parental attitudes and behaviors toward child's oral health. This is supported by the fact that OHT also showed promise for changing variables integral to the Social Cognitive Theory, such as motivation, self-efficacy, and outcome expectations.

Feasibility studies are used to determine whether an intervention is appropriate for further testing, particularly when there are few or no published studies on a particular intervention technique [34]. Our study met the criteria for intervention feasibility outlined by Bowen et al [34]. Specifically, we demonstrated acceptability (satisfaction and perceived appropriateness), demand (participants were engaged with the program), implementation (successful execution; no technical problems with the text messages), practicality (ease and quality of implementation and low burden on patients and providers), integration (fit into clinic work flow and lack of disruption of clinical care), and limited efficacy (intended effects of the program on key variables; perceived impact).

### Limitations

The primary purpose of the study was feasibility rather than a fully powered clinical trial, so caution should be used when interpreting group differences because of the lack of power. The small sample precludes generalization to the larger population, and as the sample was mostly women and those whose income was below the poverty line, it is unclear if the program would also be acceptable to men and those having higher incomes. Generalizability was also limited by our inclusion and exclusion criteria, which included adequate ability to read health-related material, no previous or current serious mental illness, and no current alcohol or drug abuse. These limitations are offset by the strengths of our study design, which matched groups on text message dose, frequency, and features; creative engagement strategies; objective measurement of engagement; targeting a high-risk population; and implementation in a real-world setting.



## Conclusions

Dissemination of text message interventions is highly viable given the high rate of SMS text messaging and lack of disparities by income, race, or ethnicity. Text message interventions could be disseminated at low cost and are delivered exactly as

designed, resulting in 100% reliable intervention. This study provides evidence that a larger fully powered randomized controlled trial with objective outcomes (clinical exam) should be conducted. If effective, the program could be disseminated nationally to other federally qualified pediatric clinics that serve vulnerable and high-risk populations.

## Acknowledgments

This study was funded by the National Institute of Dental and Craniofacial Research UH2 DE025492 to BB (lead co-principal investigator) and MH (co-principal investigator).

## Conflicts of Interest

SW is the chief executive officer of Agile Health, the company that deployed the text messages described in this study.

This randomized study has not been registered, explained by authors with "During the time of study start-up, the rules regarding clinical trial registration were unclear for pilot feasibility studies and we were advised not to register it." The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

### Multimedia Appendix 1

Engagement and interaction with the text message program over time.

[PDF File (Adobe PDF File), 77 KB - [mhealth\\_v7i11e14247\\_app1.pdf](#) ]

### Multimedia Appendix 2

Changes in child brushing over time in the oral health text messages group.

[PDF File (Adobe PDF File), 52 KB - [mhealth\\_v7i11e14247\\_app2.pdf](#) ]

### Multimedia Appendix 3

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2819 KB - [mhealth\\_v7i11e14247\\_app3.pdf](#) ]

## References

1. Dye BA, Mitnik GL, Iafolla TJ, Vargas CM. Trends in dental caries in children and adolescents according to poverty status in the United States from 1999 through 2004 and from 2011 through 2014. *J Am Dent Assoc* 2017 Dec;148(8):550-65.e7. [doi: [10.1016/j.adaj.2017.04.013](#)] [Medline: [28619207](#)]
2. Albino J, Tiwari T. Preventing childhood caries: a review of recent behavioral research. *J Dent Res* 2016 Jan;95(1):35-42 [FREE Full text] [doi: [10.1177/0022034515609034](#)] [Medline: [26438210](#)]
3. Finlayson TL, Siefert K, Ismail AI, Delva J, Sohn W. Reliability and validity of brief measures of oral health-related knowledge, fatalism, and self-efficacy in mothers of African American children. *Pediatr Dent* 2005;27(5):422-428 [FREE Full text] [Medline: [16435644](#)]
4. Bouchery E. Utilization of dental services among medicaid-enrolled children. *Medicare Medicaid Res Rev* 2013;3(3) [FREE Full text] [doi: [10.5600/mmrr.003.03.b04](#)] [Medline: [24753978](#)]
5. Child Trends. 2018. Well-Child Visits: Key Facts About Well-Child Visits. URL: <https://www.childtrends.org/?indicators=well-child-visits> [accessed 2019-03-15]
6. Duggan M. Pew Research Center. 2012. Cell Phone Activities 2012. URL: <http://www.pewinternet.org/2012/11/25/cell-phone-activities-2012/> [accessed 2019-03-15]
7. Pew Research Center. 2019. Mobile Fact Sheet. URL: <http://www.pewinternet.org/fact-sheet/mobile/> [accessed 2019-03-15]
8. Guerriero C, Cairns J, Roberts I, Rodgers A, Whittaker R, Free C. The cost-effectiveness of smoking cessation support delivered by mobile phone text messaging: Txt2Stop. *Eur J Health Econ* 2013 Oct;14(5):789-797 [FREE Full text] [doi: [10.1007/s10198-012-0424-5](#)] [Medline: [22961230](#)]
9. Chen YF, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. *Health Technol Assess* 2012;16(38):1-205, iii-v [FREE Full text] [doi: [10.3310/hta16380](#)] [Medline: [23046909](#)]

10. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2016 May 20;4(2):e49 [FREE Full text] [doi: [10.2196/mhealth.5436](https://doi.org/10.2196/mhealth.5436)] [Medline: [27207211](https://pubmed.ncbi.nlm.nih.gov/27207211/)]
11. Saffari M, Ghanizadeh G, Koenig HG. Health education via mobile text messaging for glycemic control in adults with type 2 diabetes: a systematic review and meta-analysis. *Prim Care Diabetes* 2014 Dec;8(4):275-285. [doi: [10.1016/j.pcd.2014.03.004](https://doi.org/10.1016/j.pcd.2014.03.004)] [Medline: [24793589](https://pubmed.ncbi.nlm.nih.gov/24793589/)]
12. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. *PLoS One* 2014 Feb 5;9(2):e88166 [FREE Full text] [doi: [10.1371/journal.pone.0088166](https://doi.org/10.1371/journal.pone.0088166)] [Medline: [24505411](https://pubmed.ncbi.nlm.nih.gov/24505411/)]
13. Stockwell MS, Kharbanda EO, Martinez RA, Vargas CY, Vawdrey DK, Camargo S. Effect of a text messaging intervention on influenza vaccination in an urban, low-income pediatric and adolescent population: a randomized controlled trial. *J Am Med Assoc* 2012 Apr 25;307(16):1702-1708. [doi: [10.1001/jama.2012.502](https://doi.org/10.1001/jama.2012.502)] [Medline: [22535855](https://pubmed.ncbi.nlm.nih.gov/22535855/)]
14. Hashemian TS, Kritz-Silverstein D, Baker R. Text2Floss: the feasibility and acceptability of a text messaging intervention to improve oral health behavior and knowledge. *J Public Health Dent* 2015;75(1):34-41. [doi: [10.1111/jphd.12068](https://doi.org/10.1111/jphd.12068)] [Medline: [25091471](https://pubmed.ncbi.nlm.nih.gov/25091471/)]
15. Sharma R, Hebbal M, Ankola AV, Murugabupathy V. Mobile-phone text messaging (SMS) for providing oral health education to mothers of preschool children in Belgaum City. *J Telemed Telecare* 2011;17(8):432-436. [doi: [10.1258/jtt.2011.110416](https://doi.org/10.1258/jtt.2011.110416)] [Medline: [22025742](https://pubmed.ncbi.nlm.nih.gov/22025742/)]
16. Makvandi Z, Karimi-Shahanjarini A, Faradmal J, Bashirian S. Evaluation of an oral health intervention among mothers of young children: a clustered randomized trial. *J Res Health Sci* 2015;15(2):88-93 [FREE Full text] [Medline: [26175290](https://pubmed.ncbi.nlm.nih.gov/26175290/)]
17. American Academy on Pediatric Dentistry, American Academy of Pediatrics. Policy on early childhood caries (ECC): classifications, consequences, and preventive strategies. *Pediatr Dent* 2008;30(7 Suppl):40-43. [Medline: [19216381](https://pubmed.ncbi.nlm.nih.gov/19216381/)]
18. American Academy on Pediatric Dentistry Council on Clinical Affairs. Policy on early childhood caries (ECC): unique challenges and treatment option. *Pediatr Dent* 2008;30(7 Suppl):44-46. [Medline: [19216382](https://pubmed.ncbi.nlm.nih.gov/19216382/)]
19. Borrelli B, Henshaw M, Rosen P, Endrighi R, Heeren T, Diamond C. Mixed Methods to Develop Parent-targeted Pediatric Oral Health Text Messages. In: Proceedings of the International Association for Dental Research. 2017 Presented at: IADR'17; March 22-25, 2017; San Francisco, CA.
20. Bandura A. *Social Foundations Of Thought And Action: A Social Cognitive Theory*. Upper Saddle River, NJ: Prentice Hall; 1986.
21. Rimer BK, Glanz K. *Theory at a Glance: A Guide for Health Promotion Practice*. Second edition. Washington, DC: US Department of Health and Human Services; 2005.
22. Bandura A, Freeman WH, Lightsey R. *Self-Efficacy: The Exercise of Control*. New York: Worth Publishers; 1997.
23. Healthy Children. 2019. The AAP Parenting Website. URL: <https://www.healthychildren.org/English/Pages/default.aspx> [accessed 2019-03-15]
24. Bright Futures - AAP. 2019. URL: <https://brightfutures.aap.org/Pages/default.aspx> [accessed 2019-03-15]
25. Morris NS, MacLean CD, Chew LD, Littenberg B. The single item literacy screener: evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract* 2006 Mar 24;7:21 [FREE Full text] [doi: [10.1186/1471-2296-7-21](https://doi.org/10.1186/1471-2296-7-21)] [Medline: [16563164](https://pubmed.ncbi.nlm.nih.gov/16563164/)]
26. Bush K, Kivlahan DR, McDonnell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory care quality improvement project (ACQUIP). Alcohol use disorders identification test. *Arch Intern Med* 1998 Sep 14;158(16):1789-1795. [Medline: [9738608](https://pubmed.ncbi.nlm.nih.gov/9738608/)]
27. 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 Nov 3;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
28. Wilson AR, Brega AG, Campagna EJ, Braun PA, Henderson WG, Bryant LL, et al. Validation and impact of caregivers' oral health knowledge and behavior on children's oral health status. *Pediatr Dent* 2016;38(1):47-54 [FREE Full text] [doi: [10.1016/j.pcd.2013.10.018](https://doi.org/10.1016/j.pcd.2013.10.018)] [Medline: [26892215](https://pubmed.ncbi.nlm.nih.gov/26892215/)]
29. Braun PA, Lind KE, Batliner T, Brega AG, Henderson WG, Nadeau K, et al. Caregiver reported oral health-related quality of life in young American Indian children. *J Immigr Minor Health* 2014 Oct;16(5):951-958 [FREE Full text] [doi: [10.1007/s10903-013-9870-0](https://doi.org/10.1007/s10903-013-9870-0)] [Medline: [23857123](https://pubmed.ncbi.nlm.nih.gov/23857123/)]
30. Hoelt KS, Barker JC, Shiboski S, Pantoja-Guzman E, Hiatt RA. Effectiveness evaluation of Contra Caries Oral Health Education Program for improving Spanish-speaking parents' preventive oral health knowledge and behaviors for their young children. *Community Dent Oral Epidemiol* 2016 Dec;44(6):564-576 [FREE Full text] [doi: [10.1111/cdoe.12250](https://doi.org/10.1111/cdoe.12250)] [Medline: [27517458](https://pubmed.ncbi.nlm.nih.gov/27517458/)]
31. Tiwari T, Quissell DO, Henderson WG, Thomas JF, Bryant LL, Braun PA, et al. Factors associated with oral health status in American Indian children. *J Racial Ethn Health Disparities* 2014 Sep 1;1(3):148-156 [FREE Full text] [doi: [10.1007/s40615-014-0017-3](https://doi.org/10.1007/s40615-014-0017-3)] [Medline: [25232515](https://pubmed.ncbi.nlm.nih.gov/25232515/)]

32. Albino J, Tiwari T, Gansky SA, Henshaw MM, Barker JC, Brega AG, et al. The basic research factors questionnaire for studying early childhood caries. *BMC Oral Health* 2017 May 19;17(1):83 [FREE Full text] [doi: [10.1186/s12903-017-0374-5](https://doi.org/10.1186/s12903-017-0374-5)] [Medline: [28526003](https://pubmed.ncbi.nlm.nih.gov/28526003/)]
33. Resnicow K, Soler R, Braithwaite RL, Ahluwalia JS, Butler J. Cultural sensitivity in substance use prevention. *J Community Psychol* 2000 May;28(3):271-290. [doi: [10.1002/\(sici\)1520-6629\(200005\)28:3<271::aid-jcop4>3.0.co;2-i](https://doi.org/10.1002/(sici)1520-6629(200005)28:3<271::aid-jcop4>3.0.co;2-i)]
34. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. *Am J Prev Med* 2009 May;36(5):452-457 [FREE Full text] [doi: [10.1016/j.amepre.2009.02.002](https://doi.org/10.1016/j.amepre.2009.02.002)] [Medline: [19362699](https://pubmed.ncbi.nlm.nih.gov/19362699/)]

## Abbreviations

**CWT:** child wellness text messages  
**GEE:** Generalized Estimating Equations  
**MARS:** Mobile Application Rating Scale  
**OHT:** oral health text messages  
**OR:** odds ratio  
**SMS:** short message service

*Edited by G Eysenbach; submitted 03.04.19; peer-reviewed by L Jamir, S Smith, A Card; comments to author 14.06.19; revised version received 06.08.19; accepted 18.08.19; published 11.11.19.*

### *Please cite as:*

*Borrelli B, Henshaw M, Endrighi R, Adams WG, Heeren T, Rosen RK, Bock B, Wertz S  
An Interactive Parent-Targeted Text Messaging Intervention to Improve Oral Health in Children Attending Urban Pediatric Clinics:  
Feasibility Randomized Controlled Trial  
JMIR Mhealth Uhealth 2019;7(11):e14247  
URL: <https://mhealth.jmir.org/2019/11/e14247>  
doi: [10.2196/14247](https://doi.org/10.2196/14247)  
PMID: [31710306](https://pubmed.ncbi.nlm.nih.gov/31710306/)*

©Belinda Borrelli, Michelle Henshaw, Romano Endrighi, William G Adams, Timothy Heeren, Rochelle K Rosen, Beth Bock, Scott Wertz. Originally published in *JMIR Mhealth and Uhealth* (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR mhealth and uhealth*, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Impact of Social Determinants of Health and Demographics on Refill Requests by Medicare Patients Using a Conversational Artificial Intelligence Text Messaging Solution: Cross-Sectional Study

Rena Brar Prayaga<sup>1</sup>, MA, JD; Ridhika Agrawal<sup>1,2</sup>; Benjamin Nguyen<sup>1,2</sup>, BA; Erwin W Jeong<sup>3</sup>, PharmD; Harmony K Noble<sup>4</sup>, MSHCA; Andrew Paster<sup>1</sup>, BS; Ram S Prayaga<sup>1</sup>, MS

<sup>1</sup>mPulse Mobile, Inc, Encino, CA, United States

<sup>2</sup>Grinnell College, Grinnell, IA, United States

<sup>3</sup>Medicare Medication Therapy Management & Medication Adherence programs, Kaiser Permanente Southern California, Downey, CA, United States

<sup>4</sup>Regional Pharmacy Clinical Operations, Kaiser Permanente Southern California, Downey, CA, United States

**Corresponding Author:**

Rena Brar Prayaga, MA, JD

mPulse Mobile, Inc

16530 Ventura Blvd

Suite 500

Encino, CA, 91436

United States

Phone: 1 888 678 5735

Email: [rena@mpulsemobile.com](mailto:rena@mpulsemobile.com)

## Abstract

**Background:** Nonadherence among patients with chronic disease continues to be a significant concern, and the use of text message refill reminders has been effective in improving adherence. However, questions remain about how differences in patient characteristics and demographics might influence the likelihood of refill using this channel.

**Objective:** The aim of this study was to evaluate the efficacy of an SMS-based refill reminder solution using conversational artificial intelligence (AI; an automated system that mimics human conversations) with a large Medicare patient population and to explore the association and impact of patient demographics (age, gender, race/ethnicity, language) and social determinants of health on successful engagement with the solution to improve refill adherence.

**Methods:** The study targeted 99,217 patients with chronic disease, median age of 71 years, for medication refill using the mPulse Mobile interactive SMS text messaging solution from December 2016 to February 2019. All patients were partially adherent or nonadherent Medicare Part D members of Kaiser Permanente, Southern California, a large integrated health plan. Patients received SMS reminders in English or Spanish and used simple numeric or text responses to validate their identity, view their medication, and complete a refill request. The refill requests were processed by Kaiser Permanente pharmacists and support staff, and refills were picked up at the pharmacy or mailed to patients. Descriptive statistics and predictive analytics were used to examine the patient population and their refill behavior. Qualitative text analysis was used to evaluate quality of conversational AI.

**Results:** Over the course of the study, 273,356 refill reminders requests were sent to 99,217 patients, resulting in 47,552 refill requests (17.40%). This was consistent with earlier pilot study findings. Of those who requested a refill, 54.81% (26,062/47,552) did so within 2 hours of the reminder. There was a strong inverse relationship ( $r_{10}=-0.93$ ) between social determinants of health and refill requests. Spanish speakers (5149/48,156, 10.69%) had significantly lower refill request rates compared with English speakers (42,389/225,060, 18.83%;  $X^2_1 [n=273,216]=1829.2$ ;  $P<.001$ ). There were also significantly different rates of refill requests by age band ( $X^2_6 [n=268,793]=1460.3$ ;  $P<.001$ ), with younger patients requesting refills at a higher rate. Finally, the vast majority (284,598/307,484, 92.23%) of patient responses were handled using conversational AI.

**Conclusions:** Multiple factors impacted refill request rates, including a strong association between social determinants of health and refill rates. The findings suggest that higher refill requests are linked to language, race/ethnicity, age, and social determinants of health, and that English speakers, whites, those younger than 75 years, and those with lower social determinants of health

barriers are significantly more likely to request a refill via SMS. A neural network-based predictive model with an accuracy level of 78% was used to identify patients who might benefit from additional outreach to narrow identified gaps based on demographic and socioeconomic factors.

(*JMIR Mhealth Uhealth* 2019;7(11):e15771) doi:[10.2196/15771](https://doi.org/10.2196/15771)

## KEYWORDS

text messaging; SMS; refill adherence; medication adherence; Medicare patients; conversational AI; social determinants of health; predictive modeling; machine learning; health disparities

## Introduction

### Background

Medication nonadherence is when patients are unable to follow prescribed treatment dose, time of day, and frequency [1]. There are a range of factors, including patient-related, physician-related, and health system barriers, that contribute to nonadherence [2]. Adherence has been shown to have a significant effect on treatment outcomes [3,4] and has a major impact in managing chronic conditions such as hypertension, cardiovascular disease, and diabetes. The Centers for Medicare and Medicaid Services (CMS) define an adherent patient as someone whose proportion of days covered is greater or equal to 80%. Put another way, patients are considered adherent when they refill often enough to cover 80% or more of their medication plan as prescribed by their health care provider and as agreed to by the patient [5]. Dispensing or refill data is commonly used to compute adherence levels because of the validity, relative accessibility, and inexpensiveness of such data [6].

The World Health Organization estimates a medication nonadherence rate of 50% for patients with 1 or more chronic conditions [2,7]. This staggering proportion of nonadherence is estimated to annually cost between 100 to 290 billion dollars in the United States [8]. Moreover, nonadherence is estimated to cause approximately 125,000 deaths and at least 10% of hospitalizations every year [9,10].

### Medicare Patients and Adherence

Individuals suffering from multiple chronic conditions and taking multiple medications are more likely to be nonadherent [11]. Patients eligible for Medicare, who are individuals older than 65 years and/or who have a disability, fit this description of patients at greater risk of nonadherence. Older patients and those with disabilities have more chronic conditions and are usually on multiple medications. Of 586 Medicare recipients offered medication therapy management, 575 (98.1%) completed a survey that asked questions relating to adherence. Among those who responded, 406 (69.2%) reported that they took their medication regularly and as prescribed. Of the remaining 169 (30%), 123 identified forgetfulness as an issue, 18 (11%) mentioned side effects, and 17 (10%) said the medication was not needed. Lower adherence rates were associated with difficulty paying for medication. Finally, subsidy recipients and non-English speakers were significantly less likely to be counseled about drug side effects [12].

### Use of Mobile Technology for Adherence

In a 2018 press release, the CMS committed to supporting modern and virtual methods of health care [13]. Furthermore, in a multinational survey conducted in 2018, 77% of respondents said the ability to request prescription refills via text message would increase their likelihood of choosing a health care provider. This percent is a 10-point increase over a 3-year span [14]. The Deloitte Center for Health Solutions conducted a nationally representative survey in which approximately a third of individuals indicated interest in receiving text messages for nutrition, exercise, sleep, and stress management [15]. These trends represent a changing societal landscape, and the health care field is poised to address this identified need. A recent interactive mobile solution for appointment reminders within the Department of Veterans Affairs (VEText) has been used to send SMS text message reminders to over 6 million veterans [16,17]. Text messages can also provide links to resources and reminders toward adopting healthier behaviors. Several meta-analyses corroborate the effectiveness of SMS for medication adherence [18-20]. An earlier study by the authors [21] measured the impact of SMS text reminders on refill rates of nonadherent and partially adherent Medicare patients with chronic disease. They found that text reminders increased refill rates by 14 percentage points compared with those who did not receive these reminders.

Important to the success of any intervention is its implementation, scalability, and sustainability [22]. Text messaging presents an effective, affordable, and scalable tool [21] that can use conversational AI to greatly impact health outcomes. More specifically, conversational AI (or conversational agents) can encourage health care consumers to engage with systems that imitate human conversations using text [23,24]. For example, a fully automated conversational AI system has been used to promote weight loss among overweight and obese diabetic patients [25]. As conversational agents can learn over time, interventions with thousands of users can be used to inform and improve the quality of the conversations, often within days or weeks. However, there is limited research currently available on the use of conversational AI within SMS (and not app-based) messaging for refill adherence.

### Social Determinants of Health

The World Health Organization has defined social determinants of health (SDOH) as the conditions or circumstances in which people are born, grow, live, work, and age [26]. The most commonly identified SDOH in the United States are housing, income, food, transportation, education, race/ethnicity, and unemployment [27,28]. Despite notable improvements in overall

health over the past few decades, inequalities of SDOH contribute to persistent disparities in life expectancy and health outcomes [29,30].

Social determinants have also been linked to nonadherence, and a recent study using data from the National Health Interview Survey found that half of the adults with diabetes perceived financial stress, while one-fifth reported financial insecurity and food insecurity [30]. Since SDOH are not always recognized, they might be overlooked by a clinician in a medical setting. In a study of patients with chronic disease, two-thirds of those who reported not taking medications as prescribed due to cost never shared this with their physician [30,31]. The National Academy of Medicine has published a framework for educating health care professionals on the importance of SDOH [32], but there is still limited research studying the specific ways in which SDOH pathways interact to impact adherence, particularly in older populations.

### Objectives

A pilot study examining Medicare Part D patients over a period of 3 months supported the value of using SMS text message refill reminders to increase medication refill rates [21]. This study increases the sample size of those receiving SMS refill reminders to 99,217 Medicare recipients, includes both English and Spanish language text messages (the pilot used only English messages), and expands the duration to a 2-year follow-up period. This analysis focuses on a few different questions:

- First, were the results from the pilot study replicable with a much larger population using an enhanced version of the text messaging solution with improved conversational AI?
- Second, is there a relationship between SDOH and refill request rates, how large is this association, and do SDOH attenuate the association?
- Third, how do other variations in patient characteristics (age, gender, ethnicity, and language) moderate and predict

likelihood of requesting a refill using a text message solution?

## Methods

### Participants

The SMS refill reminder program began as a 3-month pilot in December 2016 and was expanded to include multiple regions within a large integrated health system. The analysis covers a 2-year period from December 2016 to February 2019. It includes a population of 99,217 English- and Spanish-speaking patients (median age 71 years) targeted for medication refill by Kaiser Permanente, Southern California. The Kaiser Permanente, Southern California, Institutional Review Board determined that this program did not require review and was exempt.

All patients had Medicare Part D as their pharmacy benefits and had one or more chronic conditions (diabetes, hypertension, high cholesterol, and/or anticoagulation). Patients in this program refilled one or more of the following 4 classes of drugs: oral diabetes medications, blood pressure medications (renin-angiotensin system antagonists), statins, and direct oral anticoagulants (DOAC).

Targeted patients were shared by Kaiser Permanente, Southern California, in a weekly file, and they had varying levels of nonadherence. The total number of patients targeted for refill ranged from 1000 to 9000 patients per month. Note that these patients were not distinct each month because they could be on the list several times in a year. Patient records included first name, date of birth (DOB), gender, spoken language, address, race/ethnicity, mobile phone number, opt-in status, and refill drug(s). These fields, when available, were used for all analysis.

Tables 1 and 2 provide age and race/ethnicity breakdowns for this group.

**Table 1.** Age of text messaging group.

Age band (years)	Patients, n (%)	Reminders, n (%)
<60	6344 (6.39)	20,883 (7.64)
60-65	4502 (4.54)	13,806 (5.05)
65-70	29,600 (29.83)	79,349 (29.03)
70-75	27,201 (27.42)	75,902 (27.77)
75-80	15,039 (15.16)	42,221 (15.44)
80-85	7899 (7.96)	22,247 (8.14)
>85	5458 (5.50)	14,385 (5.26)
Unspecified	3174 (3.20)	4563 (1.67)
Total	99,217 (100)	273,356 (100)

**Table 2.** Race/ethnicity of text messaging group.

Race/ethnicity	Patients, n (%)	Reminders, n (%)
White	30,683 (30.93)	81,544 (29.83)
Hispanic/Latino	21,841 (22.01)	67,266 (24.60)
Black/African American	9124 (9.20)	28,365 (10.38)
Asian	8705 (8.77)	23,870 (8.73)
Other/mixed	1372 (1.38)	3812 (1.39)
Unspecified/unknown	27,492 (28.71)	68,499 (25.06)
Total	99,217 (100)	273,356 (100)

## Procedure

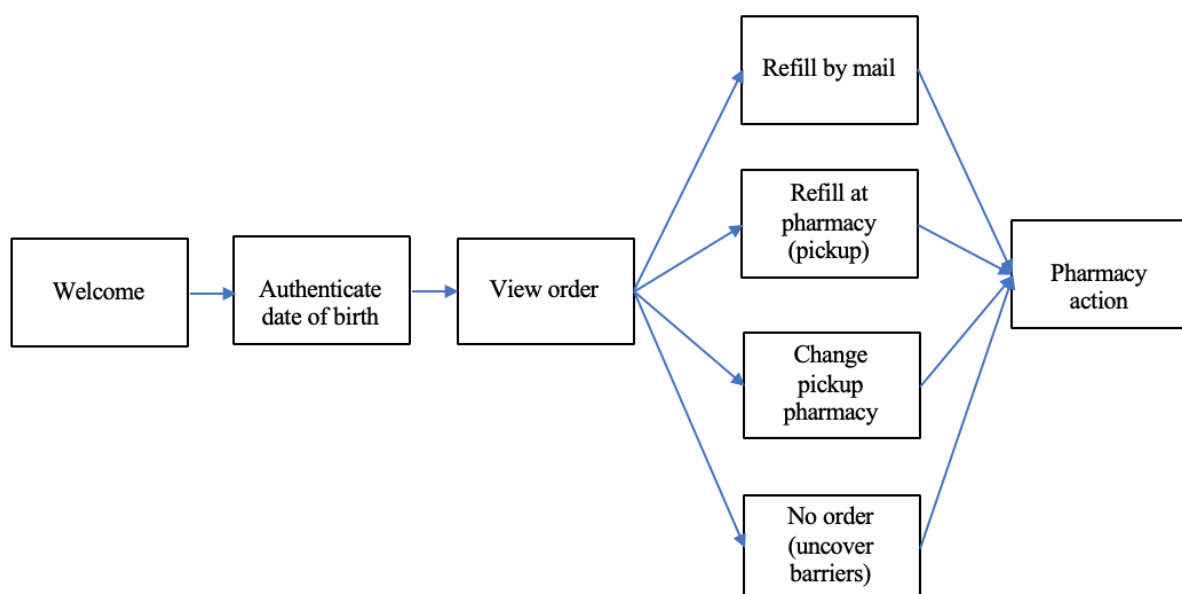
The intervention used the mPulse Mobile platform to deliver SMS text messages to patients. Patients in the text messaging group received a refill reminder dialogue that consisted of a series of messages written at a sixth-grade readability level.

Text message refill reminders were sent out on a weekly basis to patients who were due for a refill. The first message was a greeting, reminding patients that they were due for a refill. They were then prompted to validate their DOB (to ensure the person was the intended recipient of the reminder) by choosing from one of five options. If the patient validated their DOB successfully, they could then view their medication(s) due for refill and the last filling pharmacy. Patients did not get a second chance if they did not validate correctly. In those cases, a member from the pharmacy staff would reach out to discuss barriers to nonadherence and/or complete the refill by phone. As part of the refill workflow, patients could select whether to receive their medication by mail (this was added as an option

in May 2018) or to pick up their medication(s) at a Kaiser Permanente outpatient pharmacy. They could also change their default pickup pharmacy. Pharmacy staff who were located at a central location were involved with managing the responses that came back from the patients. These patient responses were categorized as “Refill request,” “Barriers,” “Date of birth issue,” “Free text response,” “Side effects,” “Change Pharmacy,” and “Help.”

As in the pilot study [21], we used mPulse Mobile’s Engagement Console to support the pharmacy staff. This Web-based user interface allows users to quickly identify and prioritize subgroups of patients (as described above) for quicker follow-up.

Figure 1 provides a view of the refill reminder message flow and the various steps or options within this dialogue. The initial step reminds the patients to refill the medication(s) and requests that they respond with the *structured options* “1” to continue or “2” to end. However, in some instances, patients respond with *unstructured responses* such as “I’ve already refilled” or “No thanks.”

**Figure 1.** Overview of message flow within refill dialogue.

A patient would receive a maximum of 3 messages if they did not respond (initial reminder, 2-hour reminder, and 24-hour reminder). The patient could opt out at any point during the messaging flow, and no subsequent messages would be sent.

Conversational AI was developed and used within the solution to also automatically process the following types of unstructured responses and patient requests:

- Patient wants to unsubscribe from texting program but does not reply “STOP” or 7867 (patients with older phones could use number keys to correspond to letters).
- Patient confirms intent to request a refill by using an unexpected phrase.
- Patient is experiencing side effects and might require medical attention.
- Patient wants to change pharmacy location where they wish to pick up the refill.
- Patient does not want to refill and provides a reason for not refilling.
- Patient requests help or wants additional information.
- Patient provides correct DOB instead of selecting a numeric option.
- Patient wants to switch language (English to Spanish or vice-versa).

For purposes of the qualitative analysis, each response was strictly coded as structured or unstructured. A response was considered *structured* if it exactly matched any of the following strings: 0, 1, 2, 3, 4, 5, 6, 7867, and case insensitive versions of AYUDA, HELP, MAIL, RESUB, STOP, and STOPALL). All other responses were coded as unstructured.

Two coders classified each conversational AI rule using the description within the solution (eg, DOB validation and change order). There were a total of 75 conversational AI rules. These individual rules were then combined into 13 broader “rule type” categories: *initial reply, refill process, change pharmacy, DOB validation, barriers, payment, subscription, language change, member information, acknowledgment, feedback, help, and did not understand*. All rules belonged to the 13 rule types, and any ambiguity about rule type were resolved by discussion and agreement between the two coders.

An internally developed SDOH index was used to understand how unmet needs might impact patient refill behavior. When patient address was available, an SDOH index was computed. [Multimedia Appendix 1](#) outlines the factors that were used to compute the SDOH index and to create low, medium, and high SDOH clusters.

A neural network multilayer perceptron (MLP) model was used to perform predictive analysis on factors that might impact refill requests.

## Results

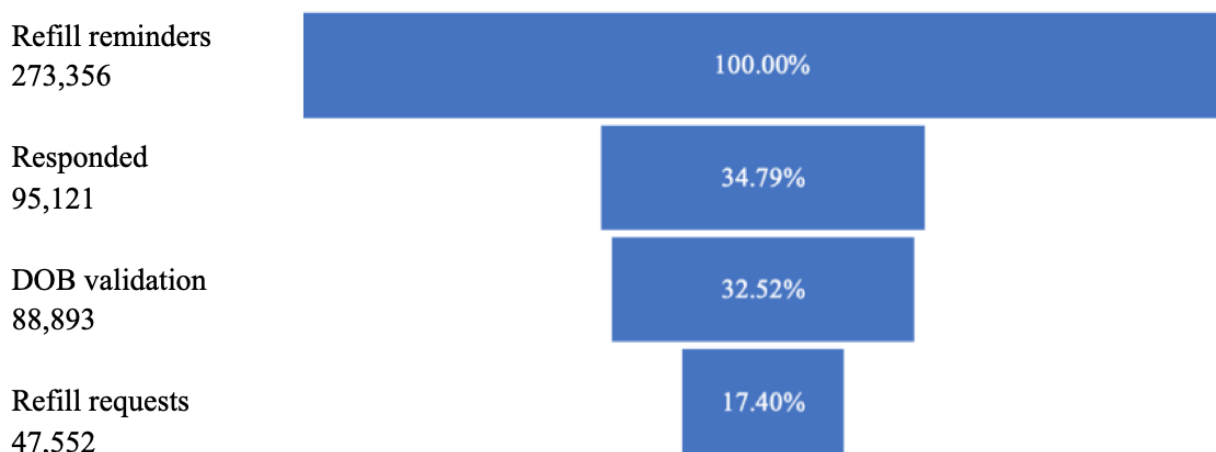
The results of the scaled intervention are summarized in 3 parts: (1) a replication of the analysis performed in the pilot study; (2) results of subgroup exploratory analysis, including the use of an SDOH index and a predictive model; and (3) a qualitative analysis of the use and value of conversational AI and interactivity.

### Part 1: Analysis of Scaled Program

A total of 273,356 SMS reminders were sent over a 2-year period to 99,217 Medicare Part D patients who had opted in to texting. In response, 17.40% (47,552/273,356) refills were requested. [Figure 2](#) shows the conversion funnel from reminder to refill request.

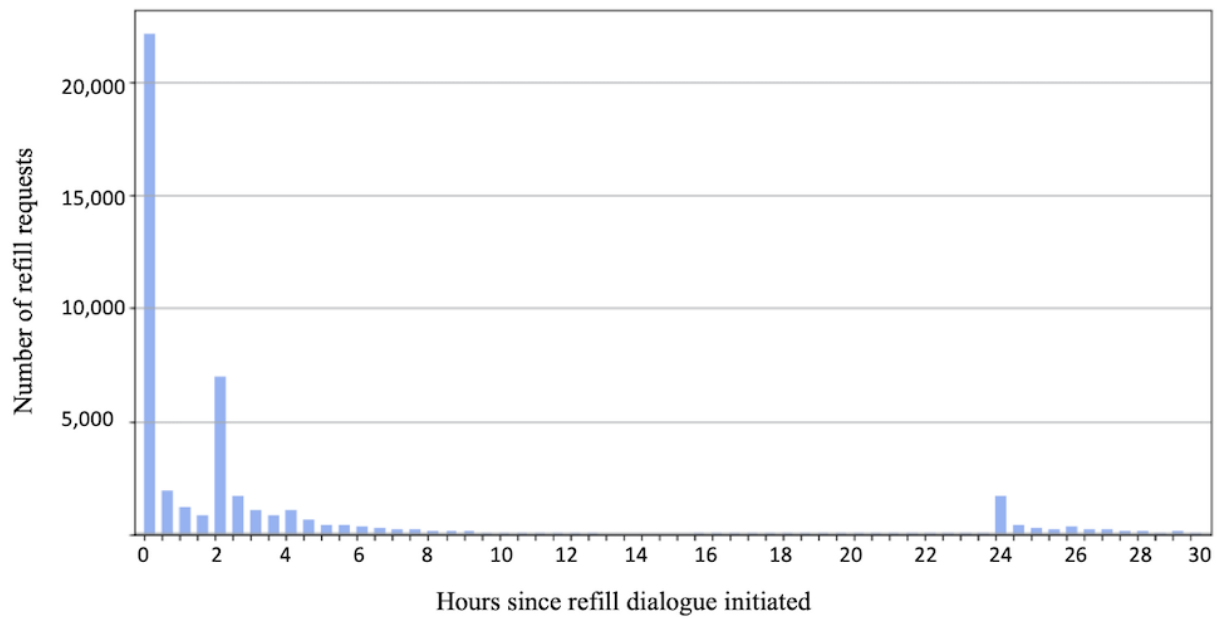
DOB validation was a necessary step to view the refill information, and this step resulted in a drop-off (DOB validation failures or did not attempt) of 6.55% (6288/95,121). Of those who requested a refill, 54.81% did so within 2 hours after receiving the initial reminder (N=26,062/47,552). As displayed in [Figure 3](#), there are spikes in refill activity immediately after the initial message (“0”), after the 2-hour reminder (“2”), and the 24-hour reminder (“24”).

**Figure 2.** Conversion of refill reminder to refill request. DOB: date of birth.





**Figure 3.** Refills requests by hour from initial reminder.



**Part 2: Exploratory Subgroup Analysis**

We present the following subgroup analysis using 4 variables: SDOH, language, race/ethnicity, and age (gender did not have a significant moderating effect on refill rates).

**Social Determinants of Health Analysis**

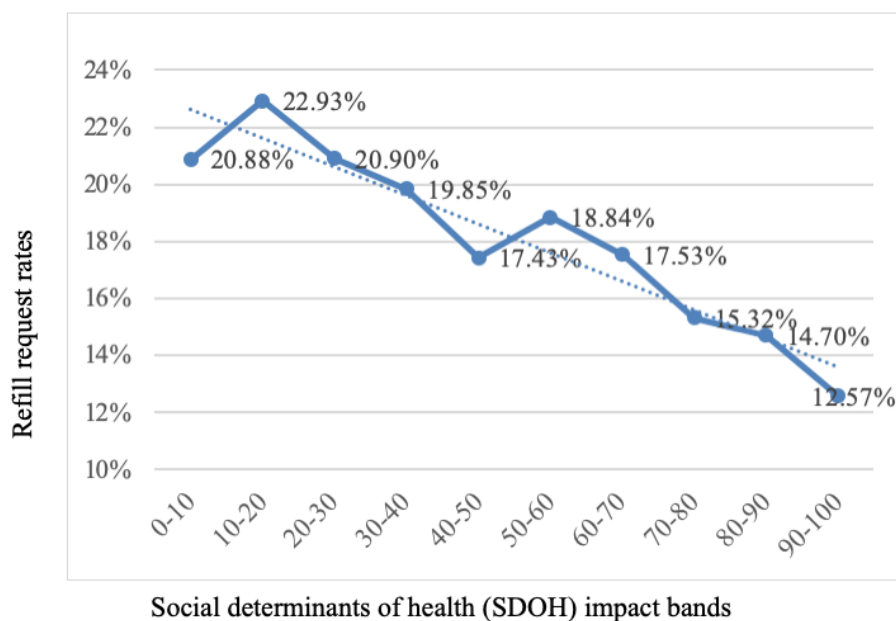
Patients were grouped into 10 evenly spaced SDOH bands from 0 to 100. Refill requests were very highly inversely correlated with SDOH bands ( $r=-0.93$ ), as shown in Figure 4.

To further understand the impact of SDOH on refill process, we grouped the SDOH bands further into 3 SDOH clusters

(high, medium, and low) using k-means clustering as described further in Multimedia Appendix 1.

As can be seen in Table 3 and Figure 5, the negative correlation of refill request rates to SDOH index is driven primarily by the initial response rates (ie, after receiving the “Welcome” message in Figure 1, the patient confirms their intent to move forward in the dialogue). The difference in average SDOH between those who reply and do not reply was statistically significant ( $t_{252,834}=-55.07$ ;  $P<.001$ ), but there was no impact of SDOH for refill requests after the patient engages with the initial text message ( $t_{87,234}=1.71$ ;  $P=.09$ ). In other words, once a patient is willing to engage with the texting program, they request refills at the same rate regardless of SDOH levels.

**Figure 4.** Refill rates versus social determinants of health bands.



**Table 3.** Refill request rate for text message group by social determinants of health level.

Social determinants of health	Refill dialogues, N	Percent who responded, n (%)	Percent of responders who requested refill, n (%)
Low (0-52.8)	124,423	47, 949 (38.54)	23,623 (49.27)
Medium (52.8-86.1)	94,489	30,755 (32.55)	15,464 (50.28)
High (86.1-100)	33,924	8532 (25.15)	4357 (51.07)

**Figure 5.** Social determinants of health impact on response rates and percentage of responders who request refill. SDOH: social determinants of health.



### Spanish Versus English

Spanish-speaking patients had significantly lower refill request rates (5149/48,156, 10.69%) compared with English-speaking patients (42,389/225,060, 18.83%;  $X^2_1, [n=273,216]=1829.2; P<.001$ ). As with SDOH impact, the initial response rates also vary by spoken language, where Spanish-speaking patients (8984/48,156, 18.66%) were far less likely to engage with text messaging than English-speaking patients (86,105/225,060, 38.26%;  $X^2_1, [n=225,060]=6716.9; P<.001$ ).

Interestingly, a difference in refill request rates by language after the patient engaged with the reminder shows that Spanish-speaking patients request refills at a higher rate compared with English-speaking patients once they engage (5149/8984, 57.31% vs 42,386/86,105, 49.22%;  $X^2_1 [n=95,089]=212.5; P<.001$ ).

We used a pointwise biserial correlation (point biserial correlation  $r_{pb}=0.27; P<.001; N=252,696$ ) to find that higher SDOH values are correlated with Spanish language preference. Spanish speakers ( $N=44,869$ ) had a higher average SDOH of 71.78 as compared with English speakers' ( $N=207,827$ ) average of 55.65 ( $t=151.39; P<.001$ ).

### Age

Younger patients were significantly more likely to reply and request refills compared with older patients ( $t_{83,415}=-43.30; P<.001$ ). The older age group (75 years and older) responded at a rate of 29.84%, whereas patients younger than 45 years responded at a rate of 47.81%. There were also significantly different rates of refill requests by age band ( $X^2_7 [n=268,793]=1460.3; P<.001$ ), with younger patients requesting refills at a higher rate, as shown in Table 4. We do see a spike in refill requests in the 85+ years group, and this suggests that caregivers or family members might be more actively assisting patients in this age band.

**Table 4.** Response and refill request rates by age.

Age band (years)	Refill dialogues, N	Responded, n	Date of birth validation, n	Refills requested, n	Request rate, %
<60	20,883	9388	8929	5277	25.27
60-65	13,806	5444	5061	2808	20.34
65-70	79,349	29,625	27,787	15,014	18.92
70-75	75,902	25,707	23,949	12,458	16.41
75-80	42,221	127,732	11,781	6103	14.45
80-85	22,247	6256	5780	3039	13.66
>85	14,385	4538	4242	2351	16.34
Unspecified	4563	1431	1364	501	10.98
Total	273,356	95,121	88,893	47,552	17.40

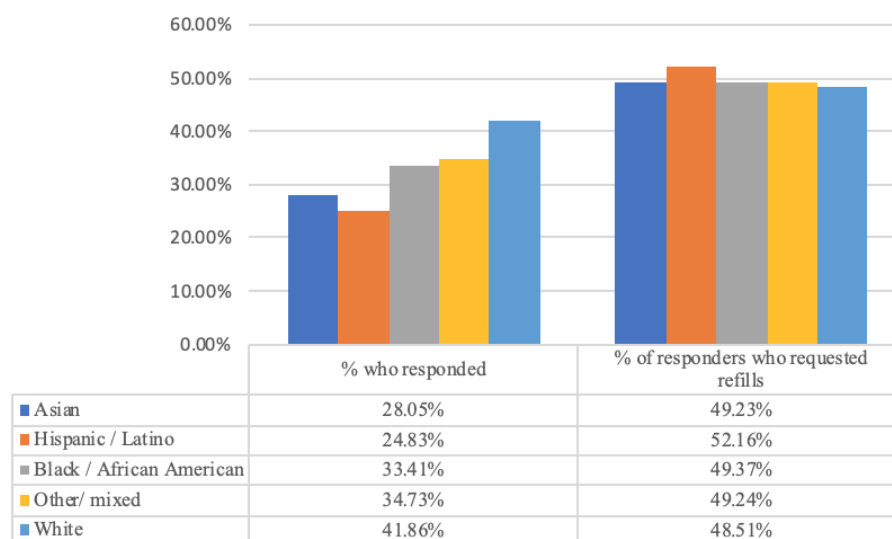
### Does Race/Ethnicity Have an Influence on Engagement and Refill Request Rates?

Patient race/ethnicity had a significant effect on initial response rates ( $X^2_4 [n=204,857]=5282.40$ ;  $P<.001$ ). Patients who identified as white responded at the highest rate (34,134/81,544,

41.86%), whereas Hispanic/Latino patients had the lowest response rate (16,700/67,266, 24.83%). Once someone did respond, race/ethnicity did not influence whether they refilled or not ( $X^2_4 [n=68,329]=2.37$ ;  $P=.07$ ), as reflected in Table 5 and Figure 6.

**Table 5.** Response and refill request rates by race/ethnicity.

Race/ethnicity	Refill dialogues, N	Responded, n	Date of birth validation, n	Refills requested, n	Request rate, %
Asian	23,870	6695	6233	3296	13.81
Hispanic/Latino	67,266	16,700	15,430	8711	12.95
Black/African American	28,365	9476	8751	4678	16.50
Other/mixed	3812	1324	1216	652	17.10
White	81,544	34,134	32,181	16,557	20.30

**Figure 6.** Impact of race/ethnicity on response rates and percentage of responders who request refill.

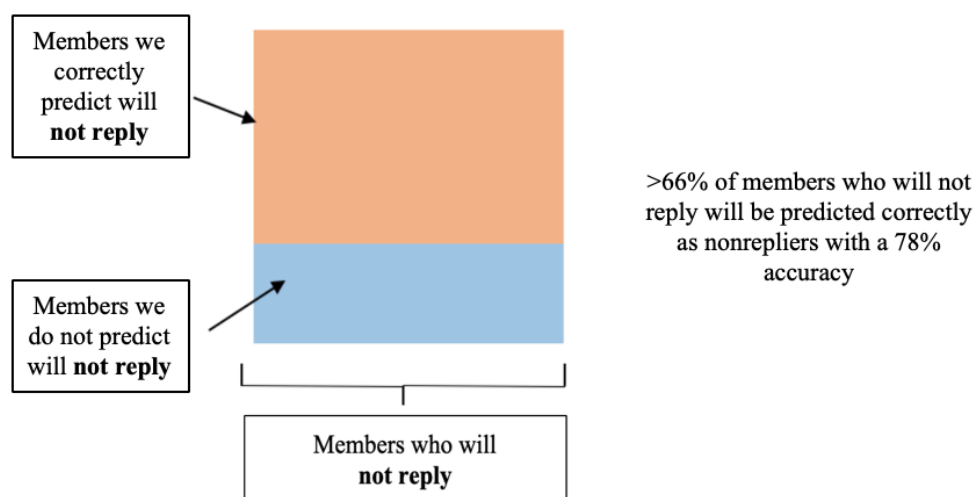
### Predictive Model to Improve Refill Adherence

An in-depth analysis of the results revealed that the *reply rate* was the most important variable that we could influence to drive refill rates. After replying to the reminder, about half of the

patients continue the conversation to validate their DOB and then request a refill of their medication(s), while the other half drop off and do not engage further.

As we were uncertain of the interaction of age, gender, SDOH, language, and race/ethnicity and how they moderate reply rates, we built a machine learning model using these factors as inputs to predict those least likely to reply. We trained a supervised model to predict reply likelihood based on known attributes. The model consisted of a neural network multilayer perceptron (MLP) with 24 input units, 10 hidden units, and a single output unit (to represent reply likelihood). The training set consisted of 70% of the available data, and model validation was done with the remaining 30%. The input vector consisted of features representing age, race/ethnicity, gender, SDOH, language, drug class, etc. We selected the most discriminative features to avoid overfitting and to reduce multicollinearity and redundancy in the feature space, and we used one-hot encoding to ensure uniform scale across features. Finally, we excluded rows with incomplete values. Each data row represented an instance of a unique series of events after a reminder was sent to a patient. This meant that if a member had been contacted 3 times to refill a drug and they only replied twice, the “reply” value would now be 0.66. This solved the problem of contradictory data and also converted the reply variable into a continuous variable representing reply likelihood.

**Figure 7.** Minimizing the number of false positives.



### Part 3: Qualitative Analysis of the Usefulness of Conversational Artificial Intelligence

Our solution incorporated conversational AI (CAI) and natural language understanding (NLU) to provide a robust and successful interactive experience that ensures that the patient is able to request their refill as quickly and conveniently as possible.

We performed a qualitative analysis of all patient responses to evaluate whether the more complex capabilities of CAI and NLU were necessary and helpful for a better patient experience. As part of this analysis, we coded 307,484 responses as either structured or unstructured (as described in the Methods section).

Of the 307,484 responses that we received during the study period, only 7.77% (n=23,886) were not understood by the CAI. Table 6 provides a breakdown of structured and unstructured

### Optimizing the Model to Predict Those Least Likely to Engage and Request Refill

To develop an approach to impact refill adherence, we wanted to first maximize the prediction accuracy for patients who were not likely to reply at all. We were less concerned about prediction accuracy for patients who were likely to engage and request refills. For the model to have value in an applied setting, we wanted to capture as many patients who were not likely to engage and might require additional support to request a refill. Figure 7 contains a visual representation of the confusion matrix, which highlights only those predictions that relate to nonreplying members.

To address these dual objectives (accurately predicting those who require outreach while maximizing the number of people who require outreach), we found a cut-off of 37% predicted likelihood of replying as a good threshold. This means that a patient whose predicted likelihood is less than 37% should fall into a category of requiring additional outreach. In summary, we can identify over 66% of those requiring outreach (because they will not reply) at a model accuracy of 78%.

responses as well as the steps in the refill reminder dialogue where they typically occurred. All unstructured responses shown in Table 6 were understood by the CAI engine, which triggered the appropriate replies.

There were several instances when asked to provide structured response (eg, text 1 to view the medication), a member replies with additional information (eg, “1 – I only took this medication for 1 day and it caused great muscle pain. In the meantime, my cholesterol levels are now below 200”). Similarly, when asked to validate DOB by choosing from 1 of the 3 options, members will choose a number but also input the DOB as in “1 - 1950-05-30.” The system was largely successful in recognizing responses and accurately categorizing them. In Table 7, we present additional examples where the CAI was successful.

Textbox 1 includes a few sample responses that we failed to understand but could have handled appropriately if the patient

had responded when expected or provided more context. In summary, our results indicate that despite the overall high accuracy (over 92%) of handling patient responses by the CAI, there are a number of instances (almost 8%) that were not handled at all, and in these cases, the patient was informed that the system was unable to understand their message.

**Table 6.** Counts of the type of responses that were handled by the conversational artificial intelligence (CAI).

Role of CAI	Structured responses	Unstructured responses	Total, n (%)
Supported by CAI	272,364	11,234	284,598 (92.23)
CAI could not handle	3001	20,885	23,886 (7.77)
Total	275,365 (89.55)	32,119 (10.45)	307,484 (100)

**Table 7.** Examples of unstructured patient messages when the conversational artificial intelligence successfully understood the message.

Sample patient responses	Response category
"I-I only took this medication for 1 day and it caused great muscle pain. In the meantime my cholesterol levels are now below 200"	Side effect barrier
"1/2 tablet twice or thrice weekly. Changed on 6/25/18 due to recurring muscle pain and Dr XXXX concurred"	Side effect barrier
"Both kinds of pain. I can't walk very far if i take too many so I've cut it to 1/4th and then when it geta to bad i stop it for a few days."	Side effect barrier
"Caused pain on calvrs so bad i could not walk"	Side effect barrier
"Doctor took me off medication because of too much muscle pain"	Side effect barrier
"I cant take a Statin It gives me terrible muscle pain and I cant sleep !!!! No."	Side effect barrier
"I have reported to my previous primary and his medical assistant...that the dosage prescribed gave me leg cramps and pain...I agreed to take half a pi"	Side effect barrier
"Do not have the money right now to fill it will refill it a when I have the funds"	Cost barrier
"Don't have money right now"	Cost barrier
"Don't have the extra money till the first."	Cost barrier
"Don't have the money yet to refill but plan on refilling soon"	Cost barrier
"Filling px on the military base at no cost"	Cost barrier
"1 (one)"	"1"
"1 (sent with Invisible Ink)"	"1"
"1 proceed to refill"	"1"
"1 refill"	"1"
"1-"	"1"

**Textbox 1.** Sample of unstructured patient messages when the conversational AI responded to patients that it did not understand their message.

Sample patient responses
"0 - This has already been ordered last weekend!"
"2 cuando lo recojo?"
"2 my body can't tolerate this drug"
"A refill is not needed until the end of march."
"Already ordered online"
"Cancel rx change"
"Doctor has lowered dosage from prescribed amount. Doctor needs to update prescription."
"I am in San Diego now"
"I am out of the country"
"I need talk to Sam body"
"I want to talk to my doctor first."
"Random"
"Still have full bottle of 60. Will order later."
"What medication??"

## Discussion

### Principal Findings

We found that the refill request rates reported in the pilot study with English speakers [21] could be replicated at scale. The results in this study confirm and replicate the pilot results and reveal that, even after significant expansion of the population (scaled to over 7 times the initial group size and with the addition of Spanish speakers), the solution was effective in moving patients through the text dialogue to quickly complete a refill request. Pilot refill request rates of 18.1% were closely mirrored (17.40%), or even higher when limited to requests by English speakers (18.83%). Expanding the capabilities to include Spanish speakers was an important feature that is not commonly available in other texting solutions.

On the basis of the results in the previous section, we found that patients are not equally likely to request a refill. The findings suggest that good refill adherence is linked to language, race/ethnicity, age, and SDOH and that English speakers, white patients, those younger than 75 years, and those with lower SDOH barriers have significantly higher odds of requesting a refill via SMS. We studied these associations and developed a tool and approach that can be used for future outreach to narrow identified gaps based on demographic and socioeconomic factors and to increase overall refill adherence.

Finally, we wanted to evaluate the impact of the conversational AI engine, which is not typically found in other health care texting solutions. The results indicate that patients needed conversational AI as they traversed the refill reminder dialogue, and as health care consumers become more comfortable and familiar with artificial intelligence-based agents and chatbots, this expectation will only increase.

To our knowledge, no other published study (other than the study by Brar Prayaga et al [21]) has reported a refill adherence

solution with results at scale [20]. The closest comparable solution with a large volume of patients is the VEText system for appointment reminders offered by the Department of Veterans Affairs, and we look forward to an in-depth study of that solution.

### Understanding Who Engages and Requests a Medication Refill

As our results indicate, we have identified several associations between factors that impact a patient's likelihood to reply to the reminder and continue on to request a refill via SMS text messaging. These include language, SDOH, age, and race/ethnicity. For example, Spanish speakers were much less likely to engage with the reminder at all and, therefore, requested refills at a significantly lower rate of 10.69%, compared with English speakers at 18.83%. Note that for purposes of this text message solution, we messaged all patients in English (including those with a preferred language of Tagalog, Mandarin, etc in their health record) unless they had a preferred language of Spanish. Patients who identified as Hispanic/Latino had the lowest refill request rates of 12.95%, followed by Asian (13.81%), black (16.50%), and white (20.30%). Interestingly, neighborhood-based SDOH levels were highly correlated with patient language (English/Spanish), and patients with high SDOH barriers had significantly lower refill request rates of 12.84%, compared with medium SDOH (16.37%) and low SDOH levels (18.99%), as shown in Figure 5. Finally, younger patients were significantly more likely to reply and request refills compared with older patients. We believe these associations and effects should be addressed and have developed a predictive tool to help improve overall refill adherence rates for Medicare D patients.

### Using a Predictive Model to Assign Resources

A key goal is to increase the initial reply rate, and we were most interested in uncovering the population (roughly two-thirds) who were *unlikely to engage at all*. We found that the predictive

model is able to accurately (>78%) pinpoint a high percentage (>66%) of patients who will not engage with an SMS text reminder to request a refill. This can be used as a valuable tool by a health provider or pharmacy to proactively communicate with populations who are least likely to complete a refill request. While current outreach methods to encourage adherence, such as phone calls by pharmacy staff or automated interactive voice response calls, are typically more costly and time consuming [21,33], a targeted approach using the predictive model could optimize limited staff resources.

### Addressing Barriers to Improve Refill Adherence

Another recommendation is to reduce the *initial barriers* or unwillingness to engage with the program. For example, lack of familiarity with texting or mistrust of the channel, especially among older patients, could be addressed with more tailored versions of the initial reminder that alleviate possible unease with using SMS text messaging to complete a refill transaction.

Similarly, health plans and providers could provide supplemental resources (such as an informational video to explain the process for requesting a refill via text message and to address any specific concerns of Spanish speakers), and the text dialogue could link to these resources. It might also be beneficial to add multilanguage support to expand beyond English and Spanish. In addition, as cost can be a barrier to refill for patients, especially for those with high SDOH levels, the text message dialogue could remind patients that, for instance, mail order refills are incentivized (eg, “did you know that if you order by mail you can get a 3-month supply of medication for the cost of a 2-month supply?”).

We also found that for a large percentage of refill reminders (15.75%, n=14,005), patients start the refill process (ie, view medication and validate their DOB) and then share a barrier or other concern that causes them to drop out of the process. There were an additional 10,192 instances where patients shared a reason for not refilling in other parts of the conversational flow. These barriers include a wide range of topics such as cost, side effects, already refilled, still have sufficient medication, do not want to refill, and taking differently than prescribed. We are sharing this response data with Kaiser Permanente, Southern California, and they are continuing to find ways to address these issues and follow-up with patients as required.

### The Role of Conversational Artificial Intelligence

Conversational AI was helpful in moving patients through the refill dialogue using mostly structured inputs almost 90% of the time. Although complex conversational AI was not essential for driving outcomes (validating DOB, completing refill, etc), it played an important role in keeping the conversation going when patients engaged using unprompted or unstructured messages (“Caused pain on calvrs so bad I could not walk [sic]”). In this example, it is recognizing the patient’s expressed concern as a *side effect*, thereby allowing the member to respond as they see fit to continue the conversation instead of being constrained by a rigid structure. This model of primarily relying on prompted responses and also understanding those cases where users want to go outside a closed response set allows the user to control the flow of the conversation in keeping with Grice’s

Maxim of Manner [34]. Press releases relating to the recent launch of VEText [16] suggest a significant impact on appointment no-show rates at scale (there is no peer-reviewed study on the solution currently available, but the press release reported a no-show reduction from 13.7% to 11.7%, N not reported). The system requires users to only use prompted and structured alphanumeric responses, such as R2 and K3 [17]. While the details of the VEText solution are unclear, support for unstructured responses (estimating 10% based on our results) using conversational AI would likely improve user experience as well as overall outcomes. We believe this hybrid approach of supporting dual modes of interaction will support a fluid and frictionless conversation to enable task completion and allows a more empathetic exchange with the health care partner. Finally, while our accuracy rate of conversational AI was over 92%, we continue to explore ways to improve the system. A review of the literature using conversational AI reveals that the focus area is limited in scale and tends to be restricted to apps [24] and not SMS.

It is unclear why there is limited adoption of conversational AI within SMS text messaging as this is a channel with potential to reach all segments of the population. This study is unique in analyzing a text messaging and conversational AI solution at scale that allows elderly populations to easily and conveniently request their medication refills.

### Limitations

The findings of this study have to be seen in light of some limitations. The study was not a randomized controlled trial, and there is the possibility of selection bias. Due to regulations within the wireless communication industry and the Telephone Consumer Protection Act, we must have prior consent before messaging patients, and this constraint applies to any automated text messaging solution. As a result, the study targeted only those patients who had already opted into digital engagement. As the messaging program requires that patients have a mobile number with a texting plan, patients with only landline numbers were excluded. The nontext group received phone call outreach as part of the standard of care. However, we have already demonstrated the value and incremental benefit of SMS text reminders as compared with phone reminders [21], and this was not a focus of the study.

Using bivariate analyses can inflate the type 1 error, and this is a limitation. At the same time, we attempted to address this issue by using a neural network predictive model, which is a form of multivariate regression and can reduce the impact of multicollinearity. In addition, we did not address the impact of multiple reminders over time—that is, does a patient who requests a refill after receiving a reminder, and later becomes nonadherent again, continue this positive behavior upon receiving a future reminder?

Finally, the solution described in this study is a commercial system offered by mPulse Mobile with a licensing fee. We are unable to share financial benefits of using the system, such as operational efficiencies, health savings, and reimbursement revenue from improved adherence, as this information is proprietary and confidential. At the same time, we cannot disclose solution costs (such as messaging costs, license fees,

and implementation costs). However, the overall financial gains from using the refill reminder solution were greater than system costs.

## Conclusions

Overall, this study indicates that there are sharp differences in likelihood to reply to a refill reminder and complete a refill request via SMS based on demographic and socioeconomic factors. We found a strong association between refill request rates and patient language, age, race/ethnicity, and SDOH levels, and these differences may contribute to health disparities and

impact health outcomes in Medicare patients. Using a predictive and innovative model to target patients least likely to engage with the SMS solution and crafting a tailored mobile communication and conversational AI strategy could reduce these inequalities and improve refill adherence. We will continue to refine our solution and optimize our predictive model to validate our results and hope to be able to address disparities and drive even stronger outcomes. Finally, we believe that, to ensure the success of a text messaging solution and yield similar results, message tone and content, ease of use, level of tailoring, and quality of conversational AI are important considerations.

## Acknowledgments

The authors would like to thank the following individuals for their contributions to this project: Joyce Lin for data translation; Bhumi Gupta for support with conversational AI; Emily Haag for review of member responses; Kayvon Moradi, Michael Steinmetz, and Heather Forst for overall project support; and Carlton Segbefia for help with code optimization and analysis.

## Conflicts of Interest

RBP, AP, and RSP are current employees of mPulse Mobile, Inc, which is a vendor for Kaiser Permanente. RA and BN contributed to this study as part of their internship at mPulse Mobile, Inc.

## Multimedia Appendix 1

Calculating the Social Determinants of Health Index.

[[PDF File \(Adobe PDF File\), 93 KB - mhealth\\_v7i11e15771\\_app1.pdf](#)]

## References

1. Hugtenburg J, Timmers L, Elders PJ, Vervloet M, van Dijk L. Definitions, variants, and causes of nonadherence with medication: a challenge for tailored interventions. *Patient Prefer Adherence* 2013;7:675-682 [[FREE Full text](#)] [doi: [10.2147/PPA.S29549](https://doi.org/10.2147/PPA.S29549)] [Medline: [23874088](https://pubmed.ncbi.nlm.nih.gov/23874088/)]
2. de Geest S, Sabaté E. *Adherence to Long-term Therapies: Evidence for Action*. Geneva, Switzerland: World Health Organization; 2003.
3. Modi AC, Quittner AL. Barriers to treatment adherence for children with cystic fibrosis and asthma: what gets in the way? *J Pediatr Psychol* 2006 Sep;31(8):846-858. [doi: [10.1093/jpepsy/jsj096](https://doi.org/10.1093/jpepsy/jsj096)] [Medline: [16401680](https://pubmed.ncbi.nlm.nih.gov/16401680/)]
4. DiMatteo MR, Haskard-Zolnierok KB, Martin LR. Improving patient adherence: a three-factor model to guide practice. *Health Psychol Rev* 2012;6(1):74-91 [[FREE Full text](#)] [doi: [10.1080/17437199.2010.537592](https://doi.org/10.1080/17437199.2010.537592)]
5. Centers for Medicare & Medicaid Services. 2018. Medicare-Medicaid Plan Performance data Technical Notes. URL: [https://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/FinancialAlignmentInitiative/Downloads/MMPPerformanceDataTechNotesCY2018\\_04252018.pdf](https://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/FinancialAlignmentInitiative/Downloads/MMPPerformanceDataTechNotesCY2018_04252018.pdf) [accessed 2019-06-13]
6. Torres-Robles A, Wiecek E, Cutler R, Drake B, Benrimoj SI, Fernandez-Llimos F, et al. Using dispensing data to evaluate adherence implementation rates in community pharmacy. *Front Pharmacol* 2019;10:130 [[FREE Full text](#)] [doi: [10.3389/fphar.2019.00130](https://doi.org/10.3389/fphar.2019.00130)] [Medline: [30863308](https://pubmed.ncbi.nlm.nih.gov/30863308/)]
7. Brown MT, Bussell JK. Medication adherence: WHO cares? *Mayo Clin Proc* 2011 Apr;86(4):304-314 [[FREE Full text](#)] [doi: [10.4065/mcp.2010.0575](https://doi.org/10.4065/mcp.2010.0575)] [Medline: [21389250](https://pubmed.ncbi.nlm.nih.gov/21389250/)]
8. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. *BMJ Open* 2018 Jan 21;8(1):e016982 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2017-016982](https://doi.org/10.1136/bmjopen-2017-016982)] [Medline: [29358417](https://pubmed.ncbi.nlm.nih.gov/29358417/)]
9. Viswanathan M, Golin CE, Jones CD, Ashok M, Blalock SJ, Wines RC, et al. Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic review. *Ann Intern Med* 2012 Dec 4;157(11):785-795. [doi: [10.7326/0003-4819-157-11-201212040-00538](https://doi.org/10.7326/0003-4819-157-11-201212040-00538)] [Medline: [22964778](https://pubmed.ncbi.nlm.nih.gov/22964778/)]
10. Iuga AO, McGuire MJ. Adherence and health care costs. *Risk Manag Healthc Policy* 2014;7:35-44 [[FREE Full text](#)] [doi: [10.2147/RMHP.S19801](https://doi.org/10.2147/RMHP.S19801)] [Medline: [24591853](https://pubmed.ncbi.nlm.nih.gov/24591853/)]
11. Williams A, Kenyon K. eHealth Initiative. 2017. Issue brief: Medication adherence and health IT. URL: [https://www.ehidc.org/sites/default/files/resources/files/medicationadherence\\_and\\_hit\\_issue\\_brief\\_ONC.pdf](https://www.ehidc.org/sites/default/files/resources/files/medicationadherence_and_hit_issue_brief_ONC.pdf) [accessed 2019-06-13]
12. Carr-Lopez SM, Shek A, Lastimoso J, Patel R, Woelfel J, Galal S, et al. Medication adherence behaviors of Medicare beneficiaries. *Patient Prefer Adherence* 2014;8:1277-1284 [[FREE Full text](#)] [doi: [10.2147/PPA.S64825](https://doi.org/10.2147/PPA.S64825)] [Medline: [25258521](https://pubmed.ncbi.nlm.nih.gov/25258521/)]



13. Centers for Medicare & Medicaid Services. 2018 Jul 12. CMS Proposes Historic Changes to Modernize Medicare and Restore the Doctor-Patient Relationship. URL: <https://www.cms.gov/newsroom/press-releases/cms-proposes-historic-changes-modernize-medicare-and-restore-doctor-patient-relationship> [accessed 2019-06-13]
14. Safavi K, Webb K, Kalis B. Accenture. 2019. Accenture 2019 Digital Health Consumer Survey. URL: <https://www.accenture.com/acnmedia/PDF-94/Accenture-2019-Digital-Health-Consumer-Survey.pdf> [accessed 2019-06-13]
15. Betts D, Korend L. Deloitte. 2018. Inside the Patient Journey: Three Key Touch Points for Consumer Engagement Strategies. URL: <https://www2.deloitte.com/insights/us/en/industry/health-care/patient-engagement-health-care-consumer-survey.html> [accessed 2019-06-13]
16. MDedge. 2019. VEText 1 Year Later - Still Growing. URL: <https://www.mdedge.com/fedprac/article/196293/health-policy/vetext-1-year-later-still-growing> [accessed 2019-10-07]
17. Veterans Affairs. 2019. VEText FAQs. URL: [https://www.va.gov/HEALTH/VEText\\_FAQs.asp](https://www.va.gov/HEALTH/VEText_FAQs.asp) [accessed 2019-10-07]
18. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. *JAMA Intern Med* 2016 Mar;176(3):340-349. [doi: [10.1001/jamainternmed.2015.7667](https://doi.org/10.1001/jamainternmed.2015.7667)] [Medline: [26831740](https://pubmed.ncbi.nlm.nih.gov/26831740/)]
19. Vervloet M, Linn AJ, van Weert JC, de Bakker DH, Bouvy ML, van Dijk L. The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. *J Am Med Inform Assoc* 2012;19(5):696-704 [FREE Full text] [doi: [10.1136/amiainjnl-2011-000748](https://doi.org/10.1136/amiainjnl-2011-000748)] [Medline: [22534082](https://pubmed.ncbi.nlm.nih.gov/22534082/)]
20. Schwebel FJ, Larimer ME. Using text message reminders in health care services: a narrative literature review. *Internet Interv* 2018 Sep;13:82-104 [FREE Full text] [doi: [10.1016/j.invent.2018.06.002](https://doi.org/10.1016/j.invent.2018.06.002)] [Medline: [30206523](https://pubmed.ncbi.nlm.nih.gov/30206523/)]
21. Prayaga RB, Jeong EW, Feger E, Noble HK, Kmiec M, Prayaga RS. Improving refill adherence in medicare patients with tailored and interactive mobile text messaging: pilot study. *JMIR Mhealth Uhealth* 2018 Jan 30;6(1):e30 [FREE Full text] [doi: [10.2196/mhealth.8930](https://doi.org/10.2196/mhealth.8930)] [Medline: [29382623](https://pubmed.ncbi.nlm.nih.gov/29382623/)]
22. Zullig LL, Deschodt M, Liska J, Bosworth HB, de Geest S. Moving from the trial to the real world: improving medication adherence using insights of implementation science. *Annu Rev Pharmacol Toxicol* 2019 Jan 6;59:423-445. [doi: [10.1146/annurev-pharmtox-010818-021348](https://doi.org/10.1146/annurev-pharmtox-010818-021348)] [Medline: [30125127](https://pubmed.ncbi.nlm.nih.gov/30125127/)]
23. Laranjo L, Dunn AG, Tong H, Kocaballi AK, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc* 2018 Sep 1;25(9):1248-1258 [FREE Full text] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
24. Bibault J, Chaix B, Nectoux P, Pienkowsky A, Guillemasse A, Brouard B. Healthcare ex Machina: Are conversational agents ready for prime time in oncology? *Clin Transl Radiat Oncol* 2019 May;16:55-59 [FREE Full text] [doi: [10.1016/j.ctro.2019.04.002](https://doi.org/10.1016/j.ctro.2019.04.002)] [Medline: [31008379](https://pubmed.ncbi.nlm.nih.gov/31008379/)]
25. Stein N, Brooks K. A fully automated conversational artificial intelligence for weight loss: longitudinal observational study among overweight and obese adults. *JMIR Diabetes* 2017 Nov 1;2(2):e28 [FREE Full text] [doi: [10.2196/diabetes.8590](https://doi.org/10.2196/diabetes.8590)] [Medline: [30291087](https://pubmed.ncbi.nlm.nih.gov/30291087/)]
26. Commission on Social Determinants of Health. World Health Organization. 2008. Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health. URL: [https://www.who.int/social\\_determinants/final\\_report/csdh\\_finalreport\\_2008.pdf](https://www.who.int/social_determinants/final_report/csdh_finalreport_2008.pdf) [accessed 2019-07-09]
27. Office of Disease Prevention and Health Promotion. Healthy People 2020. Social Determinants of Health. URL: <http://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health> [accessed 2019-07-05]
28. Artiga S, Hinton E. Kaiser Family Foundation. 2018 May 10. Beyond Health Care: The Role of Social Determinants in Promoting Health and Health Equity. URL: <http://www.kff.org/disparities-policy/issue-brief/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/> [accessed 2019-07-05]
29. Singh GK, Daus GP, Allender M, Ramey C, Martin E, Perry C, et al. Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935-2016. *Int J MCH AIDS* 2017;6(2):139-164 [FREE Full text] [doi: [10.21106/ijma.236](https://doi.org/10.21106/ijma.236)] [Medline: [29367890](https://pubmed.ncbi.nlm.nih.gov/29367890/)]
30. Patel MR, Piette JD, Resnicow K, Kowalski-Dobson T, Heisler M. Social determinants of health, cost-related nonadherence, and cost-reducing behaviors among adults with diabetes: findings from the national health interview survey. *Med Care* 2016 Aug;54(8):796-803 [FREE Full text] [doi: [10.1097/MLR.0000000000000565](https://doi.org/10.1097/MLR.0000000000000565)] [Medline: [27219636](https://pubmed.ncbi.nlm.nih.gov/27219636/)]
31. Piette JD, Heisler M, Wagner TH. Cost-related medication underuse among chronically ill adults: the treatments people forgo, how often, and who is at risk. *Am J Public Health* 2004 Oct;94(10):1782-1787. [doi: [10.2105/ajph.94.10.1782](https://doi.org/10.2105/ajph.94.10.1782)] [Medline: [15451750](https://pubmed.ncbi.nlm.nih.gov/15451750/)]
32. National Academies of Sciences, Engineering, and Medicine. A Framework to Address the Social Determinants of Health. Washington, DC: The National Academies Press; 2016.
33. Reidel K, Tamblyn R, Patel V, Huang A. Pilot study of an interactive voice response system to improve medication refill compliance. *BMC Med Inform Decis Mak* 2008 Oct 9;8:46 [FREE Full text] [doi: [10.1186/1472-6947-8-46](https://doi.org/10.1186/1472-6947-8-46)] [Medline: [18845004](https://pubmed.ncbi.nlm.nih.gov/18845004/)]
34. Grice HP. Logic and conversation. In: Cole P, Morgan JL, editors. *Speech Acts*. New York: Academic Press; 1975:41-48.

## Abbreviations

**AI:** artificial intelligence  
**CAI:** conversational artificial intelligence  
**CMS:** Centers for Medicare and Medicaid Services  
**DOB:** date of birth  
**MLP:** multilayer perceptron  
**NLU:** natural language understanding  
**SDOH:** social determinants of health

*Edited by G Eysenbach; submitted 05.08.19; peer-reviewed by B Smith, K Davison, J Cornelius, S Brabyn, S Badawy, L McGoron; comments to author 27.08.19; revised version received 05.09.19; accepted 22.10.19; published 18.11.19.*

*Please cite as:*

*Brar Prayaga R, Agrawal R, Nguyen B, Jeong EW, Noble HK, Paster A, Prayaga RS  
Impact of Social Determinants of Health and Demographics on Refill Requests by Medicare Patients Using a Conversational Artificial Intelligence Text Messaging Solution: Cross-Sectional Study  
JMIR Mhealth Uhealth 2019;7(11):e15771  
URL: <http://mhealth.jmir.org/2019/11/e15771/>  
doi: [10.2196/15771](https://doi.org/10.2196/15771)  
PMID: [31738170](https://pubmed.ncbi.nlm.nih.gov/31738170/)*

©Rena Brar Prayaga, Ridhika Agrawal, Benjamin Nguyen, Erwin W Jeong, Harmony K Noble, Andrew Paster, Ram S Prayaga. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 18.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Development of MyTeen Text Messaging Program to Support Parents of Adolescents: Qualitative Study

Joanna Ting Wai Chu<sup>1</sup>, PhD; Angela Wadham<sup>1</sup>, BA; Yannan Jiang<sup>1</sup>, PhD; Robyn Whittaker<sup>1</sup>, PhD; Karolina Stasiak<sup>2</sup>, PhD; Matthew Shepherd<sup>3</sup>, PhD; Christopher Bullen<sup>1</sup>, MD

<sup>1</sup>National Institute for Health Innovation, University of Auckland, Auckland, New Zealand

<sup>2</sup>Department of Psychological Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

<sup>3</sup>School of Psychology, Massey University, Auckland, New Zealand

**Corresponding Author:**

Joanna Ting Wai Chu, PhD

National Institute for Health Innovation, University of Auckland

261 Morrin Road, Glen Innes

Auckland, 1072

New Zealand

Phone: 64 3737599

Email: [jt.chu@auckland.ac.nz](mailto:jt.chu@auckland.ac.nz)

## Abstract

**Background:** Parents play an important role in the lives of adolescents, and supporting and addressing the needs of families continue to be the focus of many researchers and policy makers. Mobile health interventions have great potential for supporting parents at a population level because of their broad reach and convenience. However, limited evidence exists for such interventions for parents of adolescents. This study reports on the formative work conducted with parents and/or primary caregivers to identify their needs and preferences for the development of MyTeen—an SMS text messaging program on promoting parental competence and mental health literacy for parents of adolescents (aged 10-15 years).

**Objective:** The aim of this qualitative study was to explore parents and/or primary caregivers' perspectives around youth well-being, parenting, and parenting support and their input on the development of MyTeen SMS text messaging parenting intervention.

**Methods:** A total of 5 focus groups (n=45) were conducted with parents or primary caregivers of adolescents aged 10 to 15 years between October and December 2017 in New Zealand. A semistructured interview guideline and prompts were used. Data were audiotaped, transcribed, and analyzed using inductive thematic analysis.

**Results:** Participants were concerned about youth mental health (ie, stigma and increasing demand on adolescents), and a number of parenting challenges (ie, social expectations, time, impact of technology, changes in family communication pattern, and recognizing and talking about mental health issues) were noted. Importantly, participants reported the lack of services and support available for families, and many were not aware of services for parents themselves. A number of recommendations were given on the style, content, and frequency of developing the text messaging program.

**Conclusions:** Findings from this qualitative work informed the development of MyTeen, an SMS text messaging program designed to increase parental competence and improve mental health literacy for parents of adolescents.

(*JMIR Mhealth Uhealth* 2019;7(11):e15664) doi:[10.2196/15664](https://doi.org/10.2196/15664)

**KEYWORDS**

programs; mHealth; adolescents; parents; text messaging

## Introduction

Depressive disorder is a major health issue among adolescents [1]. In New Zealand, prevalence rates are around 4% to 8% at 15 years, increasing to 17% to 18% by 18 years [2]. The serious developmental consequences of adolescent depression and

associated treatment challenges once problems have developed underscore the need for programs aimed at prevention [3]. Parents play a significant role in shaping processes associated with adolescent well-being [4,5]. Relatedly, parenting programs aimed at strengthening parenting skills and increasing knowledge on adolescent development have led to positive

effects on parent-adolescent relationships, parent well-being, and adolescent well-being [6,7].

There have been increasing calls to adopt a public health approach to parenting support [8,9]. The aim of the approach is to enhance parenting practices, competence, and adjustment for all parents and thus produce multiple beneficial health and developmental outcomes for young people at the population level [9,10]. However, poor participation by parents stands as the greatest barrier to widespread effective implementation of such programs [9,11]. This implies that a large segment of the population is failing to receive the benefits of such programs [8]. Many parents of adolescents are left without the support and information on the normative changes and parenting issues of adolescence that they need. Studies conducted with a diverse population suggest that all parents report a desire to learn about parenthood and that many families would welcome additional support on the task of parenting [12-15].

Mobile health (mHealth) interventions have great potential for supporting parents at a population level because of their broad reach and convenience [16]. mHealth offers a wide range of potential benefits over traditional approaches such as programs that can be delivered anywhere at any time; programs are more proactive (initiated by the service) than traditional services; they are flexible and can be personalized and tailored to specific cultural, age group, and health needs; reach is increased because the barriers of face-to-face contact (such as time, cost, and travel) are removed; and disparities in access across the socioeconomic status gradient are decreased because of the high penetration of mobile phones across these groups [13]. Specifically, mobile-based messaging programs may be particularly useful because of their simplicity and minimal burden for participants. Evidence suggest that SMS text messaging programs are effective in promoting parenting behavior change for parents of young children (eg, decreasing the likelihood of abuse and neglect, increasing childhood vaccinations, and encouraging healthy pregnancies) [17-21]. The interventions were well received by parents of various population, including those that are socially deprived [18,21].

Although there is evidence for SMS text messaging programs for parents with young children, its application among adolescent parent populations are unknown. There is considerable potential in New Zealand to supplement and enhance existing services for families by leveraging mHealth technologies in delivering programs for parents. A total of 92% of New Zealand households have access to a mobile phone with no differences in internet access or mobile phone ownership by ethnicity or education and few differences by age below 65 years. As such, we developed a program of text messages adapted from the Parenting Strategies Program, a set of evidence-based parenting guidelines developed through (1) a systematic review and meta-analysis of parental factors associated with adolescent depression and/or anxiety and (2) a Delphi study of international expert consensus about actionable strategies parents can use to reduce their child's risk of depression and anxiety. The guideline is Web-based and freely available [22]. Of the 13 domains, 9 (helping your child to deal with problems, helping your child to deal with anxiety, minimizing conflict with your child, encouraging supportive

relationships, encouraging good health habits, supporting your child's autonomy, avoiding criticizing your child, establishing family rules and consequences, and improving your relationship with your child) were used to guide the development of the text messages. These domains were chosen as it has been found to be particularly useful for parents of adolescents. Key information from each domain were edited to a maximum of 160 characters (maximum length of text). Moreover, 1 or 2 sample text messages were derived from each domain.

We report on formative research conducted with parents of teenagers (aged 10-15 years) to seek their perspectives around parenting adolescents and parenting support. Specifically, we were interested in parents' views of their needs in supporting their adolescent's well-being, the extent to which they were aware of support and services available to them, and their input for the development of a mobile-based SMS text messaging parenting intervention.

## Methods

### Overview

This study is part of a wider project that aims to develop and evaluate the effectiveness of MyTeen, an SMS text messaging intervention program that promotes parental competence and mental health literacy for parents of adolescents. The protocol of the randomized controlled trial is published elsewhere [23]. To inform the development of the SMS text messaging program, 5 focus groups (n=45) were conducted from October to December 2017.

### Recruitment

Efforts (ie, recruitment via youth rugby sports club, cultural organizations, word of mouth via personal networks, and individuals were also encouraged to forward information about the intervention to others who might be interested) were made to recruit a diverse sample, including Māori (indigenous people of New Zealand) and Pacific parents, who are considered as hard to reach. To maximize recruitment effort, 1 of the research assistant who identified as Māori actively engaged with ethnic minorities via her own network and promoted the visibility of the study within the Māori community. Promotion materials directed interested participants to contact a member of the project team to find out more about participation. Eligibility for the study was ascertained at that time. Individuals were eligible to participate if they were a parent/primary caregiver of a teenager aged 10 to 15 years, able to speak English, and able to attend a focus group. Eligible participants then received information about the study, and those who agreed to participate were offered a time and place to join a focus group.

### Ethics Approval

Ethics approval was obtained from the University of Auckland Human Participant Ethics Committee (UAHPEC, Ref 019659). Participation was voluntary. Written consent was collected, and a questionnaire on sociodemographic characteristics was completed before the start of the focus group.

## Data Collection

Each focus group lasted between 60 and 90 min and was conducted by 2 trained moderators experienced in conducting focus groups. Group size ranged from 3 to 15 participants. Sessions were audio-recorded with the consent of the participants. At the beginning of each focus group, participants were given information on the overall aim of the study including the development and trialing of the SMS text messaging program. A topic guide and prompts were developed based on a standardized questioning technique to cover a range of key issues related to the research questions. A total of 5 questions were asked: what are your perceptions about mental health and well-being among youth?, what are your perceptions on parenting teens?, how accessible is mental health support at school or in the local community?, what resources or support would you like to see for youth and families in your community?, and what are your thoughts around an SMS-based mobile intervention for supporting parents?. In addition, a list of sample text messages from the SMS text messaging program was shown to parents to comment and provide feedback. The focus groups were conducted with flexibility to allow unanticipated themes to emerge. Refreshments were provided, and each participant received a NZD \$50 voucher in appreciation of his or her time and transportation cost.

## Data Analysis

Transcripts were imported into NVivo (Version 11.0; QSR International) to allow for electronic coding and retrieval of data. The data were analyzed using thematic analysis [24]. This method was deemed appropriate for summarizing the data and identifying patterns in the data to provide an interpretation. The 6 steps outlined by Braun and Clarke [24] were followed. The first author and an experienced research assistant familiarized themselves with the data by thoroughly reading, rereading, and annotating the material with preliminary ideas. Initial themes were generated based on inferences from the data. The data were then categorized into these themes. Subthemes were developed to more clearly capture the data.

As a mean of assessing credibility of the coded themes and to minimize bias, 2 researchers performed the coding analyses. A small number of differences were found and were resolved through discussion. A third researcher reviewed the transcripts to further ensure that the themes identified reflected the data. Field notes were also constantly referred to as supplementary materials. Peer debriefing was also used and involved ongoing discussion between researchers during the course of focus groups. Finally, informal member checking occurred during data collection when the moderator reflected back on her understanding of participant responses and sought feedback regarding whether this understanding was accurate.

## Results

### Participants

Of the 45 participants, 27 were mothers, 7 were fathers, 3 were grandparents, and 8 identified as others (siblings and relatives). Over half (24/45, 53%) of the participants were identified as married, 29 (13/45) as single, and 18% (8/45) as divorced.

Participants were mostly Māori (3/45, 62%), 16% (7/45) were European, 13% were Pacific (6/45), and 9% (4/45) identified as others. Majority (32/45, 71%) of the participants reported completing at least high school education. Several themes emerged from the focus groups and are reported below.

### Themes

#### Concerns Around Youth Mental Health

Participants were concerned about the high rates of youth depression and suicide in New Zealand. Sadly, many of our participants knew of a young person (directly or indirectly) that had either experienced depression, self-harm, or suicide:

*When you have a look at the subject (suicide) itself, talk about mental health, there wouldn't be anyone that hasn't been touched. It's just something that, you know, the stats back it up. It's terrible.*

#### Stigma Around Youth Mental Health

The stigma attached to youth mental health was a prominent theme. Participants discussed the terminology itself as problematic and come with a negative connotation:

*When you look at the way kids use the word mental, you know, it's someone that's done something really stupid, you're mental, you're deranged or whatever. It's got a stigmatism that comes with it.*

Participants also reported having experienced stigma associated with asking for help as it was seen as a sign of weakness or an indication that their family was not coping. Participants acknowledged that for a very long time, people did not talk openly about mental health issues and for a number of reasons such as concerns over being judged or treated differently, burden to the family, and denial:

*I think the biggest problem is that people are afraid to admit it.*

#### Increasing Demands on Adolescents

Participants felt that many of today's young people are experiencing high levels of stress, stemming from high expectations for academic achievements, jobs, social media, and peer pressure:

*Just seems to me the kids are under a lot of stress, a lot of pressure, social media stuff, bullying, social media bullying. Peer pressure.*

*I think there's a lot of pressures these days, more for academic success than in our days. The pressure to have a job and stuff as well. I mean its money, it's everything. The society today is just one big pressure for them all.*

#### Increasing Challenges for Parents

Parents expressed that at times they are overwhelmed by the expectations and social pressure to be a "good parent" and raise competent adolescents. Time and juggling multiple demands throughout the day were frequently mentioned as a parenting challenge:

*I think society's expectations on what your child should be and what it shouldn't be and that's really hard to juggle and challenge.*

*Most parents I think are guilty of this, is making sure that I actually allocate proper quality time but work and life takes over and you get so caught up in everything else.*

Technology contributed to the challenges parents experienced, which were not present in previous generations. For example, parents spoke about cyberbullying, problematic screen use, and adolescent access to internet content. Parents noted that these issues all had an impact on adolescent well-being:

*Social media has a huge impact (on their wellbeing).*

*They can just Google any word. It will bring up porn, it will bring up conspiracy theories, it will bring up anything and it just impacts their view straight away.*

### **Changes in Family Communication Pattern**

All participants agreed that family communication pattern has changed significantly. Participants felt that the increased use of technology devices created a communication barrier within the family. The quality of communication has also changed, with adolescents using text or social media instead of face-to-face conversations with parents:

*I find it frustrating that they, they just can't communicate anymore, like they can't speak to a person. It's all done with phones or messages, no one seems to be able to pick up a phone or confront somebody directly.*

*I had my granddaughter around home with her friends, and they are all sitting there, head down. And then suddenly they'll all laugh at the same time. What was the point of you all being in here?*

### **Challenge With Recognizing and Talking About Mental Health**

A number of our participants noted that they did not know how to best respond to and assist their adolescents if they were experiencing problems. For some, the challenge was recognizing the symptoms:

*The thing for me is I know a few people that have committed suicide, and the people that have been in contact with them that day, or even the week leading up to it, they seem to be the happiest they've been in a long time. They often set a date that they're going to do it, and they're happy because they're going to be out of whatever misery. If the signs were obvious then you will be able to tackle it, but it's not.*

Participants felt that it was difficult to talk with their adolescents on mental health issues and expressed their fears about what and how much to say. Others felt uncertain or incompetent about what information to give and where to direct their teens for support:

*That's what I said to my wife, coz we said we need to talk about this, and the first thing I said was does she even know what suicide is, or what that means when*

*a person does that? But we had to, you know, and that was my fear, are we giving them information of an option, which was a fear.*

Participants also noted that their adolescents did not want to talk to them about mental health issues:

*Even once you notice that something is up with your child, the challenge then is to open that channel of communication with them. You do get that (from them), "No, I am alright, there's nothing wrong, I am okay."*

*They'll tell you there's nothing wrong with them, and yet what I've found, my moko (grandchild) committed suicide the other day. She was 10, these are things that we should know about.*

### **Lack of Mental Health Support From Schools**

When asked about the level of support in schools and in their community, many of our participants were dissatisfied. In particular, concerns were expressed about the quality of school-based counseling services. All participants agreed that the resources at schools were insufficient, including the student-counselor ratio, lack of cultural sensitivity, poor parent-teacher communication, and the barriers associated with accessing school-based services. In New Zealand, school deciles are a measure of the socioeconomic position of a school's student community relative to other schools throughout the country. Lower decile rating indicates higher proportion of students from low socioeconomic communities:

*My school are decile one and two so we get nothing.*

*My son, his school's a decile nine and there's not even a social worker or anything there.*

One parent spoke about her experience and disappointment with the school:

*She was going downhill in all of her mahi (work) and all of that. And the teacher gave me a leaving form, he didn't offer any kind of support. This was the dean, so no, (school services) not the greatest.*

Another parent described the concerns around confidentiality and the stigma associated with one's reputation when seen walking into a counselor's office at school:

*I couldn't take my daughter to the school because if they got called up to the counsellor at the school she uses a particular coloured piece of paper. So all the kids would know where you're going, like they can tell by whatever slip they're given. So there was no way she would speak to a counsellor through the school.*

### **Lack of Mental Health Services for Adolescents**

Apart from schools, parents were only able to name 1 or 2 services that they knew of for supporting youth. These included Child Youth Services, WINZ, and Youthline. A few participants who knew about it from personal experience were concerned that these services were only directed to those already in crisis and stressed the frustration in having to wait too long before a professional contact could be established. Only 2 of our

participants knew about other supportive services (eg, Te Kawa te Rangatahi); these participants knew about this as they worked in social services:

*I had a problem with my teenagers and what I did was I went to organisation after organisation until I got help from them. I just kept going and going until I got help.*

*When it comes to this age group, the teenager, there's not many (services). There's mental services but they are too young to get into a lot of the services. There's Youthline, but yeah. So Dr Google, GP, and schools.*

### **Lack of Services for Parents**

Even less was known about services available to parents. Participants were unable to identify any formal parenting support services beyond their family doctor, Plunket, and Well Child Tamariki Ora, with the latter 2 being Ministry of Health services provided to New Zealand families and are directed at parents of young children (aged 0-5 years). Parents indicated that they do not know where to access information, ask questions, or seek advice or formal support when needed. Parents often turned to family, friends, and the internet for advice:

*I'll be honest, I wouldn't know where to go if my kid had any issues. And I think a lot of it is just left up to us (as parents) these days, because the systems just aren't there for us.*

### **Recommendation for Youth and Parent Services**

Overall, participants expressed a need for more accessible services for both adolescent and parents:

*Get the information out there, just flood with information about feelings and emotions and the normality of them.*

*Education, we need to know more about everything, like you know teen emotions, mental health.*

Participants noted that it was important for adolescents themselves to recognize the mental health issue and seek support when needed:

*How are we to support them when they don't even understand themselves what mental health is and what support? It's really them that should be creating their own supports in schools and communities. But then again, there's a lot of schools that wouldn't allow it (suicide) to be talked about. So how are we supposed to educate our tamariki (children) if the schools can't even educate them properly on mental health, and what that means.*

Participants suggested several support services for adolescents. This ranged from technology-based programs (eg, apps, SMS text messaging, and Web-based) designed specifically for adolescents to incorporating mental health education within schools, increasing school counseling resources (eg, full-time dedicated trained staff that are culturally sensitive and after school drop in), and public mental health promotion campaigns. Nonetheless, many acknowledged the challenges with their suggestions (eg, lack of funding and resources):

*Our (suicide) rate being very high, they should try and have better counselling/support services at all schools. Going to secondary school, there's not much help. I think there should be a compulsory thing happening at secondary schools, because that's when most of our teens are taking their lives.*

*The Education Board's not going to pay too much for more social workers in the schools. It comes down to money in the end.*

Parents also spoke about support that they wanted as parents. This includes easily accessible information; parenting programs that were brief and not time consuming, or that they can do at their own pace; and parenting apps.

### **Text Messaging Parenting Program**

Following discussion on accessing and increasing support, participants were reminded about the wider project to develop and evaluate the SMS text messaging program that promotes parental competence and mental health literacy for parents of adolescents. Participants were given sample text messages to review, and their input on the style, content, and frequency of the intervention were sought. Overall, parents were enthusiastic about receiving parenting information via text messages. Participants noted that it was a valid approach to engagement, as mobile phone usage is widespread among older generations (eg, grandparents). They also thought that the text messages could be shared with other family members and serve as a catalyst to engage in conversations.

All participants agreed that the intervention would be best delivered in English, and other languages were optional. Participants were aware that the program was proposed to be brief and preventative; hence, personally tailored messages did not matter for many of our participants. They were also aware of the wording limit (160 characters) in text messages, and many of the preferred to receive simple, jargon-free, nonjudgemental, and direct information. A majority of participants mentioned that the text messages should be fun to read, including the use of emoji and images:

*Keep it simple with the words. Don't make them too long, coz then you switch off.*

*Use emoji, sometimes it's easier to just send a heart, than to actually type I love you.*

Provision of practical parenting tips and strategies, knowledge about adolescent development, symptoms of mental health problems, and where to access information were viewed as important content to include in the program. This was consistent across focus groups:

*Where to find information, where to get help.*

*Brain development and stuff, if it's normal, the behaviour that they're doing because of their brain development.*

Parents also wanted messages on self-care and acknowledgment of their parenting effort. One parent communicated the importance of positivity. This sentiment was echoed across focus groups, with participants frequently mentioning the need for encouraging messaging. Participants provided sample

wordings such as “you’re not alone,” “you are doing a good job as a parent,” and “just breathe.” Many parents wanted the program to emphasize that they were not alone and support is available when needed:

*Sometimes you just need to see a little ray of sunshine.*

*Supportive little notes for parents.*

A small number of parents mentioned the importance of including spirituality and cultural relevance in the text messages:

*Reminding us the connection with the land, with our forest, or with our sea. Brings you so much peace.*

There was no general consensus about how often the text messages should be delivered and for how long. Some participants preferred a daily text message, whereas others thought 2 to 3 messages per week was sufficient. Similarly, some participants thought the text messages should be ongoing beyond the intended 1 month; others felt that the duration was sufficient or that it depended on the individual receiving the program.

## Discussion

### Principal Findings

In this study, we examined parents’ perspectives on youth well-being, parenting, and parenting support and their input on the development of MyTeen text messaging parenting program. There are several findings from our focus groups that are particularly noteworthy.

The need to support parents of adolescents was apparent. Many of the challenges (eg, social pressure to be a “good parent,” lack of time, the impact of technology, and parent-adolescent communication) reported are universal and consistent with other studies [25-27]. New technology, in particular, pushes boundaries of parenting and creates new challenges for today’s parents [28,29]. For example, the internet creates informational freedom and changes the nature of social interactions among individuals [30]. Several studies have reported negative association between internet use and the quality of parent-adolescent relationships. Although outside the scope of our study, more research is needed to examine parenting with respect to adolescents’ use of technology and its impact on family dynamics. For example, a child’s greater familiarity and proficiency with technology could undermine parental authority and potentially adversely affect parenting and parent-child relationships [31].

It was evident that many of our participants lacked competence on initiating discussion about mental health with their child. Parents are one of the most influential sources by which adolescents learn to label, identify, and interpret emotions. Parents also act as an important change agent in assisting young

people to access help when needed [32]. Equipping parents with knowledge and tools to help them foster emotional competence and identify mental health problems in young people, raising awareness and understanding of the professional help that is available, is particularly salient in preventing and mitigating adolescent mental health problems. Increasing parents’ self-efficacy can further play a role in buffering their adolescents from mental health problems [33-35].

There is a lack of available services for parents and adolescents. Barriers related to not knowing where or how to seek help was particularly salient, suggesting the need for strategies to raise awareness of available services. Majority of services to support youth mental health focus on the adolescent themselves, with scant knowledge of, or attention to, the unique role and needs of their parents [36]. Parenting interventions can promote parents’ self-efficacy and mental well-being and improve knowledge and skills to support their adolescent well-being [6].

Parents had many suggestions for services directed at teenagers, including technology-based programs (eg, apps, SMS text messaging, and Web-based), mental health education within schools, school counseling resources, and public mental health promotion campaigns. Importantly, the idea of an SMS text messaging program to support them as parents was well received. Time was constantly stated as a barrier to accessing parenting support; hence, text messages that are proactive (initiated by the service) and do not require attendance by the participant were appealing for our participants. Although SMS text messaging may not be considered a “novel” mobile phone app it remains the most widely used [37]. Parents offered suggestions for the SMS text messaging program including the message tone, content and length, as well as the frequency of text messages delivery. As noted previously, parents perceived challenges with adolescent use of technology; however, technology has become an integral part of everyday life, including the parents themselves. Much more research is needed to understand how we can balance the benefits and negative impact of technology in families.

On the basis of the qualitative findings, we refined the sample text messages derived from the Parenting Strategies Program and developed additional text messages based on parents’ needs around self-care and providing resources for help seeking when needed. Some of the domains (eg, helping child to deal with anxiety, avoid criticizing your child, and establish family rules) were omitted as (1) it appeared to be of lesser importance to parents than other domains and (2) the intervention was short in duration; hence, not all domains can be included. Positive tone, emoji, and jargon-free wordings were all taken into consideration when designing our program. Sample text messages from MyTeen are shown in [Table 1](#).



**Table 1.** Sample text messages from MyTeen.

Domains	Sample text message
Establish and maintain a good relationship with your teenager	Show your teen love and affection, talk with them and have fun with them! By doing so you're making a huge contribution to their emotional wellbeing.
Be involved and support increasing autonomy	Spend quality time together. Don't get caught up in to-do lists & screens. Have a meal, take a walk, play a game, or make a list of things to do together.
Minimize conflict in the home	Minimise conflict in the home and remember you are the role model for your teen! Stay calm & encourage them to find solutions to problems/conflict.
Encourage good health habits	Sleep is really important for your teen! Encourage them to switch off/reduce the time they spend on their phone or devices a few hours before bedtime.
Knowledge and risk of depression	Depression = feelings of sadness & irritability lasting longer than 2 weeks, affect daily life and stop people from taking part in things they used to enjoy.
Encourage professional help seeking when needed	If u need some help or support, don't hesitate to ask. Local GP, Parent Helpline: 0800 568 856, Family Services: 0800 211 211, <a href="http://www.commonground.org.nz">http://www.commonground.org.nz</a>
Parental self-care	Take care of yourself!!! Take a break, and get support for yourself too. If your own needs are met, its easier to be patient, consistent & available to your teen.

## Limitations

There are limitations in our study. The overall number of participants was small and not intended to be representative of the wider parent population. Nonetheless, we were successful in recruiting a high proportion of hard-to-reach parents (ie, Māori and Pacific parents). Moreover, many ideas and themes were repeated as the focus groups progressed, indicating that saturation might have been reached.

The decision for developing an SMS text messaging program (rather than an app/Web-based program) was made before conducting the focus group and thus may have influenced the discussion on parenting support. Nonetheless, our participants considered text messages as feasible and were keen to try the program when available.

## Conclusions

Despite the limitations, our study reinstates the need for accessible and family-friendly parenting outreach and services.

Wider access to advice and information is now feasible by utilizing technology such as mobile phone and Web-based platforms [16]. The widespread use of mobile phone suggests that parenting programs can incorporate mHealth technology with minimal economic burden. Although a brief SMS text messaging program may not meet the need of families requiring more intensive intervention or replace face-to-face therapy, it may offer support for prevention and early identification of emerging issues to many more parents who otherwise would have had none. Findings from this qualitative work informed the development of MyTeen, an SMS text messaging program designed to increase parental competence and improve mental health literacy for parents of adolescents [23]. Future studies may wish to examine the perspectives of adolescents themselves and how family can be better supported. The effectiveness of this program has been evaluated in a randomized controlled trial, and the main results will be reported elsewhere.

## Acknowledgments

The authors acknowledge the ongoing commitment of the project staff and participants taking part in the study. This study was funded by The National Science Challenge: A Better Start/Cure Kids (Project grant number: 3713711).

## Authors' Contributions

JTWC is the primary investigator of this study and wrote the first draft of the manuscript. AW oversaw the management and day-to-day running of the project. JTWC, MS, and a research assistant reviewed and conducted the analysis. YJ, CB, RW, KS, and MS contributed to the design of the study and were involved in revising the manuscript. All authors read and approved the final manuscript.

## Conflicts of Interest

None declared.

## References

1. McLeod GF, Horwood LJ, Fergusson DM. Adolescent depression, adult mental health and psychosocial outcomes at 30 and 35 years. *Psychol Med* 2016 May;46(7):1401-1412. [doi: [10.1017/S0033291715002950](https://doi.org/10.1017/S0033291715002950)] [Medline: [26818194](https://pubmed.ncbi.nlm.nih.gov/26818194/)]

2. Clark T, Fleming T, Fleming T, Bullen P, Crengle S, Denny S, et al. Youth'12 Prevalence Tables: The Health and Wellbeing of New Zealand Secondary School Students in 2012. Auckland, New Zealand: University of Auckland. Faculty of Medical and Health Sciences; 2013.
3. Cummings CM, Caporino NE, Kendall PC. Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychol Bull* 2014 May;140(3):816-845 [FREE Full text] [doi: [10.1037/a0034733](https://doi.org/10.1037/a0034733)] [Medline: [24219155](https://pubmed.ncbi.nlm.nih.gov/24219155/)]
4. Steca P, Bassi M, Caprara GV, Fave AD. Parents' self-efficacy beliefs and their children's psychosocial adaptation during adolescence. *J Youth Adolesc* 2011 Mar;40(3):320-331. [doi: [10.1007/s10964-010-9514-9](https://doi.org/10.1007/s10964-010-9514-9)] [Medline: [20204688](https://pubmed.ncbi.nlm.nih.gov/20204688/)]
5. Restifo K, Bögels S. Family processes in the development of youth depression: translating the evidence to treatment. *Clin Psychol Rev* 2009 Jun;29(4):294-316. [doi: [10.1016/j.cpr.2009.02.005](https://doi.org/10.1016/j.cpr.2009.02.005)] [Medline: [19356833](https://pubmed.ncbi.nlm.nih.gov/19356833/)]
6. Wade C, Cann W, Matthews J. Introduction to special issue: parenting interventions and the mental health of children and parents. *Adv Ment Health* 2019 Mar 22;17(1):1-5. [doi: [10.1080/18387357.2019.1585647](https://doi.org/10.1080/18387357.2019.1585647)]
7. Kuhn ES, Laird RD. Family support programs and adolescent mental health: review of evidence. *Adolesc Health Med Ther* 2014;5:127-142 [FREE Full text] [doi: [10.2147/AHMT.S48057](https://doi.org/10.2147/AHMT.S48057)] [Medline: [25177156](https://pubmed.ncbi.nlm.nih.gov/25177156/)]
8. Chu JT, Farruggia SP, Sanders MR, Ralph A. Towards a public health approach to parenting programmes for parents of adolescents. *J Public Health (Oxf)* 2012 Mar;34(Suppl 1):i41-i47. [doi: [10.1093/pubmed/fdr123](https://doi.org/10.1093/pubmed/fdr123)] [Medline: [22363030](https://pubmed.ncbi.nlm.nih.gov/22363030/)]
9. Sanders MR, Kirby JN, Tellegen CL, Day JJ. The Triple P-Positive Parenting Program: a systematic review and meta-analysis of a multi-level system of parenting support. *Clin Psychol Rev* 2014 Jun;34(4):337-357. [doi: [10.1016/j.cpr.2014.04.003](https://doi.org/10.1016/j.cpr.2014.04.003)] [Medline: [24842549](https://pubmed.ncbi.nlm.nih.gov/24842549/)]
10. Sanders MR. Triple P-Positive Parenting Program as a public health approach to strengthening parenting. *J Fam Psychol* 2008 Aug;22(4):506-517. [doi: [10.1037/0893-3200.22.3.506](https://doi.org/10.1037/0893-3200.22.3.506)] [Medline: [18729665](https://pubmed.ncbi.nlm.nih.gov/18729665/)]
11. Mytton J, Ingram J, Manns S, Thomas J. Facilitators and barriers to engagement in parenting programs: a qualitative systematic review. *Health Educ Behav* 2014 Apr;41(2):127-137. [doi: [10.1177/1090198113485755](https://doi.org/10.1177/1090198113485755)] [Medline: [23640123](https://pubmed.ncbi.nlm.nih.gov/23640123/)]
12. Gardner F, Hutchings J, Bywater T, Whitaker C. Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *J Clin Child Adolesc Psychol* 2010;39(4):568-580. [doi: [10.1080/15374416.2010.486315](https://doi.org/10.1080/15374416.2010.486315)] [Medline: [20589567](https://pubmed.ncbi.nlm.nih.gov/20589567/)]
13. Sayal K, Tischler V, Coope C, Robotham S, Ashworth M, Day C, et al. Parental help-seeking in primary care for child and adolescent mental health concerns: qualitative study. *Br J Psychiatry* 2010 Dec;197(6):476-481. [doi: [10.1192/bjp.bp.110.081448](https://doi.org/10.1192/bjp.bp.110.081448)] [Medline: [21119154](https://pubmed.ncbi.nlm.nih.gov/21119154/)]
14. Keller J, McDade K. Cultural diversity and help-seeking behavior. *J Multicult Soc Work* 1997 May 15;5(1-2):63-78. [doi: [10.1300/j285v05n01\\_06](https://doi.org/10.1300/j285v05n01_06)]
15. Gibson L, Hanson V. Digital Motherhood: How Does Technology Help New Mothers? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI'13; April 27 - May 2, 2013; Paris, France p. 313-322. [doi: [10.1145/2470654.2470700](https://doi.org/10.1145/2470654.2470700)]
16. MacDonell KW, Prinz RJ. A review of technology-based youth and family-focused interventions. *Clin Child Fam Psychol Rev* 2017 Jun;20(2):185-200 [FREE Full text] [doi: [10.1007/s10567-016-0218-x](https://doi.org/10.1007/s10567-016-0218-x)] [Medline: [27787701](https://pubmed.ncbi.nlm.nih.gov/27787701/)]
17. Militello LK, Melnyk BM, Hekler E, Small L, Jacobson D. Correlates of healthy lifestyle beliefs and behaviors in parents of overweight or obese preschool children before and after a cognitive behavioral therapy intervention with text messaging. *J Pediatr Health Care* 2016;30(3):252-260. [doi: [10.1016/j.pedhc.2015.08.002](https://doi.org/10.1016/j.pedhc.2015.08.002)] [Medline: [26429638](https://pubmed.ncbi.nlm.nih.gov/26429638/)]
18. Sharifi M, Dryden EM, Horan CM, Price S, Marshall R, Hacker K, et al. Leveraging text messaging and mobile technology to support pediatric obesity-related behavior change: a qualitative study using parent focus groups and interviews. *J Med Internet Res* 2013 Dec 6;15(12):e272 [FREE Full text] [doi: [10.2196/jmir.2780](https://doi.org/10.2196/jmir.2780)] [Medline: [24317406](https://pubmed.ncbi.nlm.nih.gov/24317406/)]
19. Hofstetter AM, DuRivage N, Vargas CY, Camargo S, Vawdrey DK, Fisher A, et al. Text message reminders for timely routine MMR vaccination: a randomized controlled trial. *Vaccine* 2015 Oct 26;33(43):5741-5746 [FREE Full text] [doi: [10.1016/j.vaccine.2015.09.042](https://doi.org/10.1016/j.vaccine.2015.09.042)] [Medline: [26424607](https://pubmed.ncbi.nlm.nih.gov/26424607/)]
20. Bigelow KM, Carta JJ, Lefever JB. Txt u ltr: using cellular phone technology to enhance a parenting intervention for families at risk for neglect. *Child Maltreat* 2008 Nov;13(4):362-367. [doi: [10.1177/1077559508320060](https://doi.org/10.1177/1077559508320060)] [Medline: [18794303](https://pubmed.ncbi.nlm.nih.gov/18794303/)]
21. Kharbanda EO, Stockwell M, Fox H, Rickert VI. Text4Health: a qualitative evaluation of parental readiness for text message immunization reminders. *Am J Public Health* 2009 Dec;99(12):2176-2178. [doi: [10.2105/AJPH.2009.161364](https://doi.org/10.2105/AJPH.2009.161364)] [Medline: [19833982](https://pubmed.ncbi.nlm.nih.gov/19833982/)]
22. Parenting Strategies. 2013. How to Reduce Your Child's Risk of Depression and Clinical Anxiety. URL: <https://www.parentingstrategies.net/depression/media/pdf/Guidelines.pdf> [accessed 2018-02-02]
23. Chu JT, Whittaker R, Jiang Y, Wadham A, Stasiak K, Shepherd M, et al. Evaluation of MyTeen - a SMS-based mobile intervention for parents of adolescents: a randomised controlled trial protocol. *BMC Public Health* 2018 Oct 26;18(1):1203 [FREE Full text] [doi: [10.1186/s12889-018-6132-z](https://doi.org/10.1186/s12889-018-6132-z)] [Medline: [30367613](https://pubmed.ncbi.nlm.nih.gov/30367613/)]
24. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
25. Chand N, Farruggia S, Dittman C, Sanders M, Chu J. Promoting positive youth development: through a brief parenting intervention program. *Youth Stud Aust* 2013;32(1):29.

26. Chu JT, Bullen P, Farruggia SP, Dittman CK, Sanders MR. Parent and adolescent effects of a universal group program for the parenting of adolescents. *Prev Sci* 2015 May;16(4):609-620. [doi: [10.1007/s11121-014-0516-9](https://doi.org/10.1007/s11121-014-0516-9)] [Medline: [25373684](https://pubmed.ncbi.nlm.nih.gov/25373684/)]
27. Scharf M, Shulman S. Intergenerational transmission of experiences in adolescence: the challenges in parenting adolescents. In: *Parenting Representations: Theory, Research, and Clinical Implications*. New York: Cambridge University Press; 2006:319-351.
28. Doty J, Dworkin J. Parents' of adolescents use of social networking sites. *Comput Hum Behav* 2014 Apr;33:349-355. [doi: [10.1016/j.chb.2013.07.012](https://doi.org/10.1016/j.chb.2013.07.012)]
29. Floros G, Siomos K. The relationship between optimal parenting, internet addiction and motives for social networking in adolescence. *Psychiatry Res* 2013 Oct 30;209(3):529-534. [doi: [10.1016/j.psychres.2013.01.010](https://doi.org/10.1016/j.psychres.2013.01.010)] [Medline: [23415042](https://pubmed.ncbi.nlm.nih.gov/23415042/)]
30. Swickert RJ, Hittner JB, Harris JL, Herring JA. Relationships among Internet use, personality, and social support. *Comput Hum Behav* 2002 Jul;18(4):437-451. [doi: [10.1016/s0747-5632\(01\)00054-1](https://doi.org/10.1016/s0747-5632(01)00054-1)]
31. Ramirez ER, Norman GJ, Rosenberg DE, Kerr J, Saelens BE, Durant N, et al. Adolescent screen time and rules to limit screen time in the home. *J Adolesc Health* 2011 Apr;48(4):379-385 [FREE Full text] [doi: [10.1016/j.jadohealth.2010.07.013](https://doi.org/10.1016/j.jadohealth.2010.07.013)] [Medline: [21402267](https://pubmed.ncbi.nlm.nih.gov/21402267/)]
32. Yap MB, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: a systematic review and meta-analysis. *J Affect Disord* 2014 Mar;156:8-23. [doi: [10.1016/j.jad.2013.11.007](https://doi.org/10.1016/j.jad.2013.11.007)] [Medline: [24308895](https://pubmed.ncbi.nlm.nih.gov/24308895/)]
33. Hoagwood KE, Cavaleri MA, Olin S, Burns BJ, Slaton E, Gruttadaro D, et al. Family support in children's mental health: a review and synthesis. *Clin Child Fam Psychol Rev* 2010 Mar;13(1):1-45. [doi: [10.1007/s10567-009-0060-5](https://doi.org/10.1007/s10567-009-0060-5)] [Medline: [20012893](https://pubmed.ncbi.nlm.nih.gov/20012893/)]
34. Yap MB, Jorm AF. Parents' beliefs about actions they can take to prevent depressive disorders in young people: results from an Australian national survey. *Epidemiol Psychiatr Sci* 2012 Mar;21(1):117-123. [doi: [10.1017/s2045796011000667](https://doi.org/10.1017/s2045796011000667)] [Medline: [22670419](https://pubmed.ncbi.nlm.nih.gov/22670419/)]
35. Yap MB, Martin PD, Jorm AF. Online parenting guidelines to prevent adolescent depression and anxiety: evaluating user characteristics and usefulness. *Early Interv Psychiatry* 2018 Oct;12(5):951-958. [doi: [10.1111/eip.12503](https://doi.org/10.1111/eip.12503)] [Medline: [29076246](https://pubmed.ncbi.nlm.nih.gov/29076246/)]
36. Reardon T, Harvey K, Baranowska M, O'Brien D, Smith L, Creswell C. What do parents perceive are the barriers and facilitators to accessing psychological treatment for mental health problems in children and adolescents? A systematic review of qualitative and quantitative studies. *Eur Child Adolesc Psychiatry* 2017 Jun;26(6):623-647 [FREE Full text] [doi: [10.1007/s00787-016-0930-6](https://doi.org/10.1007/s00787-016-0930-6)] [Medline: [28054223](https://pubmed.ncbi.nlm.nih.gov/28054223/)]
37. Research New Zealand. 2014. New Zealanders' Use of Mobile Electronic Devices. URL: <http://www.researchnz.com/special-reports.html#mobile> [accessed 2018-02-02]

## Abbreviations

**mHealth:** mobile health

*Edited by G Eysenbach; submitted 26.07.19; peer-reviewed by S Brown, M Glattacker; comments to author 28.08.19; revised version received 30.08.19; accepted 02.09.19; published 20.11.19.*

*Please cite as:*

Chu JT, Wadham A, Jiang Y, Whittaker R, Stasiak K, Shepherd M, Bullen C

Development of MyTeen Text Messaging Program to Support Parents of Adolescents: Qualitative Study

*JMIR Mhealth Uhealth* 2019;7(11):e15664

URL: <http://mhealth.jmir.org/2019/11/e15664/>

doi: [10.2196/15664](https://doi.org/10.2196/15664)

PMID: [31746767](https://pubmed.ncbi.nlm.nih.gov/31746767/)

©Joanna Ting Wai Chu, Angela Wadham, Yannan Jiang, Robyn Whittaker, Karolina Stasiak, Matthew Shepherd, Christopher Bullen. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 20.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# An Interactive Mobile Phone App (SMART 5-A-DAY) for Increasing Knowledge of and Adherence to Fruit and Vegetable Recommendations: Development and Pilot Randomized Controlled Trial

Katherine Marie Appleton<sup>1</sup>, BSc, PhD; David Passmore<sup>1</sup>, BSc; Isobel Burn<sup>1</sup>, BSc; Hanna Pidgeon<sup>1</sup>, BSc; Philippa Nation<sup>1</sup>, BSc; Charlotte Boobyer<sup>1</sup>, BSc; Nan Jiang<sup>1</sup>, BSc, PhD

Bournemouth University, Poole, United Kingdom

**Corresponding Author:**

Katherine Marie Appleton, BSc, PhD

Bournemouth University

Poole House, Fern Barrow

Poole, BH12 5BB

United Kingdom

Phone: 44 01202965985

Email: [k.appleton@bournemouth.ac.uk](mailto:k.appleton@bournemouth.ac.uk)

## Abstract

**Background:** Fruit and vegetable consumption is important for health, but many individuals fail to consume adequate amounts for health benefits. Although many individuals are aware of current fruit and vegetable consumption recommendations, research suggests that adherence to these is hampered by low knowledge of the details of these recommendations.

**Objective:** This paper reports the development and details of a pilot randomized controlled test of a novel interactive mobile phone app for addressing low knowledge of the UK 5-a-day fruit and vegetable recommendations.

**Methods:** Requirements for the app were first defined by researchers and potential end users and prioritized using the MoSCoW (Must have, Should have, Could have, Won't have) method. Second, a prototype mobile phone app was developed using an agile approach. Third, the prototype app was tested in a randomized controlled pilot trial for impacts on knowledge and intake of fruit and vegetables. Volunteers were randomized to either receive (n=50) or not receive the app (n=44) for 2 or 4 weeks, and fruit and vegetable knowledge, intake, and behavior were assessed at the beginning of the study and after 1 and 2 weeks or after 2 and 4 weeks, respectively. App usage and qualitative feedback were also investigated. All findings then informed the development of a final app.

**Results:** Low knowledge of consumption recommendations centered around portion sizes and the need for variety, and an interactive mobile phone app was considered a suitable tool for improving this knowledge in a practical manner that would be available both at time of consumption and outside of these times. The pilot test revealed improved behavior after 2 weeks compared with baseline in volunteers who received the app, but improvements in knowledge on fruit and vegetable recommendations were found in both groups, and no improvements in fruit and vegetable intakes were found in formal measures. Patterns of app usage and qualitative feedback also suggested a number of modifications. The resultant final app incorporates several behavior change techniques (goal-setting, self-monitoring, and personalized feedback) as well as aiming to improve knowledge.

**Conclusions:** A novel interactive mobile phone app was successfully developed based on requirements, and when tested in a pilot randomized controlled trial, this app was found to have some impacts on fruit and vegetable outcomes. Although benefits from the app were small, impacts will likely increase as a result of recent modifications. The final SMART 5-A-DAY app is available in the Google Play Store and now needs testing in the target population.

**Trial Registration:** ClinicalTrials.gov NCT02779491; <https://www.clinicaltrials.gov/ct2/show/NCT02779491>

(*JMIR Mhealth Uhealth* 2019;7(11):e14380) doi:[10.2196/14380](https://doi.org/10.2196/14380)

**KEYWORDS**

fruit; vegetables; diet therapy; knowledge; questionnaires; portion sizes

## Introduction

### Background

A high consumption of fruit(s) and vegetable(s) (FV) is associated with reduced risk of a number of global health concerns [1-10]. Resulting from these health benefits, the World Health Organization currently recommends consumption of at least 400 g FV per day [3-6], and governments around the world have operationalized these recommendations as recommended consumption of a number of portions of FV per day. Campaigns promoting these FV recommendations are easily available, but despite the campaigns, population FV intakes in Europe, the United States, and across the world remain low [11-14].

Populations do seem largely aware of FV consumption recommendations [15-21], and awareness of the recommendations has been associated with improved FV consumption [15,16,18,22,23]. Difficulties are reported, however, with the details of the recommendations. Consumers report confusion and poor knowledge around the foods that can be included as FV [24,25]; the amount of FV required in portion sizes [24-28], or contributing to portion sizes when portion sizes may be small, for example, for small fruits or in composite dishes [25]; the number of portions needed per day [28]; the need for a variety of FV [24,25]; and the benefits of a high FV consumption [26,27]. Furthermore, recent work of ours reported not only low knowledge of the details of the recommendations, but also a direct association between low knowledge of these details and low FV consumption [15]. These findings suggest that FV consumption would benefit from increasing knowledge of the details of FV recommendations.

Nutrition-related knowledge has previously been related to FV consumption [29-31] and is traditionally increased through educational campaigns and classes [1,31-33]. Educational campaigns, however, can be limited in scope, and classes can be limited in reach [32,33]. Furthermore, educational campaigns typically demonstrate success for improving very limited knowledge, whereas educational classes can achieve increases in knowledge and can be particularly valuable for teaching practical knowledge and for encouraging engagement, knowledge retention, and future use, but these are impractical for population-wide change [32,33].

This study sought to increase knowledge of the details of the UK FV recommendations both in a practical sense by providing details of the FV recommendations at the time of consumption to aid appropriate FV intakes, and by providing these details in an engaging, useful, and personally relevant manner, such that individuals would remember and benefit from those details also at a later time point [31-33]. To fulfill this aim, a mobile phone app was developed. Mobile phone apps can provide information to consumers at the time of food purchase and consumption, as well as outside of these times, and can offer an interactive platform encouraging practical use, personal relevance, and practical benefit to encourage information retention and future use. Of specific relevance to this study, an app was initially considered suitable for developing knowledge on FV recommendations because an app could allow users to input FV and receive immediate feedback on inclusion or not in the FV

recommendations; allow users to input any amount of FV, regardless of contribution to an official portion, and receive immediate feedback on portion sizes; store and add inputted FV to provide a running total; relate this total to recommendations; incorporate the need for variety as part of the portion size and running total function; provide all information quickly, with minimal effort for the user; utilize attractive and colorful visual displays; and an app could be mobile and so could address concerns at the time of purchase and at the time of consumption, as well as outside of these times. The potential value of mobile phone apps is also aided by rapidly growing numbers of mobile phone users with penetration rates of 68.4% in North America and 64.7% in Western Europe and estimations of use by over a third of the world's population [34].

Mobile phone apps for encouraging healthy eating are widely available, and some also focus specifically on encouraging FV intakes [35-40]. Although these existing FV apps largely focus on encouraging intakes and changing behavior [35-40], the focus of our study was to facilitate knowledge of FV consumption recommendations, such that this would result in increased adherence to recommendations and intakes. Previous work demonstrates particular confusions with FV recommendations and suggests that clarification of this knowledge may encourage FV consumption. Our aim was primarily to impart knowledge and facilitate retention and future use of that knowledge, such that FV intakes would benefit both at the time of app use and in the future without the need for ongoing app use.

### Objectives

This paper reports the development of a prototype app, the results of a subsequent pilot trial to test the app for improvements in FV recommendations knowledge and intake, and suggested improvements. First, the requirements for the app were defined from the literature and potential end users and prioritized using the Must have, Should have, Could have, Won't have (MoSCoW) method [41]. Then, a prototype app was designed and developed using an agile approach based on Google's Material Design Guidelines and best industrial practice [42]. The prototype app was then tested in a randomized controlled trial, where end users also provided qualitative feedback, and finally, an amended version of the app was developed. The app was developed specifically for a UK audience; thus, current UK FV recommendations were used. These recommendations specify the consumption of 5 80-g portions of different FV per day—the 5-a-day FV recommendations [43,44].

## Methods

### Stage 1: Defining and Prioritizing the App Requirements

#### *Defining the App Requirements*

App requirements were defined based on previous published research and engagement with potential end users. Previous published work by us and others has investigated the confusion

and concerns of individuals regarding FV recommendations [15-20,23-28].

Further engagement with potential end users was also undertaken at 4 public engagement workshops in Bournemouth, United Kingdom, in July 2014 and July 2015. These workshops were undertaken as part of Bournemouth University's Festival of Learning 2014 and 2015 and were entitled *The 5-a-day fruit and vegetable message* and marketed for the general public. The workshops detailed current FV recommendations for the United Kingdom, asked consumers for their knowledge and confusions, addressed these confusions, and provided advice for increasing intakes. Finally, participants were asked for the appropriateness of an app to help solve their confusions and encourage intakes. A total of 4 workshops were held at a number of different times in the day to allow attendance by a range of different individuals. Each workshop was run by the project PI (KMA) and either audio-recorded and transcribed or notes of all suggestions were taken at the time by an additional researcher. Each workshop followed the same format. All transcriptions and notes were subsequently analyzed using thematic analysis.

### **Prioritizing the App Requirements**

Suggested requirements for the app from both the literature and the public engagement workshops were then discussed and prioritized by the principal researchers (KMA and NJ) using MoSCoW principles. The MoSCoW method [41] is a technique used in software development to prioritize the importance of the delivery of all identified requirements. Requirements are categorized as *must have*, *should have*, *could have*, and *won't have*, based on importance, and then prioritized during the development process in this order. Requirements identified as *must have* are considered central to project success; those identified as *should have* are considered important, but not necessary; those identified as *could have* are considered desirable but not necessary; and those identified as *won't have* are considered least important [41]. Consideration was also given to the UK FV recommendations. For example, the UK recommendations stipulate that 5 different FV must be consumed per day; thus, additional consumption of eaten FV would not contribute to the 5-a-day total, and that fruit juices/smoothies can contribute to total FV consumption but can only count as 1 portion regardless of variety and quantity consumed [43,44].

### **Stage 2: Designing and Developing the App**

A prototype app was developed to include all requirements identified as *must have* and *should have* and avoid requirements identified as *won't have*. The app was developed for Android (Google) mobile phones following Google's Material Design Guidelines and industrial best practices, with reference to the adapted technology acceptance model (TAM) [45-47]. The adapted TAM proposes that technology usage is positively predicted by *perceived usefulness* ("the degree to which a person believes that using a particular system would enhance his or her performance") [45], *perceived ease of use* ("the degree to which a person believes that using a particular system would be free of effort") [45], *perceived enjoyment* ("the extent to which the activity of using the [technology] is perceived to be

enjoyable in its own right, apart from any performance consequences that may be anticipated") [46], and *perceived visual attractiveness* (the degree to which a person believes that the [technology] is aesthetically pleasing to the eye) [47].

### **Stage 3: Testing of the Prototype App**

Evaluation of the app was undertaken using a randomized controlled pilot trial, where volunteers were randomized to receive or not receive the app for either 2 or 4 weeks, and FV knowledge, FV intakes, and FV behavior were assessed and compared at baseline and after either 1 and 2 weeks, or after 2 and 4 weeks.

### **Volunteers**

Volunteers to test the app were recruited from the staff and students of Bournemouth University, United Kingdom, from November 2015 to March 2016, June 2016 to August 2016, and from November 2016 to March 2017. We aimed to recruit 100 volunteers in total—50 to test the app and 50 to act as controls. No earlier research was available to allow power calculations; thus, 50 volunteers were considered sufficient to gain feedback and assess potential impacts of the app, while ensuring the work would remain ethical should few impacts be found. Adult volunteers (aged 18 years and over) were required to own an Android mobile phone (as the app was only developed for Android platforms), and there were no other inclusion/exclusion criteria to maximize the generalizability of the study. Volunteers were recruited for a study to *test a novel mobile phone app for encouraging healthy behaviors*. Volunteers were thus aware at the study start, that they may or may not receive an app to test, but they were informed that the app may target one of a number of health behaviors, such as healthy eating, stress reduction, or exercising.

### **Intervention/Control**

Volunteers were randomized to receive the app (intervention) or not receive the app (control). Randomization was undertaken on study entry by drawing lots (participants selected 1 of 2 colored dice from a bag), and recruitment stopped once 50 individuals had been randomized to test the app. All volunteers who received the app were asked to download the app onto their phones, to register with the app to set up a user profile, and to use the app as often as they wished for either a 2-week or a 4-week period. Duration of the test period for 2 or 4 weeks was undertaken to estimate effects following very short- and longer-term use. Various evidence suggests that apps can have limited effects on behavior because an initial high use typically fades [38,39]. Initial download and access to the app were undertaken in the presence of the researcher where possible to ensure correct download. No additional information on the FV recommendations or on FV intakes was provided as part of the study to either group. The only difference between the intervention and control group was the receipt of the app (intervention group) or not (control group). The app was tested for 2 weeks from November 2015 to March 2016 and from November 2016 to March 2017, and for 4 weeks from June 2016 to August 2016.

## Outcomes

Awareness of the 5-a-day FV recommendations, FV knowledge, FV intakes, and FV behavior were assessed as outcomes. Awareness of the recommendations, FV knowledge, and self-reported FV intakes were assessed using a questionnaire previously developed by us [15]. The questionnaire consists of 2 questions on awareness of the 5-a-day message, 4 questions on knowledge of the details of the message (which foods are included, portion sizes, the need for variety, and reasons for consumption), and 2 questions on FV intake. Self-reported FV intake was also assessed using a validated Food Frequency Questionnaire (FFQ)—the Leeds Food and Nutrition Survey [48]. FV behavior was assessed using a behavioral measure of complementary drink choice. Demographic and lifestyle characteristics that have previously been associated with FV consumption and dietary knowledge [11,15-17,21,28] were also assessed as potential confounders. All volunteers (intervention and control) completed all outcome assessments in the same manner. To maximize the data collected in the study period, data were collected from those in the 2-week study at baseline, week 1, and week 2, and from those in the 4-week study at baseline, week 2, and week 4. The 2 self-report questionnaires used [15,48] are discussed briefly further and provided in [Multimedia Appendix 1](#).

### Awareness of the Recommendations

Awareness of the recommendations were assessed using 2 open-response questions: *Are you aware of the 5-a-day fruit and vegetable message?* and *What do you think it means?*

### Fruit and Vegetable Knowledge

FV knowledge was assessed using 4 structured closed-response questions on (1) the FV that are included in the UK recommendations; (2) the portion sizes that are required for the recommendations; (3) the variety of FV that is required for the UK recommendations; and (4) the reasons for FV consumption. These questions include (1) a number of foods; (2) a number of different portions of FV; (3) a number of combinations of FV to be consumed in a day; and (4) a number of different health conditions, respectively, and respondents were asked to report (1) inclusion in the recommendations or not; (2) contribution to the recommendations based on portion sizes; (3) number of FV portions consumed in the day; and (4) the impact of FV on each health condition, respectively. For all questions, a correct response, based on current recommendations from the UK Government [44] is scored +1, an incorrect response is scored -1, and *don't know/not sure* is scored 0.

### Self-Reported Fruit and Vegetable Intake

FV intake was assessed using 1 single open-response question, 1 structured open-response question, and a validated FFQ [48]. The open-response question asked for estimated number of portions of FV consumed per day, to provide a measure of *Estimated FV*. The structured open-response question requested household amounts (eg, tablespoons) of all FV consumed at various time points (before breakfast, breakfast, morning, lunch, afternoon, evening meal, and evening) on a typical weekday and on a typical weekend day. This questionnaire was used to calculate portions of FV consumed per day, to provide a measure

of *calculated FV*. The validated FFQ [48] requests frequency of consumption for 65 different foods using the response format *2 or more times a day, every day, 3 to 5 times a week, 1 to 2 times a week, 1 to 3 times a month, and rarely/never*, which are subsequently scored 2, 1, 0.5, 0.21, 0.07, and 0, respectively, to provide a measure of frequency of consumption per day. The questionnaire was validated in adults at the time of development. A total of 10 questions on FV are provided, and responses to these 10 questions were then converted to consumption per day and summed, to give a measure of *FFQ FV*.

### Fruit and Vegetable Behavior

FV intake was also assessed using a behavioral measure. Volunteers were offered a drink while completing all questionnaires and given the choice of tea, coffee, water, or fruit smoothie. The UK 5-a-day recommendations include fruit juice and fruit smoothies as FV [43,44]; thus, selections of the fruit smoothie were considered an FV choice, whereas all other drinks were considered a non-FV choice. No drink was also a permitted option.

### Demographic and Lifestyle Characteristics

Demographic and lifestyle characteristics also assessed were gender, age, marital status, living status, number of years of education, smoking habits, alcoholic drinking habits, dietary supplement taking habits, and height and weight (to calculate body mass index).

### App Feedback

Number of uses were requested from volunteers who received the app and downloaded from the app itself. Volunteers who received the app were also asked to give feedback on their experiences and offer suggestions for the app. This feedback was requested as part of the study debrief. Participants were free to offer as many or as few comments as they wished in a written or verbal form.

### Additional Measures

To encourage a perception that the study was investigating the impacts of a number of apps for a variety of health behaviors, some additional measures, for example, questions on physical activity and stress, were also undertaken. These data were not analyzed.

### Procedure

Volunteers undertook all outcome assessments at the Eating Behaviours Laboratory, Bournemouth University, United Kingdom. On each assessment occasion, volunteers completed all questionnaires using a Web-based platform (Qualtrics), were offered a drink, and had every opportunity to ask questions. One researcher randomized all volunteers and dealt with all queries, whereas another researcher oversaw all outcome assessments; thus, this researcher was blind to treatment (intervention/control).

The study was given ethical approval by the Research Ethics Committee of Bournemouth University before commencement and was registered as a clinical trial on ClinicalTrials.gov (NCT02779491). Methods were undertaken as detailed in the trial registration with the exception that a behavioral measure

of FV intake was added to the study before commencement, and a measure of FV attitudes was cut. The original study proposal included a measure of attitudes toward FV, but these were decided against before the study start to reduce demand characteristics given the extensive FV knowledge questionnaire. All participants provided written informed consent before starting the study.

### Analysis

Quantitative data were analyzed on an intention-to-treat basis, where missing data were completed using multiple imputation [49], based on gender, age, study period, and baseline measures. Demographic and lifestyle variables and all measures at baseline were first described and compared using 2-tailed *t* tests, on the basis of study duration and intervention/control grouping. To investigate impacts of the app with time, all FV knowledge and intake outcomes were analyzed using analysis of variance (ANOVA) for differences between baseline and week 2, and baseline and week 4. A covariate of study duration was also added to the ANOVA for the 2-week data, to accommodate differences between those studied for 2 weeks and those studied for 4 weeks. Thus, effects at week 2 were investigated using a 2 (intervention/control) × 2 (baseline/week 2) mixed analysis of covariance, and effects at week 4 were investigated using a 2 (intervention/control) × 2 (baseline/week 4) mixed ANOVA. Our behavioral measure of FV intake—choice of fruit drink or nonfruit drink was analyzed using chi-square tests. All data are reported as means and standard deviations. Significance was set at  $P < .05$ . Qualitative comments were analyzed using thematic analysis.

### Stage 4: Development of the Final App

Finally, the results and feedback from the users of the pilot randomized controlled trial were used to suggest amendments to create a final version of the app.

## Results

### Stage 1: Defining and Prioritizing the App Requirements

#### App Requirements

Previous published work reveals confusion around the foods that are included in the recommendations; the amount of FV required for a portion, particularly where large items, small items, and composite dishes do not always contribute complete portions; the number of portions needed per day; and the need for a variety of FV [15-20,23-28].

The 4 workshops were attended by 32 members of the population of Bournemouth. We did not measure any demographic variables, but individuals were noticeably of both genders, aged from 18 years to old age, and based on their questions or self-disclosures were students, mothers of young children, working professionals, and retired individuals.

These participants voiced similar confusions to those found in the literature and suggested that an app would potentially be appropriate to aid with these concerns. A total of 5 key themes emerged from analysis of the workshop discussions.

### 1. Useful for Portion Sizes

Participants expressed particular difficulties over the differing portion sizes required for differing FV, and valued an idea that amount consumed could be entered into an app using household measures, for example, spoonfuls, and converted into portion sizes for them:

*Oh yeah, that would be cool, so I can type in like 10 grapes, and it tells me, yeah, that's one portion...or that's only half a portion, or whatever...yeah, that would be handy.*

*If it could tell me my stew gives me two portions, when I have loads of veggies in it, just all in pieces,...then that would be handy.*

### 2. Useful Monitor

The app was considered likely to be useful for keeping track of FV consumption, particularly for small amounts of FV, for example, in composite dishes:

*I like the idea that I might be getting 5 a day already but I just don't know it...but to have a little machine to keep track of it in the day for me, and then I can check at the end, that would be helpful.*

### 3. Useful Target

The calculator function was also considered useful for telling users how close they were to a daily target:

*If you could have some sort of bar to tell you how close you were to the 5 a day, that would be useful...you know, a man who gradually fills up, or something similar.*

### 4. Useful to Have It Mobile

Potential users also liked the idea that the app would be with them whenever they needed it; thus, they could use it in the evening to recap at the end of a day, but they could use it also, at point of purchase or point of consumption:

*So you could use it in the shop or in the canteen and just try, you know, if I had the salad I would have 3 half portions, but if I have the hot meal and two veg [vegetable portions] instead of the chips, that would be two portions—that would be better. I would never think like that normally.*

### 5. Possible Negative Monitor

A few reservations were also expressed around the feedback that users may receive following their use of the app and the possibility that this may be negative:

*I think it's a neat idea, but I wouldn't want anything telling me I was bad, or not eating well enough...I wouldn't use it in that case—it needs to be nice to me!*

### Requirement Priorities

The priorities for the app based on MoSCoW principles are presented in Table 1.



**Table 1.** Must have, Should have, Could have, Won't have (MoSCoW) requirements for the app.

Serial number	Requirement
<b>Must have</b>	
1	Allow users to input FV <sup>a</sup> consumed at any time and using household amounts, for example, number of items and number of spoonfuls.
2	Provide users with a list of all FV for selection, as opposed to requiring manual input.
3	Categorize FV (eg, fruits, vegetables, and salad items) to avoid overly long lists of FV items for inputting.
4	Allow users to input part items/units, where only part items have been consumed, for example, in composite dishes.
5	Provide immediate feedback on inclusion or not of the FV in the UK 5-a-day recommendations.
6	Calculate contribution to a portion for the UK 5-a-day recommendations based on amount consumed.
7	Allow fractions of portions in these calculations, but do not allow multiple portions of the same FV in any one day.
8	Provide immediate feedback on contribution of the portion to the UK 5-a-day recommendations.
9	Sum contributions of portions to provide a running daily FV total.
10	Relate this running daily total to the recommendations of 5 FV per day.
11	Provide immediate feedback on the daily FV consumption per day.
12	Require users to set up an account to allow FV to be tracked on a personal basis.
13	Ensure users data are retained on their own device, to ensure data protection and privacy.
<b>Should have</b>	
14	Provide FV items using colored picture icons as well as FV names.
15	Display total daily FV consumed in a graphical manner allowing representation also of the target, for example, using a filled bar.
16	Provide constructive feedback to highlight if the amount consumed is insufficient to amount to a whole portion, for example, <i>an additional spoonful of xxx would provide a full portion.</i>
<b>Could have</b>	
17	Store daily running totals over time to allow users to view their history.
18	Provide a signal when the 5-a-day target was met, for example, applause sound.
19	Provide a reward when the 5-a-day target was met, for example, a token to be traded for material gain.
<b>Won't have</b>	
20	Provide instructive advice based on user inputs, for example, <i>you need to eat more of xxx.</i>
21	Require users to input additional information, for example, time and place.
22	Allow users to amend FV consumption in the past

<sup>a</sup>FV: fruit(s) and vegetable(s).

## Stage 2: Designing and Developing the App

The app was developed using an agile approach as described by Google's Material Design Guidelines and industrial best practices [42]. A user journey map was first created to visualize the timeline of interactions with the potential app from the landing page. Wireframes of each app screen were then produced using Balsamiq. These wireframes focused on app screen layout and content structure and were organized to reflect the user journey map. These wireframes were then mapped to mock-ups showing the actual visual designs for each screen. An interactive prototype was created using InVision, and from this, an Android app was developed using native Android Studio. Primary

researchers (KMA and NJ) were consulted at each step for feedback.

The prototype app consisted of a series of screens allowing consumers to input and view their daily FV intake in comparison with the UK 5-a-day recommendations. All requirements identified as *must have* and *should have* were included with the exception that picture icons were not provided for some FV items (Table 1, requirement 14). Icons were not easily available for all FV items, and although desirable, icons for all FV items were considered not necessary at the prototype stage. Names were provided for all FV. All *won't have* requirements were also avoided. Details of the app, per screen, are given in Table 2. Screenshots of screens 4, 5, 6, and 7 are given in Figures 1-6.

**Table 2.** Details of the prototype app.

Feature	Detail	Supported user actions	Requirements addressed
1. Welcome	The app name and app logo	Swipe to continue	__ <sup>a</sup>
2. Registration	Request to login or register for an account	Provide a username to allow data to be tracked	12,13
3. Daily summary	Total FV <sup>b</sup> inputted in the current day	Options to add (more) FV	11,15
4. Input categories	Lists of FV, categorized as <i>fruit</i> , <i>vegetables</i> , <i>salad</i> , and <i>drink</i>	Select relevant FV category	1-3
5. Input item	Individual FV items per category, displayed by name and icon (where available)	Select relevant FV item	1,2,4,5,14
6. Input amount	Arrows to select amount consumed, provided as items or spoons, as most commonly used	Select amount	1,4
7. Updated summary	Amount consumed provided in portions based on recommendations. Details of amount required for a full portion if <1 portion. Total FV inputted for the current day updated and displayed. Motivational or congratulatory message also displayed.	Options to add (more) FV	6-11,15,16

<sup>a</sup>No specified requirement addressed.

<sup>b</sup>FV: fruit(s) and vegetable(s).

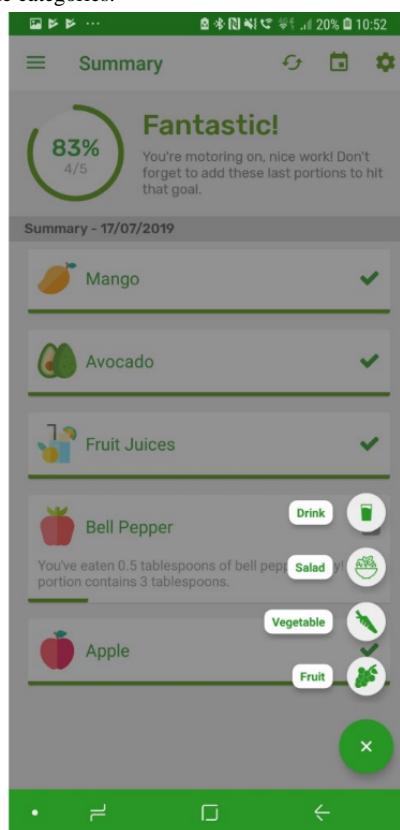
**Figure 1.** Screenshots of the app: fruit and vegetable categories.

Figure 2. Screenshots of the app: fruit icons.

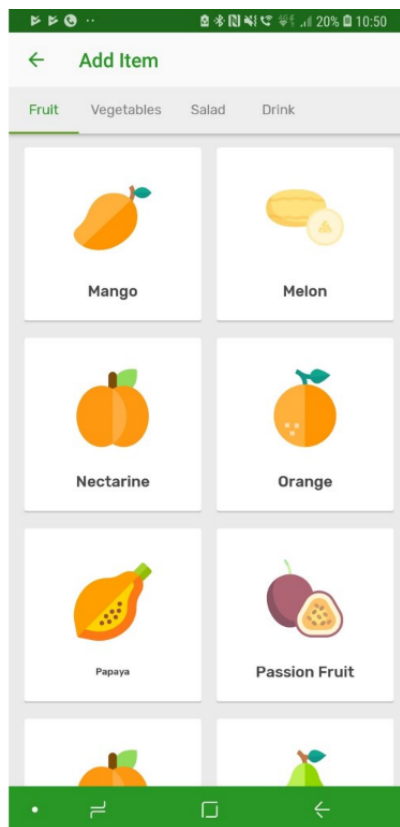


Figure 3. Screenshots of the app: vegetable icons.

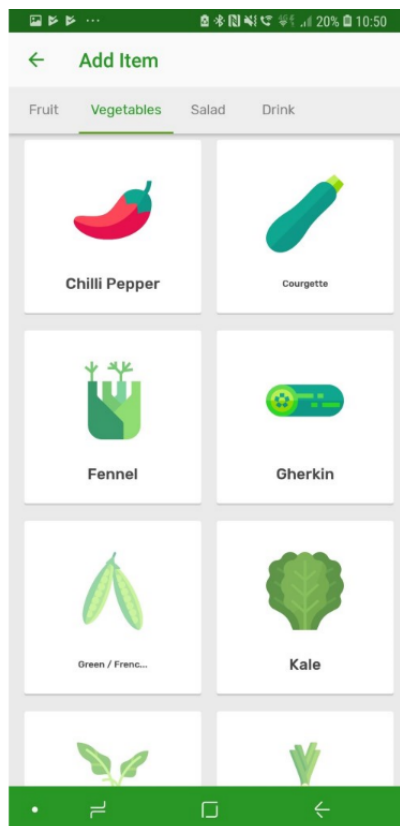


Figure 4. Screenshots of the app: fruit and vegetable selection.

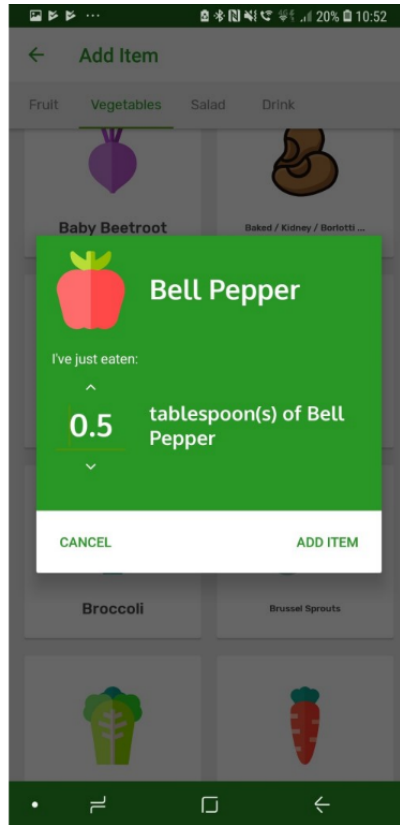
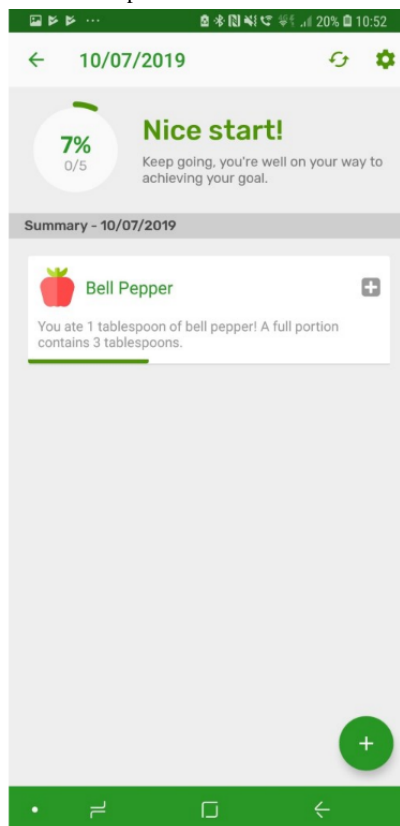
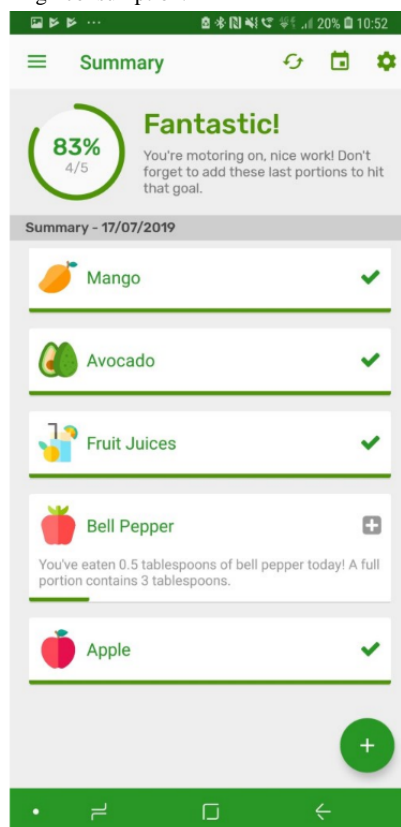


Figure 5. Screenshots of the app: Summary screen – low consumption.



**Figure 6.** Screenshots of the app: Summary screen – high consumption.

### Stage 3: Initial Testing of the Prototype App

#### Volunteers

A total of 94 volunteers took part in the randomized controlled trial—50 who received and tested the app, and 44 who acted as controls. Of these, 32 volunteers received the app for 2 weeks, 27 volunteers acted as controls; 18 volunteers received the app for 4 weeks, 17 volunteers acted as controls. Demographic and lifestyle characteristics of all participants are given in [Table 3](#). Volunteers who were studied for 2 weeks were more likely to be younger ( $t_{92}=2.52$ ;  $P=.02$ ) and less educated ( $t_{92}=4.08$ ;  $P<.001$ ), than those who were studied for 4 weeks, predominantly because volunteers were studied for a 2-week period when most of the volunteers were undergraduate students,

and for 4 weeks when most of the volunteers were postgraduate students or university staff. No differences were found between the intervention and control groups in any demographic and lifestyle variable (largest  $t_{57}=1.57$ ;  $P=.12$ ).

Adherence to the study was good. A total of 88 of 94 (94%) volunteers took part in all 3 test sessions, 1 volunteer undertook the first 2 sessions but failed to undertake the final session (control volunteer for 4 weeks), 3 volunteers undertook the first session but failed to undertake the second 2 sessions (1 volunteer received the app for 4 weeks, 1 volunteer was a control volunteer for 4 weeks, and 1 volunteer was a control volunteer for 2 weeks), and 2 volunteers undertook the first and third sessions but missed the second session (both volunteers received the app for 2 weeks). Reasons for dropout were not recorded.

**Table 3.** Demographic and lifestyle characteristics.

Characteristics	2-week study		4-week study	
	App (n=32)	Control (n=27)	App (n=18)	Control (n=17)
<b>Gender, n (%)</b>				
Male	10 (31)	10 (37)	6 (33)	5 (29)
Female	22 (69)	17 (63)	12 (67)	12 (71)
Age (years), mean (SD)	22.3 (7.7)	21.4 (5.3)	25.9 (7.9)	25.9 (7.9)
<b>Marital status, n (%)</b>				
Married	0 (0)	0 (0)	1 (6)	2 (12)
Not married	32 (100)	27 (100)	17 (94)	15 (88)
<b>Living, n (%)</b>				
Alone	2 (6)	3 (11)	2 (11)	2 (12)
With others	30 (94)	24 (89)	16 (89)	15 (88)
Education (years), mean (SD)	14.9 (1.7)	15.0 (1.5)	16.9 (2.8)	16.9 (2.8)
<b>Smoking status, n (%)</b>				
Nonsmoker	27 (85)	21 (78)	16 (89)	13 (18)
Light (0-2/day)	2 (6)	3 (11)	2 (11)	2 (12)
Moderate (2-10/day)	3 (9)	2 (7)	0 (0)	1 (6)
Heavy (10-20/day)	0 (0)	1 (4)	0 (0)	1 (6)
<b>Use of supplements, n (%)</b>				
Never	18 (56)	14 (52)	5 (28)	8 (47)
Occasionally	7 (22)	11 (41)	11 (61)	7 (41)
Regularly	7 (22)	2 (7)	2 (11)	2 (12)
<b>Alcohol consumption, n (%)</b>				
Never	4 (13)	2 (7)	6 (33)	3 (18)
Light	20 (62)	19 (70)	7 (39)	9 (53)
Moderate	7 (22)	9 (33)	4 (22)	4 (23)
Heavy	1 (3)	0 (0)	1 (6)	1 (6)
Body mass index (kg/m <sup>2</sup> ), mean (SD)	21.6 (8.4)	24.0 (5.5)	25.0 (4.4)	25.1 (4.3)
<b>Activity, n (%)</b>				
None	1 (3)	1 (4)	0 (0)	0 (0)
Standing all day	3 (9)	6 (22)	0 (0)	2 (12)
Light	4 (13)	3 (11)	3 (17)	2 (12)
Moderate	10 (31)	7 (26)	4 (22)	6 (35)
Heavy	12 (38)	7 (26)	10 (55)	6 (35)
Very heavy	2 (6)	3 (11)	1 (6)	1 (6)

### Fruit and Vegetable Outcomes

Details of all FV outcomes are given in [Table 4](#). Analyses of FV outcomes at baseline again revealed significant differences between volunteers studied for 2 weeks and those studied for 4 weeks in estimated FV consumption ( $t_{92}=3.46$ ;  $P<.001$ ) and

FFQ FV intakes ( $t_{92}=2.49$ ;  $P=.02$ ). Volunteers studied for 4 weeks estimated and reported higher FV intakes. No differences were found in FV knowledge (largest  $t_{92}=1.63$ ;  $P=.11$ ). No differences were found between the intervention and control groups at baseline (largest  $t_{92}=1.10$ ;  $P=.28$ ).

**Table 4.** Mean (SD), fruit and vegetable knowledge scores, self-reported intake, and drink choice for all volunteers at baseline and at weeks 1, 2, and 4.

Outcome Variable	App				Control			
	Baseline (n=50)	1 week (n=32)	2 weeks (n=50)	4 weeks (n=18)	Baseline (n=44)	1 week (n=27)	2 weeks (n=44)	4 weeks (n=17)
<b>FV<sup>a</sup> knowledge, mean score (SD)</b>								
Foods (scored -35 to +35)	18.4 (6.8)	22.9 (6.4)	21.5 (8.1)	19.3 (7.9)	17.9 (7.1)	18.1 (8.8)	20.5 (8.3)	21.6 (6.3)
Portion sizes (scored -27 to +27)	-6.4 (6.2)	-5.0 (8.0)	-4.4 (7.8)	-7.6 (9.5)	-7.8 (5.7)	-6.7 (5.7)	-6.4 (7.5)	-5.3 (7.4)
Variety (scored -18 to +18)	-2.8 (5.7)	-0.4 (5.6)	0.1 (6.4)	-3.3 (6.3)	-2.6 (5.5)	-1.4 (6.0)	-1.1 (5.9)	-0.8 (6.9)
Reasons (scored -25 to +25)	0.1 (5.7)	0.8 (5.7)	1.1 (4.8)	1.2 (5.6)	-0.1 (6.5)	-0.6 (5.3)	0.7 (5.2)	1.4 (5.9)
<b>FV intake, mean (SD)</b>								
FV estimated (FV portions/day),	3.3 (1.4)	2.9 (1.4)	3.4 (1.5)	4.1 (1.2)	3.2 (1.7)	3.1 (1.5)	3.3 (1.7)	3.6 (2.2)
FV calculated (FV portions/day)	4.1 (1.4)	3.7 (1.5)	4.1 (1.6)	5.3 (1.6)	3.7 (2.0)	3.6 (2.2)	3.8 (1.8)	3.9 (1.6)
FV FFQ <sup>b</sup> (daily FV intake [portions])	3.5 (1.9)	2.6 (1.1)	3.2 (1.8)	3.5 (1.9)	3.3 (2.2)	2.3 (1.3)	2.9 (2.3)	3.9 (2.9)
<b>FV behavior, n (%)</b>								
Drink choice—Fruit smoothie	13 (21)	— <sup>c</sup>	23 (45) <sup>d</sup>	4 (25)	11 (36)	—	8 (16) <sup>d</sup>	5 (31)
Drink choice—Other drink	22 (18)	—	8 (16) <sup>d</sup>	5 (31)	15 (25)	—	12 (23) <sup>d</sup>	2 (13)

<sup>a</sup>FV: fruit(s) and vegetable(s).

<sup>b</sup>FFQ: Food Frequency Questionnaire.

<sup>c</sup>Data not collected at this timepoint.

<sup>d</sup>Significant differences between the app and control groups ( $\chi^2_{1}=6.0$ ;  $P=.02$ ).

### Fruit and Vegetable Awareness

All volunteers with the exception of 2 volunteers in the 4-week study (1 who received the app and 1 who was a control) were aware of the 5-a-day FV recommendations at baseline, and at sessions 2 and 3, all volunteers were aware of the recommendations.

### Fruit and Vegetable Knowledge

Significant increases by week 2 were found for the questions on foods included in the recommendations ( $F_{1,91}=5.11$ ;  $P=.03$ ) and portion sizes ( $F_{1,91}=5.69$ ;  $P=.02$ ), and by week 4 for all FV knowledge questions (smallest  $F_{1,33}=4.65$ ;  $P=.04$ ). No differences were found between the intervention and control groups with time (largest  $F_{1,33}=1.03$ ;  $P=.32$ ).

### Fruit and Vegetable Intake

No differences were found between the intervention and control groups over time (largest  $F_{1,91}=0.44$ ;  $P=.51$ ). Significant differences based on study duration were retained in estimated FV and FV intakes assessed by FFQ (smallest  $F_{1,91}=7.83$ ;  $P=.01$ ). Correlations among all 3 FV intake measures also demonstrated comparability (smallest  $r=0.41$ ;  $P<.001$ ). No effects of time were found (largest  $F_{1,91}=1.73$ ;  $P=.19$ ).

### Fruit and Vegetable Behavior

No significant differences between groups were found at baseline ( $\chi^2_{1}=0.17$ ;  $P=.68$ ). By week 2, significantly more fruit smoothies were chosen by those in the intervention group

compared with controls ( $\chi^2_{1}=5.96$ ;  $P=.02$ ), but no effects were found at week 4 ( $\chi^2_{1}=1.17$ ;  $P=.28$ ).

### App Usage

Self-reported usage of the app was high—most participants reported using the app on most days or every other day. Recorded use of the app also suggested almost daily usage or usage every other day. Following initial access, volunteers in the 2-week study used the app for a mean of 11.4 (SD 7.2) times, ranging from 0 to 27 times, and volunteers in the 4-week study used the app for a mean of 13.7 (SD 9.2) times, ranging from 2 to 34 times. App usage was greater in the earlier part of each test period. On the majority of days on which it was used, the app was used only once. In total, the app was used once a day on 63.6% (232/365) days on which the app was used in the 2-week study and on 65.6% (162/247) days on which the app was used in the 4-week study; twice a day on 31.8% (116/365) days and 22.3% (55/247) days, respectively; 3 times a day on 3.8% (14/365 total) days and 9.7% (24/247) days, respectively; 4 times a day on 0.8% (3/365) days and 1.0% (3/247) days, respectively; and 5 times a day on 1.0% (3/247) days on which the app was used in the 4-week study.

### App Feedback

Qualitative feedback on the app was positive—almost all volunteers reported liking the app although many also reported room for improvements. Suggested improvements included an option to add FV for the previous day because these were possibly simply forgotten; an option for changing the goal from 5 a day to more than this if individuals preferred to aim higher; a need for missing FV to be added, or an option to give feedback

that FV were missing so that these could be added; a daily notification or option to add these to remind users to interact with the app; and tips or suggestions for how to increase FV consumption.

The majority of volunteers also reported that the app was useful. Almost all volunteers reported that the app was useful for keeping a record of their consumption and for making them aware of limited consumption:

*Made me conscious of what I was eating.*

*Good to have a record of how much of 5-a-day was eaten and also to know when you're short.*

Volunteers also reported increased FV intakes through a wish to engage with the app:

*I think it made me want to eat more fruit and veg because I had to write it down.*

...and adhere more fully with recommendations:

*I think it was useful in terms of realizing that I don't eat enough fruit and veg, as it has made me think about it more.*

*I would eat more at dinner if I noticed I had not eaten enough that day.*

A limited number of volunteers also felt that the app was unnecessary:

*The app was useful, but I personally don't need an app to ensure that I get my 5 a day.*

For some, it did not help them:

*Did not help as I plan meals the week before.*

#### Stage 4: Development of the Final App

On the basis of the outcomes and feedback from the pilot test, a second version of the app is under development. Amendments that have so far been completed are to include picture icons for all FV items included in the app; to ensure more FV are included on the app; to allow users to return to previous days to add additional items where desired; and to allow users to change the target FV to more than 5 if desired (the default setting is for a target of 5 FV per day). A *history* option allows users to return to a previous day to add additional items. An ability to return to previous days was initially avoided in the prototype app to discourage users from adding false information as a result of faulty recollections. Considering that the app is primarily for the benefit of the user and that false information can be added to the app at any time, requests for access to previous days has been granted and may be beneficial for some users. The option to change the intake goal is presented to users at registration and can be amended as desired as part of the user profile settings. The additional screens for the final app are given in [Table 5](#); all screens for the prototype app also remain.

**Table 5.** Additional screens of the final app.

Feature	Detail	Supported user actions	Requirement addressed
8. Personal preferences	Options for <i>History</i> to allow inputs for previous days; <i>Reports</i> to provide an overview for the week; <i>Refresh</i> to request updates; <i>Settings</i> to update goal targets and add notifications.	Select options or return to Summary (feature 3)	__ <sup>a</sup>
9. History	Calendar display	Select date, input FV <sup>b</sup> as for the current day	Historical input permitted
10. Reports	Overview of FV intake for the previous week/month (not yet enabled)	—	—
11. Refresh	Refreshes and updates total	Enabled	—
12. Settings	Options for <i>Help</i> to feedback to the developer; <i>User</i> to access details of the user and amend intake goal; <i>Notifications</i> to set alarms; <i>Devices</i> and <i>App</i> .	Select options or return to Summary (feature 3)	—
13. Help	Abilities to contact the development team (not yet enabled)	—	—
14. User	User details and user setting displayed	Option to amend intake goal	Goal amendment permitted
15. Notifications	Abilities to set up notifications (not yet enabled)	—	—
16. Devices	Device details provided	—	—
17. App	App version details provided	—	—

<sup>a</sup>No specified user action or requirement addressed.

<sup>b</sup>FV: fruit(s) and vegetable(s).

Amendments that are still under development will allow users to reduce or delete an FV item once this has been logged (this is currently not possible); allow users access to an overview of FV consumed over the previous week or month; allow users to set up notifications; and allow users to give feedback directly to the development team. Consumption totals for previous days can currently be viewed individually, but a historical overview may also be helpful. An interactive notification is intended to

demonstrate to users the further consumption required on any one day to meet the recommendations. The default setup will be for no notifications to avoid negative reactions to the app, but notification setup will also be easy if desired. Other suggestions from app users to include tips and suggestions to increase FV are not currently planned to retain the focus and simplicity of the app.



The final app is now available for download for Android mobile phones at no cost from the Google Play Store under the name of SMART 5-A-DAY. Development continues, and updated versions of the app will be released as new features are added.

## Discussion

### Principal Findings

A novel mobile phone app was conceived to increase knowledge of the details of the UK FV recommendations both in a practical sense by providing details of the FV recommendations at the time of consumption, and by providing these details in an engaging, useful, and personally relevant manner, such that individuals would remember and benefit from those details also at a later time point. A prototype app was developed and tested by 50 users as part of a randomized controlled pilot trial, for either 2 or 4 weeks. FV assessments and positive qualitative comments suggested positive impacts of the app, but reported effect sizes were small. Additional features were suggested, and a final version of the app is currently under development.

The early development work confirmed low knowledge of the details of the 5-a-day FV recommendations in consumers, as found in the published literature [15-20,22,23], and reinforced the researchers' suggestions on the suitability of an app for providing increased FV knowledge. App development was then possible as required, to result in a fully functioning interactive mobile phone app. The results of the randomized controlled pilot trial demonstrate limited impacts of the app on the questionnaire measures of FV knowledge and FV intakes, although an impact on FV behavior was found, and qualitative feedback suggested benefits. Improvements in FV knowledge were found across the study (regardless of app receipt) presumably as a result of inclusion in a study on healthy eating and the repeated assessment of FV knowledge and FV intakes, so increased awareness of these issues. The limited findings specific to those who received the app suggest that benefits of the app are small, particularly in addition to the benefits of taking part in the study, although increased FV knowledge in all study volunteers regardless of app/no app provision may have masked impacts of app use.

An impact on FV behavior was found. Given a choice of a range of available drinks, use of the app for 2 weeks resulted directly in increased FV selection and consumption. Behavioral outcomes are important, as it is only behavior that will impact on health [31-33], and we have previously suggested that spontaneous behavioral outcomes, such as those found here may be particularly valuable in an environment of plenty [50]. Small spontaneous changes in behavior such as this may also remain largely unnoticed by individuals themselves and so may go largely unreported in self-reported measures such as those also used and often included in studies such as these [51-53].

The qualitative feedback also suggested potential changes in intake, but again that these changes may be small and may go uncaptured by traditional dietary assessment methods [52]. Furthermore, the qualitative feedback goes on to suggest that these small changes may have occurred more as a result of volunteers becoming more aware of their intakes than previously

and becoming particularly aware of low intakes. Awareness of a need for change has previously been suggested as an important step toward behavior change [33,34]. The qualitative feedback also fails to suggest impacts on FV knowledge, and although the app was intended to increase knowledge, it is well recognized that recording food intake can alert consumers to eating patterns, particularly some eating patterns that are not easily recognized over a whole day, and that this realization can change behavior [51,54-56]. The importance of increasing awareness of low intakes was not anticipated, but this finding suggests an added benefit from the app.

The qualitative comments were largely positive. Negative comments centered solely around a lack of personal interest or relevance because these users were already high FV consumers. High FV consumers are not the target audience for the app.

Additional findings from the initial test also related to app usage. Around 65% of those who received the app used it initially; these figures dropped throughout the test period, and the majority of users used the app once per day. These data are comparable with those found in studies of similar apps [38,39]. Our app was intended for use as often as volunteers wished—possibly once a day for record keeping or more often to acquire knowledge or encourage good adherence. The pattern of use suggests our testers were using the app more to track intake than to gain knowledge. These findings suggest that for maximum benefit from the app, it may be useful to market the app specifically for gaining knowledge of the recommendations and for adherence to these. This would also help distinguish our app from other apps that are intended primarily for tracking and record keeping [35-39]. Increased usage at the start of the usage period is commonly found in app testing studies, and the reduced subsequent usage is frequently cited as suggestive of poor engagement. Many users also suggested an additional reminder to aid interaction with the app, or requested an ability to return to a previous day to input forgotten items. These findings suggest that motivation to use the app was quite low among our testers, but our trial was not advertised as a study on FV consumption or healthy eating (to avoid demand characteristics); thus, our testers are likely to have been less motivated than those who would be more likely to use an app on 5-a-day FV recommendations of their own volition. Importantly, furthermore, on the basis of the qualitative comments, we do not consider this reduced usage to demonstrate poor functionality of the app. The app was intended mainly to encourage users to understand and learn the FV recommendations; thus, extended use should not be necessary.

The work conducted here further demonstrates the value of the early consultative work and the randomized controlled pilot test. Positive responses to the app overall demonstrate the value of the early research and the initial consultation exercises with potential end users. The increases in FV knowledge and intakes in all trial volunteers demonstrate the value of a randomized controlled trial for testing the app. Not all apps are tested for impacts on behavior before release, and many that are tested are done so without also involving a control group. Consideration in our study of only the 50 app testers would have suggested considerable increases in FV knowledge and intakes as a result of the app, whereas the inclusion of the control

group demonstrates these impacts to probably result more from study inclusion or FV questionnaire completion.

Our randomized controlled trial was limited through the repeated assessment of FV outcomes and the repeated use of self-reported measures. These types of measures have previously been demonstrated as accurate [48,51-53], but very brief measures may have been insensitive to small changes. Our trial was also limited through the inclusion of testers who were not our intended target group. The app is intended for those who wish to improve their knowledge of the FV recommendations, probably to aid FV intakes. To avoid demand characteristics in our trial, we asked only for those who wished to try a new health-orientated app, and some of these individuals may have been unmotivated, unwilling or unable to improve FV intakes. By comparison, our testers were unsure of the apps being tested in the study, thus were unclear that FV was the focus for all users, and responses to the FV questionnaire in our volunteers did confirm low knowledge of the 5-a-day FV recommendations among the population [15-20,22,23]. Impacts based on age and education have also been found previously [11,22,23]. We could also have measured usability of the app using more formal measures, such as the System Usability Scale [57]. Considering the more comprehensive measures of app usage and app benefit in terms of knowledge and intakes in our trial, we did not collect these usability measures, but information from these measures may have allowed comparison with other apps or technological devices [57].

The positive responses and potential for changes in behavior have resulted in continued development of the app to result in an amended version. This version includes clear details of an FV consumption goal, allows users to input FV consumed and provides detailed and graphical information on how this consumption relates to the FV recommendations, provides clear personalized feedback on distance to the goal, and allows users to change their target FV goal to a goal of their choice as they wish. Our final app thus includes 3 key aspects of behavior that have previously been suggested to lead to successful behavior change, particularly for dietary behaviors, alongside increased knowledge: self-monitoring, goal-setting, and feedback in relation to goal attainment [54-56]. Other apps and interventions aiming to improve FV intakes and dietary quality also use similar behavior change techniques [35-40], and self-monitoring and goal-setting have previously been suggested as particularly important techniques by professionals [54-56] and by consumers [35,39]. It is interesting that although our app initially aimed to increase only knowledge, functions as a result of user feedback now also include established behavior change techniques.

The increased FV behavior and qualitative reports suggest that the app has potential to benefit FV intakes and health, although changes may be small. Small changes on a population-wide level, however, will have significant impacts. Increased benefit is also likely from the added features, from highlighting the knowledge component, from tests of the app in our target

audience, and from the additional features still under development.

Our amended app now also needs testing. Further testing will not only demonstrate the improved value of the app but may also demonstrate the aspects of the app of particular benefit, given our inclusion also of behavior change techniques and knowledge and the reported value of these [54-56]. Additional functionality also allows direct linkage with additional software allowing direct access to questionnaires or other research materials.

Further development of the app may also be of value. Notably, our initial discussions with potential end users suggested an interest in both immediate and longer-term rewards for reaching a target goal. Repeated work demonstrates a value for rewards for encouraging healthy food consumption, including FV consumption [58], and rewards have previously formed an integral part of many successful dietary change interventions [36,55,59]. Other studies also suggest only limited benefit from apps for behavior change and have suggested a need for strategies to ensure continued use [36,40]. One of the advantages of our app was an intention that users would gain knowledge through the app; thus, extended use should not be required. Many apps related to social activities, such as eating, also include a *share* option to allow others to view the inputs of others or allow comparisons between users or with an established norm. Feedback options for others to comment on FV inputs, through *likes*, may also facilitate motivation, and so facilitate engagement with and action based on the app. Offline and different versions of the app may also be desirable, for example, through the use of different formats, different controls, or different setups, possibly for different population groups. Adolescents and young adults are groups with low FV consumption that may be particularly inclined toward digital interventions [39]. Socially deprived consumers may also benefit from specific aspects of the app, and investigation in different cultures (using local recommendations) would be of interest. Further work discriminating between those who use and do not use the app, and between those who find and do not find the app useful, would be very valuable.

## Conclusions

In conclusion, this study details the development and early test results of a novel interactive mobile phone app for improving knowledge and implementation of the UK 5-a-day FV recommendations. SMART 5-A-DAY was developed following assessment of the existing literature and consultation with potential end users, and then tested in a randomized controlled pilot trial. The trial demonstrated increased FV behavior 2 weeks after app receipt compared with control, and resulted in positive feedback, although resultant changes in FV knowledge and intakes were small. Suggestions for amendments were also made. Development of the app is ongoing, and further testing is required.

## Acknowledgments

This study was funded by Bournemouth University, the United Kingdom. Continued support of the app has been provided by Vers Creative UK, Bournemouth, the United Kingdom.

## Conflicts of Interest

DP is the founder and CEO of Vers Creative UK, a digital solutions development company, based in Bournemouth, the United Kingdom. There are no other conflicts of interest.

### Multimedia Appendix 1

FV Knowledge Questionnaire (Appleton et al, 2017); Leeds Food and Nutrition Survey (Margetts et al, 1989).

[[DOCX File, 59 KB - mhealth\\_v7i11e14380\\_app1.docx](#) ]

### Multimedia Appendix 2

CONSORT - EHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 290 KB - mhealth\\_v7i11e14380\\_app2.pdf](#) ]

## References

1. Appleton KM, Hemingway A, Saulais L, Dinnella C, Monteleone E, Depezay L, et al. Increasing vegetable intakes: rationale and systematic review of published interventions. *Eur J Nutr* 2016 Apr;55(3):869-896 [[FREE Full text](#)] [doi: [10.1007/s00394-015-1130-8](https://doi.org/10.1007/s00394-015-1130-8)] [Medline: [26754302](https://pubmed.ncbi.nlm.nih.gov/26754302/)]
2. Woodside JV, Young IS, McKinley MC. Fruits and vegetables: measuring intake and encouraging increased consumption. *Proc Nutr Soc* 2013 May;72(2):236-245. [doi: [10.1017/S0029665112003059](https://doi.org/10.1017/S0029665112003059)] [Medline: [23324158](https://pubmed.ncbi.nlm.nih.gov/23324158/)]
3. World Health Organization. Diet, Nutrition and the Prevention of Chronic Diseases, Who Technical Report Series 797. Geneva, Switzerland: World Health Organization; 1990.
4. World Health Organization. The World Health Report 2002: Reducing Risks to Health, Promoting Healthy Life. Geneva: World Health Organization; 2002.
5. World Health Organization. Diet, Nutrition and the Prevention of Chronic Diseases. WHO Technical Report Series 916. Report of a Joint WHO/FSA Expert Consultation. Geneva: WHO; 2003.
6. World Health Organization. Fruit and Vegetables for Health Report of a Joint FAO/WHO Workshop. Geneva: WHO; 2005.
7. Bazzano LA, He J, Ogden LG, Loria CM, Vupputuri S, Myers L, et al. Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. *Am J Clin Nutr* 2002 Jul;76(1):93-99. [doi: [10.1093/ajcn/76.1.93](https://doi.org/10.1093/ajcn/76.1.93)] [Medline: [12081821](https://pubmed.ncbi.nlm.nih.gov/12081821/)]
8. Tohill BC. Dietary Intake of Fruit and Vegetables and Management of Body Weight. Background Paper for the Joint FAO/WHO Workshop of Fruit and Vegetables for Health. Geneva: WHO; 2005.
9. Harding A, Wareham NJ, Bingham SA, Khaw K, Luben R, Welch A, et al. Plasma vitamin C level, fruit and vegetable consumption, and the risk of new-onset type 2 diabetes mellitus: the European prospective investigation of cancer--Norfolk prospective study. *Arch Intern Med* 2008 Jul 28;168(14):1493-1499. [doi: [10.1001/archinte.168.14.1493](https://doi.org/10.1001/archinte.168.14.1493)] [Medline: [18663161](https://pubmed.ncbi.nlm.nih.gov/18663161/)]
10. Riboli E, Norat T. Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. *Am J Clin Nutr* 2003 Sep;78(3 Suppl):559S-569S. [doi: [10.1093/ajcn/78.3.559S](https://doi.org/10.1093/ajcn/78.3.559S)] [Medline: [12936950](https://pubmed.ncbi.nlm.nih.gov/12936950/)]
11. Appleton KM, McGill R, Woodside JV. Fruit and vegetable consumption in older individuals in Northern Ireland: levels and patterns. *Br J Nutr* 2009 Oct;102(7):949-953. [doi: [10.1017/S0007114509332122](https://doi.org/10.1017/S0007114509332122)] [Medline: [19785930](https://pubmed.ncbi.nlm.nih.gov/19785930/)]
12. Bates B, Lennox A, Swan G. Assets: How They Work. National Diet and Nutrition Survey Headline Results from Years 1 and 2 (combined) of the Rolling Programme (2008/2009 – 2009/10) URL: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/216484/dh\\_128550.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216484/dh_128550.pdf) [accessed 2019-09-25]
13. European Food Safety Authority. The EFSA Comprehensive European Food Consumption Database URL: <http://www.efsa.europa.eu/en/datexfoodcdb/datexfooddb.htm> [accessed 2019-04-12]
14. United States Department of Agriculture. Food consumption Spreadsheets. URL: <http://www.ers.usda.gov/Data/FoodConsumption/FoodAvailSpreadsheets.htm> [accessed 2019-04-12]
15. Appleton KM, Krumpalevska K, Smith E, Rooney C, McKinley MC, Woodside JV. Low fruit and vegetable consumption is associated with low knowledge of the details of the 5-a-day fruit and vegetable message in the UK: findings from two cross-sectional questionnaire studies. *J Hum Nutr Diet* 2018 Feb;31(1):121-130. [doi: [10.1111/jhn.12487](https://doi.org/10.1111/jhn.12487)] [Medline: [28573733](https://pubmed.ncbi.nlm.nih.gov/28573733/)]
16. Appleton KM, McGill R, Neville C, Woodside JV. Barriers to increasing fruit and vegetable intakes in the older population of Northern Ireland: low levels of liking and low awareness of current recommendations. *Public Health Nutr* 2010 Apr;13(4):514-521. [doi: [10.1017/S1368980009991790](https://doi.org/10.1017/S1368980009991790)] [Medline: [20230659](https://pubmed.ncbi.nlm.nih.gov/20230659/)]

17. Ashfield-Watt PAL. Fruits and vegetables, 5+ a day: are we getting the message across? *Asia Pac J Clin Nutr* 2006;15(2):245-252 [FREE Full text] [Medline: [16672211](#)]
18. Dixon H, Borland R, Segan C, Stafford H, Sindall C. Public reaction to Victoria's '2 Fruit 'n' 5 Veg Every Day' campaign and reported consumption of fruit and vegetables. *Prev Med* 1998;27(4):572-582. [doi: [10.1006/pmed.1998.0328](#)] [Medline: [9672951](#)]
19. Buyuktuncer Z, Kearney M, Ryan CL, Thurston M, Ellahi B. Fruit and vegetables on prescription: a brief intervention in primary care. *J Hum Nutr Diet* 2014 Apr;27(Suppl 2):186-193. [doi: [10.1111/jhn.12109](#)] [Medline: [23651065](#)]
20. Glasson C, Chapman K, James E. Fruit and vegetables should be targeted separately in health promotion programmes: differences in consumption levels, barriers, knowledge and stages of readiness for change. *Public Health Nutr* 2011 Apr;14(4):694-701. [doi: [10.1017/S1368980010001643](#)] [Medline: [20576194](#)]
21. Pollard CM, Miller MR, Daly AM, Crouchley KE, O'Donoghue KJ, Lang AJ, et al. Increasing fruit and vegetable consumption: success of the Western Australian Go for 2&5 campaign. *Public Health Nutr* 2008 Mar;11(3):314-320. [doi: [10.1017/S1368980007000523](#)] [Medline: [17612423](#)]
22. Erinosh TO, Moser RP, Oh AY, Nebeling LC, Yaroch AL. Awareness of the Fruits and Veggies-More Matters campaign, knowledge of the fruit and vegetable recommendation, and fruit and vegetable intake of adults in the 2007 Food Attitudes and Behaviors (FAB) Survey. *Appetite* 2012 Aug;59(1):155-160. [doi: [10.1016/j.appet.2012.04.010](#)] [Medline: [22524998](#)]
23. Pollard CM, Daly AM, Binns CW. Consumer perceptions of fruit and vegetables serving sizes. *Public Health Nutr* 2009 May;12(5):637-643. [doi: [10.1017/S1368980008002607](#)] [Medline: [18547446](#)]
24. Herbert G, Butler L, Kennedy O, Lobb A. Young UK adults and the 5 A DAY campaign: perceived benefits and barriers to eating more fruits and vegetables. *Int J Cons Studies* 2010;34:657-664. [doi: [10.1111/j.1470-6431.2010.00872.x](#)]
25. Rooney C, McKinley MC, Appleton KM, Young IS, McGrath AJ, Draffin CR, et al. How much is '5-a-day'? A qualitative investigation into consumer understanding of fruit and vegetable intake guidelines. *J Hum Nutr Diet* 2017 Feb;30(1):105-113. [doi: [10.1111/jhn.12393](#)] [Medline: [27334026](#)]
26. Carter OB, Pollard CM, Atkins JF, Milliner JM, Pratt IS. 'We're not told why--we're just told': qualitative reflections about the Western Australian Go for 2&5® fruit and vegetable campaign. *Public Health Nutr* 2011 Jun;14(6):982-988. [doi: [10.1017/S1368980010003381](#)] [Medline: [21205404](#)]
27. Dibsall LA, Lambert N, Frewer LJ. Using interpretative phenomenology to understand the food-related experiences and beliefs of a select group of low-income UK women. *J Nutr Educ Behav* 2002;34(6):298-309. [doi: [10.1016/s1499-4046\(06\)60112-7](#)] [Medline: [12556268](#)]
28. Dixon H, Mullins R, Wakefield M, Hill D. Encouraging the consumption of fruit and vegetables by older Australians: an experiential study. *J Nutr Educ Behav* 2004;36(5):245-249. [Medline: [15707547](#)]
29. Guillaumie L, Godin G, Vézina-Im LA. Psychosocial determinants of fruit and vegetable intake in adult population: a systematic review. *Int J Behav Nutr Phys Act* 2010 Feb 2;7:12 [FREE Full text] [doi: [10.1186/1479-5868-7-12](#)] [Medline: [20181070](#)]
30. Shaikh AR, Yaroch AL, Nebeling L, Yeh M, Resnicow K. Psychosocial predictors of fruit and vegetable consumption in adults a review of the literature. *Am J Prev Med* 2008 Jun;34(6):535-543. [doi: [10.1016/j.amepre.2007.12.028](#)] [Medline: [18471592](#)]
31. Spronk I, Kullen C, Burdon C, O'Connor H. Relationship between nutrition knowledge and dietary intake. *Br J Nutr* 2014 May 28;111(10):1713-1726. [doi: [10.1017/S0007114514000087](#)] [Medline: [24621991](#)]
32. Pomerleau J, Lock K, Knai C, McKee M. Interventions designed to increase adult fruit and vegetable intake can be effective: a systematic review of the literature. *J Nutr* 2005 Oct;135(10):2486-2495. [doi: [10.1093/jn/135.10.2486](#)] [Medline: [16177217](#)]
33. Thomson CA, Ravia J. A systematic review of behavioral interventions to promote intake of fruit and vegetables. *J Am Diet Assoc* 2011 Oct;111(10):1523-1535. [doi: [10.1016/j.jada.2011.07.013](#)] [Medline: [21963019](#)]
34. Statista. 2018. Smartphone User Penetration as Percentage of Total Global Population From 2014 to 2021 URL: <https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/> [accessed 2018-11-27]
35. Hendrie GA, James-Martin G, Williams G, Brindal E, Whyte B, Crook A. The development of VegEze: smartphone app to increase vegetable consumption in Australian adults. *JMIR Form Res* 2019 Mar 27;3(1):e10731 [FREE Full text] [doi: [10.2196/10731](#)] [Medline: [30916653](#)]
36. Mandracchia F, Llauradó E, Tarro L, del Bas JM, Valls R, Pedret A, et al. Potential use of mobile phone applications for self-monitoring and increasing daily fruit and vegetable consumption: a systematized review. *Nutrients* 2019 Mar 22;11(3):686 [FREE Full text] [doi: [10.3390/nu11030686](#)] [Medline: [30909484](#)]
37. Mummah SA, King AC, Gardner CD, Sutton S. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. *Int J Behav Nutr Phys Act* 2016 Aug 8;13:90 [FREE Full text] [doi: [10.1186/s12966-016-0400-z](#)] [Medline: [27501724](#)]
38. Mummah S, Robinson TN, Mathur M, Farzinkhou S, Sutton S, Gardner CD. Effect of a mobile app intervention on vegetable consumption in overweight adults: a randomized controlled trial. *Int J Behav Nutr Phys Act* 2017 Sep 15;14(1):125 [FREE Full text] [doi: [10.1186/s12966-017-0563-2](#)] [Medline: [28915825](#)]

39. Nour M, Chen J, Allman-Farinelli M. Young adults' engagement with a self-monitoring app for vegetable intake and the impact of social media and gamification: feasibility study. *JMIR Form Res* 2019 May 10;3(2):e13324 [FREE Full text] [doi: [10.2196/13324](https://doi.org/10.2196/13324)] [Medline: [31094322](https://pubmed.ncbi.nlm.nih.gov/31094322/)]
40. Zhao JB, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. *J Med Internet Res* 2016 Oct 31;18(11):e287 [FREE Full text] [doi: [10.2196/jmir.5692](https://doi.org/10.2196/jmir.5692)] [Medline: [27806926](https://pubmed.ncbi.nlm.nih.gov/27806926/)]
41. Clegg D, Barker R. *Case Method Fast-Track: A RAD Approach*. Boston MA: Addison-Wesley Longman Publishing Co Inc; 1994.
42. Google Design. *Material Design Guidelines* URL: <https://design.google/resources/> [accessed 2019-04-12]
43. Cullum A. Increasing fruit and vegetable consumption: the 5 A DAY programme. *Nutr Bulletin* 2003 Jun;28(2):159-163. [doi: [10.1046/j.1467-3010.2003.00302.x](https://doi.org/10.1046/j.1467-3010.2003.00302.x)]
44. National Health Service. 5 A Day: what counts? URL: <http://www.nhs.uk/livewell/5aday/pages/5adayhome.aspx/> [accessed 2019-04-12]
45. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
46. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manag Sci* 1989 Aug;35(8):982-1003. [doi: [10.1287/mnsc.35.8.982](https://doi.org/10.1287/mnsc.35.8.982)]
47. van der Heijden H. User acceptance of Hedonic information systems. *MIS Quarterly* 2004;28(4):695-704. [doi: [10.2307/25148660](https://doi.org/10.2307/25148660)]
48. Margetts BM, Cade JE, Osmond C. Comparison of a food frequency questionnaire with a diet record. *Int J Epidemiol* 1989 Dec;18(4):868-873. [doi: [10.1093/ije/18.4.868](https://doi.org/10.1093/ije/18.4.868)] [Medline: [2621023](https://pubmed.ncbi.nlm.nih.gov/2621023/)]
49. Li P, Stuart E. Best (but oft-forgotten) practices: missing data methods in randomized controlled nutrition trials. *Am J Clin Nutr* 2019 Mar 1;109(3):504-508. [doi: [10.1093/ajcn/nqy271](https://doi.org/10.1093/ajcn/nqy271)] [Medline: [30793174](https://pubmed.ncbi.nlm.nih.gov/30793174/)]
50. Appleton KM. Greater fruit selection following an appearance-based compared with a health-based health promotion poster. *J Public Health (Oxf)* 2016 Dec 2;38(4):731-738 [FREE Full text] [doi: [10.1093/pubmed/fdv147](https://doi.org/10.1093/pubmed/fdv147)] [Medline: [28158693](https://pubmed.ncbi.nlm.nih.gov/28158693/)]
51. Bingham SA. The dietary assessment of individuals: Methods, accuracy, new techniques and recommendation. *Nutr. Abstr. Rev. (Ser. A)* 1987;57:705-737 [FREE Full text]
52. Mainvil LA, Horwath CC, McKenzie JE, Lawson R. Validation of brief instruments to measure adult fruit and vegetable consumption. *Appetite* 2011 Feb;56(1):111-117. [doi: [10.1016/j.appet.2010.09.008](https://doi.org/10.1016/j.appet.2010.09.008)] [Medline: [20970470](https://pubmed.ncbi.nlm.nih.gov/20970470/)]
53. Wolfe WS, Frongillo EA, Cassano PA. Evaluating brief measures of fruit and vegetable consumption frequency and variety: cognition, interpretation, and other measurement issues. *J Am Diet Assoc* 2001 Mar;101(3):311-318. [doi: [10.1016/S0002-8223\(01\)00081-5](https://doi.org/10.1016/S0002-8223(01)00081-5)] [Medline: [11269609](https://pubmed.ncbi.nlm.nih.gov/11269609/)]
54. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health* 2011 Feb 18;11:119 [FREE Full text] [doi: [10.1186/1471-2458-11-119](https://doi.org/10.1186/1471-2458-11-119)] [Medline: [21333011](https://pubmed.ncbi.nlm.nih.gov/21333011/)]
55. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
56. Teasdale N, Elhoussein A, Butcher F, Piernas C, Cowburn G, Hartmann-Boyce J, et al. Systematic review and meta-analysis of remotely delivered interventions using self-monitoring or tailored feedback to change dietary behavior. *Am J Clin Nutr* 2018 Feb 1;107(2):247-256 [FREE Full text] [doi: [10.1093/ajcn/nqx048](https://doi.org/10.1093/ajcn/nqx048)] [Medline: [29529158](https://pubmed.ncbi.nlm.nih.gov/29529158/)]
57. Brooke J. *SUS: A quick and dirty usability scale*. In: *Usability Evaluation in Industry*. London: Taylor & Francis; 1996.
58. Appleton KM, Hemingway A, Rajska J, Hartwell H. Repeated exposure and conditioning strategies for increasing vegetable liking and intake: systematic review and meta-analyses of the published literature. *Am J Clin Nutr* 2018 Oct 1;108(4):842-856 [FREE Full text] [doi: [10.1093/ajcn/nqy143](https://doi.org/10.1093/ajcn/nqy143)] [Medline: [30321277](https://pubmed.ncbi.nlm.nih.gov/30321277/)]
59. Hartmann-Boyce J, Jebb SA, Fletcher BR, Aveyard P. Self-help for weight loss in overweight and obese adults: systematic review and meta-analysis. *Am J Public Health* 2015 Mar;105(3):e43-e57. [doi: [10.2105/AJPH.2014.302389](https://doi.org/10.2105/AJPH.2014.302389)] [Medline: [25602873](https://pubmed.ncbi.nlm.nih.gov/25602873/)]

---

## Abbreviations

**ANOVA:** analysis of variance

**FFQ:** food frequency questionnaire

**FV:** fruit(s) and vegetable(s)

**MoSCoW:** Must have, Should have, Could have, Won't have

**TAM:** technology acceptance model

*Edited by C Dias; submitted 13.04.19; peer-reviewed by J Woodside, M Allman-Farinelli, J Nunez-Cordoba, J Kay, J Chen; comments to author 24.05.19; revised version received 18.07.19; accepted 23.08.19; published 20.11.19.*

*Please cite as:*

*Appleton KM, Passmore D, Burn I, Pidgeon H, Nation P, Boobyer C, Jiang N*

*An Interactive Mobile Phone App (SMART 5-A-DAY) for Increasing Knowledge of and Adherence to Fruit and Vegetable Recommendations: Development and Pilot Randomized Controlled Trial*

*JMIR Mhealth Uhealth 2019;7(11):e14380*

*URL: <http://mhealth.jmir.org/2019/11/e14380/>*

*doi: [10.2196/14380](https://doi.org/10.2196/14380)*

*PMID: [31746766](https://pubmed.ncbi.nlm.nih.gov/31746766/)*

©Katherine Marie Appleton, David Passmore, Isobel Burn, Hanna Pidgeon, Philippa Nation, Charlotte Boobyer, Nan Jiang. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 20.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Evaluation of Heart Failure Apps to Promote Self-Care: Systematic App Search

Sahr Wali<sup>1,2</sup>, MSc; Catherine Demers<sup>3,4</sup>, MSc, MD, FRCPC; Hiba Shah<sup>3</sup>, MSc; Huda Wali<sup>5</sup>, MSc; Delphine Lim<sup>3</sup>, MSc; Nirav Naik<sup>3</sup>, MSc; Ahmad Ghany<sup>3</sup>, MSc; Ayushi Vispute<sup>6</sup>, MHI; Maya Wali<sup>3</sup>; Karim Keshavjee<sup>2,7</sup>, MBA, MSc, MD, CCFP

<sup>1</sup>Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

<sup>2</sup>Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

<sup>3</sup>Department of Medicine, McMaster University, Hamilton, ON, Canada

<sup>4</sup>Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada

<sup>5</sup>School of Pharmacy, University of Waterloo, Kitchener, ON, Canada

<sup>6</sup>Centennial College, Toronto, ON, Canada

<sup>7</sup>InfoClin Inc, Toronto, ON, Canada

**Corresponding Author:**

Catherine Demers, MSc, MD, FRCPC

Department of Medicine

McMaster University

237 Barton St E

Hamilton, ON, L8L 2X2

Canada

Phone: 1 905 525 9140 ext 73324

Email: [demers@hhsc.ca](mailto:demers@hhsc.ca)

## Abstract

**Background:** Heart failure (HF) is a chronic disease that affects over 1% of Canadians and at least 26 million people worldwide. With the continued rise in disease prevalence and an aging population, HF-related costs are expected to create a significant economic burden. Many mobile health (mHealth) apps have been developed to help support patients' self-care in the home setting, but it is unclear if they are suited to the needs or capabilities of older adults.

**Objective:** This study aimed to identify HF apps and evaluate whether they met the criteria for optimal HF self-care.

**Methods:** We conducted a systematic search of all apps available exclusively for HF self-care across Google Play and the App Store. We then evaluated the apps according to a list of 25 major functions pivotal to promoting HF self-care for older adults.

**Results:** A total of 74 apps for HF self-care were identified, but only 21 apps were listed as being both HF and self-care specific. None of the apps had all 25 of the listed features for an adequate HF self-care app, and only 41% (31/74) apps had the key weight management feature present. HF Storylines received the highest functionality score (18/25, 72%).

**Conclusions:** Our findings suggest that currently available apps are not adequate for use by older adults with HF. This highlights the need for mHealth apps to refine their development process so that user needs and capabilities are identified during the design stage to ensure the usability of the app.

(*JMIR Mhealth Uhealth* 2019;7(11):e13173) doi:[10.2196/13173](https://doi.org/10.2196/13173)

**KEYWORDS**

mHealth; heart failure; self-care; mobile phone

## Introduction

Heart failure (HF) is the most important cardiovascular condition leading to hospitalization and rehospitalization in older adults, and it has a significant economic burden [1]. Despite an overall

decline in HF hospitalization rates, readmission rates remain high [2,3]. A systematic review found that HF readmissions can be reduced if patients with HF adopt self-care (hazard ratio [HR] 0.80; 95% CI 0.71-0.89) [3,4]. Specifically, weight monitoring has been identified as a pivotal component of HF self-care as weight gain has been independently associated with a poor

postdischarge prognosis, considering it is the last common step before worsening of clinical outcomes (HR per kg increase 1.16; 95% CI 1.09-1.23;  $P < .001$ ) [5]. However, many older adults find daily weight monitoring and adjusting diuretics to be challenging [6-8]. In addition to managing comorbid conditions, many older adults with HF exhibit mild cognitive impairment and poor medication adherence, both of which are associated with reduced ability to self-care [9-11]. To adequately promote HF self-care, strategies should be targeted to the patient's cognitive capabilities, learning needs, and literacy and numeracy levels [11].

Mobile health (mHealth) apps have been developed to support patients with self-care [12,13]. Unfortunately, although the initial uptake of mHealth apps looked promising, the majority of individuals have stopped using them because of reasons such as a loss of interest, manual data entry burden, and hidden costs [14,15]. Older adults do not commonly use mHealth apps because of the perception that they are not suited to their needs or capabilities [16]. This may explain the shortcomings of previous programs that have failed to promote self-care and utilize the opportunity to help decrease HF-related hospitalizations, deaths, and costs to health care systems [9-12].

Previous studies have reviewed current apps for HF self-care and found that there are limited number of apps available to support disease management [12,17]. Nevertheless, these studies were unable to effectively evaluate app quality because of their lack of disease specificity within the rating scale design [18,19]. For example, the commonly used Mobile Application Rating Scale (MARS) was able to provide an overall assessment of the quality of apps with respect to engagement, functionality, aesthetics, information, and subjective opinion, but it does not evaluate the usability or effectiveness of the app features specific for the disease population [12,15,18,19]. Therefore, generic health apps (eg, WebMD) receive higher app quality scores using the MARS even if they do not have crucial app features (eg, weight management) for proper self-care [17]. The lack of disease specificity can be attributed to the absence of a reference architecture to guide the scale's development. Other chronic disease app rating scales or checklists have been developed using similar constructs as the MARS, leaving them to face the same shortcomings with their app evaluations [20,21]. This highlights the need to further evaluate the adequacy of the current HF self-care apps available.

To address this gap, we conducted a systematic search of all the apps currently available exclusively for HF self-care. We used Chindalo et al's peer-reviewed mHealth app reference architecture to define the app design requirements [19]. Contrary to other rating scales, this architecture allows us to combine the evaluative components related to the aesthetics, usability, and HF self-care to effectively evaluate whether the current HF apps are meeting the end user's self-care needs and capabilities [19]. The objective of this study was to determine the number of HF apps available and evaluate whether they met the criteria to promote HF self-care.

## Methods

### Search Strategy

We conducted an extensive search across Google Play and the App Store to identify all available apps for HF self-care. The search was facilitated with the use of following key terms: HF management, HF manager, HF self-care, HF, and HF tracker. Apps were included in the review if they (1) were HF specific and (2) contained a self-care component (ie, medication, symptom management, reminder system, and behavior tracking). Apps were excluded if they were intended for use in a conference, for education, or for reference purposes.

### App Adequacy Assessment

In accordance with Chindalo et al's reference architecture, we developed a list of 25 major functions that would promote HF self-care for older adults [19,22-24] (Table 1). These features were identified according to HF self-care and patient-engagement guidelines [19] as well as the expertise from our clinician authors (CD and KK). Before the start of the study, a design session was conducted where CD (cardiologist/HF specialist) and KK (family physician/clinical information technology architect) created individual lists for potential app features. A second design session was conducted with CD, KK, and SW to finalize the list of app functions as well as the specific functions required for app adequacy. All 25 major functions were not deemed *required* to adequately promote HF self-care but were more beneficial if included. App adequacy was determined if the following standard disease management features were included: (1) diagnosis, (2) weight, (3) behavior tracking, (4) self-care, and (5) notifications. These 5 features were chosen based on their ability to capture factors related to HF management protocol, personalized care for older adults, and health promotion [25,26].

Within each of the 25 functions, a list of descriptors was developed to help specify the components within the listed function. If the app included 1 of the descriptors, the feature was listed as present (Table 1). For example, the self-care feature consisted of 3 components including self-maintenance, self-management, and self-confidence [6,7]. Self-maintenance includes actions associated with treatment adherence, such as taking medication or following treatment regimens. Self-management includes the recognition of and response to changes in symptoms. Finally, self-confidence refers to the individual's assurance in implementing necessary decisions during the management process. Self-confidence is not an explicit self-care behavior but has been recognized as an important moderator of self-care effectiveness [6,7,27]. Thus, self-confidence would not be captured in app functionality as clearly as self-maintenance and self-management, but instead could be expressed as a series of patient experience-related questions (Figure 1). Overall, if an app included any 1 of the 3 self-care components, the feature would be counted as being present.



**Table 1.** List of app features required for an adequate heart failure self-care app.

#	App feature	App descriptors
1	Prescribed	Physician prescribed for treatment; pharmacist recommendation
2	Diagnosis <sup>a</sup>	Patient predetermined diagnosis included (acute heart failure)
3	Patient demographics	Age; sex or gender; location
4	Patient sociocultural	Literacy; numeracy; socioeconomic status; culture/ethnicity; parental history
5	Patient symptoms	Shortness of breath; dizziness; orthopnea; leg edema or general swelling; paroxysmal nocturnal dyspnea
6	Patient behaviors	Smoking; exercise; fitness/movement; salt intake
7	Patient physiological observations	Heart rate; blood pressure; elevated jugular venous pressure; chest crackles; heart murmurs
8	Weight <sup>a</sup>	Management; monitoring; tracker
9	Comorbidities	Presence of other diseases (eg, diabetes and hypertension)
10	Drug list	List of medications
11	Laboratory results	Hemoglobin and hematocrit; creatinine and estimated glomerular filtration rate; brain natriuretic peptide; thyroid stimulating hormone; lipid profile
12	Diagnostic testing	Electrocardiogram; chest x-ray; echocardiogram
13	Behavior tracking <sup>a</sup>	Diet; exercise; patient-reported experience; compliance with medications
14	Education/recommendation	Behaviorally appropriate; culturally appropriate; health literacy appropriate; accredited/credible sources; evidence based
15	Self-care <sup>a</sup>	Self-maintenance; self-management: system provides patient with recommendation if clinical condition changes (eg, if weight increases, take extra Lasix); algorithm based or physician guidance; self-confidence
16	Health system utilization	Reviewed by family doctor; reviewed by nurse clinician, practitioner, or physician assistant; visit to ED <sup>b</sup> ; hospitalization; seen by specialist
17	Notifications <sup>a</sup>	Presence of reminder or notification
18	Integrations	Integrated into personal health record and electronic medical record; integrated into other health and fitness apps
19	Social supports	Connect/share results with caregiver or family; contact caregiver or family
20	Patient reported outcome measure/patient reported experience measure	Patient experience of care; app experience; quality of life; cognitive assessment; patient progress
21	Incentives to use	Easy access to provider; gamification; social aspect—connect with others
22	Predictive analytics	Length of stay, acuity of admission, comorbidities, ED visits, (readmissions); hospital admission risk prediction
23	Outcomes	Visit to family physician, specialist, or ED; hospitalized; death
24	Safety issues	Risk of falls; worsening kidney function; hyper- or hypokalemia
25	User interface	Easy to navigate functionality; simple to screen with minimal content on each page; features for visual (font size and color), hearing (audio cues), or general accessibility

<sup>a</sup>Standard disease management feature for heart failure.

<sup>b</sup>ED: emergency department.

**Figure 1.** Sample heart failure self-care confidence questions.

SECTION C: In general, how confident are you that you can:				
	Not confident	Somewhat confident	Very confident	Extremely confident
17. Keep yourself <u>free of heart failure symptoms</u> ?	1	2	3	4
18. <u>Follow the treatment advice</u> you have been given?	1	2	3	4
19. <u>Evaluate the importance</u> of your symptoms?	1	2	3	4
20. <u>Recognize changes</u> in your health if they occur?	1	2	3	4
21. <u>Do something</u> that will relieve your symptoms	1	2	3	4
22. <u>Evaluate</u> how well a remedy works?	1	2	3	4

### App Screening and Evaluation

A total of 2 reviewers (SW and KK) completed a preliminary screening of apps available in both Google Play and the App Store using the key terms mentioned previously. Following the search, apps were reviewed according to their title and summary description (Figure 1). This screening served as a method to separate the bulk of non-HF self-care apps from further evaluation.

Once screening was completed, a calibration session was held among the 8 reviewers (SW, KK, AG, AD, HW, HS, NN, and DL). During the calibration session, each reviewer was asked to evaluate and score the same 5 apps before the start of the session (Figure 1). Reviewers had a Google Sheet created for them with the list of functions for evaluation. During the session, the 5 apps were then reviewed on a summary sheet, and any discrepancies regarding scoring were discussed and resolved. This allowed us to standardize the training among all reviewers and ensure that we had at least an 85% agreement rate when evaluating the remaining apps. Additional information and comments from the calibration session were recorded and added to the final protocol. The evaluation sheet on Google Sheets was also revised for the final app evaluations.

After the calibration session, the remaining apps were assigned for evaluation, where 2 different reviewers evaluated each app. Each evaluation was completed on the revised Google Sheet with the respective app assignments. Once all the app evaluations were completed, the data were combined into a summary sheet for review. Each app score was then reviewed

by the 2 assigned reviewers to ensure that they were within the 85% agreement rate. To determine if the scores met the 85% agreement rate, we conducted interrater agreement statistic and reviewed the kappa value. If the apps did not meet the 85% agreement threshold, the 2 reviewers completed an in-person or virtual evaluation session to review the app discrepancies. Both reviewers were required to provide evidence (ie, screenshot or quote) to support the presence of the feature, and a discussion was held until consensus was achieved. All supporting evidence was sent to SW for a final review. Following consensus, a postreview screening was then completed to filter out any apps that were not HF or self-care specific. The remaining apps within the inclusion criteria were analyzed through a descriptive analysis to assess the app search's findings.

Reviewers evaluated each app based on the description, screenshots, videos, and reviews available on each app store website. Owing to the limited resources and to ensure a consistent method of app evaluation, we did not download the apps. Our rationale for not downloading apps was also based on the premise that users decide to download an app after reviewing it externally [28-31]. Many of the guidelines assisting patients with choosing a health app have urged users to become more meticulous with the apps they install and, in turn, have provided them with a series of questions to consider before downloading [29-31]. For example, in 2 articles, they suggested users consider the following questions before downloading an app: (1) does the app consider your needs (ie, symptoms and disease management), (2) is the app made by a health care system/physician or by a controversial company (ie, pharmaceutical company), (3) does the app have positive

reviews (ie, on the Web or by physicians), (4) is the app regularly updated, and (5) are the app features relevant for you (ie, review screenshots and description of features) [29,30]. Currently, about 90% of most information for decision making about app adoption is available in its documentation [29,32,33]. Given the minimal incremental information available from the app itself, we felt that downloading the apps would not significantly change our evaluation.

In accordance with the guidelines for app review before download, reviewers also extracted the following data from each app: number of downloads, date of last update, cost, and developer [29,30]. However, considering that apps from the App Store do not publicly list their number of downloads or date of last update, reviewers omitted the number of downloads criteria for these apps and used the latest version date for the last update.

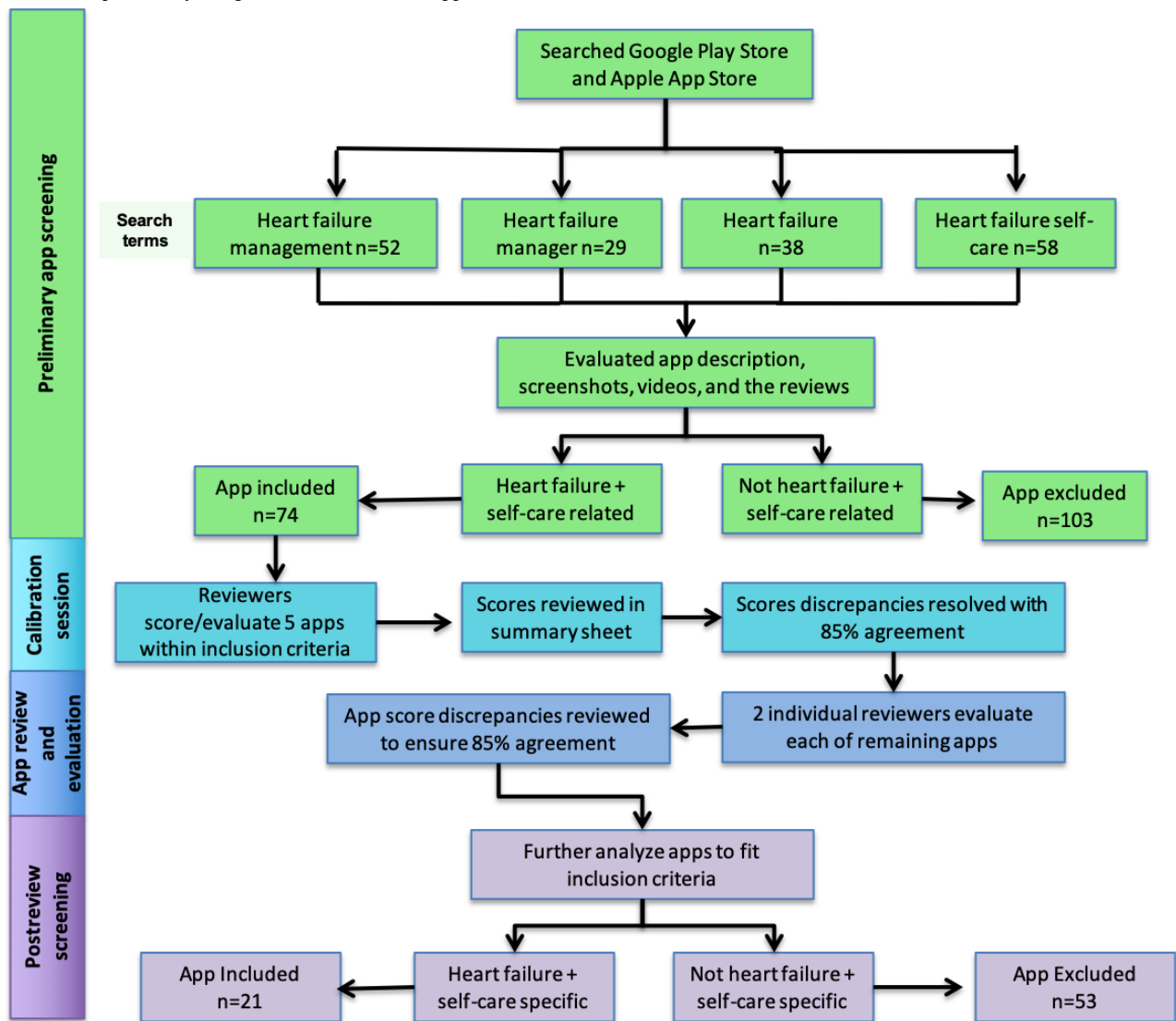
### Reviewer Training

Each reviewer selected was equipped with postsecondary experience in the electronic health or health technology field to allow them to effectively evaluate the respective apps. Reviewers were required to follow the training protocol in accordance with the mHealth design architecture as well as attend the calibration session previously described [19].

### Results

Preliminary screening identified a total of 74 apps as HF self-care apps within the combined app store searches (Figure 2). From these 74 apps, none of the apps had all the 25 listed features required to promote HF self-care, and only 32% (24/74) apps had 10 features or more present. Moreover, only 51 out of the 74 apps had a self-care feature present. Instead, the majority of the apps reviewed were used mainly for education purposes (Table 2).

Figure 2. Conceptual study design of the mobile health app review.



**Table 2.** Features present in the reviewed heart failure self-care apps (N=74).

App feature	Apps, n (%)
Education/recommendations	67 (90)
Self-care	51 (68)
User interface	51 (68)
Diagnosis	40 (54)
Notifications	33 (44)
Weight	31 (41)
Patient demographics	28 (37)
Patient symptoms	27 (36)
Patient physiological observations	26 (35)
Behavior tracking	22 (29)
Patient behaviors	19 (25)
Drug list	18 (24)
Patient reported outcome measure/patient reported experience measure	18 (24)
Incentives to use	18 (24)
Comorbidities	9 (12)
Lab results	9 (12)
Diagnostic testing	9 (12)
Social supports	9 (12)
Prescribed	8 (10)
Health system utilization	8 (10)
Outcomes	7 (9)
Predictive analytics	6 (8)
Patient socio-cultural	6 (8)
Safety issues	0 (0)

Following the postreview screening, 21 apps were listed as both HF and self-care specific. Moreover, 53 apps were excluded for the following reasons: (1) HF specific but not for self-care (n=9), (2) used for self-care but not specifically for HF (n=16), and (3) neither HF nor self-care specific but used for general cardiac education (n=28). As there is an increasing number of apps for entertainment or novelty purposes, from the 53 apps excluded, 12 were shortlisted for this reason. From the 21 HF self-care apps included, more than 50% (12/21, 57%) of the apps had 10 features or more (Table 3; Multimedia Appendix 1). The apps scores ranged from 0 to 18, where HF Storylines achieved the highest score. The most prevalent feature among these apps was diagnosis and user interface (20/21, 95%; Table 3). However, from the included apps, only 1 was found to have been evaluated by patients or investigated in a clinical setting

(Medly). All reviewer's app scores fell within the 85% agreement rate (mean kappa=0.86).

Many apps include features such as patient demographics, patient symptoms, education, self-care, notifications, and, most notably, weight. Unexpectedly, less than 50% (9/21) of the apps included a patient behavior or a behavior tracking feature, both of which are vital for adequate HF self-care (Table 3). Only 11 of the apps included a drug list within the app to keep track of medications or, more specifically, the usage of diuretics. In addition, a limited number of apps included social support (5) or were by prescription (4), and even fewer included a patient sociocultural (3) or comorbidities (2) feature. None of the reviewed apps included diagnostic testing, predictive analytics, outcomes, or safety measures (Table 3).

**Table 3.** Features present in filtered heart failure self-care apps (N=21).

App feature	Apps, n (%)
Diagnosis	20 (95)
User interface	20 (95)
Self-care	19 (90)
Notifications	18 (85)
Education/recommendations	17 (80)
Weight	17 (80)
Patient demographics	13 (61)
Patient symptoms	13 (61)
Patient physiological observations	11 (52)
Drug-list	11 (52)
Behavior tracking	10 (47)
Incentives to use	10 (47)
Patient behaviors	9 (42)
Integrations	8 (38)
Patient reported outcome measure/patient reported experience measure	8 (38)
Social supports	5 (23)
Health system utilization	4 (19)
Prescribed	4 (19)
Patient sociocultural	3 (14)
Lab results	3 (14)
Comorbidities	2 (9)
Diagnostic testing	0 (0)
Predictive analytics	0 (0)
Outcomes	0 (0)
Safety issues	0 (0)

**Table 4** displays a list of the total score of the filtered HF self-care apps with their respective app characteristics. Of the apps available, only 9 listed their number of downloads, and among these apps, the number of downloads varied with the total app scores. Surprisingly, the lowest-scoring app had a higher number of downloads compared with the highest-scoring app. This discrepancy could be linked to more effective marketing strategies, public brand awareness, or a well-regarded national health organization for certain apps [31,32].

With respect to cost, the majority of apps could be downloaded for free; however, 2 apps had an associated cost. For consistency, apps were not downloaded. As a result, we found that the 2 apps with a download cost had relatively lower scores (score of 5=US \$50 and score of 8=US \$7) and did not list their number of downloads.

Each app's last update varied from 2013 to 2019; however, most of the recently updated apps received higher scores. One of the most recently updated apps (HF Storylines) obtained the highest total app score of 18.

**Table 4.** Total score of the filtered heart failure self-care apps and their corresponding number of downloads, last updates, and cost (N=21).

App name	Total app score	Number of downloads	Last updated	Cost (US \$)	Developer
Heart Failure	2	1000-5000	December 22, 2014	0	Leon Do
HF Coach	5	__ <sup>a</sup>	April 29, 2016	50	Etectera Edutainment Inc
HF Defender	7	1000-5000	November 10, 2016	0	Cardio Fortress
HF Buddy	8	—	June 08, 2016	0	Singapore Health Services
HF Monitoring	8	—	February 22, 2017	0	Van Phuc Nguyen
HF Tracker	8	—	October 31, 2014	0	Rebecca Boxer
HF Path	8	—	January 04, 2017	7	American Heart Association
HF Log	9	—	February 09, 2016	0	Narnar LLC
HF Buddy	9	500-1000	April 08, 2016	0	—
Heart Scribe	10	50-100	July 30, 2016	0	Rohan Tanjea
Heart Failure Health	11	—	—	0	Self-Care Catalyst Inc
Health Plus	12	—	May 12, 2016	0	Hany Assaad
Heart Lessons	12	—	Apr 17, 2017	0	Palo Alto Medical Foundation for Health Care, Research and Education
My HF	12	1000-5000	September 22, 2016	0	Les Laboratoires Servier
WOW ME 200mg	12	100-500	July 24, 2013	0	AtantiCare Regional Medical Center Inc
HF Self- Management	13	—	August 11, 2015	0	—
Pulsario	14	—	October 07, 2016	0	Cardio Fortress Inc
My Heart Mate	14	—	November 14, 2016	0	Elevator Entertainment
Heart Partner	16	—	October 07, 2016	0	Novartis Pharmaceuticals
Medly	17	—	Ongoing	—	University Health Network
HF Storylines	18	500-1000	March 10, 2017	0	HF Society of America

<sup>a</sup>Not available.

## Discussion

### Principal Findings

Self-care is pivotal for HF patients to prevent worsening of HF, yet the majority of current HF apps available are neither HF or self-care specific. To our knowledge, this is the first study to identify and evaluate apps exclusively for HF self-care. We found 21 apps that were both HF and self-care specific. From the 21 apps, few contained key features such as behavior tracking. Apps that included the self-care feature were also listed as only being capable of self-maintenance. Thus, patients would be able to, at most, follow their treatment regimen but would not be able to respond to any changes. Potential features to expand on self-care could include medication titration algorithms to adjust medication doses according to weight fluctuations or the use of telehealth services to connect with a physician to modify their treatment regime [34,35]. Loop diuretics are currently used as the agent of choice for reducing symptoms of HF and controlling weight. Diuretics are traditionally adjusted by physicians. However, with self-directed medication titration becoming more commonly used for chronic disease management, this feature could be an opportunity to improve patient HF self-care in the home setting [34,35].

Our findings suggest that the current available apps are not able to support patients adequately with HF self-care; instead, are in need of further redesign or development. Many developers may have limited resources to accommodate all 25 features in a single app. Therefore, to appropriately engage patients in self-care, apps should at minimum have the following functions: (1) diagnosis, (2) weight, (3) behavior tracking, (4) self-care, and (5) notifications. However, from our systematic search, none of the apps even had these 5 functional features. Not only are these app features key for HF self-care but they can also be easily transferable to other health conditions, such as diabetes or asthma, as it captures the essential components for treatment management. The specifics detailing each feature will differ depending on the condition, but it provides a sufficient baseline category to allow the consumer, researcher, or clinician to incorporate key components for the app evaluation. The consumer, researcher, or clinician may list similar or different components within the 5 features, but with this, they are able to incorporate their perspective within 5 wider categories, while maintaining its relevance for multiple audiences. A prime example of an effective mHealth app is BlueStar from WellDoc Diabetes Management. The BlueStar app is a digital therapeutic for diabetes mellitus type 2 that serves as a virtual coach for patients, providing tailored guidance and facilitating the

coordination of diabetes care with their existing care team [36]. BlueStar is a clinically validated tool developed by endocrinologists and clinical diabetes educators and has been evaluated in a clinical trial and reviewed in over 40 publications [37,38]. Patients who used BlueStar showed significant improvements in their diabetes management and reported a high satisfaction when using the app. These improvements are strongly linked to the diabetes expertise leading the development and evaluation of this tool, as their guidance helps ensure that the intervention is aligned with self-management principles integral to patient care [36,37]. Older adults with chronic disease already face many challenges with managing their condition, and the use of technology may further contribute to their difficulties if poorly designed. To ensure apps assist with patient self-care regimens, they should be developed in a manner similar to BlueStar, where specific disease management and patient usability criteria are used to both design and evaluate app effectiveness [37].

One surprising finding from this app search was that the lowest-scoring apps had a relatively higher number of downloads compared with the highest-scoring app (Table 4). These findings ultimately question whether the inclusion of more app features or just the standard features for disease management are more appealing for the end user. Features that patients and consumers view as valuable can vary depending on their self-care abilities. However, many studies have indicated that the primary reason apps fail to maintain user activity is because of the complexity of the app as a whole or the lack of growth in app functionality in accordance with user needs [39,40]. Thus, as effective self-care promotion is the primary goal of app usage, apps would require the inclusion of the standard disease management features, but their presentation should also be modified to accommodate for the challenges older adults face with app use. In previous literature, older adults have indicated that in-app customizable considerations improved the likelihood of their continued usage as they were able to modify their preferences according to their changing needs [20,21]. This could include characteristics such as customizing screen or font sizes and incorporating text to speech, audio cues, or in-app automation features as needed [21,22].

Nevertheless, it is also important to note that the fact that lower-scoring apps had higher downloads can also be attributed to several external factors promoting public app awareness. This includes factors such as effective marketing, links to a national health body, or the use of Web-based search engine optimization [29,32]. From our findings, we found that the app with the highest downloads (My HF) had a moderate score of 12, but it was developed by a privately owned pharmaceutical company that specializes in medication for cardiological conditions (Les Laboratories Servier). Although apps built by health care systems scored the highest, they had much lower download rates (Table 4). With these discrepancies between app scores and the number of downloads, our findings display how higher

downloads may not be an appropriate representation for app effectiveness because of its potential ties to a developer with more marketing power.

### Limitations

The limitations of our study were as follows: (1) the apps were not downloaded but were reviewed based on app description, screenshots, videos, and reviews from current/past users, and there are good reasons to believe that the quality of the assessment is not severely compromised, as mentioned above; (2) descriptions for review varied in detail and quality (eg, download data not available for Apple iOS apps); (3) we were only able to review the number of downloads but could not quantify active app use; and (4) actual HF patients were not consulted to define the criteria for adequate HF self-care app.

### Future Research

Future studies should involve end users to better understand their needs with the design of an app to ensure the uptake and usability of an intervention. Specifically, the use of engagement strategies with HF patients and health care providers would be strongly desirable to ensure the findings of this study are congruent with what is experienced in reality. In our app evaluation, we incorporated app user reviews to assess user perspectives, but this method is limited to the feedback available and the quality of the responses on the Web. This study also did not evaluate the value of the 5 functional criteria for HF self-care compared with the remaining categories. Future studies should aim to understand the relative importance of each criteria in relation to patient outcomes, potentially with the use of focus groups or user testing to develop priority weightings for each function. In addition to this, as the potential rise in scale modification for disease specificity could lead to inappropriate features being selected for app shortlisting, there is a need to also evaluate the priority features in relation to other common chronic conditions (ie, diabetes and asthma). We believe there is value in the inclusion of the 5 minimum features for app evaluation, as it allows for specific components to be embedded within wider priority categories and provides a baseline mode to manage the use of multiple modified scales. However, before any scales can be managed or reviewed, future studies need to confirm the reliability of the 5 features for the management of priority app categories.

### Conclusions

In summary, our study was the first to specifically evaluate HF self-care apps according to the criteria essential to promote HF self-care for older adults. We found that there was a lack of usable apps to promote HF self-care for older adults, and this is mainly because of the lack of a patient-centered design. With a rise in the aging population, identifying features pivotal for patient self-care will be crucial to increase their user experience and ensure the longevity of the app's use.

## Acknowledgments

The authors would like to thank Everett McKay for providing his ongoing expertise throughout the design and conduct of this project.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Total features present in preliminary and postreview screening apps. PROMS: patient reported outcome measure; PREMS: patient reported experience measure.

[[PDF File \(Adobe PDF File\), 335 KB - mhealth\\_v7i11e13173\\_app1.pdf](#)]

## References

1. Heart and Stroke Foundation of Canada. 2017. 2016 Report on the Health of Canadians: The Burden of Heart Failure URL: <https://www.heartandstroke.ca/-/media/pdf-files/canada/2017-heart-month/heartandstroke-repertonhealth-2016.ashx?la=en> [accessed 2019-09-16]
2. Gandhi S, Chen S, Hong L, Sun K, Gong E, Li C, et al. Effect of mobile health interventions on the secondary prevention of cardiovascular disease: systematic review and meta-analysis. *Can J Cardiol* 2017 Feb;33(2):219-231. [doi: [10.1016/j.cjca.2016.08.017](https://doi.org/10.1016/j.cjca.2016.08.017)] [Medline: [27956043](https://pubmed.ncbi.nlm.nih.gov/27956043/)]
3. Cajita MI, Gleason KT, Han H. A systematic review of mhealth-based heart failure interventions. *J Cardiovasc Nurs* 2016;31(3):E10-E22 [FREE Full text] [doi: [10.1097/JCN.0000000000000305](https://doi.org/10.1097/JCN.0000000000000305)] [Medline: [26544175](https://pubmed.ncbi.nlm.nih.gov/26544175/)]
4. Lycholip E, Aamodt IT, Lie I, Šimbelytė T, Puronaitė R, Hillege H, et al. The dynamics of self-care in the course of heart failure management: data from the IN TOUCH study. *Patient Prefer Adherence* 2018;12:1113-1122 [FREE Full text] [doi: [10.2147/PPA.S162219](https://doi.org/10.2147/PPA.S162219)] [Medline: [29983549](https://pubmed.ncbi.nlm.nih.gov/29983549/)]
5. Ambrosy AP, Cerbin LP, Armstrong PW, Butler J, Coles A, DeVore AD, et al. Body weight change during and after hospitalization for acute heart failure: patient characteristics, markers of congestion, and outcomes: findings from the ASCEND-HF trial. *JACC Heart Fail* 2017 Jan;5(1):1-13 [FREE Full text] [doi: [10.1016/j.jchf.2016.09.012](https://doi.org/10.1016/j.jchf.2016.09.012)] [Medline: [28034373](https://pubmed.ncbi.nlm.nih.gov/28034373/)]
6. Riegel B, Barbaranelli C, Carlson B, Sethares KA, Daus M, Moser DK, et al. Psychometric testing of the revised self-care of heart failure index. *J Cardiovasc Nurs* 2019;34(2):183-192. [doi: [10.1097/JCN.0000000000000543](https://doi.org/10.1097/JCN.0000000000000543)] [Medline: [30303894](https://pubmed.ncbi.nlm.nih.gov/30303894/)]
7. Sedlar N, Lainscak M, Mårtensson J, Strömberg A, Jaarsma T, Farkas J. Factors related to self-care behaviours in heart failure: a systematic review of European Heart Failure Self-Care Behaviour Scale studies. *Eur J Cardiovasc Nurs* 2017 Apr;16(4):272-282. [doi: [10.1177/1474515117691644](https://doi.org/10.1177/1474515117691644)] [Medline: [28168895](https://pubmed.ncbi.nlm.nih.gov/28168895/)]
8. Spaling MA, Currie K, Strachan PH, Harkness K, Clark AM. Improving support for heart failure patients: a systematic review to understand patients' perspectives on self-care. *J Adv Nurs* 2015 Nov;71(11):2478-2489. [doi: [10.1111/jan.12712](https://doi.org/10.1111/jan.12712)] [Medline: [26084885](https://pubmed.ncbi.nlm.nih.gov/26084885/)]
9. Cannon JA, Moffitt P, Perez-Moreno AC, Walters MR, Broomfield NM, McMurray JJ, et al. Cognitive impairment and heart failure: systematic review and meta-analysis. *J Card Fail* 2017 Jun;23(6):464-475. [doi: [10.1016/j.cardfail.2017.04.007](https://doi.org/10.1016/j.cardfail.2017.04.007)] [Medline: [28433667](https://pubmed.ncbi.nlm.nih.gov/28433667/)]
10. Mlynarska A, Golba KS, Mlynarski R. Capability for self-care of patients with heart failure. *Clin Interv Aging* 2018;13:1919-1927 [FREE Full text] [doi: [10.2147/CIA.S178393](https://doi.org/10.2147/CIA.S178393)] [Medline: [30349210](https://pubmed.ncbi.nlm.nih.gov/30349210/)]
11. Clark AM, Spaling M, Harkness K, Spiers J, Strachan PH, Thompson DR, et al. Determinants of effective heart failure self-care: a systematic review of patients' and caregivers' perceptions. *Heart* 2014 May;100(9):716-721. [doi: [10.1136/heartjnl-2013-304852](https://doi.org/10.1136/heartjnl-2013-304852)] [Medline: [24548920](https://pubmed.ncbi.nlm.nih.gov/24548920/)]
12. Riegel B. I forgot: memory and medication adherence in heart failure. *Circ Heart Fail* 2016 Dec;9(12):e003642. [doi: [10.1161/CIRCHEARTFAILURE.116.003642](https://doi.org/10.1161/CIRCHEARTFAILURE.116.003642)] [Medline: [27895070](https://pubmed.ncbi.nlm.nih.gov/27895070/)]
13. 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](https://doi.org/10.2196/mhealth.2737)] [Medline: [25098320](https://pubmed.ncbi.nlm.nih.gov/25098320/)]
14. Research2Guidance. 2015. mHealth App Developer Economics 2015 URL: <https://research2guidance.com/product/mhealth-developer-economics-2015/> [accessed 2018-07-15]
15. Cajita MI, Hodgson NA, Budhathoki C, Han HR. Intention to use mhealth in older adults with heart failure. *J Cardiovasc Nurs* 2017;32(6):E1-E7 [FREE Full text] [doi: [10.1097/JCN.0000000000000401](https://doi.org/10.1097/JCN.0000000000000401)] [Medline: [28248747](https://pubmed.ncbi.nlm.nih.gov/28248747/)]
16. Hägglund E, Lyngå P, Frie F, Ullman B, Persson H, Melin M, et al. Patient-centred home-based management of heart failure. Findings from a randomised clinical trial evaluating a tablet computer for self-care, quality of life and effects on knowledge. *Scand Cardiovasc J* 2015 Aug;49(4):193-199. [doi: [10.3109/14017431.2015.1035319](https://doi.org/10.3109/14017431.2015.1035319)] [Medline: [25968968](https://pubmed.ncbi.nlm.nih.gov/25968968/)]



17. 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](https://doi.org/10.2196/mhealth.5882)] [Medline: [27302310](https://pubmed.ncbi.nlm.nih.gov/27302310/)]
18. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
19. Chindalo P, Keshavjee K, Karim A, Brahmhatt R, Saha N. Health apps by design: a reference architecture for mobile engagement. *Int J Handheld Comput Res* 2016;7(2):34-43. [doi: [10.4018/IJHCR.2016040103](https://doi.org/10.4018/IJHCR.2016040103)]
20. Anderson K, Burford O, Emmerton L. App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management. *JMIR Res Protoc* 2016 Nov 4;5(4):e204 [FREE Full text] [doi: [10.2196/resprot.6194](https://doi.org/10.2196/resprot.6194)] [Medline: [27815233](https://pubmed.ncbi.nlm.nih.gov/27815233/)]
21. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP health informatics unit checklist. *Clin Med (Lond)* 2015 Dec;15(6):519-521 [FREE Full text] [doi: [10.7861/clinmedicine.15-6-519](https://doi.org/10.7861/clinmedicine.15-6-519)] [Medline: [26621937](https://pubmed.ncbi.nlm.nih.gov/26621937/)]
22. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. *PLoS One* 2016;11(5):e0156164 [FREE Full text] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]
23. Hundert AS, Huguet A, McGrath PJ, Stinson JN, Wheaton M. Commercially available mobile phone headache diary apps: a systematic review. *JMIR Mhealth Uhealth* 2014 Aug 19;2(3):e36 [FREE Full text] [doi: [10.2196/mhealth.3452](https://doi.org/10.2196/mhealth.3452)] [Medline: [25138438](https://pubmed.ncbi.nlm.nih.gov/25138438/)]
24. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. User-centered adaptation of an existing heart failure telemonitoring program to ensure sustainability and scalability: qualitative study. *JMIR Cardio* 2018 Dec 6;2(2):e11466. [doi: [10.2196/11466](https://doi.org/10.2196/11466)]
25. Woods LS, Duff J, Roehrer E, Walker K, Cummings E. Patients' experiences of using a consumer mhealth app for self-management of heart failure: mixed-methods study. *JMIR Hum Factors* 2019 May 2;6(2):e13009 [FREE Full text] [doi: [10.2196/13009](https://doi.org/10.2196/13009)] [Medline: [31045504](https://pubmed.ncbi.nlm.nih.gov/31045504/)]
26. Lefler LL, Rhoads SJ, Harris M, Funderburg AE, Lubin SA, Martel ID, et al. Evaluating the use of mobile health technology in older adults with heart failure: mixed-methods study. *JMIR Aging* 2018 Dec 4;1(2):e12178 [FREE Full text] [doi: [10.2196/12178](https://doi.org/10.2196/12178)] [Medline: [31518257](https://pubmed.ncbi.nlm.nih.gov/31518257/)]
27. Riegel B, Lee S, Hill J, Daus M, Baah FO, Wald JW, et al. Patterns of adherence to diuretics, dietary sodium and fluid intake recommendations in adults with heart failure. *Heart Lung* 2019;48(3):179-185. [doi: [10.1016/j.hrtlng.2018.12.008](https://doi.org/10.1016/j.hrtlng.2018.12.008)] [Medline: [30638609](https://pubmed.ncbi.nlm.nih.gov/30638609/)]
28. Mobile Application Development Companies UK. 2016. How Do People Discover And Decide To Download Your App? URL: <https://www.overpass.co.uk/people-discover-decide-download-app/> [accessed 2019-09-16]
29. Aterja A. HuffPost. 2018. How to Choose the Best Health App for You URL: [https://www.huffpost.com/entry/how-to-choose-the-best-health-app-for-you\\_b\\_5a54d736e4b0ee59d41c0e01](https://www.huffpost.com/entry/how-to-choose-the-best-health-app-for-you_b_5a54d736e4b0ee59d41c0e01) [accessed 2019-09-16]
30. Armstrong S. Which app should I use? *Br Med J* 2015 Sep 9;351:h4597. [doi: [10.1136/bmj.h4597](https://doi.org/10.1136/bmj.h4597)] [Medline: [26353800](https://pubmed.ncbi.nlm.nih.gov/26353800/)]
31. Visser BJ, Bouman J. There's a medical app for that. *Br Med J* 2012 Apr 18;344:e2162. [doi: [10.1136/sbmj.e2162](https://doi.org/10.1136/sbmj.e2162)]
32. Tjongson J. Think With Google. 2015. Mobile App Marketing Insights: How Consumers Really Find and Use Your Apps URL: <https://www.thinkwithgoogle.com/consumer-insights/mobile-app-marketing-insights/> [accessed 2019-09-16]
33. Basilico A, Marceglia S, Bonacina S, Pinciroli F. Advising patients on selecting trustful apps for diabetes self-care. *Comput Biol Med* 2016 Apr 1;71:86-96. [doi: [10.1016/j.combiomed.2016.02.005](https://doi.org/10.1016/j.combiomed.2016.02.005)] [Medline: [26897071](https://pubmed.ncbi.nlm.nih.gov/26897071/)]
34. Riegel B, Moser DK, Anker SD, Appel LJ, Dunbar SB, Grady KL, et al. State of the science. *Circulation* 2009 Sep 22;120(12):1141-1163. [doi: [10.1161/circulationaha.109.192628](https://doi.org/10.1161/circulationaha.109.192628)]
35. Kuljis J, Money AG, Perry M, Barnett J, Young T. Technology-assisted self-testing and management of oral anticoagulation therapy: a qualitative patient-focused study. *Scand J Caring Sci* 2017 Sep;31(3):603-617. [doi: [10.1111/scs.12374](https://doi.org/10.1111/scs.12374)] [Medline: [27500499](https://pubmed.ncbi.nlm.nih.gov/27500499/)]
36. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. *Diabetes Technol Ther* 2008 Jun;10(3):160-168. [doi: [10.1089/dia.2008.0283](https://doi.org/10.1089/dia.2008.0283)] [Medline: [18473689](https://pubmed.ncbi.nlm.nih.gov/18473689/)]
37. Desveaux L, Shaw J, Saragosa M, Soobiah C, Marani H, Hensel J, et al. A mobile app to improve self-management of individuals with type 2 diabetes: qualitative realist evaluation. *J Med Internet Res* 2018 Mar 16;20(3):e81 [FREE Full text] [doi: [10.2196/jmir.8712](https://doi.org/10.2196/jmir.8712)] [Medline: [29549070](https://pubmed.ncbi.nlm.nih.gov/29549070/)]
38. WellDoc Inc. Bluestar Digital Therapeutics Assistant | WellDoc Wellness URL: <https://www.welldoc.com/product/> [accessed 2017-07-12]
39. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. *JMIR Mhealth Uhealth* 2015 May 13;3(2):e40 [FREE Full text] [doi: [10.2196/mhealth.4283](https://doi.org/10.2196/mhealth.4283)] [Medline: [25972309](https://pubmed.ncbi.nlm.nih.gov/25972309/)]
40. Liew MS, Zhang J, See J, Ong YL. Usability challenges for health and wellness mobile apps: mixed-methods study among mhealth experts and consumers. *JMIR Mhealth Uhealth* 2019 Jan 30;7(1):e12160 [FREE Full text] [doi: [10.2196/12160](https://doi.org/10.2196/12160)] [Medline: [30698528](https://pubmed.ncbi.nlm.nih.gov/30698528/)]

## Abbreviations

**HF:** heart failure

**HR:** hazard ratio

**mHealth:** mobile health

**MARS:** Mobile Application Rating Scale

*Edited by G Eysenbach; submitted 17.12.18; peer-reviewed by E McKay, M Nomali, K Carlsen, M Valero, K Anderson, H Mehdizadeh, M Jones, R Webster, V Strotbaum; comments to author 17.04.19; revised version received 11.06.19; accepted 19.08.19; published 11.11.19.*

*Please cite as:*

*Wali S, Demers C, Shah H, Wali H, Lim D, Naik N, Ghany A, Vispute A, Wali M, Keshavjee K*

*Evaluation of Heart Failure Apps to Promote Self-Care: Systematic App Search*

*JMIR Mhealth Uhealth 2019;7(11):e13173*

*URL: <https://mhealth.jmir.org/2019/11/e13173>*

*doi: [10.2196/13173](https://doi.org/10.2196/13173)*

*PMID: [31710298](https://pubmed.ncbi.nlm.nih.gov/31710298/)*

©Sahr Wali, Catherine Demers, Hiba Shah, Huda Wali, Delphine Lim, Nirav Naik, Ahmad Ghany, Ayushi Vispute, Maya Wali, Karim Keshavjee. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study

Miguel Angel Mayer<sup>1\*</sup>, MPH, PhD, MD; Octavi Rodríguez Blanco<sup>2\*</sup>, RNC, MSci; Antonio Torrejon<sup>2\*</sup>, RNC, BSciS, MSci

<sup>1</sup>Research Programme on Biomedical Informatics. Hospital del Mar Medical Research Institute, Universitat Pompeu Fabra, Barcelona, Spain

<sup>2</sup>Col·legi Oficial d'Infermeres i Infermers de Barcelona, Barcelona, Spain

\*all authors contributed equally

**Corresponding Author:**

Antonio Torrejon, RNC, BSciS, MSci  
Col·legi Oficial d'Infermeres i Infermers de Barcelona  
Pujades 350  
Barcelona, 08019  
Spain  
Phone: 34 932128108  
Email: [atorrejon@coib.cat](mailto:atorrejon@coib.cat)

**Related Article:**

This is a corrected version. See correction statement: <https://mhealth.jmir.org/2020/4/e18516/>

## Abstract

**Background:** In the last few years, the number of mobile apps for health professionals has increased exponentially. Nevertheless, there is a lack of knowledge about the professional use, training requirements, and quality perception of these apps among health care professionals such as nurses. Considering that the nursing profession is the largest segment of health care workforce in many countries such as Spain, the impact of the use of health apps by these professionals can be critical to the future of modern health care.

**Objective:** The main objective of this study was to determine if nurses were using health apps professionally and what types of apps they were using. The secondary objectives were (1) to find out if, among nurses, there is a need for training in the use of health apps and (2) to explore nurses' perceptions of health professional apps, determining whether there is a need for a certification process for health apps and the type of institution or organization that should review and validate these apps for professional use.

**Methods:** After an initial piloting survey, all registered nurses at the Nursing Association of Barcelona were invited to participate in a 34-item online survey. Eventually, 1293 nurses participated in the survey; however, 52 did not complete the survey properly, omitting both age or gender information, and they were excluded from the analysis.

**Results:** About half of the respondents (600/1241, 48.35%) had health professional apps installed on their devices and were included for analysis. Most participants in the survey were women (474/600, 79.0%) and the remaining were men (126/600, 21.0%). The most popular types of apps used and installed among nurses were related to drug information, health calculators, and health guidelines. Overall, 97.0% (582/600) of nurses thought that the health apps should be certified, and 80.0% (480/600) agreed that the certification process should be carried out by professional or health institutions. Furthermore, 14.5% (87/600) of participants mentioned that they were asked by their patients to prescribe a health app and only 6.5% (28/430) recommended them often. Most nurses (354/433, 81.8%) who answered the question about the importance of receiving specific training on using and prescribing health apps considered this point a very relevant issue.

**Conclusions:** About half of the nurses in Catalonia use health apps for professional purposes, and they believe that these types of tools should be validated and certified by health or professional institutions before using them in clinical environments. Although the prescription of health apps in clinical environments is infrequent among nurses, they would be willing to prescribe apps if they were certified by a health organization. Finally, among nurses, there is a need for training in using and prescribing health apps for health care purposes.

**KEYWORDS**

nurse's role; smartphones; mobile phone; mobile apps; mHealth

## Introduction

### Background

In the last few years, the number of mobile health apps has increased exponentially. The potential of several advantages of using health apps by professionals in clinical environments has been discussed in many forums [1-3]. Currently, there are a lot of apps available to assist health professionals with many different tasks such as health and drug information, communication with patients and other professionals, patient management and monitoring, clinical decision making, and health education and training [4]. In addition, smartphone use by health care professionals is a growing market [5,6].

Several surveys have been carried out to know the use of health apps among general public [7,8], patient associations [9], and some health professionals in hospital settings [10]. In addition, there is a pilot study with the prescription of mobile apps by medical doctors in primary care environments [11]. Moreover, there are also many studies and publications that analyze the quality of health apps, and in most cases, a lack of quality was detected [12-16]. In the recent years, coinciding with the growing number of health apps, diverse organizations and professional institutions have been setting up specific guidelines on the use of health apps addressed to health professionals and students, and above all, among medical doctors [16-20]. This reflects the importance and concerns of the correct use of health apps from legal and ethical points of view.

Nursing practice is developed in many different types of settings such as hospitals, primary care, schools, long-term care facilities, community and public health centers, and homes [21]. Considering that the nursing profession is the largest segment of health care workforce in most countries in the world, including Spain [22,23], the impact of the use of health apps by these professionals is highly likely to be critical to the future of modern health care [24]. Therefore, there is a need for more knowledge about the real use of health apps among nurses in many different clinical environments, including their training requirements and perceptions on the matter. On the basis of this information, health and professional institutions can establish specific measures to improve the use and management of apps in health care settings.

### Objectives

The main objective of this study was to determine if nurses were using health apps professionally and the types of apps they were using. The secondary objectives were (1) to find out if, among nurses, there is a need for training in the use of health apps and (2) to explore nurses' perceptions of health professional apps, determining whether there is a need for a certification process for health apps and the type of institution or organization that should review and validate these apps for professional use. To the best of our knowledge no previous studies of these characteristics have been carried out focusing exclusively on

nurses in Spain. This information will allow us to have a better knowledge of the current use of these technologies by nurses in clinical environments.

## Methods

### Study Design

The study is a descriptive cross-sectional study based on an internet survey of 1293 nurses in Spain. LimeSurvey: GmbH [25], an open source platform, was used to manage the survey online. In October 2017, a pilot study was carried out by sending the survey to 38 nurses to check its proper functioning and to find out if the respondents understood the questions well. The participants in this pilot survey were selected by a convenience method among colleagues in the authors' workplace. On the basis of the pilot survey, no additional relevant changes were required in the final version of the survey, apart from some typing errors detected in some questions. In March 2018, an email, including a link to the website with the survey, was sent to the 26,907 registered nurses who had a contact email address in the Nursing Association of Barcelona (COIB) of Catalonia [26]. COIB is the second largest nursing association in Spain, with 34,327 registered nurses, after the Nursing Council of Madrid, representing more than 11% of all the registered nurses in this country. The email included information about the purpose of the survey and that the information gathered would be analyzed anonymously. After the first email, 2 more emails were sent as reminders within the following 30 days.

### Survey Items

The list of questions was generated based on the previous experience of the authors in the elaboration of questionnaires addressed to health professionals, such as nurses, physicians, and pharmacists [27,28], and the review of published surveys related to the use of smartphones and health professional apps [7,10].

The survey comprised 34 questions organized in 3 parts: (1) questions related to sociodemographic and professional information such as age, gender, workplace, and professional activity or specialty; (2) information about the use of health apps for professional purposes such as the frequency of use, the type and number of apps installed, and their use; and finally (3) questions and considerations about the quality and certification of health apps and training needs for using health apps for professional purposes. Descriptive statistics were applied for all the items in the survey. Pearson chi-square tests examined the differences in questions with 3 or more options, comparing differences in categorical data among groups, and a Student *t* test was used for comparing continuous variables such as age. Statistical significance was taken at a level of  $P < .05$ . Questions with open-ended responses or free text, such as the mention of the 3 most frequent apps used or the reasons why they did not prescribe apps, were examined manually, annotating its content.

Owing to the exploratory nature of the study, no inferential statistics were applied.

The study was approved by the governing board of the COIB. The analysis of the data was performed using the R programming language for statistical computing and graphics, version 3.4.2 (R Development Core Team).

## Results

### Demographic and Professional Characteristics

The number of respondents to the survey was 1293 (1293/26,907, 4.80%); however, 52 did not provide complete basic information in the survey, such as age or gender, and they were not included in the analysis. With regard to the use of health apps, 51.65% (641/1241) of nurses answered that they did not have health professional apps installed, and the main reasons for not installing these types of apps were (1) lack of knowledge of health professional apps and (2) as they did not use them or they considered that the health apps were not useful for health purposes. Other less frequent reasons were the fact that their professional activity did not involve contact with patients, existence of technical limitations with their mobile

phones, or mistrust of health apps. In this group, 85.6% (549/641) were women and 14.4% (92/641) were men. The mean age was 44.96 (SD 13.45) years.

Overall, 48.35% (600/1241) of the respondents had health professional apps installed on their devices. In this group, 79.0% (474/600) were women and 21.0% (126/600) were men, with statistically significant differences with regard to the group of nurses who had not health professional apps installed ( $P < .001$ ). The average age of the respondents who had installed health apps was 43.12 (SD 11.32) years, and there were no significant differences between the ages of men and women ( $t_{160.51} = -1.5078$ ;  $P = .13$ ). The distribution of age was 16% baby boomers (>56 years), 58% generation X (35-56 years), and 26% millennials (<35 years). With regard to the professional area, most nurses came from hospital care (288/600, 48.0%) and primary care (144/600, 24.0%). With regard to the specialization, the most frequent groups were general nursing (318/600, 53.0%) and medical-surgical nursing (84/600, 14.0%). The complete distribution of gender, age, nursing specialties, professional activity, and the area of work of the nurses who had installed and were using health professional apps can be seen in [Table 1](#).

**Table 1.** Demographics of surveyed nurses who had health professional apps installed, in terms of gender, age, specialty, and professional area (N=600).

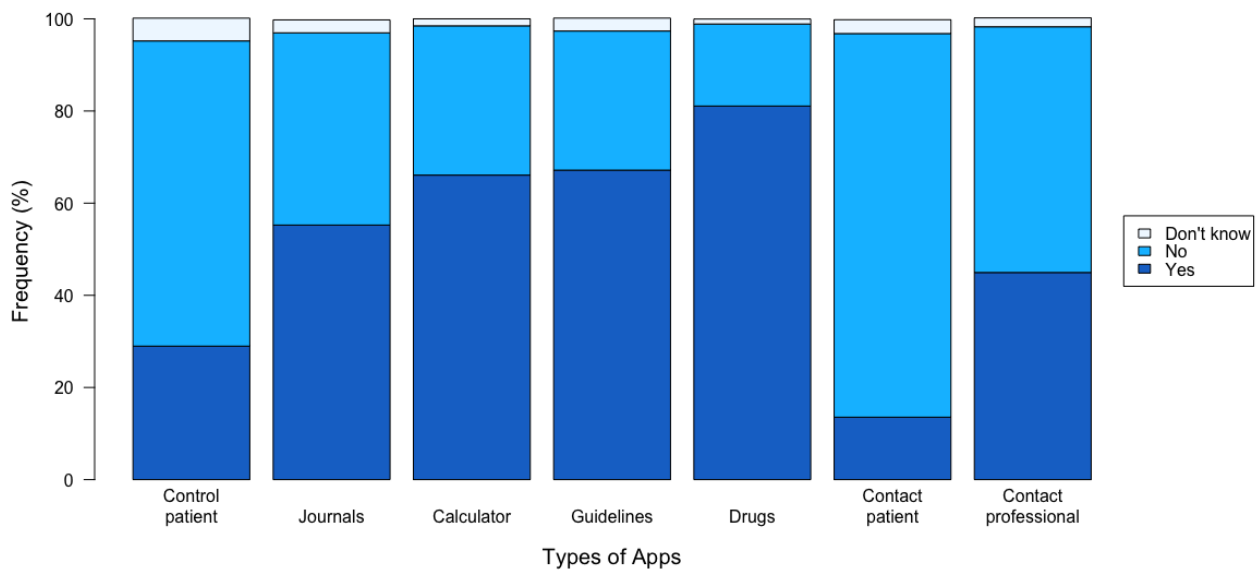
Demographics	Values, n (%)
<b>Gender</b>	
Male	126 (21.0)
Female	474 (79.0)
<b>Age (years)</b>	
Millennials (<35)	156 (26.0)
Generation X (35-56)	348 (58.0)
Baby boomers (>57)	96 (16.0)
<b>Nursing specialties</b>	
General	318 (53.0)
Midwifery	30 (5.0)
Mental health	18 (3.0)
Pediatrics	48 (8.0)
Family and community	42 (7.0)
Occupational health care	36 (6.0)
Medical-surgical	84 (14.0)
Geriatrics	30 (5.0)
<b>Professional activity</b>	
Hospital care	288 (48.0)
Primary care	144 (24.0)
Social health care	42 (7.0)
Management	30 (5.0)
Prehospital care	24 (4.0)
Teaching/research	18 (3.0)
Private practice	18 (3.0)
Other	42 (7.0)
<b>Area of work</b>	
Urban	565 (94.2)
Rural	35 (5.8)

## Health Apps Usage

Overall, nurses prefer smartphones (489/536, 91.2%) instead of tablets (47/536, 8.8%) for managing apps professionally. Most devices were less than 2 years old (413/536, 77.0%), and 42.7% (229/536) of the devices were less than a year old. As for the operating system, 51.1% (274/536) had Android and 45.7% (245/536) had iOS. Furthermore, 57.1% (300/525) of the nurses had organized health professional apps in specific folders in their mobile phone.

Most nurses had installed between 2 and 5 apps; 65.4% (344/526) used 1 health professional app at least once a week, 30.6% (161/526) used between 2 and 5 apps, and 3.2% (17/526) used more than 5 health apps. The number of apps installed

depended on the gender of the participants; men had more health apps installed than women ( $\chi^2_2=12.65$ ;  $P=.002$ ), and at the same time, they used apps more frequently than women ( $\chi^2_2=9.30$ ;  $P=.009$ ). Age groups did not significantly differ with regard to the number of apps installed on the devices ( $\chi^2_4=7.77$ ;  $P=.10$ ), although baby boomers tended to have more professional apps installed than the other age groups. Age group was not a factor that determined the number of health apps used ( $\chi^2_2=5.21$ ;  $P=.27$ ). The most popular types of apps installed and used were related to drug information among 81.1% of nurses (377/465), health guidelines among 67.0% (313/467), health calculators among 66.1% (308/466), and communication with patients among 13.5% (63/465; see [Figure 1](#)).

**Figure 1.** Most frequent types of apps used by nurses.

### Prescription of Health Apps

Overall, 97.0% (582/600) of nurses thought that health apps should be certified and 80.0% (480/600) agreed that the certification process should be carried out by professional or health institutions. Most nurses who answered the question related to the prescription of apps never prescribed a health professional app (266/430, 61.9%). If the apps were certified by a health or scientific organization, 50.4% (134/266) of nurses who answered this question would be willing to prescribe them. The willingness to prescribe certified apps was not different

between genders ( $\chi^2_2=3.66$ ;  $P=.16$ ) and among age groups ( $\chi^2_4=3.04$ ;  $P=.55$ ). The recommendation of apps by nurses to colleagues did not show significant differences among age groups ( $\chi^2_4=4.28$ ;  $P=.37$ ) or gender ( $\chi^2_4=3.96$ ;  $P=.14$ ). [Table 2](#) shows the list of questions and responses related to the features and use of health apps for professional purposes, and [Table 3](#) shows the questions related to different aspects of the prescription of apps (see [Multimedia Appendix 1](#) for the English version of the full list of items and responses).

**Table 2.** Questions and answers, including information about the features and use of health apps by nurses for professional purposes.

Survey question and response options	Answers received, n (%)
<b>How many apps do you have installed on your smartphone? (n=522)</b>	
1	154 (29.5)
2-5	316 (60.5)
>5	52 (10.0)
<b>How many apps do you use regularly (at least once a week)? (n=522)</b>	
1	344 (65.9)
2-5	161 (30.8)
>5	17 (3.3)
<b>You are using health apps for:</b>	
<b>As a tool for managing patients (n=466)</b>	
Yes	135 (29.0)
No	308 (66.1)
I don't know	23 (4.9)
<b>Access to scientific journals (n=465)</b>	
Yes	258 (55.5)
No	194 (41.7)
I don't know	13 (2.8)
<b>As a calculator for doses and scales (n=466)</b>	
Yes	308 (66.1)
No	151 (32.4)
I don't know	7 (1.5)
<b>Clinical guidelines and protocols (n=467)</b>	
Yes	313 (67.0)
No	141 (30.2)
I don't know	13 (2.8)
<b>Information about drugs (n=465)</b>	
Yes	377 (81.1)
No	83 (17.8)
I don't know	5 (1.1)
<b>Communication with patients (n=465)</b>	
Yes	63 (13.5)
No	388 (83.4)
I don't know	14 (3.0)
<b>Communication with professionals (n=466)</b>	
Yes	209 (44.8)
No	248 (53.2)
I don't know	9 (1.9)
<b>In relation with the use of health apps, indicate the option that is closest to your opinion (n=466)</b>	
<b>They help me solve professional doubts</b>	
Never	14 (3.0)
Sometimes	226 (48.5)
Often	226 (48.5)



Survey question and response options	Answers received, n (%)
<b>In general, they are easy to use</b>	
Never	11 (2.4)
Sometimes	139 (29.8)
Often	316 (67.8)
<b>They facilitate my professional tasks</b>	
Never	16 (3.4)
Sometimes	236 (50.6)
Often	214 (45.9)
<b>They are a support tool for managing my patients</b>	
Never	161 (34.5)
Sometimes	188 (40.3)
Often	117 (25.1)
<b>Do you recommend any apps to your colleagues for professional use? (n=430)</b>	
Never	62 (14.4)
Sometimes	283 (65.8)
Often	85 (19.8)

**Table 3.** Questions and answers about the prescription of health professional apps.

Survey question and response options	Answers received, n (%)
<b>Currently, do you prescribe apps to your patients? (n=430)</b>	
Never	266 (61.9)
Sometimes	136 (31.6)
Often	28 (6.5)
<b>If that's the case, what is the aim of prescribing apps to your patients? (n=336)</b>	
Health promotion	129 (38.4)
Health information, educational support	98 (29.2)
Patient self-management and control	64 (19.0)
Monitoring patients	38 (11.3)
Others	7 (2.1)
<b>Would you prescribe apps to your patients if they were certified by a health organization? (n=266)</b>	
Yes	134 (50.4)
No	50 (18.8)
I don't know	82 (30.8)
<b>Did your patients ask for an app to be prescribed to them? (n=430)</b>	
Never	297 (69.1)
Sometimes	128 (29.8)
Often	5 (1.2)
<b>Do you think that some health organization, scientific or professional association should certify health apps (both for patients and professionals)? (n=430)</b>	
Yes	392 (91.2)
No	11 (2.5)
I don't know	27 (6.3)
<b>Do you think being trained to use and prescribe health apps could be of interest to you? (n=430)</b>	
Yes	354 (82.3)
No	27 (6.3)
I don't know	49 (11.4)
<b>Do you think the health or professional organization should promote recommendations for developing health apps (addressed to developers)? (n=430)</b>	
Yes	377 (87.7)
No	4 (0.9)
I don't know	49 (11.4)

The reasons why surveyed nurses had not recommended health apps to patients were because of the *lack of knowledge* about the type and use of these apps, although there were other factors such as the type of patients, patients with limited capacity to use apps (elderly people), or related to the type of professional activity such as in intensive care units, working in a night shift or geriatrics department, no patient contact, and lack of time. Very few participants (16/266, 6.0%) who answered the question about the reasons why they did not prescribe health apps considered them unnecessary. Furthermore, 59.8% (359/600) answered the question about the 3 most frequent apps they used, and the first 3 most frequent apps mentioned were *Infermera virtual* (an app developed and promoted by the COIB),

*Vademecum International* (an app containing a handbook for drugs consultation), and *061 CatSalut Respon* (an app promoted by the Catalan government for contacting health care services and obtaining health advice). In addition, the most frequent apps addressed to specific diseases were *Trata la upp* (an app for the management of pressure ulcers), *Pukono* (an app that includes information and diet recommendations for patients with hypertension), *Diabetes a la carta* (an app for managing diabetes), *SocialDiabetes* (an app for managing diabetes), and *S'acabó* (an app for smoking cessation). [Table 4](#) includes the 16 most frequent apps used by nurses among the 238 apps they mentioned in the survey.

**Table 4.** The 16 most frequently used apps by nurses in Catalonia.

Name of the app	Description	Mentions, n
Infermera virtual	This app is promoted by the Nursing Association of Barcelona and provides a collection of resources, services, and useful information for nurses	92
Vademecum Internacional	This app contains a handbook for drugs consultation and is promoted by Vidal Vademecum Spain	84
061 CatSalut	An app promoted by the Catalan government for contacting health care services and obtaining health advice and information for the general public and health professionals	46
Enfermeria blog	App of a blog with useful information for nurses	24
HealthScience	This app includes the most recent abstracts of scientific publications for health professionals	22
NandaNocNic	This app includes the codes of the North American Nursing Diagnosis Association diagnosis, results, and interventions	21
Medicamentos via parenteral	An app that includes information for the management and administration of parenteral drugs	19
iVacunes Gencat	An app promoted by TICSalut of the Catalan government that includes information about vaccines and other related resources	17
aempsCIMA	This app offers information about all the authorized drugs in Spain and is promoted by the Spanish Medicines Agency	15
DosisPedia	An app for the management, administration, and doses of drugs	13
Trata la upp	An app for the management of pressure ulcers	12
ECG Práctico	An app of the Spanish Cardiology Society with guidelines for the interpretation of the electrocardiogram,	11
Babymecum	An app for the management of drugs doses and administration in pediatrics	10
GuiaSalud	This app includes the most important clinical guidelines and protocols of the Spanish National Health Service	10
iDoctus	This app includes information on drugs and diseases, several dosage calculators, and drug-drug interactions	8
eMerMed Science	An app that includes the most recent abstracts of scientific publications for emergency medicine professionals	7

Only 14.5% (87/600) of the participants mentioned that they were asked by their patients to prescribe a health app to them. The most frequent apps prescribed were related to diet and exercise (19/87, 21.8%), diabetes (13/87, 14.9%), health tips (10/87, 11.5%), hypertension (8/87, 9.2%), and smoking cessation (8/87, 9.2%). With regard to the importance of receiving specific training in using and prescribing health apps, 72.2% (433/600) of nurses answered this question. Most of them (354/433, 81.8%) answered that it was necessary to receive specific training in the matter, without significant differences among age groups ( $\chi^2_4=3.40$ ;  $P=.35$ ) or gender.

## Discussion

### Principal Findings

First, just under half of the nurses who participated in the survey used health professional apps. The main reasons for not using health apps were the lack of knowledge or interest. In only a few cases did they consider these apps unnecessary for carrying out their jobs. It is worth noting that the most frequent apps used by nurses were promoted by health institutions of reference such as the Health Department of Catalonia and the COIB and that these apps are used to support their professional activity and patient advice. It is surprising that older nurses (baby boomers) tend to have more apps installed than the younger nurses, considering that in general, young people are more used to incorporating new information and communication

technologies than older people. Baby boomers, who have more professional experience, usually did not require the help of apps for drug information and for calculating doses in clinical care. In addition, on the basis of their professional experience, they were willing to use health professional apps for patient control or to be in touch with them. On the contrary, younger nurses consult health professional apps more frequently to calculate drug posology and less frequently to establish communication with patients. Perhaps this is because of a lack of confidence in using apps for monitoring patients. Considering the number of apps installed by gender, okmen had more apps installed on their devices than women. Baby boomers tended to have more professional apps installed than the other age groups.

Second, the prescription of health apps is not common among health professionals, and one reason for that could be the fact that this activity is not yet totally regulated. Nevertheless, nurses would be willing to prescribe apps more frequently if they were certified by a health or scientific organization. This is the first study that explores the opinion of health professionals, particularly nurses, about the need of certifying health apps. The prescription of apps in clinical settings requires several considerations; on the one hand the importance of using health apps of high quality and those that have been previously tested are among the most relevant ones. On the other hand, it is necessary to guarantee their technical and organizational integration in a health care system, based on specific protocols [11]. If health apps are to be adopted, patients and health

professionals must be sure that they are safe and effective [29]. With regard to the types of apps prescribed, this aspect may be explained considering that the conditions for which these apps are prescribed are among the most common chronic diseases, and at the same time, there are a lot of apps in the market focusing on smoking, exercise, and diabetes control.

Finally, our findings also suggest that currently the prescription and use of health apps is not frequent among certain patients that health professionals in general are taking care of, particularly elderly people. In this sense, the current aging population requires strategies to promote the use of health apps among this population. One of the possible solutions is to facilitate the training of the elderly in the use of mobile technologies. This population and health professionals can benefit from the use of health apps for monitoring and managing these patients suffering from chronic diseases [30]. In addition, the survey provides us with some insights related to relevant aspects such as the need to train nurses in the use and prescription of health professional apps in clinical environments and also the importance of having some guidelines to be able to select apps of high quality [31-33]. This implies that evaluation tools and protocols should be applied before deciding on the introduction of apps in clinical settings. Professional and health institutions should have a key role in these types of activities, promoting training programs and the evaluation of the quality and usefulness of specific health apps in clinical environments.

In the last few years, the evolution of mobile phones and other devices, as well as the apps developed, has happened at an unprecedented rate. Nurses, and health professionals in general, must be trained to use information and communication technologies such as mobile apps, smart wearable technologies, or telemedicine apps [4,24,34,35]. For this reason, it is critical to have information about how these professionals use and view the utility of these apps. This study examined the health app usage among registered nurses from the second most important nursing association of Spain.

### Strengths and Limitations of the Study

On the one hand this study has several strengths. First, this is the first time that this type of survey has been performed among nurses in Spain, the third largest country in Western Europe,

with a health system that is almost universal (covering 99.1% of the population) and the highest life expectancy in the European Union [36]. Second, although the rate of participation is low in absolute terms, the number of participants is quite high, and included all the different nursing specialties. Finally, original aspects related to the use and certification of health professional apps were analyzed. On the other hand, this study presents some limitations. The number of participants in the survey was very low compared with the number of nurses who received the survey. In addition, the number of men was significantly higher than the group of nurses who did not respond to the survey. For these reasons, the results could overestimate or underestimate the real use and the interest of nurses in health apps and the influence of some factors such as gender, introducing a potential selection bias, which is very common in this type of online survey. We should take into account that the survey was sent to nurses who had an email address; however, 21.73% (7471/34,378) of the registered nurses did not have a contact email address or did not use a COIB email address, and they did not receive the questionnaire.

### Conclusions

The use of health professional apps was not generalized among nurses at the time of the survey and it was very unusual for patients to ask for a health app. The most popular apps are used to support the professional activity of nurses, and most of them are promoted by professional and health institutions. In addition, the prescription of health apps is not common among nurses, but almost all nurses believed that health apps should be certified by health institutions or professional associations. According to the results of the survey, the certification of health apps would be a very important deciding factor in the prescription of apps to patients in clinical environments. Moreover, nurses expressed their interest in and their need for specific training in the use and prescription of health apps, indicating the potential impact of the introduction of these technologies in clinical environments. On the basis of the results of the survey, COIB will consider the possibility of giving nurses some training in the use and prescription of health apps in clinical environments. On the basis of the concerns about the quality of apps that nurses expressed, COIB will consider setting up some specific guidelines to rate the quality of health apps.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

Health professional apps survey.

[PDF File (Adobe PDF File), 158 KB - [mhealth\\_v7i11e15195\\_app1.pdf](#)]

---

### References

1. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. *Pediatr Diabetes* 2013 Jun;14(4):231-238 [FREE Full text] [doi: [10.1111/pedi.12034](#)] [Medline: [23627878](#)]
2. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. *J Med Internet Res* 2012 Sep 27;14(5):e128 [FREE Full text] [doi: [10.2196/jmir.1994](#)] [Medline: [23017375](#)]

3. Barton AJ. The regulation of mobile health applications. *BMC Med* 2012 May 8;10:46 [FREE Full text] [doi: [10.1186/1741-7015-10-46](https://doi.org/10.1186/1741-7015-10-46)] [Medline: [22569114](https://pubmed.ncbi.nlm.nih.gov/22569114/)]
4. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *Pharm Therap* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](https://pubmed.ncbi.nlm.nih.gov/24883008/)]
5. Wyatt T, Krauskop P. E-Health and nursing: using smartphones to enhance nursing practice. *Online J Nurs Inform* 2012;16(2):10 [FREE Full text]
6. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med* 2013 Jun;18(3):90-92. [doi: [10.1136/eb-2012-100885](https://doi.org/10.1136/eb-2012-100885)] [Medline: [22923708](https://pubmed.ncbi.nlm.nih.gov/22923708/)]
7. Krebs P, Duncan DT. Health app use among us mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov 4;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
8. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SS, et al. Health app possession among smartphone or tablet owners in hong kong: population-based survey. *JMIR Mhealth Uhealth* 2017 Jun 5;5(6):e77 [FREE Full text] [doi: [10.2196/mhealth.7628](https://doi.org/10.2196/mhealth.7628)] [Medline: [28583905](https://pubmed.ncbi.nlm.nih.gov/28583905/)]
9. Mosconi P, Radrezza S, Lettieri E, Santoro E. Use of health apps and wearable devices: survey among Italian associations for patient advocacy. *JMIR Mhealth Uhealth* 2019 Jan 15;7(1):e10242 [FREE Full text] [doi: [10.2196/10242](https://doi.org/10.2196/10242)] [Medline: [30664455](https://pubmed.ncbi.nlm.nih.gov/30664455/)]
10. Mobasheri MH, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: a multicentre survey study. *BMJ Innov* 2015;1(4):174-181. [doi: [10.1136/bmjinnov-2015-000062](https://doi.org/10.1136/bmjinnov-2015-000062)]
11. Segui FL, Bufill C, Gimenez N, Roldan J, Cuyas F. The prescription of mobile apps by primary care teams: a pilot project in Catalonia. *JMIR Mhealth Uhealth* 2018 Jun 21;6(6):e10701 [FREE Full text] [doi: [10.2196/10701](https://doi.org/10.2196/10701)] [Medline: [29929948](https://pubmed.ncbi.nlm.nih.gov/29929948/)]
12. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. *J Med Internet Res* 2011 Sep 22;13(3):e65 [FREE Full text] [doi: [10.2196/jmir.1874](https://doi.org/10.2196/jmir.1874)] [Medline: [21979293](https://pubmed.ncbi.nlm.nih.gov/21979293/)]
13. Wolf JA, Moreau JF, Akilov O, Patton T, English JC, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA Dermatol* 2013 Apr;149(4):422-426 [FREE Full text] [doi: [10.1001/jamadermatol.2013.2382](https://doi.org/10.1001/jamadermatol.2013.2382)] [Medline: [23325302](https://pubmed.ncbi.nlm.nih.gov/23325302/)]
14. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. *J Med Internet Res* 2013 Jan 4;15(1):e1 [FREE Full text] [doi: [10.2196/jmir.2301](https://doi.org/10.2196/jmir.2301)] [Medline: [23291245](https://pubmed.ncbi.nlm.nih.gov/23291245/)]
15. Samples C, Ni Z, Shaw RJ. Nursing and mHealth. *Online J Nurs Sci* 2014 Dec;1(4):330-333. [doi: [10.1016/j.ijnss.2014.08.002](https://doi.org/10.1016/j.ijnss.2014.08.002)]
16. Göransson C, Eriksson I, Ziegert K, Wengström Y, Langius-Eklöf A, Brovall M, et al. Testing an app for reporting health concerns-experiences from older people and home care nurses. *Int J Older People Nurs* 2018 Jun;13(2):e12181 [FREE Full text] [doi: [10.1111/opn.12181](https://doi.org/10.1111/opn.12181)] [Medline: [29210218](https://pubmed.ncbi.nlm.nih.gov/29210218/)]
17. Mayer MA, Leis A, Mayer A, Rodriguez-Gonzalez A. How medical doctors and students should use social media: a review of the main guidelines for proposing practical recommendations. *Stud Health Technol Inform* 2012;180:853-857. [doi: [10.3233/978-1-61499-101-4-853](https://doi.org/10.3233/978-1-61499-101-4-853)] [Medline: [22874313](https://pubmed.ncbi.nlm.nih.gov/22874313/)]
18. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
19. Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. *JMIR Mhealth Uhealth* 2015 Sep 14;3(3):e87 [FREE Full text] [doi: [10.2196/mhealth.4532](https://doi.org/10.2196/mhealth.4532)] [Medline: [26369346](https://pubmed.ncbi.nlm.nih.gov/26369346/)]
20. van Velsen L, Beaujean DJ, van Gemert-Pijnen JE. Why mobile health app overload drives us crazy, and how to restore the sanity. *BMC Med Inform Decis Mak* 2013 Feb 11;13:23 [FREE Full text] [doi: [10.1186/1472-6947-13-23](https://doi.org/10.1186/1472-6947-13-23)] [Medline: [23399513](https://pubmed.ncbi.nlm.nih.gov/23399513/)]
21. Institute of Medicine, Committee on the Robert Wood Johnson Foundation Initiative on the Future of Nursing, at the Institute of Medicine. *The Future of Nursing: Leading Change, Advancing Health*. Washington (DC): National Academies Press; 2011.
22. Instituto Nacional de Estadística. 2019. Registered Health Professionals Statistics, Year 2018 URL: [http://www.ine.es/en/prensa/epsc\\_2018\\_en.pdf](http://www.ine.es/en/prensa/epsc_2018_en.pdf) [accessed 2019-05-01]
23. Crisp N, Chen L. Global supply of health professionals. *N Engl J Med* 2014 Mar 6;370(10):950-957. [doi: [10.1056/NEJMra1111610](https://doi.org/10.1056/NEJMra1111610)] [Medline: [24597868](https://pubmed.ncbi.nlm.nih.gov/24597868/)]
24. Billings DM. There's an 'app' for that: tips for preparing nurses for roles in mobile health. *J Contin Educ Nurs* 2015 Sep;46(9):390-391. [doi: [10.3928/00220124-20150821-14](https://doi.org/10.3928/00220124-20150821-14)] [Medline: [26352040](https://pubmed.ncbi.nlm.nih.gov/26352040/)]
25. Professional online surveys with LimeSurvey. URL: <https://www.limesurvey.org> [accessed 2018-07-01]
26. Col·legi Oficial d'Infermers i Infermeres de Barcelona (COIB). URL: <http://www.coib.cat> [accessed 2018-07-20]

27. Mayer MA. Evaluación de los sistemas de acreditación de webs sanitarias: la experiencia de Web Médica Acreditada. Barcelona: Universitat Pompeu Fabra; 2006.
28. Lupiáñez-Villanueva F, Mayer MA, Torrent J. Opportunities and challenges of Web 2.0 within the health care systems: an empirical exploration. *Inform Health Soc Care* 2009 Sep;34(3):117-126. [doi: [10.1080/17538150903102265](https://doi.org/10.1080/17538150903102265)] [Medline: [19670002](https://pubmed.ncbi.nlm.nih.gov/19670002/)]
29. Bates DW, Landman A, Levine DM. Health apps and health policy. What is needed? *J Am Med Assoc* 2018 Nov 20;320(19):1975-1976. [doi: [10.1001/jama.2018.14378](https://doi.org/10.1001/jama.2018.14378)] [Medline: [30326025](https://pubmed.ncbi.nlm.nih.gov/30326025/)]
30. Ferreira C, Silva JC, Silva JL. Mobile Applications for Active Aging. In: *Proceedings of the World Conference on Information Systems and Technologies*. 2018 Presented at: WorldCist'18; March 27-29, 2018; Naples, Italy p. 1067-1073. [doi: [10.1007/978-3-319-77712-2\\_102](https://doi.org/10.1007/978-3-319-77712-2_102)]
31. Orchard J, Lowres N, Freedman SB, Ladak L, Lee W, Zwar N, et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): a feasibility study. *Eur J Prev Cardiol* 2016 Oct;23(2 suppl):13-20. [doi: [10.1177/2047487316670255](https://doi.org/10.1177/2047487316670255)] [Medline: [27892421](https://pubmed.ncbi.nlm.nih.gov/27892421/)]
32. McNeil P, McArthur EC. Evaluating health mobile apps: information literacy in undergraduate and graduate nursing courses. *J Nurs Educ* 2016 Aug 1;55(8):480. [doi: [10.3928/01484834-20160715-12](https://doi.org/10.3928/01484834-20160715-12)] [Medline: [27459438](https://pubmed.ncbi.nlm.nih.gov/27459438/)]
33. Peck JL. Social media in nursing education: responsible integration for meaningful use. *J Nurs Educ* 2014 Mar;53(3):164-169. [doi: [10.3928/01484834-20140219-03](https://doi.org/10.3928/01484834-20140219-03)] [Medline: [24530130](https://pubmed.ncbi.nlm.nih.gov/24530130/)]
34. Sprenger M, Mettler T, Osma J. Health professionals' perspective on the promotion of e-mental health apps in the context of maternal depression. *PLoS One* 2017;12(7):e0180867 [FREE Full text] [doi: [10.1371/journal.pone.0180867](https://doi.org/10.1371/journal.pone.0180867)] [Medline: [28704442](https://pubmed.ncbi.nlm.nih.gov/28704442/)]
35. Doswell WM, Braxter B, DeVito Dabbs A, Nilsen W, Klem ML. mHealth: technology for nursing practice, education, and research. *J Nurs Educ Pract* 2013;3(10):99-109. [doi: [10.5430/jnep.v3n10p99](https://doi.org/10.5430/jnep.v3n10p99)]
36. Bernal-Delgado E, Garcia-Armesto S, Oliva J, Sanchez Martinez FI, Repullo JR, Pena-Longobardo LM, et al. Spain: Health system review. *Health Syst Transit* 2018 May;20(2):1-179 [FREE Full text] [Medline: [30277216](https://pubmed.ncbi.nlm.nih.gov/30277216/)]

## Abbreviations

**COIB:** Nursing Association of Barcelona

*Edited by G Eysenbach; submitted 26.06.19; peer-reviewed by M Bauer, H Devan, C Reis, B Worm, L Sheets; comments to author 19.07.19; revised version received 25.08.19; accepted 27.09.19; published 01.11.19.*

*Please cite as:*

Mayer MA, Rodríguez Blanco O, Torrejon A

Use of Health Apps by Nurses for Professional Purposes: Web-Based Survey Study

*JMIR Mhealth Uhealth* 2019;7(11):e15195

URL: <https://mhealth.jmir.org/2019/11/e15195>

doi: [10.2196/15195](https://doi.org/10.2196/15195)

PMID: [31682587](https://pubmed.ncbi.nlm.nih.gov/31682587/)

©Miguel Angel Mayer, Octavi Rodríguez Blanco, Antonio Torrejon. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 01.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mhealth and Uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Characteristics and Usage Patterns Among 12,151 Paid Subscribers of the Calm Meditation App: Cross-Sectional Survey

Jennifer Huberty<sup>1</sup>, PhD; Ana-Maria Vranceanu<sup>2,3</sup>, PhD; Colleen Carney<sup>4</sup>, PhD, CPsych; Michael Breus<sup>5</sup>, ABSM, PhD; Michael Gordon, MFA; Megan Elizabeth Puzia<sup>6</sup>, MS

<sup>1</sup>College of Health Solutions, Arizona State University, Phoenix, AZ, United States

<sup>2</sup>Integrated Brain Health Clinical and Research Program, Massachusetts General Hospital, Boston, MA, United States

<sup>3</sup>Harvard Medical School, Boston, MA, United States

<sup>4</sup>Department of Psychology, Ryerson University, Toronto, ON, Canada

<sup>5</sup>Mindworks, LLC, Manhattan Beach, CA, United States

<sup>6</sup>Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

**Corresponding Author:**

Jennifer Huberty, PhD

College of Health Solutions

Arizona State University

550 3rd St

Phoenix, AZ, 85004

United States

Phone: 1 6028272456

Email: [jhuberty@asu.edu](mailto:jhuberty@asu.edu)

## Abstract

**Background:** Meditation has become increasingly popular due to its health benefits; however, barriers to delivering meditation programs in traditional group-based formats limit the accessibility of these benefits. Smartphone-based meditation may increase the availability of these programs to larger, more diverse audiences; however, research on subscriber characteristics and usage patterns in meditation mobile apps is lacking.

**Objective:** This study aimed to describe the demographics, clinical characteristics, and usage patterns of a convenience sample of Calm subscribers and explore the relationship between self-reported app usage and changes in health, stress, and sleep.

**Methods:** Participants were 12,151 paying Calm subscribers (response rate=12.08%, 12,151/100,594) who completed an anonymous Web-based survey with 11 quantitative questions related to user engagement, reasons for starting Calm, and changes after using the app. Demographic characteristics, chronic health diagnoses, and sleep difficulties were also assessed. Chi-square tests were used to examine differences in app usage. Logistic regression models were used to examine demographic and health characteristics that may predict changes in health, stress, and sleep.

**Results:** Respondents were 18-96 years old (mean 48.57 [SD 13.79]), primarily female (79.94%, 8778/10,981), white (81.41%, 8959/11,005), and most reported a chronic health diagnosis (56.86%, 6289/11,061). Mental health diagnoses (41.13%, 4549/11,061) were more common than physical health diagnoses (32.19%, 3560/11,061). Most respondents (76.31%, 8684/11,360) reported difficulties falling or staying asleep. On average, respondents had been using Calm for 11.49 months (SD 10.49), and 60.03% (7281/12,129) used it 5 or more times per week. Meditations (used by 80.02%, 9497/11,841) and Sleep Stories (55.66%, 6591/11,841) were the most popular components. The frequency of using Calm was associated with incremental increases in the likelihood of noticing changes in mental health ( $\chi^2_2=136.8$ ;  $P<.001$ ), physical health ( $\chi^2_2=102.8$ ;  $P<.001$ ), stress ( $\chi^2_2=128.1$ ;  $P<.001$ ), and sleep ( $\chi^2_2=141.4$ ;  $P<.001$ ). Respondents who had used Calm longer were also more likely to notice changes in mental health (OR 1.06 [95% CI 1.05 to 1.06]), physical health (OR 1.01 [95% CI 1.01 to 1.02]), stress (OR 1.04 [95% CI 1.04 to 1.05]), and sleep (OR 1.004 [95% CI 1.00 to 1.01]). Subscribers with sleep difficulties used Calm more frequently ( $\chi^2_2=11.5$ ;  $P=.003$ ), were more likely to use Sleep Stories ( $\chi^2_1=1590.2$ ;  $P<.001$ ), and were more likely to notice changes in their physical health ( $\chi^2_1=49.2$ ;  $P<.001$ ) and sleep ( $\chi^2_1=2391.1$ ;  $P<.001$ ).

**Conclusions:** Results highlight important demographic characteristics and usage patterns among a self-selected sample of Calm subscribers. Mental health concerns and sleep appear to be top reasons for downloading Calm. Sleep Stories and meditations are the most popular app components. The frequency of using Calm was associated with incremental changes in outcomes. Findings support future randomized controlled trials testing the efficacy of Calm for health, stress, and sleep. Studies should also explore strategies to attract a more diverse sample of subscribers.

(*JMIR Mhealth Uhealth* 2019;7(11):e15648) doi:[10.2196/15648](https://doi.org/10.2196/15648)

## KEYWORDS

health; psychological stress; sleep; mindfulness; meditation; consumer behavior; mobile health; mhealth; digital health

## Introduction

### Meditation Benefits

Meditation is a technique for resting the mind and attaining a state of consciousness that is different from the waking state [1]. Mindfulness meditation is the process of openly attending, with nonjudgmental awareness, to one's present moment experience [2]. Over the last 30 years, mindfulness meditation has become increasingly popular in Western society owing to its documented mental and physical health benefits [3]. Mindfulness meditation has been linked to reductions in depressive symptoms, stress, anxiety [4,5], and improved sleep [6], through biopsychosocial mechanisms such as alterations in brain structure, improved attention, increased emotional regulation, and better immune function [4]. Given these benefits, mindfulness meditation has been incorporated routinely in psychotherapy, school-based programs, corporate settings, prisons, and military and is an important part of hospital preventative or disease management programs [3].

### Meditation Accessibility

Although the evidence for the use of mindfulness meditation is slowly growing, there are multiple barriers to its delivery in traditional group-based formats [7], thus limiting its reach and its potential to be used by individuals seeking to manage mood, stress, or symptoms of a chronic illness. First, participation requires people to travel to a public location with a trained meditation instructor at a specific time of day [5]. This is often not feasible for individuals with diagnoses such as depression, anxiety, or posttraumatic stress for whom attrition rates are as high as 50% [7,8]. Caregivers, parents, and individuals with multiple responsibilities and high stress have busy schedules and may not be able to prioritize lengthy (eg, 60-90 min) weekly classes for extended period [7]. Individuals who live in rural areas or those with chronic disease may find travel to clinics for meditation sessions burdensome [9-11]. To overcome these barriers to participation, we need novel methods to deliver mindfulness meditation.

### Meditation and Smartphone Apps

Smartphone apps may be an effective way to deliver mindfulness meditation to geographically diverse and dispersed individuals that bypasses barriers associated with the traditional format of mindfulness meditation programs [5]. Meditation delivered on a smartphone app may retain the elements of traditional meditation programs (ie, a trained instructor delivering high-quality meditation), while allowing participants to tailor the type and length of meditations, use behavioral components

to increase adherence (eg, trackers, reminders, and social support communities), and choose when and where to engage with the app. Furthermore, there is relatively less cost associated with meditation delivered on an app. Approximately 81% of US adults report having a smartphone [12] and, around the world, smartphone usage is rapidly increasing [13]. In a cross-sectional survey of 1604 mobile phone users in the United States, 58% had downloaded a health-related mobile app that they used daily.

Despite their promise, the growing popularity of mindfulness-based apps, and efforts that are currently being made to evaluate and improve the science behind these apps, knowledge gaps remain [14]. To date, efficacy studies have had small samples, inadequate control groups, and short follow-up periods [5,15]. More importantly, data related to usage patterns of mindfulness meditation apps, demographics, clinical characteristics of users, reasons for downloading health apps, barriers to use, and consistency of use are lacking [16]. Furthermore, we do not know which personal characteristics predispose one to benefit more from specific app content [14]; this information is pivotal to building content and determining the app's effects. Unfortunately, app developers do not publish information on their users or the extent to which consumers continue to use apps over time [14]. When reports are available to nonindustry audiences, they are accompanied by fees, making them less accessible to the general public and researchers.

### Study Aims

The aim of this study was 2-fold. First, we reported demographics, clinical characteristics, and usage patterns in a self-selected sample of paying Calm subscribers. Second, we explored the relationship between the self-reported frequency of Calm use and the app's individual components with self-reported changes in health (mental and physical), stress, and sleep. We hypothesized that subscribers who use Calm more frequently will be more likely to report changes in their health, stress, and sleep.

## Methods

### Ethics Approval

The Institutional Review Board at Arizona State University (STUDY00009725) approved the study. All participants provided electronic consent before participating in the survey. The datasets generated or analyzed during the study are available from the corresponding author upon request.



## Study Design/Recruitment

This study was cross-sectional. Participants were paying subscribers to the mobile app, Calm (see [Multimedia Appendix 1](#) for screenshots). We used the mobile app, Calm, owing to its popularity. According to a 2015 systematic review of mindfulness apps, there are a total of 700 apps in Apple and Google Play app stores [17]. Calm was the number 1 downloaded health and fitness app between June 2018 and 2019 [18]. In an August 2019 review, CNET (a top website for technology news and editorials) confirmed that Calm was among the most highly rated meditation apps in the Apple and Google Play app stores (all above 4 stars) [19].

Calm guides users through mindfulness meditation. The 7 Days of Calm (ie, introductory course) introduces the user to mindfulness meditation and provides education about beginning a meditation practice. Users receive a new 10-min daily meditation (ie, *The Daily Calm*) and have access to a variety of meditation content (eg, reducing anxiety and depression, increasing compassion and gratitude) that ranges from 3 to 35 min. Calm offers a variety of meditations for sleep similar to other mindfulness-based apps [20]; however, unlike other apps, Calm offers Sleep Stories. Sleep Stories are narrated fictional stories using traditional storytelling methods, helping to immerse the user in their senses to improve sleep. Other content that Calm offers includes Calm Breathe (breathing training exercises), Calm Music (relaxing, soothing music and nature sounds), Calm Masterclass (monthly series of in-depth audio classes, delivered by experts that teach skills to deal with common stressors and improve well-being), and Calm Body (video lessons on slow, mindful movement routines). Calm also offers a variety of theory-based behavioral strategies to help the user be successful at regular participation including reminders to meditate, ability to track time spent meditating, and opportunities to share progress on social media.

Participants were recruited in April 2019. Calm subscribers were eligible if they were at least 18 years of age, their subscription expiration was at least two months away, they had opened at least two emails from Calm in the last 30 days, and they had completed at least two Calm sessions in the last 30 days.

Subscribers received an email inviting them to answer a series of quantitative and qualitative questions related to their usage of Calm and were informed that their answers would help improve their experience with the app, and results of this study may be used in reports, presentations, or publications. Approximately, 4744 subscribers were contacted on April 9, 54,193 were contacted on April 10, and 41,657 were contacted on April 11. Of the 100,594 subscribers who were contacted, 12,151 (12.08%, 12,151/100,594) participated in the survey. At the end of the survey, participants had the option of providing an email address to be contacted for future studies or to be entered into a draw to win 1 of 2 US \$99 Amazon gift cards. No other identifying information was collected, and responses were not linked to in-app usage data.

## Survey

The survey was developed by 3 doctoral-level researchers/clinicians in the field of behavior change, complementary approaches, and sleep. The survey was Web-based (delivered using Qualtrics) and took most participants between 5 and 10 min to complete (median=6.9; interquartile range=5.7). All responses were anonymous. Participants completed 11 multiple choice quantitative questions (not including demographics) about (1) their engagement with the Calm app (4 questions), (2) why they started using Calm (1 question), (3) clinical characteristics at the time they downloaded the app (2 questions), (4) their connections with other Calm users (3 questions), and (5) whether they noticed changes in their mental health, physical health, stress, or sleep after using Calm (1 yes or no question, separated into 4 subquestions about mental and physical health, stress, and sleep; [Textbox 1](#)). Questions varied slightly based on answers to frequency of use (see [Multimedia Appendix 2](#) for all quantitative/multiple choice questions). At the end of the survey, participants completed 4 questions about demographic characteristics. Although the survey included 3 additional open-ended questions, the scope of this paper was limited to discussion of findings from the quantitative data; qualitative analyses will be included in future publications.

**Textbox 1.** App usage questions from the Calm user survey.

Question:

1. How often do you use the Calm app?
2. Why did you start using the Calm app?
3. When you downloaded Calm, had you been diagnosed with any of the following chronic conditions? (eg, anxiety and high blood pressure)
- 4a. Which components of the app do you use most often? (eg, meditations and Sleep Stories)
- 4b. (For selected components only) How often do you use (component) on the Calm app?
5. When you first started using Calm, did you have difficulty falling asleep or staying asleep?
6. Have you used any of the following to help increase how frequently you use the app? (eg, reminders and Calm Facebook community)
7. For how many months have you been using Calm?
8. Have you noticed any changes in your (sleep/stress/mental health/physical health) after using Calm?
9. Research shows that it is easier to use an app if friends and family are also using the same app and can share their progress with each other. Do any of your friends and/or family members use Calm?
10. (If friends and/or family members use Calm) Do you communicate with your friends/family about using the Calm app?
11. Would you want to be connected with other app users?

### Statistical Analysis

Data were analyzed using IBM SPSS 25.0. All variables were categorical, with the exception of age. For questions about health diagnoses, sleep difficulties, reasons for starting Calm, the components of Calm used, and tools used to support engagement, participants were able to endorse multiple items from a list of options; selected options were treated as endorsements (eg, used meditations), and unselected options were treated as nonendorsements (eg, did not use meditations). As not all participants answered every question, sample sizes differed across analyses. Usage frequency was categorized as ordinal, reflecting use of the app or component of the app as nonendorsed, 1 to 2 times per week, 3 to 4 times per week, or 5 or more times per week. Variables for noticing changes in health, stress, and sleep were dichotomous.

Chi-square tests were used to evaluate differences in the use of tools for engagement and frequency of using Calm. To examine the relationships between demographic characteristics and the

frequency of using Calm and its components, and we used chi-square tests (sex differences) and binomial logistic regressions (age). Differences in the frequency of using Sleep Stories were further explored using ordinal logistic regressions controlling for sleep difficulties. Chi-square tests were used to analyze differences in the frequency of using Calm and the components of Calm used based on health diagnoses and sleep difficulties. We also used chi-square tests to investigate the relationship between the frequency of using Calm and noticing changes in health, stress, sleep and the relationship between sleep difficulties and noticing changes in health, stress, and sleep. In all cases, significant chi-square tests were followed up with  $z$  tests of column proportions.  $P$  values were adjusted for multiple comparisons using the Bonferroni correction.

## Results

### Demographic Characteristics

A total of 12,151 subscribers participated in the survey. Sample demographics are presented in [Table 1](#).

**Table 1.** Demographics of Calm subscribers (N=12,151).

Category	Values, n (%)
<b>Age (years; N=11,010; mean 48.6 [SD 14.0])</b>	
18-24	279 (2.53)
25-44	4230 (38.42)
45-64	4904 (44.54)
≥65	1597 (14.50)
<b>Gender (N=10,981)</b>	
Female	8778 (79.94)
Male	2177 (19.83)
Other	26 (0.24)
<b>Race (N=11,005)</b>	
White, European American, or Caucasian	8959 (81.41)
Asian or Asian American	413 (3.75)
Black, African American, or Native African	153 (1.39)
Other	1480 (13.45)
<b>Annual income (US \$; N=9474)</b>	
<21,000	665 (7.02)
21,000-60,000	2536 (26.77)
61,000-100,000	2539 (26.80)
>100,000	3734 (39.41)

### Clinical Characteristics

More than half (56.86%, 6289/11,061) of respondents had a chronic health diagnosis. Mental health diagnoses (ie, depression, anxiety, or posttraumatic stress disorder; 41.13%, 4549/11,061) were more common than physical health diagnoses (ie, high blood pressure, pain, high cholesterol, asthma, arthritis, cancer, heart disease, emphysema or chronic obstructive pulmonary disease, or other lung disease; 32.19%, 3560/11,061). The most commonly reported mental health diagnosis was

anxiety (33.22%, 3675/11,061), and the most common physical health diagnosis was high blood pressure (12.68%, 1402/11,061; [Table 2](#)).

Most respondents (76.49%, 8704/11,380) reported sleep difficulties (ie, difficulty falling or staying asleep; [Table 3](#)), and 13.90% (1537/11,061) reported insomnia diagnoses ([Table 2](#)). Sleep difficulties were significantly more common in females (80.09%, 7009/8751) than in males (61.77%, 1336/2163;  $\chi^2_{1}=355.9$ ;  $P<.001$ ).

**Table 2.** Self-reported mental and physical health diagnoses in Calm subscribers (N=11,061).

Diagnosis	Values, n (%)
Anxiety	3675 (33.22)
Depression	2720 (24.59)
Insomnia	1537 (13.90)
High blood pressure	1402 (12.68)
Pain	1006 (9.10)
High cholesterol	905 (8.18)
Asthma	773 (6.99)
Posttraumatic stress disorder	772 (6.98)
Arthritis	698 (6.31)
Diabetes	330 (2.98)
Cancer	265 (2.40)
Heart	206 (1.86)
Emphysema or chronic obstructive pulmonary disease	50 (0.45)
Other lung disease	49 (0.44)
Other chronic condition	1076 (9.73)
None	4772 (43.14)

**Table 3.** Self-reported sleep difficulties in Calm subscribers (N=11,380).

Type of sleep difficulty	Values, n (%)
Difficulty falling asleep only	2664 (23.41)
Difficulty staying asleep only	2112 (18.56)
Difficulty both falling and staying asleep	3928 (34.52)
No difficulty falling or staying asleep	2676 (23.51)

## Usage Patterns

The most common reasons for starting to use Calm were to improve sleep (62.97%, 7475/11,870) and to improve stress (62.11%, 7373/11,870), followed by reducing depression or anxiety (54.47%, 6465/11,870) and improving overall health (40.05%, 4754/11,870; see [Multimedia Appendix 3](#) for less frequently reported reasons).

## Overall Usage of Calm and Its Components

Respondents had been using Calm for an average of 11.49 months (SD 10.49), with 55.89% (6042/10,810) having used the app for less than 1 year. Most of them used Calm 5 or more times per week (60.03%, 7281/12,129), with 27.45% (3330/12,129) using 3 to 4 times per week and 12.52% (1518/12,129) using 1 to 2 times per week. Meditations were the most popular components (80.20%, 9497/11,841), followed by Sleep Stories (55.66%, 6591/11,841; see [Multimedia Appendix 4](#) for usage of less popular components).

Approximately, half (51.09%, 6111/11,962) of respondents used at least one tool to help, motivate, or encourage themselves to use Calm. Tracking (36.62%, 4380/11,962) and reminders (15.81%, 1891/11,962) were the most common tools. The least common tools were the Calm Facebook community (3.33%,

398/11,962) and sharing app usage on social media (2.40%, 287/11,962).

Low-frequency users (ie, using Calm 1 or 2 times per week) were significantly more likely to use the reminders ( $\chi^2_{1}=84.2$ ;  $P<.001$ ), whereas high-frequency users (ie, using Calm at least three times per week) were more likely to use meditation tracking ( $\chi^2_{1}=56.2$ ;  $P<.001$ ) and the Calm Facebook community ( $\chi^2_{1}=7.5$ ;  $P=.01$ ). When asked what encouraged or motivated high-frequency users to use the app, 42.34% (4493/10,611) reported that they used Calm before bed, and 53.33% (5659/10,611) were *committed* to using Calm.

A total of 40% (4500/11,249) of respondents had friends or family members that also used Calm, and, of those, 75.00% (3375/4500) talked with their friends or family about using the app. However, when participants in the overall sample were asked whether they would like to be connected with other Calm users, 81.59% (9178/11,249) declined. Respondents with mental health diagnoses were significantly more likely to want to connect with other Calm users (22.37%, 961/4295) than those without mental health diagnoses (15.24%, 1051/6895;  $\chi^2_{1}=91.3$ ;  $P<.001$ ). Desire to connect with other Calm users was not related

to physical health diagnoses (18.35%, 622/3390 vs 17.82%, 1390/7800 without;  $\chi^2_{1}=0.4$ ;  $P=.50$ ).

### Differences in Usage by Demographics

There were no significant gender differences in how often respondents used Calm ( $\chi^2_{2}=4.3$ ;  $P=.12$ ). Males were significantly more likely to use meditations (88.38%, 1924/2177) than females (77.87%, 6835/8778;  $\chi^2_{1}=120.3$ ;  $P<.001$ ), whereas females (60.42%, 5304/8778) were more likely to use Sleep Stories than males (35.42%, 771/2177;  $\chi^2_{1}=441.6$ ;  $P<.001$ ). Females also used Sleep Stories more frequently, even when controlling for sleep difficulties (OR 0.56 [95% CI 0.49 to 0.63];  $P<.001$ ). Older users tended to use the app more frequently (OR 1.02 [95% CI 1.01 to 1.02]) and were more likely to use Sleep Stories (OR 1.01 [95% CI 1.0106 to 1.011]); younger respondents were more likely to use meditations (OR 0.99 [95% CI 0.98 to 0.99]).

### Calm Usage Based on Health Diagnoses

Participants were separated into 4 groups based on diagnoses when they downloaded Calm: (1) mental health diagnoses only

(ie, no comorbid physical health diagnoses), (2) physical health diagnoses only (ie, no comorbid mental health diagnoses), (3) both physical and mental health diagnoses, and (4) no chronic health diagnoses (Table 4).

Respondents with physical health diagnoses only or with both mental and physical health diagnoses used Calm more frequently (ie, 5 or more times per week) than respondents with only mental health diagnoses or with no chronic health diagnoses ( $\chi^2_{6}=77.3$ ;  $P<.001$ ; Table 5). Respondents with only mental health diagnoses were most likely to use meditations, followed by those with no chronic health diagnoses ( $\chi^2_{3}=38.9$ ;  $P<.001$ ); conversely, those with no chronic health diagnoses were least likely to use Sleep Stories, whereas those with physical health and mental health diagnoses were most likely to use Sleep Stories ( $\chi^2_{3}=52.0$ ;  $P<.001$ ).

Subscribers with sleep difficulties used Calm more frequently (ie, 5 or more times per week;  $\chi^2_{8}=11.5$ ;  $P=.003$ ) and were significantly more likely to use Sleep Stories ( $\chi^2_{1}=1590.2$ ;  $P<.001$ ; Table 6). Those without sleep difficulties were more likely to use meditations ( $\chi^2_{1}=273.2$ ;  $P<.001$ ).

**Table 4.** Types of self-reported health diagnoses in Calm subscribers (N=11,061).

Diagnosis type	Values, n (%)
Mental health <sup>a</sup> only	2729 (24.67)
Physical health <sup>b</sup> only	1740 (15.73)
Mental and physical health	1820 (16.45)
No chronic health diagnoses	4772 (43.14)

<sup>a</sup>Mental health diagnoses were depression, anxiety, and posttraumatic stress disorder.

<sup>b</sup>Physical health diagnoses were high blood pressure, pain, high cholesterol, asthma, arthritis, cancer, heart disease, emphysema or chronic obstructive pulmonary disease, and other lung disease.

**Table 5.** Differences in the frequency of using Calm based on health diagnoses (N=11,061).

Category	Only mental health diagnoses (N=2729), n (%)	Only physical health diagnoses (N=1740), n (%)	Mental and physical health diagnoses (N=1820), n (%)	No chronic health diagnoses (N=4772), n (%)
<b>Frequency of using Calm (times per week)</b>				
1-2 <sup>a</sup>	324 (11.87)	172 (9.89)	200 (10.99)	609 (12.76)
3-4 <sup>a</sup>	778 (28.51)	411 (23.6)	440 (24.18)	1407 (29.48)
5+ <sup>a</sup>	1627 (59.62)	1157 (66.49)	1180 (64.84)	2756 (57.75)
<b>Components used</b>				
Meditations <sup>a</sup>	2279 (83.51)	1325 (76.15)	1434 (78.79)	3808 (79.80)
Sleep Stories <sup>a</sup>	1515 (55.51)	1021 (58.68)	1116 (61.32)	2495 (52.28)

<sup>a</sup>Chi-square was statistically significant;  $P<.05$ .

**Table 6.** Differences in the frequency of using Calm, frequency of using meditations and Sleep Stories, and reasons for starting to use Calm based on reported sleep difficulties (N=11,360).

Category	Sleep difficulties (N=8684), n (%)	No sleep difficulties (N=2676),n (%)
<b>Frequency of using Calm (times per week)</b>		
1-2	1020 (11.75)	367 (13.71)
3-4	2323 (26.75)	748 (27.95)
5+ <sup>a</sup>	5361 (61.73)	1561 (58.33)
<b>Components used</b>		
Meditations <sup>a</sup>	6677 (76.89)	2431 (90.84)
Sleep stories <sup>a</sup>	5589 (64.36)	755 (28.21)

<sup>a</sup>Chi-square test was statistically significant;  $P < .05$ .

### ***Relationship Between Frequency of Calm Use and Noticing Changes in Health, Stress, and Sleep***

The frequency of using Calm was associated with incremental increases in the likelihood of noticing changes in mental health ( $\chi^2_2=136.8$ ;  $P < .001$ ), physical health ( $\chi^2_2=102.8$ ;  $P < .001$ ), stress

( $\chi^2_2=128.1$ ;  $P < .001$ ), and sleep ( $\chi^2_2=141.4$ ;  $P < .001$ ; [Table 7](#)). In addition, participants who had been using Calm longer were more likely to notice changes in mental health (OR 1.06 [95% CI 1.05 to 1.06];  $P < .001$ ), physical health (OR 1.01 [95% CI 1.01 to 1.02];  $P < .001$ ), stress (OR 1.04 [95% CI 1.04 to 1.05];  $P < .001$ ), and sleep (OR 1.004 [95% CI 1.00 to 1.01];  $P = .03$ ).

**Table 7.** Differences in noticing changes in mental health, physical health, stress, and sleep based on frequency of using Calm.

Reported changes	Did not use regularly, n (%)	Used 1-2 times per week, n (%)	Used 3-4 times per week, n (%)	Used 5+ times per week, n (%)
<b>Calm</b> (N=12,192)	— <sup>a</sup>	1518 (12.97)	3330 (27.31)	7281 (59.72)
Mental health <sup>b</sup>	—	932 (72.08)	2325 (76.56)	5734 (83.83)
Physical health <sup>b</sup>	—	594 (46.19)	1558 (51.35)	4045 (59.13)
Stress <sup>b</sup>	—	875 (68.15)	2299 (75.97)	5550 (81.46)
Sleep <sup>b</sup>	—	745 (58.16)	2011 (66.02)	5028 (73.27)
<b>Meditations</b> (N=11,189)	2248 (20.09)	1889 (16.88)	2832 (25.31)	4220 (37.72)
Mental health <sup>b</sup>	1209 (54.56)	1490 (78.46)	2430 (85.96)	3862 (91.34)
Physical health <sup>b</sup>	1053 (47.14)	973 (51.51)	1598 (56.47)	2573 (61.15)
Stress <sup>b</sup>	1192 (53.94)	1428 (75.76)	2386 (84.61)	3718 (88.36)
Sleep <sup>b</sup>	1890 (84.07)	1304 (69.03)	1812 (63.98)	2778 (65.83)
<b>Sleep Stories</b> (N=11,189)	4959 (44.32)	2103 (18.80)	1866 (16.68)	2261 (20.21)
Mental health <sup>b</sup>	4272 (85.78)	1696 (80.46)	1381 (74.69)	1642 (73.53)
Physical health <sup>b</sup>	2721 (54.84)	1113 (52.92)	1016 (54.65)	1347 (60.21)
Stress <sup>b</sup>	4181 (84.24)	1607 (77.00)	1341 (72.53)	1595 (71.72)
Sleep <sup>b</sup>	2551 (50.84)	1545 (73.47)	1641 (87.94)	2077 (91.86)

<sup>a</sup>Cells containing — indicate that the response option in that column was not applicable to that question.

<sup>b</sup>Chi-square test was statistically significant;  $P < .05$ .

Frequency of using meditations was associated with significant, incremental increases in the percentage of respondents who noticed changes in their physical health ( $\chi^2_3=130.9$ ;  $P < .001$ ), mental health ( $\chi^2_3=1325.0$ ;  $P < .001$ ), and stress ( $\chi^2_3=1100.5$ ;  $P < .001$ ; [Table 7](#)). The frequency of using meditation was not associated with changes in sleep (OR 1.04 [95% CI 0.99 to 1.01];  $P = .12$ ) after controlling for sleep difficulties.

Frequency of using Sleep Stories was associated with significant, incremental increases in the percentage of respondents who noticed changes in their sleep ( $\chi^2_3=1665.3$ ;  $P < .001$ ; [Table 7](#)), and respondents who used Sleep Stories 5 or more times per week were significantly more likely to notice changes in physical health ( $\chi^2_3=27.2$ ;  $P < .001$ ). However, there were significant negative associations between the frequency of using Sleep Stories and noticing changes in mental health ( $\chi^2_3=197.3$ ;

$P < .001$ ) and stress ( $\chi^2_3=199.0$ ;  $P < .001$ ). Specifically, Sleep Stories users were less likely to notice changes in their mental health and stress, and low-frequency users (ie, 1-2 times per week) were more likely to notice changes in mental health and stress than more frequent users; see [Multimedia Appendix 5](#) for changes associated with other Calm components.

### **Relationship Between Sleep Difficulties and Noticing Changes in Health, Stress, and Sleep**

Respondents with sleep difficulties were more likely to notice changes in their physical health ( $\chi^2_1=49.2$ ;  $P < .001$ ) and sleep ( $\chi^2_1=2391.1$ ;  $P < .001$ ) after using Calm ([Table 8](#)), whereas those without sleep difficulties were more likely to notice changes in their mental health ( $\chi^2_1=33.1$ ;  $P < .001$ ) and stress ( $\chi^2_1=17.0$ ;  $P < .001$ ).

**Table 8.** Differences in noticing changes in mental health, physical health, stress, and sleep based on reported sleep difficulties (N=11,131).

Reported changes	Sleep difficulties (N=8527), n (%)	No sleep difficulties (N=2604), n (%)
Mental health <sup>a</sup>	6755 (79.30)	2212 (84.40)
Physical health <sup>a</sup>	4890 (57.35)	1290 (49.54)
Stress <sup>a</sup>	6576 (77.57)	2126 (81.36)
Sleep <sup>a</sup>	6961 (81.35)	808 (31.02)

<sup>a</sup>Chi-square test was statistically significant;  $P < .05$ .

## Discussion

### Summary

The aim of this paper was to report demographic characteristics, clinical characteristics, and usage patterns in a self-selected sample of paying subscribers of a consumer-based mobile app (ie, Calm). In addition, we explored the relationship between self-reported frequency of using Calm and its components and noticing changes in health (mental and physical), stress, and sleep disturbance. This is the first paper to report the characteristics of Calm subscribers in addition to the relationship between Calm use and reported changes in health, stress, and sleep.

Respondents were mostly white and mostly female. More than half reported a mental or physical health diagnosis, and approximately three-fourth reported sleep difficulties. The most common reason for starting to use Calm was for sleep, followed by stress and depression. Most respondents used Calm 5 or more times per week.

Meditations and Sleep Stories were the most commonly used components. It is not surprising that meditation was the most frequently used Calm component as Americans are more consistently including meditation as part of their lifestyle [21]. Meditation is the fifth most common type of complementary approach to health practiced by adults [22]. In addition, Calm was originally marketed as a mindfulness meditation app, and after many iterations and changes to the app over time, it now offers added components such as Sleep Stories.

### Gender Differences

Even though our sample was mostly women, men were slightly more likely to use meditation. Historically, women report higher use of meditation when compared with men. In a 2019 analysis of 2012 data from the National Health Interview Survey [21], 5.7 million men reported using meditation as compared with 12.2 million women. Both men and women reported meditation being helpful to reduce stress. In future studies, it would be interesting to know why men use the meditations more than the many other components of the app. This information may impact the way in which consumer-based meditation apps consider marketing content based on gender. For example, advertising Calm with pictures of men meditating may help to encourage men to subscribe to the app.

Women were more likely than men to use the Sleep Stories. This may be partially attributed to women reporting higher levels of sleep difficulties. It is well known that women have a

higher frequency of sleep disturbance as compared with men owing to hormonal changes over the life span [23]. Sleep meditations may represent a manageable, nonpharmacological strategy for women (and men) to improve sleep. However, further analyses suggest that sleep difficulties do not fully account for women's more frequent use of Sleep Stories. Future research using surveys and qualitative data from consumers (both men and women) about what they specifically enjoy about the Sleep Stories would be useful information for future content development, consumer engagement, and clinical trials.

### Tools for Engagement

Tracking and reminders were the most common in-app tools used to support engagement, and using the app at bedtime (using Calm at bedtime) appeared to be the most effective adherence strategy to maintain engagement with the app over time. Research on the efficacy of mobile apps in impacting behavior shows that tracking behavior and feedback are among the most successful strategies for behavior change [24]. Tracking is also one of the most popular technology strategies used by mobile apps for health and wellness [25]. Almost 40% of our sample reported using tracking as a tool to help them continue to engage in using Calm. Currently, Calm tracks how many sessions and minutes the user has participated in over time. Recent evidence suggests that tracking plus feedback may be more impactful for improvements in behavior change as compared with tracking alone [26-28]. For example, in a randomized clinical trial testing the effects of an automated tracking-texting intervention on physical activity (mActive), coupling text messages containing automated personalized feedback with physical activity tracking led to twice as many participants achieving their daily step goal as compared with other groups [26]. Providing the subscriber feedback, based on the use of the app, could thus be an effective strategy for consumer-based mindfulness meditation apps to keep subscribers engaged once they sign up as a paying member. In addition, finding other ways to track and provide feedback outside of sessions and minutes of meditation could provide more value to a meditation app. For example, the users' perception of their stress, mindfulness, or happiness after a session could be tracked in conjunction with patterns of use. Providing this information as feedback would allow users to observe the relationship between using Calm and changes in their moods, thoughts, and behaviors. Future research in this area is warranted.

When we asked those who engaged in the app more often (more than 3 times per week) precisely how they stay engaged, many reported that using Calm at bedtime (eg, Sleep Stories, Daily Calm, and other sleep meditations) contributed to adherence.



It is well known that there are barriers to beginning and maintaining health behaviors (eg, physical activity, diet, and meditation) [29]. However, using Calm (and other mindfulness-based apps) at bedtime/for sleep may have fewer barriers (eg, time and motivation) than participating in another health behavior. For example, Calm Sleep Stories require the user to begin the Sleep Story by lying in bed and assuming a comfortable position with the goal of helping the user to fall asleep. This does not require added time or motivation (beyond starting the Sleep Story) and may be an easier task, thus users reported using Calm at bedtime/for sleep as a way of using the app more consistently.

An interesting finding from our survey was that almost half of our sample had friends or family that used Calm, and many talked with their friends or family about Calm. However, overall, respondents reported not wanting to be connected with other Calm users. When exploring mental or physical health diagnoses and the relationship to reporting the desire to be connected to other Calm users, those with mental health diagnoses wanted to connect with other Calm users. Recent research suggests that individuals with mental health problems are increasingly turning to online communities for social support [30,31]. Although more than half of the survey sample started using Calm for mental health reasons (ie, anxiety and depression), and almost half reported a mental health diagnoses (ie, anxiety, depression, and posttraumatic stress); only 3% used the Calm Facebook community to maintain app engagement. This is surprising given the strong presence of Calm on Facebook and other social media platforms. It may be that this convenience sample of highly engaged users already felt well connected in terms of social support, and thus, their feedback on additional support might not be generalizable. It is also possible that concerns related to privacy and stigma deter users with mental health diagnoses from disclosing personal or emotional information to a larger community [32]. These individuals may prefer to connect specifically with other users who share common difficulties [33,34]. Future investigation is necessary to determine what the best way is to connect users and to facilitate social support for subscribers of these apps.

### Changes in Health, Stress, and Sleep

Participants using Calm more frequently were incrementally more likely to notice changes in both mental and physical health, stress, and sleep, as were those who reported using Calm for a longer duration. This is not surprising. Participants are likely becoming more mindful and reap the benefits associated with mindfulness [35,36]. Present literature on the amount of meditation practice needed to observe changes in a range of health outcomes is varied and limited [37]. However, there is promising evidence for greater effects on health, stress, and sleep as time spent (eg duration, frequency) increases [37,38]. This information is important because regardless of the component (ie, general use of Calm, specifically meditation, and specifically Sleep Stories), participants who used the app more frequently clearly noticed changes in their health. It is well known that individuals are more likely to participate in a behavior if they enjoy it or if they believe it benefits them [39,40]. However, it is important to note that survey questions about changes in all aspects of health were nondirectional, and

respondents' reports of noticing changes were interpreted with the assumption that changes reflect improvements. In addition, future research should explore the correlation between self-reported *noticing changes* and directional changes on validated outcome measures. It is possible that, retrospectively, users believe that there have been changes, but when completing a fully validated measure, scores remain the same over time. The relationship between actual and perceived outcomes and time spent (duration in minutes and length of membership) using the app would also be worth exploring in future research. Future research should also explore the optimal dose of mindfulness meditation on a mobile app to improve health outcomes. For example, it is possible that total time spent using a meditation app is less important than frequency of use it (eg, 10 min a day for 3 days could be better than 50 min in 1 day).

Respondents who used the Sleep Stories more frequently were more likely to report noticing changes in their sleep and physical health. Consumer sleep technologies have become quite popular to help individuals improve or self-monitor their sleep [20]. In a review by Ko in 2015 on consumer sleep technologies, a number of the top-rated sleep apps helped users track their *sleep trends* (movements that represent presence or absence of sleep), recorded ambient sounds (sleep talking and snoring), alerted users when to go to bed, and tracked their sleep habits over time. Although some of the apps were developed to facilitate sleep onset using visual graphics, relaxing music, and nature sounds, none incorporated Sleep Stories. Uniquely, Calm has Sleep Stories, similar to meditation, that help the user to practice moment-to-moment awareness, experience greater attentional control, and decrease ruminating thoughts, which may lead to less anxiety about sleeplessness [41,42]. The stories bring listeners focus into the present moment, helping them to disengage with negative thoughts/beliefs about sleeplessness, and provide cognitive distraction, which may lower presleep anxiety or arousal and help users sleep better.

It is important to note that those who used the Sleep Stories more frequently reported noticing changes in their sleep and physical health only, whereas those using meditations more frequently noticed changes in their mental and physical health, as well as stress. Although there is a large body of research demonstrating the relationship between sleep and physical health [43], many studies observe similar or stronger relationships between sleep and mental health [44,45]. Given this, it is surprising that using Sleep Stories was not associated with changes in mental health or stress. There may be more varying long-term benefits of using Calm for meditation as compared with sleep. Little research has been conducted to determine the effects of sleep components of mindfulness-based apps, and no studies have specifically assessed the effects of Calm Sleep Stories in improving sleep and health outcomes. Considering the increases in sleep components to mindfulness meditation apps, future research on the extent to which these components impact sleep is warranted.

### Limitations

There were a number of limitations, inherent in descriptive, survey-based studies. First, although this was a convenience sample of Calm users that opened an email from Calm and

decided to complete the survey, engagement was still varied (60.03% of respondents used the app 5 or more times per week, 27.45% used 3-4 times per week, and 12.52% used only once or twice weekly), and there may be a viable sample of individuals who vary in their use of Calm. In addition, individuals who use Calm more frequently may have more favorable opinions of the app and have benefitted more from using it. Those who do not find Calm beneficial may have elected not to participate. In addition, the survey did not include the option to endorse using Calm less often than 1 to 2 times per week, which may have decreased participation rates of infrequent users. Finally, our sample is mostly white, female, and with higher income. Future studies should use stratified recruitment strategies to engage a more diverse sample in terms of age, gender, income, race/ethnicity, and pattern of usage. Second, as this was a cross-sectional, nonexperimental study, we cannot infer causality with regard to the impact of Calm; rather, analyses rely on subscribers' recollections of app usage, perceptions of changes over time, and their beliefs about how Calm has impacted them. Future studies should explore these relationships prospectively, and randomized clinical trials are needed to determine extent to which changes can be attributed to Calm usage. Third, the survey items were designed specifically for this study, and thus, there is no previous research on their validity. In addition, questions asking participants to notice changes in their health, stress, and sleep after using Calm were dichotomous (yes/no), and results were interpreted with the assumption that changes reflect improvements. Future studies should extend the findings of this report by using validated questionnaires to assess directional changes in mental

health, physical health, stress, and sleep outcomes. Finally, as many respondents used both meditations and Sleep Stories, it is difficult to disentangle the unique effects of those components. Future dismantling studies could randomize participants to only use 1 component of the app and look at the benefits of each component on its own. In addition, future surveys should collect more detailed information regarding the specific aspects of meditations and Sleep Stories that users enjoy, and this information should be presented alongside qualitative data, as such data could inform future experimental studies.

## Conclusions

This is the first study to report the demographic and clinical characteristics and patterns of usage among a large sample of paid subscribers to a top-rated health and fitness app, Calm. Results showed that most subscribers used meditation and Sleep Stories at least three to four times per week, and more frequent use was associated with perceived changes in their physical and mental health, stress, and sleep. We have also identified the most popular tools within the app preferred by subscribers. This paper provides information for Calm and other mindfulness-based apps about strategies to tailor content and features (eg, emphasize benefits for sleep, provide tracking with feedback, and incorporate social support that is private/anonymous) and potentially increase subscribers and retention rates. From a scientific perspective, these strategies can help participants engage in studies and facilitate feasibility. This study sets the stage for future clinical trials aimed at determining the efficacy of Calm in improving physical and mental health, stress, and sleep.

---

## Conflicts of Interest

JH is currently the Director of Science at Calm. JH has been conducting research with Calm as a partner almost 5 years before becoming the Director of Science and the Scientific Advisory Board (SAB). A-MV, CC, and MB are members of Calm's SAB and are independent from Calm leadership. Their role is to ensure the quality of Calm's science. There are no financial incentives from the growth of Calm to any author.

---

### Multimedia Appendix 1

Screenshots of the home screen and meditation screen in Calm app.

[[PNG File , 3981 KB - mhealth\\_v7i11e15648\\_app1.png](#) ]

---

### Multimedia Appendix 2

Questions from Calm user engagement survey.

[[PDF File \(Adobe PDF File\), 60 KB - mhealth\\_v7i11e15648\\_app2.pdf](#) ]

---

### Multimedia Appendix 3

Additional details on statistical analyses of reasons for starting Calm.

[[PDF File \(Adobe PDF File\), 49 KB - mhealth\\_v7i11e15648\\_app3.pdf](#) ]

---

### Multimedia Appendix 4

Additional details on statistical analyses of Calm component usage.

[[PDF File \(Adobe PDF File\), 66 KB - mhealth\\_v7i11e15648\\_app4.pdf](#) ]

---

### Multimedia Appendix 5

Additional details on statistical analyses of changes in health stress and sleep after using Calm.

[[PDF File \(Adobe PDF File\), 73 KB - mhealth\\_v7i11e15648\\_app5.pdf](#) ]

## References

1. Kabat-Zinn J, Hanh TN. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. United States: Bantam Books; 2009.
2. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. *Clin Psychol* 2003;10(2):144-156. [doi: [10.1093/clipsy/bpg016](https://doi.org/10.1093/clipsy/bpg016)]
3. Burke A, Lam CN, Stussman B, Yang H. Prevalence and patterns of use of mantra, mindfulness and spiritual meditation among adults in the United States. *BMC Complement Altern Med* 2017 Jun 15;17(1):316 [FREE Full text] [doi: [10.1186/s12906-017-1827-8](https://doi.org/10.1186/s12906-017-1827-8)] [Medline: [28619092](https://pubmed.ncbi.nlm.nih.gov/28619092/)]
4. Creswell JD. Mindfulness Interventions. *Annu Rev Psychol* 2017 Jan 3;68:491-516. [doi: [10.1146/annurev-psych-042716-051139](https://doi.org/10.1146/annurev-psych-042716-051139)] [Medline: [27687118](https://pubmed.ncbi.nlm.nih.gov/27687118/)]
5. Flett JAM, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness* 2019;10(5):863-876. [doi: [10.1007/s12671-018-1050-9](https://doi.org/10.1007/s12671-018-1050-9)]
6. Rusch HL, Rosario M, Levison LM, Olivera A, Livingston WS, Wu T, et al. The effect of mindfulness meditation on sleep quality: a systematic review and meta-analysis of randomized controlled trials. *Ann N Y Acad Sci* 2019 Jun;1445(1):5-16. [doi: [10.1111/nyas.13996](https://doi.org/10.1111/nyas.13996)] [Medline: [30575050](https://pubmed.ncbi.nlm.nih.gov/30575050/)]
7. Wahbeh H, Svalina MN, Oken BS. Group, one-on-one, or internet? Preferences for mindfulness meditation delivery format and their predictors. *Open Med J* 2014;1:66-74 [FREE Full text] [doi: [10.2174/1874220301401010066](https://doi.org/10.2174/1874220301401010066)] [Medline: [27057260](https://pubmed.ncbi.nlm.nih.gov/27057260/)]
8. Kracen AC, Mastnak JM, Loaiza KA, Matthieu MM. Group therapy among OEF/OIF veterans: treatment barriers and preferences. *Mil Med* 2013 Jan;178(1):e146-e149. [doi: [10.7205/MILMED-D-12-00213](https://doi.org/10.7205/MILMED-D-12-00213)] [Medline: [23764162](https://pubmed.ncbi.nlm.nih.gov/23764162/)]
9. Huberty J, Eckert R, Larkey L, Kurka J, de Jesús SA, Yoo W, et al. Smartphone-based meditation for myeloproliferative neoplasm patients: feasibility study to inform future trials. *JMIR Form Res* 2019 Apr 29;3(2):e12662 [FREE Full text] [doi: [10.2196/12662](https://doi.org/10.2196/12662)] [Medline: [31033443](https://pubmed.ncbi.nlm.nih.gov/31033443/)]
10. Lau DT, Kirby JB. The relationship between living arrangement and preventive care use among community-dwelling elderly persons. *Am J Public Health* 2009 Jul;99(7):1315-1321 [FREE Full text] [doi: [10.2105/AJPH.2008.151142](https://doi.org/10.2105/AJPH.2008.151142)] [Medline: [19443817](https://pubmed.ncbi.nlm.nih.gov/19443817/)]
11. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. *Gen Hosp Psychiatry* 2013;35(4):332-338 [FREE Full text] [doi: [10.1016/j.genhosppsy.2013.03.008](https://doi.org/10.1016/j.genhosppsy.2013.03.008)] [Medline: [23664503](https://pubmed.ncbi.nlm.nih.gov/23664503/)]
12. Silver L, Smith A, Johnson C, Taylor K, Jiang J, Anderson M, et al. Pew Research Center. 2019 Mar 7. Mobile Connectivity in Emerging Economies URL: [https://www.pewinternet.org/wp-content/uploads/sites/9/2019/03/PI\\_2019.03.07\\_Mobile-Connectivity\\_FINAL.pdf](https://www.pewinternet.org/wp-content/uploads/sites/9/2019/03/PI_2019.03.07_Mobile-Connectivity_FINAL.pdf) [accessed 2019-09-25]
13. Taylor K, Silver L. Pew Research Center. 2019 Feb 5. Smartphone Ownership is Growing Rapidly Around the World, but Not Always Equally URL: [https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center\\_Global-Technology-Use-2018\\_2019-02-05.pdf](https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center_Global-Technology-Use-2018_2019-02-05.pdf) [accessed 2019-09-25]
14. Elavsky S, Smahel D, Machackova H. Who are mobile app users from healthy lifestyle websites? Analysis of patterns of app use and user characteristics. *Transl Behav Med* 2017 Dec;7(4):891-901 [FREE Full text] [doi: [10.1007/s13142-017-0525-x](https://doi.org/10.1007/s13142-017-0525-x)] [Medline: [28929368](https://pubmed.ncbi.nlm.nih.gov/28929368/)]
15. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. *JMIR Mhealth Uhealth* 2013 Nov 1;1(2):e24 [FREE Full text] [doi: [10.2196/mhealth.2733](https://doi.org/10.2196/mhealth.2733)] [Medline: [25099314](https://pubmed.ncbi.nlm.nih.gov/25099314/)]
16. Krebs P, Duncan DT. Health app use among us mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov 4;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
17. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. *JMIR Mhealth Uhealth* 2015 Aug 19;3(3):e82 [FREE Full text] [doi: [10.2196/mhealth.4328](https://doi.org/10.2196/mhealth.4328)] [Medline: [26290327](https://pubmed.ncbi.nlm.nih.gov/26290327/)]
18. Chan J. SensorTower. 2019 Jul 8. Top Health & Fitness Apps in the US for June 2019 by Downloads URL: <https://sensortower.com/blog/top-health-and-fitness-apps-us-june-2019-downloads> [accessed 2019-09-25]
19. Capritto A. CNET: Product reviews, how-tos, deals and the latest tech news. 2019 Oct 10. 11 Meditation Apps for Better Sleep and Less Stress URL: <https://www.cnet.com/how-to/10-best-meditation-apps-of-2019/> [accessed 2019-09-25]
20. Ko PT, Kientz JA, Choe EK, Kay M, Landis CA, Watson NF. Consumer sleep technologies: a review of the landscape. *J Clin Sleep Med* 2015 Dec 15;11(12):1455-1461 [FREE Full text] [doi: [10.5664/jcsm.5288](https://doi.org/10.5664/jcsm.5288)] [Medline: [26156958](https://pubmed.ncbi.nlm.nih.gov/26156958/)]
21. Upchurch DM, Johnson PJ. Gender differences in prevalence, patterns, purposes, and perceived benefits of meditation practices in the United States. *J Womens Health (Larchmt)* 2019 Feb;28(2):135-142. [doi: [10.1089/jwh.2018.7178](https://doi.org/10.1089/jwh.2018.7178)] [Medline: [30543475](https://pubmed.ncbi.nlm.nih.gov/30543475/)]
22. Clarke TC, Black LI, Stussman BJ, Barnes PM, Nahin RL. Trends in the use of complementary health approaches among adults: United States, 2002-2012. *Natl Health Stat Report* 2015 Feb 10(79):1-16 [FREE Full text] [Medline: [25671660](https://pubmed.ncbi.nlm.nih.gov/25671660/)]
23. Pengo MF, Won CH, Bourjeily G. Sleep in women across the life span. *Chest* 2018 Jul;154(1):196-206 [FREE Full text] [doi: [10.1016/j.chest.2018.04.005](https://doi.org/10.1016/j.chest.2018.04.005)] [Medline: [29679598](https://pubmed.ncbi.nlm.nih.gov/29679598/)]
24. Fitzgerald M, McClelland T. What makes a mobile app successful in supporting health behaviour change? *Health Educ J* 2017;76(3):373-381. [doi: [10.1177/0017896916681179](https://doi.org/10.1177/0017896916681179)]

25. Orji R, Moffatt K. Persuasive technology for health and wellness: state-of-the-art and emerging trends. *Health Informatics J* 2018 Mar;24(1):66-91. [doi: [10.1177/1460458216650979](https://doi.org/10.1177/1460458216650979)] [Medline: [27245673](https://pubmed.ncbi.nlm.nih.gov/27245673/)]
26. Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, et al. mActive: A randomized clinical trial of an automated mHealth intervention for physical activity promotion. *J Am Heart Assoc* 2015 Nov 9;4(11) [FREE Full text] [doi: [10.1161/JAHA.115.002239](https://doi.org/10.1161/JAHA.115.002239)] [Medline: [26553211](https://pubmed.ncbi.nlm.nih.gov/26553211/)]
27. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. *Am J Prev Med* 2007 Oct;33(4):336-345 [FREE Full text] [doi: [10.1016/j.amepre.2007.05.007](https://doi.org/10.1016/j.amepre.2007.05.007)] [Medline: [17888860](https://pubmed.ncbi.nlm.nih.gov/17888860/)]
28. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. *JMIR Mhealth Uhealth* 2015 May 14;3(2):e42 [FREE Full text] [doi: [10.2196/mhealth.4160](https://doi.org/10.2196/mhealth.4160)] [Medline: [25977197](https://pubmed.ncbi.nlm.nih.gov/25977197/)]
29. Kelly S, Martin S, Kuhn I, Cowan A, Brayne C, Lafortune L. Barriers and facilitators to the uptake and maintenance of healthy behaviours by people at mid-life: a rapid systematic review. *PLoS One* 2016;11(1):e0145074 [FREE Full text] [doi: [10.1371/journal.pone.0145074](https://doi.org/10.1371/journal.pone.0145074)] [Medline: [26815199](https://pubmed.ncbi.nlm.nih.gov/26815199/)]
30. Naslund JA, Aschbrenner KA, McHugo GJ, Ünützer J, Marsch LA, Bartels SJ. Exploring opportunities to support mental health care using social media: a survey of social media users with mental illness. *Early Interv Psychiatry* 2019 Jun;13(3):405-413. [doi: [10.1111/eip.12496](https://doi.org/10.1111/eip.12496)] [Medline: [29052947](https://pubmed.ncbi.nlm.nih.gov/29052947/)]
31. Ridout B, Campbell A. The use of social networking sites in mental health interventions for young people: systematic review. *J Med Internet Res* 2018 Dec 18;20(12):e12244 [FREE Full text] [doi: [10.2196/12244](https://doi.org/10.2196/12244)] [Medline: [30563811](https://pubmed.ncbi.nlm.nih.gov/30563811/)]
32. Zhang X, Liu S, Chen X, Wang L, Gao B, Zhu Q. Health information privacy concerns, antecedents, and information disclosure intention in online health communities. *Inform Manag* 2018 Jun;55(4):482-493. [doi: [10.1016/j.im.2017.11.003](https://doi.org/10.1016/j.im.2017.11.003)]
33. de Choudhury M, De S. Mental Health Discourse on Reddit: Self-Disclosure, Social Support, and Anonymity. In: *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. 2014 Presented at: ICWSM'14; June 1-4, 2014; Ann Arbor, Michigan, USA p. 71-80 URL: <https://pdfs.semanticscholar.org/2db7/15a479c8961d3020fe906f7bedfa0311b937.pdf>
34. Kingod N, Cleal B, Wahlberg A, Husted GR. Online peer-to-peer communities in the daily lives of people with chronic illness: a qualitative systematic review. *Qual Health Res* 2017 Jan;27(1):89-99. [doi: [10.1177/1049732316680203](https://doi.org/10.1177/1049732316680203)] [Medline: [27956659](https://pubmed.ncbi.nlm.nih.gov/27956659/)]
35. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. A meta-analysis. *J Psychosom Res* 2004 Jul;57(1):35-43. [doi: [10.1016/S0022-3999\(03\)00573-7](https://doi.org/10.1016/S0022-3999(03)00573-7)] [Medline: [15256293](https://pubmed.ncbi.nlm.nih.gov/15256293/)]
36. Victorson D, Kentor M, Maletich C, Lawton RC, Kaufman VH, Borrero M, et al. Mindfulness meditation to promote wellness and manage chronic disease: a systematic review and meta-analysis of mindfulness-based randomized controlled trials relevant to lifestyle medicine. *Am J Lifestyle Med* 2015;9(3):185-211. [doi: [10.1177/1559827614537789](https://doi.org/10.1177/1559827614537789)]
37. Zeng X, Chio FH, Oei TP, Leung FY, Liu X. A systematic review of associations between amount of meditation practice and outcomes in interventions using the four immeasurables meditations. *Front Psychol* 2017;8:141 [FREE Full text] [doi: [10.3389/fpsyg.2017.00141](https://doi.org/10.3389/fpsyg.2017.00141)] [Medline: [28220101](https://pubmed.ncbi.nlm.nih.gov/28220101/)]
38. Schoormans D, Nyklíček I. Mindfulness and psychologic well-being: are they related to type of meditation technique practiced? *J Altern Complement Med* 2011 Jul;17(7):629-634. [doi: [10.1089/acm.2010.0332](https://doi.org/10.1089/acm.2010.0332)] [Medline: [21711203](https://pubmed.ncbi.nlm.nih.gov/21711203/)]
39. Jekauc D. Enjoyment during exercise mediates the effects of an intervention on exercise adherence. *Psychol* 2015;6(1):48-54. [doi: [10.4236/psych.2015.61005](https://doi.org/10.4236/psych.2015.61005)]
40. Lewis BA, Williams DM, Frayeh A, Marcus BH. Self-efficacy versus perceived enjoyment as predictors of physical activity behaviour. *Psychol Health* 2016;31(4):456-469 [FREE Full text] [doi: [10.1080/08870446.2015.1111372](https://doi.org/10.1080/08870446.2015.1111372)] [Medline: [26541890](https://pubmed.ncbi.nlm.nih.gov/26541890/)]
41. Harris K, Spiegelhalter K, Espie CA, MacMahon KM, Woods HC, Kyle SD. Sleep-related attentional bias in insomnia: a state-of-the-science review. *Clin Psychol Rev* 2015 Dec;42:16-27. [doi: [10.1016/j.cpr.2015.08.001](https://doi.org/10.1016/j.cpr.2015.08.001)] [Medline: [26284598](https://pubmed.ncbi.nlm.nih.gov/26284598/)]
42. Lee CW, Ree MJ, Wong MY. Effective insomnia treatments: investigation of processes in mindfulness and cognitive therapy. *Behav Change* 2018;35(2):71-90. [doi: [10.1017/bec.2018.10](https://doi.org/10.1017/bec.2018.10)]
43. Itani O, Jike M, Watanabe N, Kaneita Y. Short sleep duration and health outcomes: a systematic review, meta-analysis, and meta-regression. *Sleep Med* 2017 Apr;32:246-256. [doi: [10.1016/j.sleep.2016.08.006](https://doi.org/10.1016/j.sleep.2016.08.006)] [Medline: [27743803](https://pubmed.ncbi.nlm.nih.gov/27743803/)]
44. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS-29 v2.0 profile physical and mental health summary scores. *Qual Life Res* 2018 Jul;27(7):1885-1891 [FREE Full text] [doi: [10.1007/s11136-018-1842-3](https://doi.org/10.1007/s11136-018-1842-3)] [Medline: [29569016](https://pubmed.ncbi.nlm.nih.gov/29569016/)]
45. Lallukka T, Sivertsen B, Kronholm E, Bin YS, Øverland S, Glozier N. Association of sleep duration and sleep quality with the physical, social, and emotional functioning among Australian adults. *Sleep Health* 2018 Apr;4(2):194-200. [doi: [10.1016/j.sleh.2017.11.006](https://doi.org/10.1016/j.sleh.2017.11.006)] [Medline: [29555134](https://pubmed.ncbi.nlm.nih.gov/29555134/)]

---

## Abbreviations

**SAB:** Scientific Advisory Board

*Edited by G Eysenbach; submitted 25.07.19; peer-reviewed by Z Ma, K Kaipainen, V Rocío, A Enrique, DJ Dutcher; comments to author 19.08.19; revised version received 05.09.19; accepted 28.09.19; published 03.11.19.*

*Please cite as:*

*Huberty J, Vranceanu AM, Carney C, Breus M, Gordon M, Puzia ME*

*Characteristics and Usage Patterns Among 12,151 Paid Subscribers of the Calm Meditation App: Cross-Sectional Survey*

*JMIR Mhealth Uhealth 2019;7(11):e15648*

*URL: <https://mhealth.jmir.org/2019/11/e15648>*

*doi: [10.2196/15648](https://doi.org/10.2196/15648)*

*PMID: [31682582](https://pubmed.ncbi.nlm.nih.gov/31682582/)*

©Jennifer Huberty, Ana-Maria Vranceanu, Colleen Carney, Michael Breus, Michael Gordon, Megan Elizabeth Puzia. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 03.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Improving Medication Information Presentation Through Interactive Visualization in Mobile Apps: Human Factors Design

Don Roosan<sup>1</sup>, PharmD, PhD; Yan Li<sup>2</sup>, PhD; Anandi Law<sup>1</sup>, PhD; Huy Truong<sup>3</sup>, BSc; Mazharul Karim<sup>1</sup>, BPharm; Jay Chok<sup>3</sup>, PhD; Moom Roosan<sup>4</sup>, PharmD, PhD

<sup>1</sup>Western University of Health Sciences, College of Pharmacy, Pomona, CA, United States

<sup>2</sup>Claremont Graduate University, Claremont, CA, United States

<sup>3</sup>Keck Graduate Institute, School of Pharmacy, Claremont, CA, United States

<sup>4</sup>Chapman University, School of Pharmacy, Irvine, CA, United States

**Corresponding Author:**

Don Roosan, PharmD, PhD

Western University of Health Sciences

College of Pharmacy

309 E 2nd St

Pomona, CA,

United States

Phone: 1 909 469 8778

Fax: 1 909 469 5428

Email: [droosan@westernu.edu](mailto:droosan@westernu.edu)

## Abstract

**Background:** Despite the detailed patient package inserts (PPIs) with prescription drugs that communicate crucial information about safety, there is a critical gap between patient understanding and the knowledge presented. As a result, patients may suffer from adverse events. We propose using human factors design methodologies such as hierarchical task analysis (HTA) and interactive visualization to bridge this gap. We hypothesize that an innovative mobile app employing human factors design with an interactive visualization can deliver PPI information aligned with patients' information processing heuristics. Such an app may help patients gain an improved overall knowledge of medications.

**Objective:** The objective of this study was to explore the feasibility of designing an interactive visualization-based mobile app using an HTA approach through a mobile prototype.

**Methods:** Two pharmacists constructed the HTA for the drug risperidone. Later, the specific requirements of the design were translated using infographics. We transferred the wireframes of the prototype into an interactive user interface. Finally, a usability evaluation of the mobile health app was conducted.

**Results:** A mobile app prototype using HTA and infographics was successfully created. We reiterated the design based on the specific recommendations from the usability evaluations.

**Conclusions:** Using HTA methodology, we successfully created a mobile prototype for delivering PPI on the drug risperidone to patients. The hierarchical goals and subgoals were translated into a mobile prototype.

(*JMIR Mhealth Uhealth* 2019;7(11):e15940) doi:[10.2196/15940](https://doi.org/10.2196/15940)

## KEYWORDS

visual perception; adverse drug event; human factors design; mobile health

## Introduction

### Background

Adverse drug reactions (ADRs) account for 4.2% to 30% of hospital admissions in the United States, costing up to 30.1 billion dollars annually [1,2]. For decades, the US Food and

Drug Administration (FDA) and drug manufacturers have administered detailed patient package inserts (PPIs) with prescription drugs to communicate crucial information for patient safety. In order to comply with FDA regulations, pharmaceutical manufacturers developed PPIs to assist health care professionals and patients in identifying any potential health conditions that may arise when consuming the medications.

PPIs provide important information, including warnings, precautions, and lists of adverse reactions and drug interactions [3].

PPIs play a pivotal role for patient safety. The current format for presenting important information in the PPIs does not engage patients in an effective manner. As a result, patients are often unable to identify crucial warnings, the consequences of which are partially manifested in the dramatic increase in ADR-related hospitalizations over the past decade (ie, 117% increase in medication- and drug-related hospitalization from 1997-2008) [4,5]. Moreover, lawsuits and litigation have resulted from off-label use of drugs and drug side effects even when the information to prevent these problems was included in the PPIs.

Recent research suggests minimal patient engagement with PPIs. For example, a recent study explored self-reported drug risk reading by comparing the results of eye tracking and actual information recall. Although the majority of participants claimed to have read the risk information, eye-tracking measures revealed no risk reading, and actual information recall was minimal recall [6]. One plausible explanation, as suggested by the cognition literature, is that patients do not retain medication information due to complex cognitive processes.

The cognition literature offers two perspectives for such behavior: information avoidance and familiarity. Information avoidance refers to not wanting to know information that will cause uncomfortable conflict in the individual's mind [7]. Familiarity increases with experience frequency and reduces the likelihood that laypeople will look for information. With no negative consequences for repeated usage, individuals become less concerned about the risks associated with the product [8]. Innovative interactive visualizations, such as infographics, have been frequently used to deliver health information due to their ability to present complex data in a simple and clear manner [9,10]. Presenting PPI information using interactive visualization can actively engage patients to reduce information avoidance due to a sense of perceived familiarity and empower patients to retain more information [11].

There is a clear need to present the PPI information in a logical and quick-to-find manner to effectively educate patients about pertinent and crucial life-saving information. Visualizing information using heuristics has proven to help clinicians and patients understand complex scientific data and improve performance decision outcomes [12,13]. Many studies have confirmed that visualization can improve a person's ability to remember and recall information [14,15]. Heuristics helps reduce cognition overload and cultivates faster ways of processing information by human minds [10]. Thus, it is crucial to organize information in the PPIs in a methodical manner to facilitate heuristic reasoning.

Current estimates suggest more than 40,000 mobile health (mHealth) apps are in use today [16]. Delivering PPI information through the most common platform can reach most patients and maximize the impact of our study. However, mHealth apps proliferate with little evidence for their effectiveness and little support for understanding how to best design these apps [17]. Most mHealth apps present medication information using static texts, and if visualization is used, it is not interactive.

Furthermore, evidence from recent studies shows that many of the mHealth apps do not use human factors design methodologies [18-20]. Without a clear understanding of the end-user requirements, crucial information about health can be presented in such a way that it can be misleading or misread. Among different human factors design methodologies, hierarchical task analysis (HTA) has been widely used for systems design in many different fields and shown to improve overall user satisfaction [21-23]. HTA focuses on the concept of the goal as a unit of behavior in terms of its objectives, which are decomposed into hierarchical subgoals [24]. Although HTA has been rarely used in mHealth app design, many successful fields such as aviation and the military have used HTA for their systems design and decision support.

This research seeks to bridge the gap between drug information presentation and mHealth app development. More specifically, we propose an interactive visualization approach to deliver PPI information to patients in a manner that aligns with their information processing heuristics through a mobile app that uses human factors design methodologies.

## Objective

The objective of the paper is to explore the feasibility of designing an interactive visualization-based mobile app prototype using the HTA approach.

## Methods

### Drug Selection

For the purpose of the prototype, the research team selected a drug product, risperidone, that has applications and uses for children, adults, and the elderly. Risperidone is an atypical antipsychotic drug that is a widely used benzisoxazole derivative approved by the FDA for the treatment of schizophrenia in 1994, for the short-term treatment of the mixed and manic states of bipolar disorder in 2003, and for depression and the treatment of irritability in children with autism in 2006 [25]. Risperidone has several important side effects of which the patients must be aware that are often ignored.

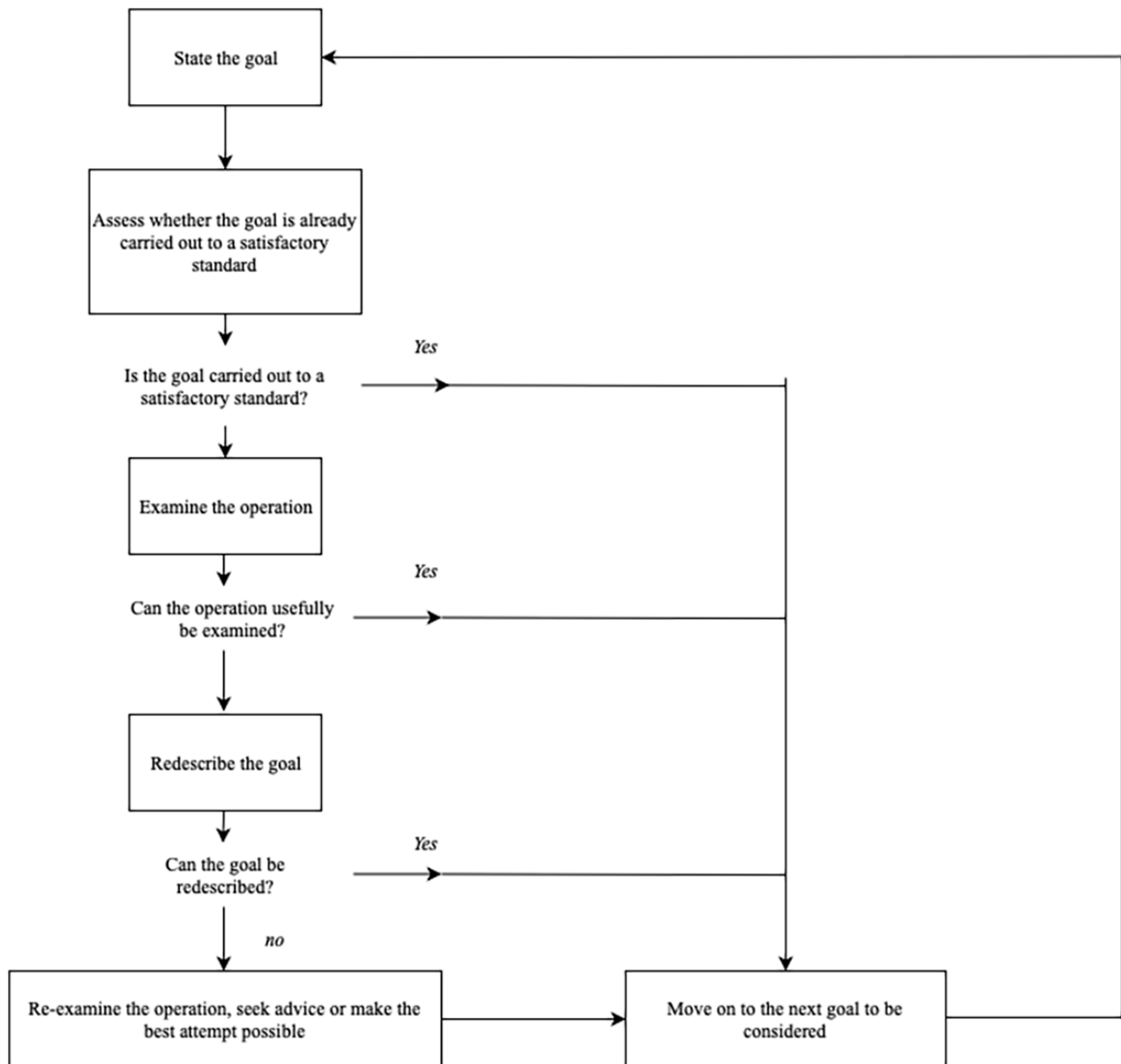
### Hierarchical Task Analysis Construction

HTA is used to understand cognitive task complexity when patients interact with the PPI information and how they would manage these tasks [11,26,27]. To carry out a task to find information, the operator has to go through a logical information scent [28]. Information scent is a term derived from the information foraging theory, which explains human information-seeking and sense interface. Information scent refers to information seekers following hints as a form of either visual or textual context in search of the desired information [23,29,30]. A strong information scent can convince users that they will find what they are looking for at the end of the journey. Information scent is the subjective perception of the cost and value of the sources obtained from proximal cues such as icons (eg, links representing content sources). Thus, placing cues that correspond with the goals can help the information seekers to find relevant information quickly and easily. The output of HTA gives app designers a better understanding of how to place information cues for better information scent.

We developed 6 high-level information-seeking goals for risperidone based on the 6 types of crucial information two licensed pharmacists decided patients should know about their medications including uses, warnings and precautions, how to take, side effects, how to store, and dosage information contents [31]. Two researchers with backgrounds in pharmacy constructed an HTA map using the goals for each level of crucial information using the HTA methodology. The main strategies

for HTA construction are described in Figure 1. After each HTA step was constructed in detail, the two researchers met, discussed, and iterated through the subgoals until consensus was reached. If the two researchers could not agree on any subgoals, the conflict was resolved by the third researcher. The three researchers would then meet and discuss the goals, subgoals, and necessary changes until the action plan for all tasks was clear.

**Figure 1.** The basic decision-making process cycle for hierarchical task analysis.



### Interactive Visualization Creation

Once the HTA goals, subgoals, and steps were created, the three researchers met and created the infographics for each step. Each step embeds important medication information related to the 6 goals. The main objective of the infographics is to present critical medication information to patients that is easily understood and does not require a substantial amount of reading. The researchers used a Gestalt approach for the colors, sizes, and fonts of the design [32].

### App Creation

We used Axure RP version 9.0669 (Axure Software Solutions) to prototype and wireframe our ideas into infographics and iterated through the design cycle until the design was approved by the research team [33]. We created infographics for the 6 subgoals and embedded relevant medication information.

Once the prototyping was developed, the functional requirements were coded using Java and PHP: Hypertext Preprocessor. The functionalities were tested several times to ensure proper functioning of the graphics and that links lead to



the desired locations. In addition, we assessed our links and hyperlinks to ensure all functionalities were appropriately executed.

### Usability Evaluation of the App

We conducted a usability evaluation of the app with 24 pharmacy students. An exemption was obtained from the Keck Graduate Institute institutional review board and consent forms were signed by the participants. The usability evaluations included two steps. For the first step, we used the concurrent thinking aloud technique to understand and measure the initial reactions of the participants. Each participant was asked first to navigate through the site for 3 to 5 minutes. After initial navigation, participants were asked to think aloud while surfing the interface. All verbal responses were recorded and transcribed. Then two researchers coded recommendations for future iterative design based on the transcripts. These recommendations were further organized under a common theme. For the second step, a System Usability Scale (SUS) survey was administered to 6 participants. SUS is a simple Likert-based scale with 10 statements that examine the global view of the subjective assessment of a user interface. A final SUS score represents a composite measure of the overall usability of the system [34].

## Results

### Hierarchical Task Analysis Construction

Six extensive steps were developed using the HTA methodology. These steps were later translated into functional requirements in Axure for mockups using an infographics approach. The final

prototype was developed after the mockups were verified and reiterated with refinements. We describe two steps that were created using the HTA method in Figure 2. For example, if the high-level goal is to find information on the dosage for an 11-year-old patient, steps would include opening the app, defining the age, and finding the information. After the information requirements are completed, users may close the app and apply the information (eg, giving the medication). After defining the age, subgoals can be found through the app. The information scent requirement leads users to the how to take stage, where 3 indications for the drug (mania, schizophrenia, and depression) are displayed. If the goal is to find the dosage for schizophrenia, the steps would be identifying patient dosage and displaying this information. Similarly, if the goal is to produce information on how to store the medication, plan 2.5 (Figure 2) will show how different interactions can help users find information using a trial-and-error approach. Additionally, it satisfies one of our subgoals of providing feedback in an adaptive fashion (ie, gives user feedback based on their response).

Explicating steps for goals and subgoals helped the design of step-by-step infographics, as shown in Figure 3, that were instantiated in the mobile prototype. For example, a user who takes Risperdal (risperidone) in Fig 3d is asked to select the proper storage method (in this case, the correct answer is in the upper cabinet at room temperature). If the user selects any other answer, the screen tells the user that the selected answer is wrong and provides an option to choose again until the user chooses the correct answer. This scenario shows how users can interact with our app.

Figure 2. Steps of hierarchical task analysis.

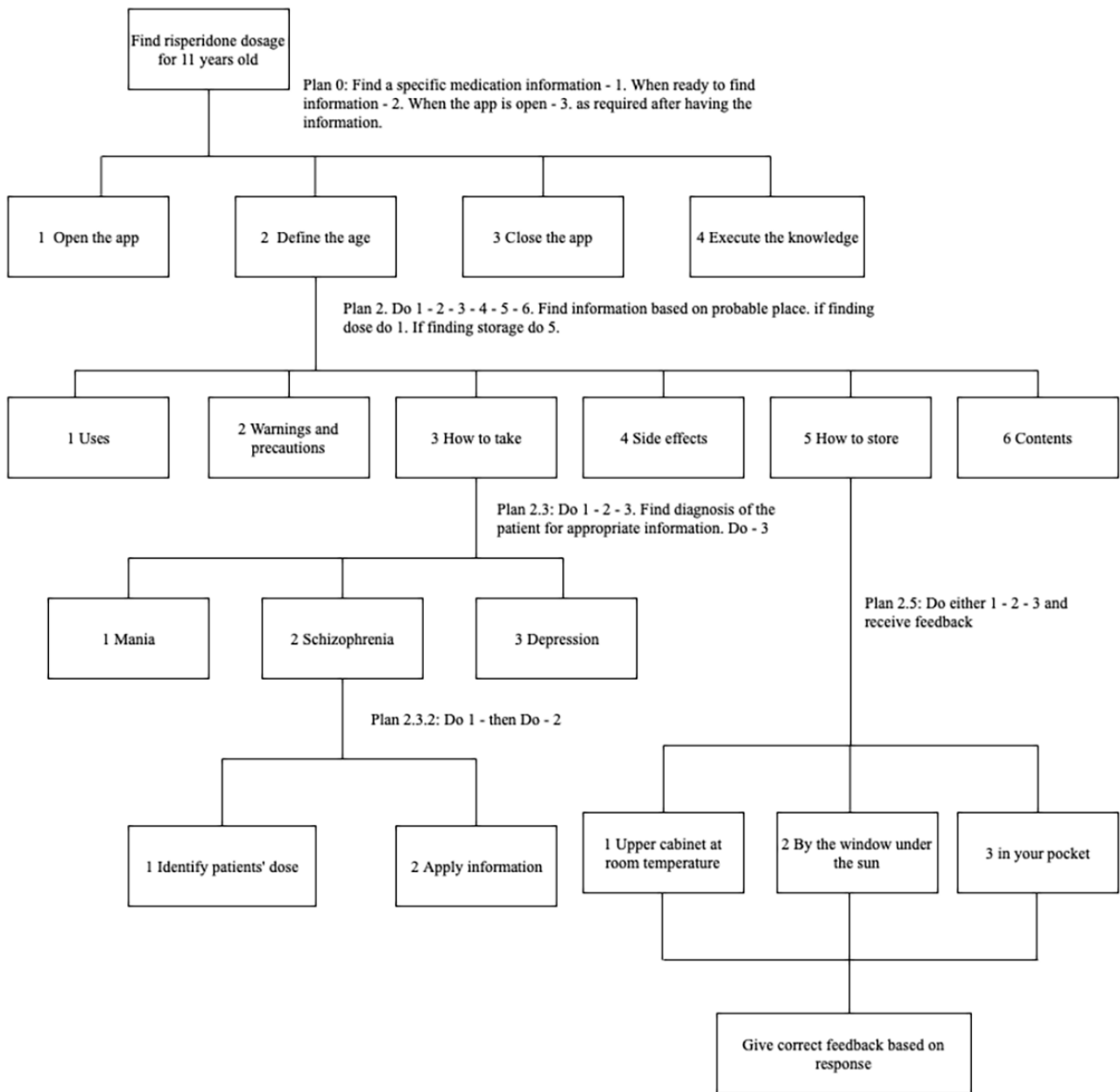
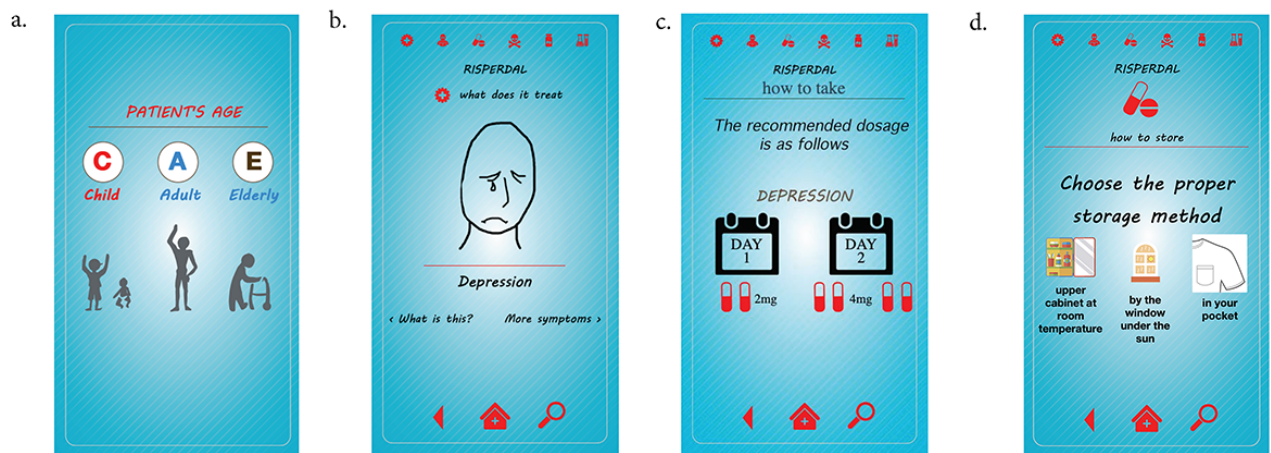


Figure 3. Screenshots from the mobile app: (a) screen showing age-based differences, (b) graphical representation of depression, (c) dosage for depression, (d) screen in which users are asked to choose the proper storage method for the medication.



## Usability Evaluations

The specific usability recommendations coded by researchers were categorized into 5 themes: initial impression, problems with page navigation, information presentation, convenience of finding information and significant changes needed. The results are summarized in [Textbox 1](#).

**Textbox 1.** System usability themes and recommendations.

<p>Initial impression:</p> <ul style="list-style-type: none"> <li>• Create a name for the app, Medinfo, and add a description</li> <li>• Create a logo that brands it as a medication information app</li> <li>• Create consistent color coding</li> </ul> <p>Problems with page navigation and surfing the site:</p> <ul style="list-style-type: none"> <li>• Add search function</li> <li>• Add file crawling options</li> <li>• Change pictures to one pill and keep dosing directions</li> </ul> <p>Information presentation, relevance, and positive impression:</p> <ul style="list-style-type: none"> <li>• In the how to store section, make the interaction quiz more obvious</li> <li>• Consider live emoji functions</li> </ul> <p>Convenience of finding information:</p> <ul style="list-style-type: none"> <li>• Different font sizes used throughout the app to focus on relevance</li> </ul> <p>Significant changes needed:</p> <ul style="list-style-type: none"> <li>• More interactions with the app</li> <li>• Relevance of warnings and side effects should be clearer</li> </ul>
--

## System Usability Score Analysis Units

The system usability score was computed based on the SUS survey (described in the Methods section) responses. The results ([Multimedia Appendix 1](#)) show the composite measure of usability is 74.5% (SD 3%).

## Discussion

### Principal Findings

Existing designs for drug medication information presentation usually include a snapshot of all related information in simple text format, which often causes information overload [35]. In contrast, a stepwise approach to include safety information has been shown to improve medication information recall for patients [36]. Our innovative approach using HTA to decompose information-seeking tasks and incorporate innovative infographics provides a unique perspective on how to present medication information to patients. More importantly, our research on using HTA corresponds with previous research in which the human factors design approaches proved to be effective [21,37-39].

The three design goals of our mobile app prototype were interaction, information overload reduction, and an enjoyable experience for the patient. Patients should be able to interact with the information. For example, if medicine X should be stored in the refrigerator, patients would be able to click on a picture of the refrigerator under the storage function and get positive feedback based on their response. In this way, patients will feel empowered to seek information. To reduce information overload, only relevant information will be displayed. For example, patients can select only certain aspects of medication

information such as how to use to understand different ways to take the medication. The interactive infographics are designed to provide a fun and enjoyable experience for the patient. Instead of using conventional images, we designed the mobile prototype using infographics to visualize the medication information, how to take the medication, and the storage site of the medication. It allowed us to deliver very complex medication information in a fast and understandable way. Additionally, we chose illustrations such as cartoon-type characters and animations, which have been shown to have positive effects on information retention [14].

The study provides novel insights on designing future medication information delivery systems for private and government organizations. Our design can provide insights into digital decision-support design also [40]. It addresses the limitation of the current medication information delivery system design, which does not ensure patient understanding and retention of crucial medication information. According to prior research, the primary causes for missed medications are forgetfulness, discordance between the patient and physician, inability to recall information, and unfavorable side effects [41]. Our research attempts to improve the information recall of patients through the interactive visualization design. We assume that by improving patient medication information recall, the system will improve medication adherence and reduce

serious/severe side effects (including adverse events) that burden our health care systems. We thus hypothesize that if patients are more aware of the potential side effects and adverse events, there is an increased likelihood of better adherence to medication management. Future research will test this hypothesis using our prototype system in an experimental setting to evaluate usability features.

Our design has several implications for PPI design. The infographics content we used can be developed for individual medications using the HTA methodology. Companies may be able to have their customers and patients focus more on the life-saving drug label information that otherwise gets ignored. Thus, this process can help with reducing adverse events, monitoring for side effects, and saving industry millions of dollars in future litigations.

Several improvements for our mobile app development have been planned based on the usability recommendations from [Textbox 1](#). One significant change recommended is to make information about warnings and side effects clearer. While currently there is not a standard way to present medication information in a graphic format, one prominent study created an iconic language called Visualization of Concepts in Medicine (VCM) to present medical concepts graphically [42]. The VCM language was designed to present textual information described in drug monographs using only a small number of graphical primitives and combinatory rules. Although the VCM language was initially designed for medical practitioners to remember drug properties, it could be extended to represent warnings and side effects in medication. More specifically, simple sentences about drug warnings, interactions, and side effects could be built with VCM icons.

Second, a reminder function will be added to create more interactions between patients and the app. To be effective, reminders should combine different modalities, including subtle status bar notifications, and should allow users to select alert types that suit their needs depending on their capabilities and social context [43,44]. Design and implementation of the reminder function could continue the same human factors approach we adopted for the current version of the prototype system.

Last, we plan to investigate other ways that would give users more interaction with the app. Designing a game-based app that is generic and aims to change the way patients think and the decisions they make about their health care can be useful [45].

Limited research has examined gamification and its impact on medication adherence. We assume built-in gamification within our mobile app would provide users more time to process the medication information, which would lead to improved medication information recall and better medication adherence.

### Limitations

Our study has several limitations. First, the prototype was designed only for an oral medication. Creating similar designs for other dosage forms such as injectables may be different. Second, the prototype needs to be tested in a real-world setting to verify its effectiveness. In this study, we created a prototype that is not for actual deployment. Our goal in the future is to include verified VCM image icons in the final design during actual implementation. Third, current prototype features such as color, background, and infographics are designed based on the feedback from the research team experts and a limited number of end users. We have demonstrated that it is feasible to design an interactive visualization-based mobile app using the HTA approach. In the next step, we plan to develop the actual mobile app using human factors design with iterations, incremental feedback, and robust testing. Therefore, our final interface design may have different and improved color contrast and background. Fourth, although interacting with simple graphical information may improve critical information recall, we have not tested information recall in patients. However, we plan to investigate other ways to increase patient interaction with the medication information and test recall in future.

### Future Work

Once our mobile app is improved and deployed, we plan to conduct an experiment that compares medication information recall between patients who interact with PPI information through our mobile app versus patients who receive PPI in a paper format. The experiment would allow us to demonstrate the effectiveness of delivering PPI information via interactive infographics on a mobile app.

### Conclusions

In this study, our goal was to design an interactive infographics-based medication information delivery system to reduce information overload and improve medication information recall. Using the HTA methodology, we successfully created a mobile prototype for delivering PPI for the drug risperidone. The hierarchical goals and subgoals were translated into a mobile prototype.

---

### Acknowledgments

This project was supported by an internal seed grant from the Keck Graduate Institute, Chapman University, School of Pharmacy and Western University of Health Sciences, College of Pharmacy. We are also very thankful for students and resources that made this project possible from Keck Graduate Institute and Claremont Graduate University.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1  
System Usability Scale scores.

[PNG File, 309 KB - [mhealth\\_v7i11e15940\\_app1.png](#)]

## References

1. Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. *J Pharmacol Pharmacother* 2013 Dec;4(Suppl 1):S73-S77 [FREE Full text] [doi: [10.4103/0976-500X.120957](#)] [Medline: [24347988](#)]
2. Jonville-Béra AP, Béra F, Autret-Leca E. Are incorrectly used drugs more frequently involved in adverse drug reactions? A prospective study. *Eur J Clin Pharmacol* 2005 May;61(3):231-236. [doi: [10.1007/s00228-004-0881-6](#)] [Medline: [15824914](#)]
3. Cunningham M. BU Well. 2016. Revolutionizing the patient package insert with infographics URL: <https://www.digitalcommons.butler.edu/cgi/viewcontent.cgi?article=1008&context=buwell> [accessed 2019-11-08]
4. Chakraborty S, Boudier F. The future of risk communication and the role of the pharmaceutical industry. *Curr Drug Saf* 2013 Feb;8(1):4-10. [Medline: [23656442](#)]
5. Edwards B, Chakraborty S. Risk communication and the pharmaceutical industry: what is the reality? *Drug Saf* 2012 Nov 01;35(11):1027-1040. [doi: [10.2165/11635220-000000000-00000](#)] [Medline: [23061779](#)]
6. Hoy MG, Levenshus AB. A mixed-methods approach to assessing actual risk readership on branded drug websites. *J Risk Res* 2016 Aug 27;21(5):521-538. [doi: [10.1080/13669877.2016.1223160](#)]
7. Dunwoody S, Griffin RJ. Risk information seeking and processing model. In: Cho H, Reimer T, McComas KA, editors. *SAGE Handbook of Risk Communication*. Thousand Oaks: Sage Publishing; 2015.
8. Wright P, Creighton P, Threlfall SM. Some factors determining when instructions will be read. *Ergonomics* 2007 Apr 05;25(3):225-237. [doi: [10.1080/00140138208924943](#)]
9. Wilkinson JL, Strickling K, Payne HE, Jensen KC, West JH. Evaluation of diet-related infographics on Pinterest for use of behavior change theories: a content analysis. *JMIR Mhealth Uhealth* 2016 Dec 08;4(4):e133 [FREE Full text] [doi: [10.2196/mhealth.6367](#)] [Medline: [27932316](#)]
10. Comello MLG, Qian X, Deal AM, Ribisl KM, Linnan LA, Tate DF. Impact of game-inspired infographics on user engagement and information processing in an eHealth program. *J Med Internet Res* 2016 Sep 22;18(9):e237 [FREE Full text] [doi: [10.2196/jmir.5976](#)] [Medline: [27658469](#)]
11. Roosan D, Del Fiore G, Butler J, Livnat Y, Mayer J, Samore M, et al. Feasibility of population health analytics and data visualization for decision support in the infectious diseases domain. *Appl Clin Inform* 2017 Dec 16;07(02):604-623. [doi: [10.4338/aci-2015-12-ra-0182](#)]
12. Islam R, Weir C, Del Fiore G. Heuristics in managing complex clinical decision tasks in experts' decision making. *IEEE Int Conf Healthc Inform* 2014 Sep;2014:186-193 [FREE Full text] [doi: [10.1109/ICHI.2014.32](#)] [Medline: [27275019](#)]
13. Price MM, Crumley-Branyon JJ, Leidheiser WR, Pak R. Effects of information visualization on older adults' decision-making performance in a Medicare plan selection task: a comparative usability study. *JMIR Hum Factors* 2016 Jun 01;3(1):e16 [FREE Full text] [doi: [10.2196/humanfactors.5106](#)] [Medline: [27251110](#)]
14. Schneider S, Nebel S, Rey GD. Decorative pictures and emotional design in multimedia learning. *Learn Instr* 2016 Aug;44:65-73. [doi: [10.1016/j.learninstruc.2016.03.002](#)]
15. Saket B, Endert A, Stasko J. Beyond usability and performance: a review of user experience-focused evaluations in visualization. 2016 Presented at: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization; 2016; New York p. 133-142. [doi: [10.1145/2993901.2993903](#)]
16. IMS Institute for Healthcare Informatics. 2013 Oct. Patient apps for improved healthcare: from novelty to mainstream URL: [http://moodle.univ-lille2.fr/pluginfile.php/215345/mod\\_resource/content/0/IIHI\\_Patient\\_Apps\\_Report.pdf](http://moodle.univ-lille2.fr/pluginfile.php/215345/mod_resource/content/0/IIHI_Patient_Apps_Report.pdf) [accessed 2019-11-08]
17. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. *Am J Prev Med* 2013 Aug;45(2):228-236 [FREE Full text] [doi: [10.1016/j.amepre.2013.03.017](#)] [Medline: [23867031](#)]
18. Yetisen AK, Martinez-Hurtado JL, da Cruz Vasconcellos F, Simsekler MCE, Akram MS, Lowe CR. The regulation of mobile medical applications. *Lab Chip* 2014 Mar 07;14(5):833-840. [doi: [10.1039/c3lc51235e](#)] [Medline: [24425070](#)]
19. Hu X, Li X, Ngai EC, Zhao J, Leung VC. Health drive: mobile healthcare onboard vehicles to promote safe driving. 2015 Presented at: 48th Hawaii International Conference on System Sciences; 2015; Kauai p. 3074-3083. [doi: [10.1109/hicss.2015.371](#)]
20. Scott K, Richards D, Adhikari R. A review and comparative analysis of security risks and safety measures of mobile health apps. *AJIS* 2015 Nov 22;19. [doi: [10.3127/ajis.v19i0.1210](#)]
21. Stanton NA. Hierarchical task analysis: developments, applications, and extensions. *Appl Ergon* 2006 Jan;37(1):55-79 [FREE Full text] [doi: [10.1016/j.apergo.2005.06.003](#)] [Medline: [16139236](#)]
22. Regunath S, Raina S, Gramopadhye A. Use of HTA in establishing training content for aircraft inspection. *Proc IIE Ann Conf* 2004:2279-2282.
23. Chi EH, Pirolli P, Chen K, Pitkow J. Using information scent to model user information needs and actions and the Web. 2001 Presented at: Proceedings of the SIGCHI conference on Human factors in computing systems; 2001; Seattle p. 490-497. [doi: [10.1145/365024.365325](#)]

24. Annett J. Hierarchical task analysis. In: Hollnagel E, editor. *Handbook of Cognitive Task Design*. Mahwah: Lawrence Erlbaum Associates; 2003:17-36.
25. Germann D, Kurylo N, Han F. Risperidone. *Profiles Drug Subst Excip Relat Methodol* 2012;37:313-361. [doi: [10.1016/B978-0-12-397220-0.00008-8](https://doi.org/10.1016/B978-0-12-397220-0.00008-8)] [Medline: [22469322](https://pubmed.ncbi.nlm.nih.gov/22469322/)]
26. Islam R, Weir CR, Jones M, Del Fiol G, Samore MH. Understanding complex clinical reasoning in infectious diseases for improving clinical decision support design. *BMC Med Inform Decis Mak* 2015 Nov 30;15:101 [FREE Full text] [doi: [10.1186/s12911-015-0221-z](https://doi.org/10.1186/s12911-015-0221-z)] [Medline: [26620881](https://pubmed.ncbi.nlm.nih.gov/26620881/)]
27. Roosan D, Weir C, Samore M, Jones M, Rahman M, Stoddard GJ, et al. Identifying complexity in infectious diseases inpatient settings: an observation study. *J Biomed Inform* 2017 Jul;71S:S13-S21 [FREE Full text] [doi: [10.1016/j.jbi.2016.10.018](https://doi.org/10.1016/j.jbi.2016.10.018)] [Medline: [27818310](https://pubmed.ncbi.nlm.nih.gov/27818310/)]
28. Pirolli PL. *Information Foraging Theory: Adaptive Interaction with Information*. New York: Oxford University Press; 2007.
29. Ong K, Järvelin K, Sanderson M, Scholer F. Using information scent to understand mobile and desktop web search behavior. 2017 Presented at: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval; 2017; Shinjuku p. 295-304. [doi: [10.1145/3077136.3080817](https://doi.org/10.1145/3077136.3080817)]
30. Maxwell D, Azzopardi L. Information scent, searching and stopping. In: *Advances in Information Retrieval*. 2018 Presented at: 40th European Conference on IR Research; 2018; Grenoble p. 210-222. [doi: [10.1007/978-3-319-76941-7\\_16](https://doi.org/10.1007/978-3-319-76941-7_16)]
31. Wilson FL, Mood D, Nordstrom CK. The influence of easy-to-read pamphlets about self-care management of radiation side effects on patients' knowledge. *Oncol Nurs Forum* 2010 Nov;37(6):774-781. [doi: [10.1188/10.ONF.774-781](https://doi.org/10.1188/10.ONF.774-781)] [Medline: [21059589](https://pubmed.ncbi.nlm.nih.gov/21059589/)]
32. Luna-Nevarez C, Hyman MR. Common practices in destination website design. *J Destin Mark Manag* 2012 Nov;1(1-2):94-106. [doi: [10.1016/j.jdmm.2012.08.002](https://doi.org/10.1016/j.jdmm.2012.08.002)]
33. Axure mobile widget libraries and templates. URL: <https://www.axurethemes.com/axure-mobile> [accessed 2019-02-15]
34. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. *Int J Hum Comput Interact* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
35. Swar B, Hameed T, Reyachav I. Information overload, psychological ill-being, and behavioral intention to continue online healthcare information search. *Comput Hum Behav* 2017 May;70:416-425. [doi: [10.1016/j.chb.2016.12.068](https://doi.org/10.1016/j.chb.2016.12.068)]
36. Shyr C, Kushniruk A, Wasserman WW. Usability study of clinical exome analysis software: top lessons learned and recommendations. *J Biomed Inform* 2014 Oct;51:129-136 [FREE Full text] [doi: [10.1016/j.jbi.2014.05.004](https://doi.org/10.1016/j.jbi.2014.05.004)] [Medline: [24860971](https://pubmed.ncbi.nlm.nih.gov/24860971/)]
37. Roosan D, Law A, Karim M, Roosan M. A protocol for improving team based decision-making using data analytics and informatics. *JMIR Res Protoc* 2019. [doi: [10.2196/16047](https://doi.org/10.2196/16047)]
38. Raduma-Tomás MA, Flin R, Yule S, Close S. The importance of preparation for doctors' handovers in an acute medical assessment unit: a hierarchical task analysis. *BMJ Qual Saf* 2012 Mar 30;21(3):211-217 [FREE Full text] [doi: [10.1136/bmjqs-2011-000220](https://doi.org/10.1136/bmjqs-2011-000220)] [Medline: [22129935](https://pubmed.ncbi.nlm.nih.gov/22129935/)]
39. Islam R, Weir C, Del Fiol G. Clinical complexity in medicine: a measurement model of task and patient complexity. *Methods Inf Med* 2016;55(1):14-22 [FREE Full text] [doi: [10.3414/ME15-01-0031](https://doi.org/10.3414/ME15-01-0031)] [Medline: [26404626](https://pubmed.ncbi.nlm.nih.gov/26404626/)]
40. Roosan D, Samore M, Jones M, Livnat Y, Clutter J. Big-data based decision-support systems to improve clinicians' cognition. *IEEE Int Conf Healthc Inform* 2016:285-288 [FREE Full text] [doi: [10.1109/ICHI.2016.39](https://doi.org/10.1109/ICHI.2016.39)] [Medline: [27990498](https://pubmed.ncbi.nlm.nih.gov/27990498/)]
41. Stefaniková S, Prokop P. Do we believe pictures more or spoken words? How specific information affects how students learn about animals. *Eurasia J Math Sci Tech Ed* 2015 Aug 17;11(5):725-733. [doi: [10.12973/eurasia.2015.1380a](https://doi.org/10.12973/eurasia.2015.1380a)]
42. Lamy J, Duclos C, Bar-Hen A, Ouvrard P, Venot A. An iconic language for the graphical representation of medical concepts. *BMC Med Inform Decis Mak* 2008 Apr 24;8:16 [FREE Full text] [doi: [10.1186/1472-6947-8-16](https://doi.org/10.1186/1472-6947-8-16)] [Medline: [18435838](https://pubmed.ncbi.nlm.nih.gov/18435838/)]
43. Stawarz K, Cox AL, Blandford A. Don't forget your pill! designing effective medication reminder apps that support users' daily routines. 2014 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2014; Toronto p. 2269-2278. [doi: [10.1145/2556288.2557079](https://doi.org/10.1145/2556288.2557079)]
44. Santo K, Chow CK, Thiagalingam A, Rogers K, Chalmers J, Redfern J. MEDication reminder APPs to improve medication adherence in Coronary Heart Disease (MedApp-CHD) Study: a randomised controlled trial protocol. *BMJ Open* 2017 Oct 08;7(10):e017540 [FREE Full text] [doi: [10.1136/bmjopen-2017-017540](https://doi.org/10.1136/bmjopen-2017-017540)] [Medline: [28993388](https://pubmed.ncbi.nlm.nih.gov/28993388/)]
45. Joshi R, Joshi D, Cheriya P. Improving adherence and outcomes in diabetic patients. *Patient Prefer Adherence* 2017;11:271-275 [FREE Full text] [doi: [10.2147/PPA.S122490](https://doi.org/10.2147/PPA.S122490)] [Medline: [28243070](https://pubmed.ncbi.nlm.nih.gov/28243070/)]

## Abbreviations

- ADR:** adverse drug reaction
- FDA:** US Food and Drug Administration
- HTA:** hierarchical task analysis
- mHealth:** mobile health
- PPI:** patient package insert
- SUS:** System Usability Scale

**VCM: Visualization of Concepts in Medicine**

*Edited by G Eysenbach; submitted 20.08.19; peer-reviewed by M Nurunnabi, A Benis, A Arcia, S Davis; comments to author 18.09.19; revised version received 09.10.19; accepted 17.10.19; published 25.11.19.*

*Please cite as:*

*Roosan D, Li Y, Law A, Truong H, Karim M, Chok J, Roosan M*

*Improving Medication Information Presentation Through Interactive Visualization in Mobile Apps: Human Factors Design*

*JMIR Mhealth Uhealth 2019;7(11):e15940*

*URL: <http://mhealth.jmir.org/2019/11/e15940/>*

*doi: [10.2196/15940](https://doi.org/10.2196/15940)*

*PMID: [31763991](https://pubmed.ncbi.nlm.nih.gov/31763991/)*

©Don Roosan, Yan Li, Anandi Law, Huy Truong, Mazharul Karim, Jay Chok, Moom Roosan. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 25.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# The Use of Smartphone-Based Triage to Reduce the Rate of Outpatient Error Registration: Cross-Sectional Study

Wanhua Xie<sup>1</sup>, PhD; Xiaojun Cao<sup>2</sup>, MD; Hongwei Dong<sup>3</sup>, MD; Yu Liu<sup>1</sup>, BSN

<sup>1</sup>Outpatient Department, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

<sup>2</sup>Department of Science, Education and Data Management, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

<sup>3</sup>Hospital Office, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

**Corresponding Author:**

Wanhua Xie, PhD

Outpatient Department

Guangzhou Women and Children's Medical Center

Guangzhou Medical University

No 9 Jinhui Road

Tianhe District

Guangzhou, 510623

China

Phone: 86 13725370379

Email: [xiewanhua@126.com](mailto:xiewanhua@126.com)

## Abstract

**Background:** In many clinics, patients now have the option to make Web-based appointments but doing so according to their own judgment may lead to wrong registration and delayed medical services. We hypothesized that smartphone-based triage in outpatient services is superior to Web-based self-appointment registration guided by the medical staff.

**Objective:** This study aimed to investigate smartphone-based triage in outpatient services compared with Web-based self-appointment registration and to provide a reference for improving outpatient care under appointment registration.

**Methods:** The following parameters in Guangzhou Women and Children's Medical Center were analyzed: wrong registration rate, the degree of patient satisfaction, outpatient visits 6 months before and after smartphone-based triage, queries after smartphone-based triage, number of successful registrations, inquiry content, and top 10 recommended diseases and top 10 recommended departments after queries.

**Results:** Smartphone-based triage showed significant effects on average daily queries, which accounted for 16.15% (1956/12,112) to 29.46% (3643/12,366) of daily outpatient visits. The average daily successful registration after queries accounted for 56.14% (1101/1961) to 60.92% (1437/2359) of daily queries and 9.33% (1130/12,112) to 16.83% (2081/12,366) of daily outpatient visits. The wrong registration rate after smartphone-based triage was reduced from 0.68% (12,810/1,895,829) to 0.12% (2379/2,017,921) ( $P < .001$ ), and the degree of patient satisfaction was improved. Monthly outpatient visits were increased by 0.98% (3192/325,710) to 13.09% (42,939/328,032) compared with the same period the preceding year ( $P = .02$ ).

**Conclusions:** Smartphone-based triage significantly reduces the wrong registration rate caused by patient Web-based appointment registration and improves the degree of patient satisfaction. Thus, it is worth promoting.

(*JMIR Mhealth Uhealth* 2019;7(11):e15313) doi:[10.2196/15313](https://doi.org/10.2196/15313)

**KEYWORDS**

smartphone; triage; outpatients; personal satisfaction

## Introduction

**Background**

Medical systems worldwide face clinical, administrative, regulatory, and financial strains that require them to increase

performance while reducing costs [1,2]. Hospitals often have a limited number of support staff responsible for patient registration and for taking medical history [3]. Registration by medical staff can create a bottleneck that increases waiting time and, thus, the likelihood of registration errors [4-6].



The National Health and Family Planning Commission of the People's Republic of China issued the Action Plan for Further Improving Medical Care Services in 2015, requiring appointment rates in tertiary hospitals to be at least 50% by the end of 2017 [7]. Therefore, hospitals began to implement appointment systems, and some of them implemented full appointment systems for nonemergency registration. The appointment systems alleviate the difficulty of temporary registration to a certain extent but may lead to wrong registration because of the limited medical knowledge of the patients. Song et al [8] reported that since the implementation of the full appointment system for nonemergency pediatric patients, the wrong registration rates by self-appointment and complaints from the children's families cannot be neglected and that the most important problem faced by the patients is the selection of the correct department [8]. This is further complicated by increasingly subdivided departments [9]. Xuezheng et al [10] reported that the loss of time and money caused by wrong registration directly leads to patient dissatisfaction with hospital visits.

Wrong registration not only wastes the time of the patients but may also lead to disease aggravation and delayed treatment. Lewis et al [11] proposed an effective clinical decision support tool to perform prehospital diagnosis and triage of ruptured aortic aneurysms accurately. Adapting such algorithms into smartphone apps that are used to make appointment registration could help decrease the error rates. It is feasible to use smartphone apps to report a dental emergency [12] and burns [13]. Previous studies mainly reported the use of smartphones for emergency triage and referral for some diseases [14-16], but the added value of using smartphone-based triage for regular appointments remains undefined. The most concerning aspect of adequate appointment registration is that the parents' level of information regarding their child's illness is limited [17]. Mobile health is a promising app for appointment registration [14]. WhatsApp (WhatsApp Inc) helps reduce unnecessary referrals and outpatient visits [15]. Smartphone-based pulse oximetry has been used to evaluate pediatric patients without hypoxia [16]. Reduction in waiting time from referral to first visit for community outpatient services might contribute to better health outcomes [18]. Therefore, investigating smartphone-based triage is critical in the modern era where the smartphone is becoming a central actor in everyday life.

In many clinics in China, the patients now have the option to make Web-based appointments, but doing so according to their own judgment may lead to wrong registration and delayed medical services. On the basis of the above, we hypothesized that smartphone-based triage in outpatient services is superior to Web-based self-appointment registration guided by the medical staff.

## Objectives

Therefore, this study aimed to investigate the effect of smartphone-based triage on outpatient visits compared with

Web-based self-appointment registration, the rate of wrong registration made by the patients, the degree of patient satisfaction, and the effects on outpatient visits. This study provides a reference for improving outpatient appointment registration.

## Methods

### Study Design and Subjects

This was a cross-sectional study of patients who made appointments with the outpatient department of Guangzhou Women and Children's Medical Center between April 2018 and March 2019. This medical center is one of the hospitals that implemented a full appointment system for nonemergency appointment registration. There are more than 10,000 outpatient visits daily, and smartphone appointments among these outpatient visits accounted for a total of 77.15% (8692/11,266).

The inclusion criteria were that the participant (1) had records available during the study period in the hospital information system, (2) used a smartphone to book the appointment, and (3) had good communication skills and consented to participate. The exclusion criteria included (1) incomplete registration or appointment data or (2) multiple visits during the study period.

### Smartphone App

Our hospital and Shenzhen Tencent Technology jointly developed and implemented smartphone-based triage based on symptoms recorded since October 2018. At the time of appointment registration, the option of selecting basic symptoms was added. Patients or family members can select symptoms based on disease conditions, which then accordingly guide them in selecting the adequate departments and doctors to visit. The best department can be selected according to different symptoms in the registration symptom page. In addition, the best route guidance can be provided based on precise, intelligent navigation within the hospital.

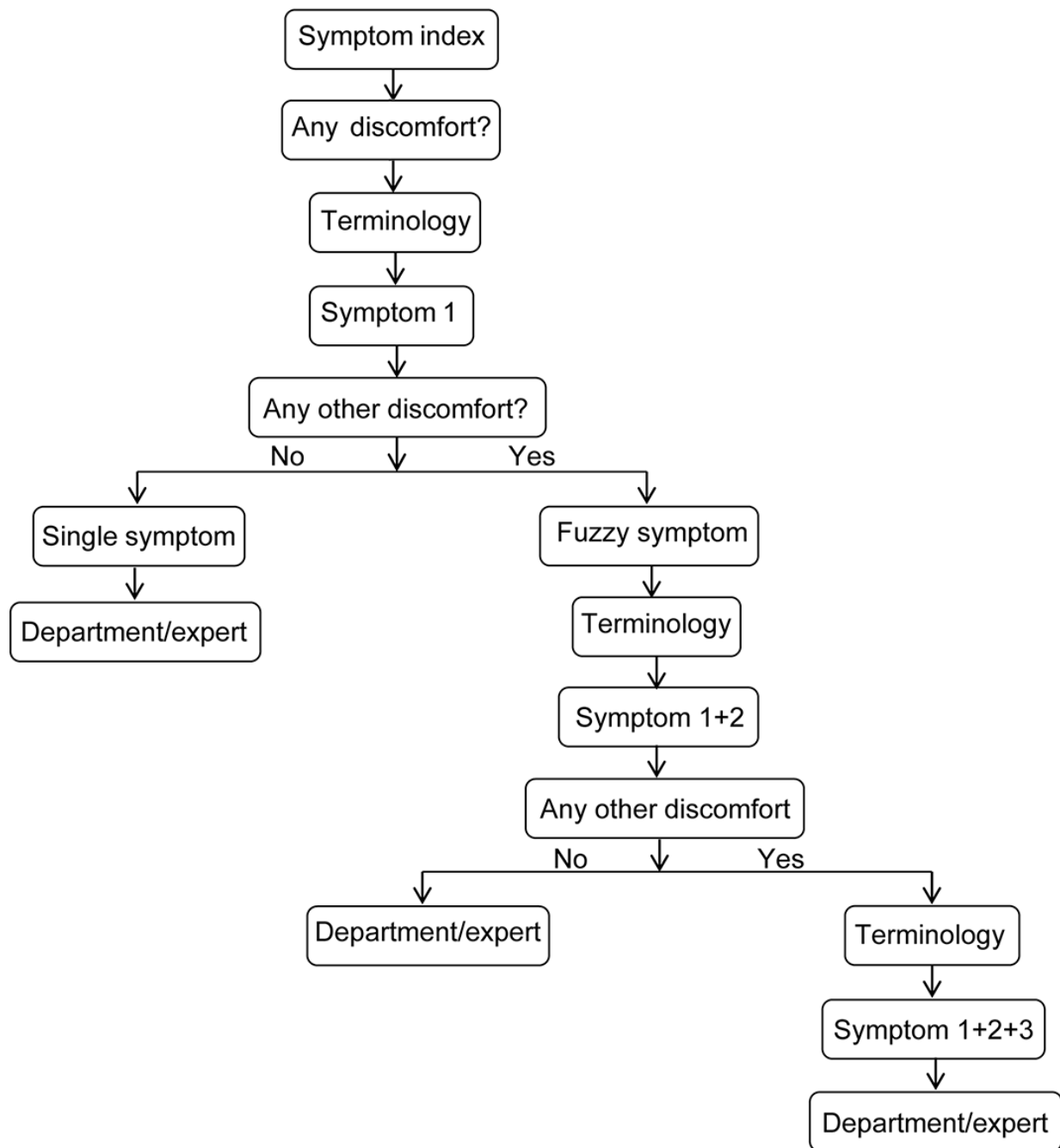
### Patient Grouping

The patients recruited from April to September 2018 made Web-based appointment registration without the smartphone system, whereas from October 2018 to March 2019, outpatients started using the smartphone app for Web-based appointments.

### Indexing Process for Smartphone Self-Appointment System

Disease indexing was performed by the patients. If the patient had a confirmed disease, the app immediately directed the patient to the appropriate department (Figure 1; Multimedia Appendix 1 shows screenshots of the app). If there was no confirmed diagnosis, the app asked a series of questions based on symptoms, discomfort, and terminology and then suggested the suitable departments to the patients. The app is completely automated.

Figure 1. Self-appointment symptom index chart.



**Data Collection**

Smartphone-based triage had storage and memory functions. The medical staff could directly obtain the content from the background system and retrieve outpatient visits from the hospital operation and decision support system (business intelligence system). The degree of satisfaction regarding the accuracy of smartphone-based triage was extracted from the degree of satisfaction survey, which spanned the period from 6 months before smartphone-based triage to 6 months thereafter.

Patients in each department of Guangzhou Women and Children’s Medical Center were enrolled. Random sampling was performed for selection, and assessment was carried out by 55 investigators who underwent unified training. The degree

of satisfaction survey was performed on site at the hospital. According to the sample size, the survey of 1850 cases indicated that the interviewees were representative. General satisfaction was the first-level indicator. In addition, there were 6 second-level indicators, including general impression, service attitude, service quality, hospital environment, price perception, and medical ethics. Furthermore, 31 third-level indicators were assessed (Multimedia Appendix 2). The responses to third-level indicators were categorized using a 5-point Likert scale. The answers were coded from 1 to 5: 5-very satisfied, 4-satisfied, 3-neutral, 2-unsatisfied, and 1-very unsatisfied. The current internationally recognized method for calculating the satisfaction degree based on public opinion was adopted for evaluating each of the third-level indicators [19,20]: 100, *very satisfied*; 80,

quite satisfied; 60, neutral; 40, dissatisfied; and 20, verydissatisfied. Unclear was assigned for missing values, which were excluded from the final analysis. A single indicator was calculated based on weighted calculation: degree of satisfaction = proportion of very satisfied  $\times$  100 + proportion of relatively satisfied  $\times$  80 + proportion of generally satisfied  $\times$  60 + proportion of not satisfied  $\times$  40 + proportion of very dissatisfied  $\times$  20. The overall degree of satisfaction was calculated by weighing each indicator score [19,20]. General satisfaction was calculated by summing the scores of all second-level indicators after multiplying them by their corresponding weights: general impression (5%), service attitude (15%), service quality (20%), hospital environment (10%), price perception (20%), and medical ethics (30%) [19,20].

### Statistical Analysis

SPSS 23.0 (IBM) was used for data analysis. Queries and outpatient visits of smartphone-based triage were expressed as frequencies. The proportion of queries and inquiry content was expressed as a percentage. A comparison of wrong registration and the degree of satisfaction was analyzed by the Student *t* test.  $P < .05$  was considered statistically significant.

### Ethical Considerations

The ethics committee of Guangzhou Women and Children's Medical Center (approval number SFE-KL-28201) approved

this study. Patients or their guardians provided signed informed consent.

## Results

### Effect of Smartphone-Based Triage

The average daily numbers of smartphone-based triage used for the 6 consecutive months were 2359, 1917, 1956, 1961, 3067, and 3643, accounting for 22.23% (2359/10,610), 16.37% (1917/11,717), 16.15% (1956/12,112), 23.20% (1961/8453), 26.09% (3067/11,754), and 29.46% (3643/12,366) of all outpatient visits, respectively. The average number of daily queries for the whole study period was 2730, which was equal to the workload of at least nine experienced precheck triage nurses and hospital guiding staff. The numbers of successful daily registrations after smartphone-based triage for the 6 consecutive months were 1437, 1101, 1130, 1101, 1736, and 2081, accounting for 60.92% (1437/2359), 57.43% (1101/1917), 57.77% (1130/1956), 56.14% (1101/1961), 56.60% (1736/3067), and 57.12% (2081/3643) of all queries, respectively, and 13.54% (1437/10,610), 9.40% (1101/11,717), 9.33% (1130/12,112), 13.02% (1101/8453), 14.77% (1736/11,754), and 16.83% (2081/12,366), respectively, of all daily outpatient visits in the respective months (Table 1).

**Table 1.** Queries of smartphone-based triage, successful registrations after query, and proportions of outpatient visits.

Time (months)	Daily outpatient visits	Proportion of daily queries to outpatient visits, n (%)	Proportion of successful daily registrations after query to outpatient visits, n (%)
1	10,610	2359 (22.23)	1437 (13.54)
2	11,717	1917 (16.37)	1101 (9.40)
3	12,112	1956 (16.15)	1130 (9.33)
4	8453	1961 (23.20)	1101 (13.02)
5	11,754	3067 (26.09)	1736 (14.77)
6	12,366	3643 (29.46)	2081 (16.83)

### Reduction in Wrong Registrations

The number of wrong registrations during the 6 months before smartphone-based triage was 12,810, accounting for 0.68% of all outpatient visits (1,895,829), that is, 2135 (SD 37) wrong registrations recorded monthly. This number was significantly reduced to 2379 after smartphone-based triage, accounting for 0.12% of all outpatient visits (2,017,921), that is, 396 (SD 17) wrong registrations recorded monthly (95% CI 1701.362-1771.68;  $P < .001$ ).

### Improved Degree of Satisfaction of the Outpatients

A total of 7400 questionnaires were released and 7188 were collected, indicating an effective recovery rate of 97.14%. The questionnaire survey data are shown in Table 2. A comparison of the first and second quarters after smartphone-based triage with pre-smartphone-based triage showed a significant improvement in the degree of patient satisfaction ( $P < .001$ ; Table 3).

**Table 2.** Basic information of the degree of patient satisfaction before and after smartphone-based triage.

Item and condition	Second quarter before triage	First quarter before triage	First quarter after triage	Second quarter after triage
<b>Valid questionnaires</b>				
Copies	1832	1838	1802	1716
<b>Sex, n (%)</b>				
Female	970 (52.95)	1112 (60.50)	1116 (61.93)	1065 (62.06)
Male	862 (47.05)	726 (39.50)	686 (38.07)	651 (37.94)
<b>Payment, n (%)</b>				
Self-paying	1280 (69.87)	1141 (62.08)	1168 (64.82)	1067 (62.18)
Medical insurance	454 (24.78)	623 (33.90)	563 (31.24)	579 (33.74)
Public expense	98 (5.35)	74 (4.02)	71 (3.94)	70 (4.08)
<b>Residence, n (%)</b>				
Local	1290 (70.41)	1317 (71.65)	1300 (72.14)	1315 (76.63)
Other city in local province	466 (25.44)	389 (21.17)	372 (20.64)	323 (18.82)
Other province	76 (4.15)	132 (7.18)	130 (7.22)	78 (4.55)
<b>Sample classification, n (%)</b>				
Patient himself	336 (18.34)	443 (24.10)	527 (29.25)	501 (29.20)
Relatives and friends	1496 (81.66)	1395 (75.90)	1275 (70.75)	1215 (70.80)
<b>Educational level, n (%)</b>				
Below junior middle school	243 (13.26)	296 (16.10)	276 (15.32)	229 (13.34)
High school or technical secondary school	432 (23.58)	432 (23.50)	421 (23.36)	340 (19.81)
Junior college or undergraduate	1073 (58.57)	1028 (55.93)	1022 (56.71)	1062 (61.89)
Master degree or above	84 (4.59)	82 (4.47)	83 (4.61)	85 (4.96)

**Table 3.** Comparison of satisfaction level before and after smartphone-based triage.

Time	After smartphone-based triage, mean (SD)	Before smartphone-based triage, mean (SD)	P value	Odds Ratio (95% CI)
First quarter	87.05 (3.49)	85.77 (3.71)	<.001	1.059(0.482-2.327)
Second quarter	87.60 (3.32)	85.27 (3.26)	<.001	0.773(0.342-1.747)

### Degree of Satisfaction Regarding the Accuracy of Smartphone-Based Triage

The survey conducted in 1850 patients about the degree of satisfaction regarding the accuracy of smartphone-based triage showed that 1601 patients were satisfied (88.68%), 144 were neutral (7.98%), 60 were dissatisfied (3.34%), and 45 were missing.

### Proportions of Inquiries in the Content of Smartphone-Based Triage

For patients inputting their diseases and symptoms in smartphone-based triage, those waiting for disease judgment accounted for 48,596 (57.68%), those looking for a doctor accounted for 14,184 (16.84%), those in the inquiry process accounted for 7260 (8.62%), those with other questions accounted for 6712 (7.97%), those inquiring for examination accounted for 3994 (4.74%), and those looking for the department accounted for 3502 (4.15%).

### Top 10 Queries of Disease Recommendation

The top 10 queries were associated with unspecified-site acute upper respiratory tract infection; other infectious diseases and unspecified etiologies, gastroenteritis, and colitis, other and unclear pyrexia, acute nasopharyngitis cold, influenza without labeled virus, cough, bronchitis unspecified as acute or chronic, pneumonia with unspecified pathogen, hand-foot-and-mouth disease, chronic disease of tonsils and adenoids.

### Top 10 Queries of Department Recommendation

The top 10 queries based on department recommendations were associated with the deputy director of outpatient internal medicine, attending internal medicine department, attending respiratory outpatient, attending auricle deformity department, chief of pediatric respiratory department, attending otolaryngology department, chief of otolaryngology department, attending dermatology department, deputy director of dermatology department and attending outpatient department of gastroenterology.

## Changes in Outpatient Visits Before and After Smartphone-Based Triage

The outpatient visits were increased by 0.98% (3192/325,710) to 13.09% (42,939/328,032) compared with the same period last year, and monthly outpatient visits after 6 months of smartphone-based triage were 328,902, 351,518, 375,458, 236,690, 364,382, and 370,971. Compared with the outpatient visits in the same period last year, the outpatient visits were increased by 0.98% (3192/325,710), 4.40% (14,815/336,703), 9.22% (31,695/343,763), 5.22% (11,742/224,948), 8.23% (27,708/336,674), and 13.09% (42,939/328,032), after smartphone-based triage (95% CI 6525.769–41,544.565;  $P=.02$ ).

## Discussion

### Principal Findings

Human-based patient registration requires money, time, and human resources and often represents a bottleneck for patient registration [1–6]. In China, where the health care system is based on specialists rather than generalists, the patients now have the option to make Web-based appointments, but doing so according to their own judgment may lead to wrong registration and delayed medical services. This study investigated the effect of implementing smartphone-based triage in the outpatient department compared with Web-based self-appointment registration, smartphone-based triage queries, rate of successful registration after query, rate of wrong registration, degree of patient satisfaction, and the top 10 recommended diseases and departments. The results indicated that smartphone-based triage has significant effects on appointment registration and is well accepted by the patients and family members, reduces wrong registration caused by patient Web-based appointment registration, and improves the degree of patient satisfaction.

This showed that patients could use smartphones to consult the doctor anytime from anywhere to be guided to the right department, thereby optimizing the visit process, reducing the wrong registration rate, and improving the degree of patient satisfaction. De Bruin et al [21] confirmed that a smartphone-based triage app was well accepted by the patients.

Previous studies reported improved effectiveness of patient registration. Nogueira et al [22] proposed the FAST-ED app for improving smartphone-based triage in acute ischemic stroke patients, and Uthoff et al [23] reported the use of smartphones in the early discovery of oral cancer lesions. Tayfur and Afacan [24] showed that the patients could perform a smartphone-based triage evaluation before seeing a doctor. Verzantvoort et al [25] stated that smartphone-based triage plays an important role in patients who are consulting doctors. Tian et al [26] used a simulation model to evaluate a mobile-based system for supporting emergency evacuation decision making. Gupta et al [27] reported technicians using a store-and-forward telemedicine device to screen for ear problems. Astarcioglu et al [28], Borve et al [29], Dahlen Gyllencreutz et al [30], and Urner et al [31] investigated the use of smartphones in skin cancer referral, time-to-reperfusion in ST-segment elevation myocardial infarction undergoing inter-hospital transfers, teledermoscopy

images acquired in primary health care and hospital settings, and accuracy in triage imaging of human papillomavirus-positive women, respectively. All these studies showed that the use of smartphone-based triage improved the wrong registration rates and patient satisfaction in specific patient populations. This study is supported by those studies and provides a higher degree of generalization as all women and pediatric patients were included, irrespective of their disease, and the department consulted. Nevertheless, those previous studies did not report how many outpatients were willing to participate in the clinical study. In contrast, this study demonstrated that queries associated with smartphone-based triage were significant, and the triage was well accepted by patients.

### Strengths and Limitations

A major advantage of the system used in this study was that no special smartphone requirement was set, and registration and smartphone-based triage could be followed after the patient had logged into the official WeChat account. Resources should be effectively allocated while using information systems [21]. In this study, smartphone-based triage took advantage of information technology to replace most of the work of routine field manual preinspection and triage by clinical service staff, effectively allocating resources. Not all patients used the smartphone app for appointment registration, but many of those who used the app did not require any assistance for the selection of the right department. This system could save time and energy of the support staff and effectively improve the utility of the precious medical resources.

Our hospital is specialized in pediatrics and woman health care; hence, more than 70% of outpatient visits were in the pediatric department. Nevertheless, the study showed that after the implementation of smartphone-based triage, the total outpatient visits increased compared with the previous year. The exact reasons for this increase remain unknown and deserve further investigation.

This study has limitations. The study site was limited to 1 hospital, and the collected data and indicators were limited. There was no randomization, and there could be a temporal bias. Multicenter trials with several indicators are required to confirm the findings. In addition, the app assessed in this study was specific to countries with a health system centered on hospital consultations by specialists, which is not the case in several other countries (eg, in Europe and America), where the patients consult a general practitioner who manages almost all diseases and refers to the proper specialists when needed. The questionnaire was not validated. Nevertheless, the 5-point Likert scale was used in this study. Finally, as the use of technology is often dependent upon age, financial resources, and education, future studies should examine the influence of those factors to optimize the use of the smartphone app for various populations of patients.

### Conclusions

In conclusion, smartphone-based triage significantly improves the rate of wrong patients' Web-based appointment registration and improves the degree of patient satisfaction. It allocates

expert resources more effectively, saving manpower for prediagnosis, triage, and patient guiding. Thus, such a smartphone approach is worth promoting. Future studies could examine its use by medical staff for inpatient management and transfer among departments. It could also be examined in general practitioners-centered settings.

## Acknowledgments

The authors are grateful for information collected by the medical staff in the outpatient department and Guangdong Situation Research Center. The study was funded by Guangzhou Science and Technology Plan Project (201806020047).

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Screenshots of smartphone-based triage.

[[PNG File , 240 KB - mhealth\\_v7i11e15313\\_app1.png](#) ]

### Multimedia Appendix 2

Satisfaction questionnaire.

[[DOC File , 85 KB - mhealth\\_v7i11e15313\\_app2.doc](#) ]

## References

1. Young RA, Burge SK, Kumar KA, Wilson JM, Ortiz DF. A time-motion study of primary care physicians' work in the electronic health record era. *Fam Med* 2018 Feb;50(2):91-99 [FREE Full text] [doi: [10.22454/FamMed.2018.184803](https://doi.org/10.22454/FamMed.2018.184803)] [Medline: [29432623](https://pubmed.ncbi.nlm.nih.gov/29432623/)]
2. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016 Dec 6;165(11):753-760. [doi: [10.7326/M16-0961](https://doi.org/10.7326/M16-0961)] [Medline: [27595430](https://pubmed.ncbi.nlm.nih.gov/27595430/)]
3. Jadczyk T, Kiwic O, Khandwalla RM, Grabowski K, Rudawski S, Magaczewski P, et al. Feasibility of a voice-enabled automated platform for medical data collection: CardioCube. *Int J Med Inform* 2019 Sep;129:388-393. [doi: [10.1016/j.ijmedinf.2019.07.001](https://doi.org/10.1016/j.ijmedinf.2019.07.001)] [Medline: [31445282](https://pubmed.ncbi.nlm.nih.gov/31445282/)]
4. Knight V, Guy RJ, Handan W, Lu H, McNulty A. It is more efficient to type: innovative self-registration and appointment self-arrival system improves the patient reception process. *Sex Transm Dis* 2014 Jun;41(6):392-394. [doi: [10.1097/OLQ.000000000000136](https://doi.org/10.1097/OLQ.000000000000136)] [Medline: [24825337](https://pubmed.ncbi.nlm.nih.gov/24825337/)]
5. Li LX, Benton WC, Leong G. The impact of strategic operations management decisions on community hospital performance. *J Oper Manag* 2002;20(4):389-408. [doi: [10.1016/s0272-6963\(02\)00002-5](https://doi.org/10.1016/s0272-6963(02)00002-5)]
6. Cao W, Wan Y, Tu H, Shang F, Liu D, Tan Z, et al. A web-based appointment system to reduce waiting for outpatients: a retrospective study. *BMC Health Serv Res* 2011 Nov 22;11:318 [FREE Full text] [doi: [10.1186/1472-6963-11-318](https://doi.org/10.1186/1472-6963-11-318)] [Medline: [22108389](https://pubmed.ncbi.nlm.nih.gov/22108389/)]
7. State Health and Family Planning Commission of the People's Republic of China. 2015. URL: <http://www.nhc.gov.cn/> [accessed 2019-11-04]
8. Song Z, Li Z, Sun L, Yan C, Li X. [Status quo analysis on reservation registration based on patient complaint in a children hospital]. *Chin Hosp* 2017;21:58-59 [FREE Full text]
9. Zhi W, Ran X, Yanfeng J. [The Research of Self-help Registration System Based on Information Index]. *Chin Figit Med* 2014;9:26-28. [doi: [10.3969/j.issn.1673-7571.2014.10.007](https://doi.org/10.3969/j.issn.1673-7571.2014.10.007)]
10. Xuezheng X, Jianping C, Yueying G. [Qualitative research on experience and psychological experience of outpatients with wrong registration]. *Nurs Rehab J* 2015;14:211-213. [doi: [10.3969/j.issn.1671-9875.2015.03.003](https://doi.org/10.3969/j.issn.1671-9875.2015.03.003)]
11. Lewis TL, Fothergill RT, Aneurysm-FILTR Study Group, Karthikesalingam A. Ambulance smartphone tool for field triage of ruptured aortic aneurysms (FILTR): study protocol for a prospective observational validation of diagnostic accuracy. *BMJ Open* 2016 Oct 24;6(10):e011308 [FREE Full text] [doi: [10.1136/bmjopen-2016-011308](https://doi.org/10.1136/bmjopen-2016-011308)] [Medline: [27797986](https://pubmed.ncbi.nlm.nih.gov/27797986/)]
12. Stein CD, Xiao X, Levine S, Schleyer TK, Hochheiser H, Thyvalikakath TP. A prototype mobile application for triaging dental emergencies. *J Am Dent Assoc* 2016 Oct;147(10):782-91.e1 [FREE Full text] [doi: [10.1016/j.adaj.2016.03.021](https://doi.org/10.1016/j.adaj.2016.03.021)] [Medline: [27206728](https://pubmed.ncbi.nlm.nih.gov/27206728/)]
13. den Hollander D, Mars M. Smart phones make smart referrals: The use of mobile phone technology in burn care - a retrospective case series. *Burns* 2017 Feb;43(1):190-194. [doi: [10.1016/j.burns.2016.07.015](https://doi.org/10.1016/j.burns.2016.07.015)] [Medline: [27575675](https://pubmed.ncbi.nlm.nih.gov/27575675/)]
14. Morse SS, Murugiah MK, Soh YC, Wong TW, Ming LC. Mobile health applications for pediatric care: review and comparison. *Ther Innov Regul Sci* 2018 May;52(3):383-391. [doi: [10.1177/2168479017725557](https://doi.org/10.1177/2168479017725557)] [Medline: [29714538](https://pubmed.ncbi.nlm.nih.gov/29714538/)]

15. Martinez R, Rogers AD, Numanoglu A, Rode H. The value of WhatsApp communication in paediatric burn care. *Burns* 2018 Jun;44(4):947-955. [doi: [10.1016/j.burns.2017.11.005](https://doi.org/10.1016/j.burns.2017.11.005)] [Medline: [29395403](https://pubmed.ncbi.nlm.nih.gov/29395403/)]
16. Tomlinson S, Behrmann S, Cranford J, Louie M, Hashikawa A. Accuracy of smartphone-based pulse oximetry compared with hospital-grade pulse oximetry in healthy children. *Telemed J E Health* 2018 Jul;24(7):527-535. [doi: [10.1089/tmj.2017.0166](https://doi.org/10.1089/tmj.2017.0166)] [Medline: [29215972](https://pubmed.ncbi.nlm.nih.gov/29215972/)]
17. Pidgeon TE, Blore CD, Webb Y, Horton J, Evans M. A patient information leaflet reduces parental anxiety before their child's first craniofacial multidisciplinary outpatient appointment. *J Craniofac Surg* 2017 Oct;28(7):1772-1776. [doi: [10.1097/SCS.0000000000003955](https://doi.org/10.1097/SCS.0000000000003955)] [Medline: [28891904](https://pubmed.ncbi.nlm.nih.gov/28891904/)]
18. Lewis AK, Harding KE, Snowdon DA, Taylor NF. Reducing wait time from referral to first visit for community outpatient services may contribute to better health outcomes: a systematic review. *BMC Health Serv Res* 2018 Nov 20;18(1):869 [FREE Full text] [doi: [10.1186/s12913-018-3669-6](https://doi.org/10.1186/s12913-018-3669-6)] [Medline: [30458761](https://pubmed.ncbi.nlm.nih.gov/30458761/)]
19. Randolph LA, Walker CK, Nguyen AT, Zachariah SR. Impact of pharmacist interventions on cost avoidance in an ambulatory cancer center. *J Oncol Pharm Pract* 2018 Jan;24(1):3-8. [doi: [10.1177/1078155216671189](https://doi.org/10.1177/1078155216671189)] [Medline: [27682600](https://pubmed.ncbi.nlm.nih.gov/27682600/)]
20. Keshvani N, Hon M, Gupta A, Brown TJ, Roy L, Marley E, et al. Reducing hospitalizations: institution of outpatient infusional epoch-based chemotherapy at a safety net hospital. *J Oncol Pract* 2019 Aug;15(8):e644-e651. [doi: [10.1200/JOP.18.00738](https://doi.org/10.1200/JOP.18.00738)] [Medline: [31206340](https://pubmed.ncbi.nlm.nih.gov/31206340/)]
21. de Bruin JS, Schuh C, Seeling W, Luger E, Gall M, Hütterer E, et al. Assessing the feasibility of a mobile health-supported clinical decision support system for nutritional triage in oncology outpatients using Arden Syntax. *Artif Intell Med* 2018 Nov;92:34-42. [doi: [10.1016/j.artmed.2015.10.001](https://doi.org/10.1016/j.artmed.2015.10.001)] [Medline: [26563776](https://pubmed.ncbi.nlm.nih.gov/26563776/)]
22. Nogueira RG, Silva GS, Lima FO, Yeh Y, Fleming C, Branco D, et al. The FAST-ED app: a smartphone platform for the Field Triage of patients with stroke. *Stroke* 2017 May;48(5):1278-1284. [doi: [10.1161/STROKEAHA.116.016026](https://doi.org/10.1161/STROKEAHA.116.016026)] [Medline: [28411260](https://pubmed.ncbi.nlm.nih.gov/28411260/)]
23. Uthoff RD, Song B, Sunny S, Patrick S, Suresh A, Kolar T, et al. Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities. *PLoS One* 2018;13(12):e0207493 [FREE Full text] [doi: [10.1371/journal.pone.0207493](https://doi.org/10.1371/journal.pone.0207493)] [Medline: [30517120](https://pubmed.ncbi.nlm.nih.gov/30517120/)]
24. Tayfur I, Afacan MA. Reliability of smartphone measurements of vital parameters: A prospective study using a reference method. *Am J Emerg Med* 2019 Aug;37(8):1527-1530. [doi: [10.1016/j.ajem.2019.03.021](https://doi.org/10.1016/j.ajem.2019.03.021)] [Medline: [30904343](https://pubmed.ncbi.nlm.nih.gov/30904343/)]
25. Verzantvoort NC, Teunis T, Verheij TJ, van der Velden AW. Self-triage for acute primary care via a smartphone application: Practical, safe and efficient? *PLoS One* 2018;13(6):e0199284 [FREE Full text] [doi: [10.1371/journal.pone.0199284](https://doi.org/10.1371/journal.pone.0199284)] [Medline: [29944708](https://pubmed.ncbi.nlm.nih.gov/29944708/)]
26. Tian Y, Zhou T, Yao Q, Zhang M, Li J. Use of an agent-based simulation model to evaluate a mobile-based system for supporting emergency evacuation decision making. *J Med Syst* 2014 Dec;38(12):149. [doi: [10.1007/s10916-014-0149-3](https://doi.org/10.1007/s10916-014-0149-3)] [Medline: [25354665](https://pubmed.ncbi.nlm.nih.gov/25354665/)]
27. Gupta N, Chawla N, Gupta D, Dhawan N, Janaki VR. Community triage of otology patients using a store-and-forward telemedicine device: A feasibility study. *Ear Nose Throat J* 2017 Jul;96(7):246-249. [Medline: [28719707](https://pubmed.ncbi.nlm.nih.gov/28719707/)]
28. 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](https://doi.org/10.1016/j.ajem.2015.07.029)] [Medline: [26299691](https://pubmed.ncbi.nlm.nih.gov/26299691/)]
29. Börve A, Gyllencreutz JD, Terstappen K, Backman EJ, Aldenbratt A, Danielsson M, et al. Smartphone teledermoscopy referrals: a novel process for improved triage of skin cancer patients. *Acta Derm Venereol* 2015 Feb;95(2):186-190 [FREE Full text] [doi: [10.2340/00015555-1906](https://doi.org/10.2340/00015555-1906)] [Medline: [24923283](https://pubmed.ncbi.nlm.nih.gov/24923283/)]
30. Gyllencreutz JD, Backman EJ, Terstappen K, Paoli J. Teledermoscopy images acquired in primary health care and hospital settings - a comparative study of image quality. *J Eur Acad Dermatol Venereol* 2018 Jun;32(6):1038-1043. [doi: [10.1111/jdv.14565](https://doi.org/10.1111/jdv.14565)] [Medline: [28850732](https://pubmed.ncbi.nlm.nih.gov/28850732/)]
31. Urner E, Delavy M, Catarino R, Viviano M, Meyer-Hamme U, Benski A, et al. A smartphone-based approach for triage of human papillomavirus-positive Sub-Saharan African women: a prospective study. *JMIR Mhealth Uhealth* 2017 May 29;5(5):e72 [FREE Full text] [doi: [10.2196/mhealth.6697](https://doi.org/10.2196/mhealth.6697)] [Medline: [28554879](https://pubmed.ncbi.nlm.nih.gov/28554879/)]

*Edited by G Eysenbach; submitted 01.07.19; peer-reviewed by F Kaliyadan, S Badawy, JR Bautista, T Clavier; comments to author 21.08.19; revised version received 16.10.19; accepted 17.10.19; published 11.11.19.*

*Please cite as:*

Xie W, Cao X, Dong H, Liu Y

*The Use of Smartphone-Based Triage to Reduce the Rate of Outpatient Error Registration: Cross-Sectional Study*

*JMIR Mhealth Uhealth* 2019;7(11):e15313

URL: <http://mhealth.jmir.org/2019/11/e15313/>

doi: [10.2196/15313](https://doi.org/10.2196/15313)

PMID: [31710300](https://pubmed.ncbi.nlm.nih.gov/31710300/)

©Wanhua Xie, Xiaojun Cao, Hongwei Dong, Yu Liu. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 11.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.



Original Paper

# Quality Awareness and Its Influence on the Evaluation of App Meta-Information by Physicians: Validation Study

Urs-Vito Albrecht<sup>1\*</sup>, Dr med, MPH; Theodor Framke<sup>2\*</sup>, Dr rer nat; Ute von Jan<sup>1\*</sup>, Dr rer biol hum

<sup>1</sup>Peter L Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

<sup>2</sup>Institute for Biometry, Hannover Medical School, Hannover, Germany

\* all authors contributed equally

**Corresponding Author:**

Urs-Vito Albrecht, Dr med, MPH

Peter L Reichertz Institute for Medical Informatics

Hannover Medical School

Carl-Neuberg-Str 1

Hannover, 30625

Germany

Phone: 49 511532 ext 3508

Email: [albrecht.urs-vito@mh-hannover.de](mailto:albrecht.urs-vito@mh-hannover.de)

## Abstract

**Background:** Meta-information provided about health apps on app stores is often the only readily available source of quality-related information before installation.

**Objective:** The purpose of this study was to assess whether physicians deem a predefined set of quality principles as relevant for health apps; whether they are able to identify corresponding information in a given sample of app descriptions; and whether, and how, this facilitates their informed usage decisions.

**Methods:** All members of the German Society for Internal Medicine were invited by email to participate in an anonymous online survey over a 6-week period. Participants were randomly assigned one app description focusing on cardiology or pulmonology. In the survey, participants were asked three times about whether the assigned description sufficed for a usage decision: they were asked (1) after giving an appraisal of the relevance of nine predefined app quality principles, (2) after determining whether the descriptions covered the quality principles, and (3) after they assessed the availability of detailed quality information by means of 25 additional key questions. Tests for significance of changes in their decisions between assessments 1 and 2, and between assessments 2 and 3, were conducted with the McNemar-Bowker test of symmetry. The effect size represents the discordant proportion ratio sum as a quotient of the test statistics of the Bowker test and the number of observation units. The significance level was set to  $\alpha=.05$  with a power of  $1-\beta=.95$ .

**Results:** A total of 441 of 724 participants (60.9%) who started the survey fully completed the questionnaires and were included in the evaluation. The participants predominantly rated the specified nine quality principles as important for their decision (approximately 80%-99% of ratings). However, apart from the practicality criterion, information provided in the app descriptions was lacking for both groups (approximately 51%-92%). Reassessment of the apps led to more critical assessments among both groups. After having familiarized themselves with the nine quality principles, approximately one-third of the participants (group A: 63/220, 28.6%; group B: 62/221, 28.1%) came to more critical usage decisions in a statistically significant manner (McNemar-Bowker test, groups A and B:  $P<.001$ ). After a subsequent reassessment with 25 key questions, critical appraisals further increased, although not in a statistically significant manner (McNemar-Bowker, group A:  $P=.13$ ; group B:  $P=.05$ ).

**Conclusions:** Sensitizing physicians to the topic of quality principles via questions about attitudes toward established quality principles, and letting them apply these principles to app descriptions, lead to more critical appraisals of the sufficiency of the information they provided. Even working with only nine generic criteria was sufficient to bring about the majority of decision changes. This may lay the foundation for aiding physicians in their app-related decision processes, without unduly taking up their valuable time.

(JMIR Mhealth Uhealth 2019;7(11):e16442) doi:[10.2196/16442](https://doi.org/10.2196/16442)

**KEYWORDS**

mobile health; evaluation studies; mobile apps; quality principles; usage decisions

## Introduction

The health app market is highly dynamic and liberally organized, which makes attempts at assessing and adequately regulating it very difficult. There is little reliable information about the actual size of the market, the manufacturers' composition, or about the products themselves [1]. Various authors lament the deficit of information regarding health apps, apart from marketing-oriented aspects, independent of the platform and area of application [2-5]. There is usually only little information on efficacy, risk profiles, manufacturing processes, and various other key aspects [6-8]. However, in order to make a well-founded decision for usage, as is ethically and legally required for physicians who want to use such health apps in their medical routine, those interested in such apps need sufficient and readily accessible information, ideally provided by the manufacturers; however, there is often a lack of such information. Our own preliminary work confirmed this assumption by means of studying a subset of apps for the cardiological and pulmonological spectrum [9]. Likewise, there are no commonly agreed-upon criteria that those interested in health-related apps, be it for their own use or other purposes such as research, could apply toward assessment of apps in a structured manner [10]; in fact, there is a wide variety of methods that are currently being used. For example, in the context of scientific evaluations, with respect to content, quality is often rated in a descriptive manner or based on surrogate outcome measures [11]. Alternatively, those interested in an app are referred to guidelines and orientation documents, which are more or less extensive catalogues of criteria or third-party certifications or quality seals. These, in turn, also commonly employ variable rating criteria (eg, Ministry of Health New Zealand [12], Canadian Medical Association [13], National Health Service [14], and Fraunhofer-Institut für Offene Kommunikationssysteme [15]; see Albrecht [16] for a more extensive listing of possible approaches). Many such methodologies may be appropriate depending on the context in which they are to be employed. However, some methodologies are possibly either too specific to a certain area of application or confront users with too large a number of criteria and aspects; this makes it difficult for them to determine if the respective approaches are indeed helpful for their use case and whether or not they should use an app.

For new technologies such as mHealth or eHealth in general, there is still a need to survey or at least involve the relevant target groups, in this case physicians, when it comes to creating and evaluating tools offered in the app context or evaluating the apps themselves in line with the requirements of users (eg, Hennemann et al [17], Martínez-Pérez et al [18], and Tarricone et al [19]). This is the case in order to gain acceptance of apps by health care professionals; for this reason, we performed a survey-based evaluation in this paper and fine-tuned our approach to evaluating quality criteria for apps in the health context. In our previous study, we surveyed medical students and demonstrated that there was a considerable discrepancy

between the high-quality requirements for health apps, as communicated by the students, and the information provided about these prerequisites in the app descriptions. The study also showed that after examining nine quality principles and 20 features and criteria, students became more critical of the information available in app descriptions. In 3 out of 4 (75%) cases examined, the students changed their opinions regarding the suitability of the app descriptions from *sufficient* or *do not know* to *insufficient*, for the purpose of making their decisions about usage.

In this study, we aimed to validate this approach among physicians. Our main objective was to investigate whether being made aware of, and working with, quality principles, as introduced previously [20], influences physicians' assessments of whether the information in app descriptions was sufficient for them to make a decision about app usage. Furthermore, the basic assessment of the relevance of quality principles for usage decisions by the physicians was also investigated. In addition, we analyzed how physicians assess the sufficiency of the information provided in the app descriptions regarding their compliance with the quality principles. Finally, the frequency of specific aspects, identified by key questions within the assigned app descriptions, was collected in a descriptive manner.

## Methods

### Overview

This study was conducted as an anonymous, investigator-blinded, randomized, and standardized Web-based survey with health care professionals who are members of the German Association for Internal Medicine (ie, Deutsche Gesellschaft für Innere Medizin [DGIM] eV). The survey was conducted at Hannover Medical School using a local installation of the SoSci Survey [21] tool, version 3.2.000 (SoSci Survey GmbH), and was open for a period of 6 weeks, between June 17 and July 29, 2019. On July 8, 2019, an email reminder was sent to all DGIM members who were initially invited to participate. Altogether, approximately 21,000 DGIM members were invited to participate using the DGIM email list. The study was approved by the Institutional Review Board of Hannover Medical School (study number: 8256\_BO\_K\_2019).

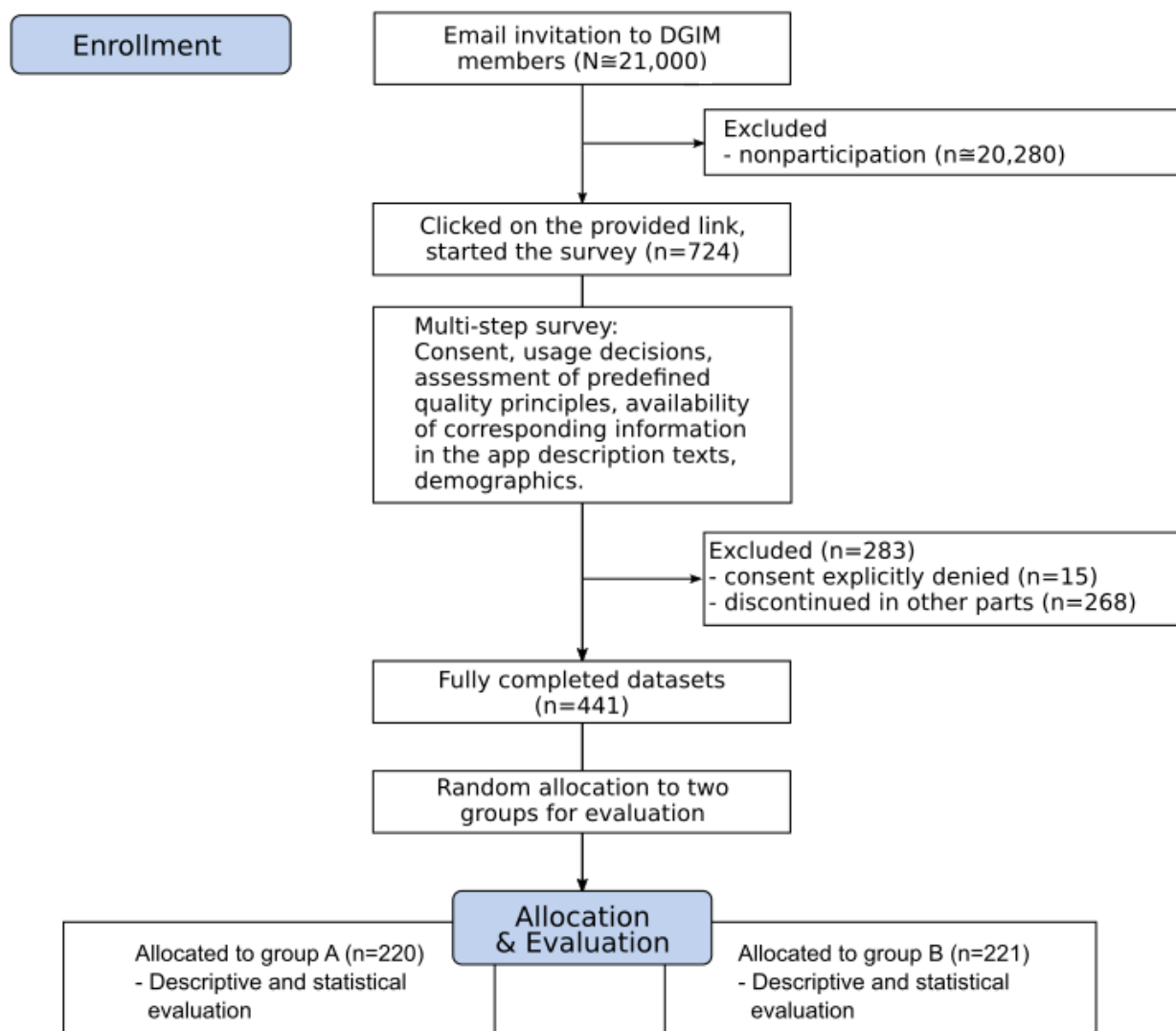
### Structure of the Survey

After a short introduction to the study and obtaining informed consent, participants were asked to evaluate one randomly assigned app description; these descriptions were taken from a selection of 126 cardiology- and pulmonology-related apps acquired from Apple's app store in 2018, which were used as basis for our previous study [9]. Initially, they were asked to assess whether, in their opinion, the information presented in the store description was sufficient for deciding for or against using the selected app (see Figure 1, assessment 1 [A1]). Subsequently, the study participants were questioned about their perception of the relevance of nine given quality principles for making a decision about usage (see Figure 1, questionnaire 1

[Q1]). This was followed by an assessment as to whether the information contained in the descriptions could be used to conclude that these quality principles were met (see Figure 1). The respondents were then asked to reassess whether the descriptive texts were sufficient for making a decision about usage (see Figure 1). Afterward, the study participants had to determine whether the relevant information was present using

closed-ended questions (see Figure 1). Thereafter, analogous to the initial assessment, the participants were prompted to reassess whether the descriptive texts were sufficient for making a decision about usage (see Figure 1). The methodology described is based on a preliminary study with medical students, which has been successfully implemented [9].

**Figure 1.** Study flowchart. DGIM: Deutsche Gesellschaft für Innere Medizin.



## Evaluation Strategy

### Overview

For validation purposes, the respondents were randomly assigned into two sample groups: A and B. For this purpose, the random number generator from R, version 3.5.2 (The R Foundation), was used to generate and assign random floating-point numbers between 1 and 441 to the fully completed datasets. The dataset rows were then sorted using these numbers. The first dataset members (N=220) were placed in group A and the rest (N=221) were placed in group B. Primary and secondary analyses were performed equally for samples A and B.

### Primary Analysis

The primary aim was to determine whether there was a change in participants' answers regarding the assessments done at

different stages of the survey (ie, assessments 1, 2, and 3 [A1, A2, and A3]).

To this end, two comparisons were made. The first comparison between the first and second assessments (ie, A1 and A2) quantified the change in assessment after getting to know the quality principles. The second comparison of the second and third assessments (ie, A2 and A3) quantified the change in assessment after participants had checked the app descriptions regarding specific (ie, operationalized) quality principles. For testing, the McNemar-Bowker test [22,23] was used, which considers the first comparison and then the second comparison in the sense of a test hierarchy. The study was regarded as successful if at least the first comparison (ie, A1 vs A2) showed a significant difference according to the research hypothesis. For each comparison, a two-sided hypothesis— $H_0: p_{ij}=p_{ji}$  for

all  $i \neq j$  versus  $H_1: p_{ij} \neq p_{ji}$  for at least one  $i \neq j$ —was used (see Table 1). The effect size represents the discordant proportion ratio

sum (DPRS) as a quotient of the test statistics of the Bowker test and the number of observation units. The significance level was set to  $\alpha = .05$ .

**Table 1.** Probabilities of the before-and-after paired decisions (eg, A1<sup>a</sup> vs A2<sup>b</sup>).

Rating before (eg, A1)	Rating after (eg, A2)			Total
	Insufficient	Do not know	Sufficient	
Insufficient	$P_{11}^c$	$P_{12}$	$P_{13}$	$P_1$
Do not know	$P_{21}$	$P_{22}$	$P_{23}$	$P_2$
Sufficient	$P_{31}$	$P_{32}$	$P_{33}$	$P_3$
Total	$P_1$	$P_2$	$P_3$	1

<sup>a</sup>A1: assessment 1.

<sup>b</sup>A2: assessment 2.

<sup>c</sup>Proportions  $p_{ij}$  are derived by dividing the counts by the total number of participants.

### Secondary Analysis

In order to analyze the influence of the individual quality aspects on the intention to make a usage decision, a logistic regression with a univariate approach was carried out with a modeled probability of *negative changes* (ie, changes to *do not know* or *insufficient*.)

Furthermore, in a descriptive form, the health care professionals' assessments of the relevance of the quality principles for their usage decision were recorded. The same procedure was used to evaluate the sufficiency of the information provided in the app descriptions in order to assess compliance with the quality principles. In addition, the frequencies of the aspects identified in the descriptions, based on the key questions, were noted.

### Sample Size Calculation

Sample size planning was done as described for the Power Analysis and Sample Size (PASS) software (NCSS) [24] and by Chow et al [25] and was carried out with R, version 3.5.2 (The R Foundation). For a preliminary study with a similar setup that was conducted previously, a DPRS of 0.13 had been

calculated. As a result of conservative planning, a DPRS of 0.1 was assumed for this project. This resulted in a sample size of 175 for  $\alpha = .05$  and  $1 - \beta = .95$  per end point.

### Tools

#### Questionnaires Used at Different Stages of the Survey

Several tools were used for the survey, which was implemented online by means of the server-based SoSci Survey tool, version 3.2.000 (SoSci Survey GmbH) [21]. Q1 and questionnaire 2 (Q2) matched those that were used in our previous study. The questions used for questionnaire 3 (Q3), however, were based on the technical report presented in Albrecht [20]. For the initial assessment (ie, Q1) of the quality principles, which included nine items, a 5-point Likert scale with an additional field for *do not know* was used and the answers were mandatory (see Table 2).

Using nine questions with possible answers of *yes*, *no*, and *do not know*—answers were mandatory; see Table 3 for the corresponding questions—participants were asked via Q2 whether there was sufficient information in the app descriptions to ascertain whether the apps met these principles.

**Table 2.** The nine quality principles and their explanations used for the initial assessment (ie, Q1<sup>a</sup>), following the definitions in Albrecht et al [9].

Quality principle ID	Explanation
Q101	Practicality: The software can be used for the intended purpose and must be as versatile as possible in order to open up the largest possible application areas and contexts.
Q102	Risk adequacy: The software provides the means to be used in a risk-appropriate manner without exposing the user or his or her environment to a disproportionate health, social, or economic risk.
Q103	Ethical soundness: Development, provision, operation, and use of the software are ethically innocuous in order to prevent discrimination and stigmatization and to facilitate fair access.
Q104	Legal conformity: Legal conformity (eg, medical device law, professional law, data protection law, and law on the advertising of therapeutic products) of the development, provision, operation, and use of the software is guaranteed for the protection of all parties involved (eg, providers, store operators, and users).
Q105	Content validity: The health-related content of the software that is presented and used is valid and trustworthy (ie, scientifically sound, up-to-date, and without conflict of interest).
Q106	Technical adequacy: Development, operation, maintenance, and use of the software correspond to the state of the art in order to enable sustainability in terms of maintainability as well as platform-independent and cross-platform use (eg, in terms of portability of the app or interoperability or compatibility with other products).
Q107	Usability: The software allows the target group to make appropriate use of it (eg, through product ergonomics, accessibility, and aesthetics), which contributes to user satisfaction.
Q108	Resource efficiency: During development of the software, elements for resource-efficient operation (eg, energy consumption) and use (eg, computing time) are taken into account.
Q109	Transparency: Complete transparency regarding the quality principles serves as a basis for evaluations of the software as well as for individual and collective usage decisions.

<sup>a</sup>Q1: questionnaire 1.

**Table 3.** Questions, via Q2<sup>a</sup>, used for determining whether the presented app descriptions contained information related to the nine quality principles, following those employed in Albrecht et al [9].

Question ID	Question
Q201	Practicality: Can you use the app description to make an assessment as to whether the app is useful?
Q202	Risk adequacy: Can you use the app description to make an assessment as to whether the app is risk adequate?
Q203	Ethical soundness: Can you use the app description to make an assessment as to whether the app is ethically safe?
Q204	Legal conformity: Can you use the app description to make an assessment as to whether the app is legally compliant?
Q205	Content validity: Can you use the app description to make an assessment as to whether the content of the app is valid?
Q206	Technical adequacy: Can you use the app description to make an assessment as to whether the app is technically appropriate?
Q207	Usability: Can you use the app description to make an assessment as to whether the app is usable?
Q208	Resource efficiency: Can you use the app description to make an assessment as to whether the app is resource efficient?
Q209	Transparency: Can you use the app description to make an assessment as to whether the app description is transparent on the above points?

<sup>a</sup>Q2: questionnaire 2.

Following this, a more detailed retrieval of information, via Q3 and using a set of 25 questions, was employed. This tool operationalizes the nine quality principles with 25 features and requirements based on the catalogue of criteria for self-declaration of health apps presented in Albrecht [20]. Answer options in this case were *yes*, *no*, *do not know*, and *not answered*. Although in this case, answers were not defined as mandatory; if any answers were missing in this part of the survey, the participants were alerted once and asked whether

they really did not want to rate the corresponding questions before being allowed to continue. There was also a free-text field for optional comments (see Table 4). The interspersed usage decision question (see steps A1-A3) had to be answered using the options *yes*, *no*, and *do not know*; answering this question was mandatory. Lastly, participant demographics (D0) were obtained using 13 items; again, answers were not considered mandatory.

**Table 4.** Operationalization of the nine quality principles using 25 detailed questions, via Q3<sup>a</sup>, according to Albrecht [20].

Question ID	Question	Contribution
Q301	Has the purpose of the app been specified?	Practicality
Q302	Is there a description of functions the app offers in order to fulfill its purpose?	Practicality
Q303	Is there a description of the app given that states which methods it employs to fulfill its purpose (eg, procedures, processes, and algorithms with which the offered functions are implemented), and are there statements regarding their suitability for this purpose?	Practicality
Q304	Is appropriate evidence cited to support the statements on fulfillment of the purpose of the app (eg, references to studies, guidelines, testing, and quality labels)?	Practicality
Q305	Are the suitability and unsuitability for certain app scenarios or user groups specified (eg, in terms of inclusion and exclusion criteria)?	Practicality
Q306	Have potential or actual risks the app poses to users or their environment been stated, with respect to health, economic, or social aspects?	Risk adequacy
Q307	Have precautions been taken to avoid health, economic, and/or social risks when using the described app?	Risk adequacy
Q308	Is there a description about the extent to which the app follows ethical principles, such as patient autonomy, equity of access, and/or professional ethics and research ethics?	Ethical soundness
Q309	Are conflicts of interest (eg, authors with affiliations to specific companies) discussed in the app description?	Ethical soundness
Q310	Is there a mention of whether the app is being provided in a research context; if so, is there a statement about whether it follows good scientific practice?	Ethical soundness
Q311	Is there a statement about whether the relevant general legal requirements, such as data protection law, telemedia law, and commercial law, have been taken into account by the manufacturer and provider of the app?	Legal conformity
Q312	Is there a statement about which requirements and regulations have been taken into account with regard to using the app in a health context, such as medical device law or medical professional law?	Legal conformity
Q313	Is there a statement about how the quality of the content has been ensured (eg, involvement of experts in the field) or which validated sources have been used (eg, consideration of current scientific findings and guidelines)?	Content validity
Q314	Is there a description of how the app is regularly adapted to new content requirements?	Content validity
Q315	Is it described to what extent the app corresponds to the current state of the art?	Technical adequacy
Q316	Is there information about how the app is regularly adapted to technical requirements?	Technical adequacy
Q317	Is there information about to what extent it is possible to switch to another operating system or device without data loss?	Technical adequacy
Q318	Is there a mention of whether the app is scalable (ie, adaptable to increasing requirements) or can be integrated into other products?	Technical adequacy
Q319	Has information been provided about proof for the app's usability (eg, usability tests)?	Usability
Q320	Is it described to what extent the function of the app has been specifically adapted to the target group, whether it is barrier-free, or whether it can be used with individual adaptations?	Usability
Q321	Is there information about the extent to which user feedback was considered for the app (eg, during the development process)?	Usability
Q322	Is there a statement about how the app ensures efficient use of the available technical resources (eg, required memory, computing power, internal or external sensors, and power consumption)?	Resource efficiency
Q323	Is the information about the app sufficient (ie, adequately specified in scope and depth of information)?	Transparency
Q324	Is valid (ie, complete and reliable) information about the app provided?	Transparency
Q325	Is the information about the app described in a manner that is adequate for the target group?	Transparency

<sup>a</sup>Q3: questionnaire 3.

### Pretest of the Questionnaires

The standardized questionnaires described above were pretested in three iterations. After the creation of the first version, the questionnaire was tested by three medical students on May 22, 2019. This was accompanied by the *thinking aloud* approach,

with utterances being used as feedback and a basis for adaptations to the survey (ie, iteration 1). The three students were also asked to paraphrase the questions contained in the survey, to ascertain whether they understood the questions correctly. This made it possible to identify and adapt unclear questions and answer options. In addition, usability issues were

addressed. The provided feedback was used for revising the questions. The follow-up version of the questionnaire was then tested by 10 public health students and two lecturers between May 23 and June 6, 2019. This was meant as a pretest under real-life conditions; we used a live version of the questionnaires in our SoSci Survey setup, albeit with an additional comment function (ie, iteration 2). Afterward, solutions for issues discovered in this second pretest, as well as any constructive remarks obtained from the participants, were incorporated into the survey. This was once again presented and discussed with the testers in a follow-up meeting on June 6, 2019 (ie, iteration 3). After any remaining minor issues that were mentioned in this final meeting had been resolved, the questionnaire was finalized.

### Software Used in the Evaluation Process

The following software programs were used at various stages of the evaluation process to generate graphics as well as for data description and statistics: IBM SPSS Statistics for Mac OS, version 26.0 (IBM Corp), and R, version 3.5.2 (The R Foundation), along with the R packages dplyr [26], ggplot2

[27], RColorBrewer [28], arsenal [29], qwraps2 [30], Hmisc [31], DescTools [32], and rcompanion [33].

## Results

### Response Rates

Out of a total of 1357 clicks on the provided survey link, 724 participants started the questionnaire and 441 of these participants (60.9%) fully completed the questionnaire (see Figure 1 and Table 5 for further details). An additional 15 participants did not consent to participation on the first page of the survey and were, thus, not shown the actual survey questions; instead, they were asked if they were willing to provide information about their reasons for not consenting. Of those who gave a reason for nonparticipation, contributing factors were lack of knowledge about apps (n=6), information about the survey being perceived as insufficient (n=3), lack of interest in the subject (n=2), and lack of time (n=2). Additionally, other reasons given in free-text form were related to being too old and/or being retired (n=3). One additional participant gave a statement to the effect of being worried that her answers would not be sufficiently anonymized.

**Table 5.** Participant dropout at different stages of the survey, from 724 participants who originally started the survey.

Survey page number	Description	Full completion of survey <sup>a</sup> , n	Dropouts, n
9	Questions about reasons for not consenting; these were for participants who had not consented to participation on page 1	15	0
8	Demographics	441	1
7	A3 <sup>b</sup> : Final usage decision	N/A <sup>c</sup>	1
6	Q3 <sup>d</sup> : In-depth evaluation based on 25 key questions	N/A	16
5	A2 <sup>e</sup> : Intermediate usage decision	N/A	45
4	Q2 <sup>f</sup> : Assessment of whether the available information suffices for assessing the app based on the nine quality principles	N/A	11
3	Q1 <sup>g</sup> : Participants' assessments of nine predefined quality principles	N/A	44
2	A1 <sup>h</sup> : Initial usage decision	N/A	59
1	Introduction and consent	N/A	91
N/A	Sum of participants	456	268

<sup>a</sup>Participants were considered to have completed the survey either by completing the final demographics questions on page 8 or, for those who had not consented to participation, by providing reasons for not giving their consent on page 9.

<sup>b</sup>A3: assessment 3.

<sup>c</sup>Not applicable.

<sup>d</sup>Q3: questionnaire 3.

<sup>e</sup>A2: assessment 2.

<sup>f</sup>Q2: questionnaire 2.

<sup>g</sup>Q1: questionnaire 1.

<sup>h</sup>A1: assessment 1.

### Demographics of the Test and Validation Samples

Baseline demographics for the participants, stratified by the randomly assigned groups A (test sample) and B (validation sample), are shown in Table 6. Table 7 provides additional data

about the participants' interests in digital topics and app usage patterns in private- and work-related areas. For all variables in these two tables, there were only statistically insignificant differences between the two groups. *P* values were calculated using the  $\chi^2$  test function provided by R [34].

**Table 6.** Baseline demographics of all participants who completed the questionnaire; demographics are stratified by gender for the randomly assigned groups A and B.

Baseline demographic (D0)	Group A (N=220), n (%)	Group B (N=221), n (%)	$\chi^2$	df	P value
<b>D002: Gender</b>			<b>3.2</b>	<b>2</b>	<b>.20</b>
Male	113 (51.4)	130 (58.8)			
Female	104 (47.3)	90 (40.7)			
Diverse	3 (1.4)	1 (0.5)			
<b>D001: Age in years</b>			<b>3.9</b>	<b>4</b>	<b>.42</b>
21-30	11 (5.0)	8 (3.6)			
31-40	60 (27.3)	79 (35.7)			
41-50	55 (25.0)	48 (21.7)			
51-60	59 (26.8)	53 (24.0)			
>60	35 (15.9)	33 (14.9)			
<b>D003: Years of work</b>			<b>7.6</b>	<b>6</b>	<b>.27</b>
<1	2 (0.9)	0 (0)			
1-5	18 (8.2)	18 (8.1)			
6-10	33 (15.0)	49 (22.2)			
11-20	66 (30.0)	57 (25.8)			
21-30	53 (24.1)	42 (19.0)			
>30	34 (15.5)	40 (18.1)			
Retired	14 (6.4)	15 (6.8)			
<b>D004: Specialty</b>			<b>28.0</b>	<b>26</b>	<b>.37</b>
Internal medicine	76 (34.5)	77 (34.8)			
Internal medicine and general medicine, without further specification	7 (3.2)	17 (7.7)			
Internal medicine and angiology	5 (2.3)	1 (0.5)			
Internal medicine and endocrinology and diabetology	6 (2.7)	4 (1.8)			
Internal medicine and gastroenterology	22 (10.0)	25 (11.3)			
Internal medicine and hematology and oncology	14 (6.4)	11 (5.0)			
Internal medicine and cardiology	17 (7.7)	27 (12.2)			
Internal medicine and nephrology	13 (5.9)	16 (7.2)			
Internal medicine and pulmonology	12 (5.5)	9 (4.1)			
Internal medicine and rheumatology	5 (2.3)	3 (1.4)			
General practitioner	28 (12.7)	19 (8.6)			
Other specialties	15 (6.8)	12 (5.4)			
<b>D005: Function at work</b>			<b>3.0</b>	<b>5</b>	<b>.71</b>
Chief physician	18 (8.2)	12 (5.4)			
Senior physician	50 (22.7)	56 (25.3)			
Junior physician	49 (22.3)	55 (24.9)			
Student	1 (0.5)	0 (0)			
Other	99 (45.0)	95 (43.0)			
No data	71 (32.3)	71 (32.1)			
<b>D006: Sector of work</b>			<b>0.7</b>	<b>3</b>	<b>.87</b>
Outpatient sector	76 (34.5)	76 (34.4)			
Clinic	120 (54.5)	116 (52.5)			



Baseline demographic (D0)	Group A (N=220), n (%)	Group B (N=221), n (%)	$\chi^2$	df	P value
Other	23 (10.5)	27 (12.2)			
No data	24 (10.9)	29 (13.1)			
<b>D007: Type of employment</b>			<b>2.3</b>	<b>3</b>	<b>.51</b>
Salaried	157 (71.4)	148 (67.0)			
Self-employed	52 (23.6)	56 (25.3)			
Other	11 (5.0)	16 (7.2)			
No data	0 (0)	1 (0.5)			
<b>D008: Highest academic degree</b>			<b>6.2</b>	<b>6</b>	<b>.41</b>
Habilitation	15 (6.8)	23 (10.4)			
Doctoral degree	126 (57.3)	133 (60.2)			
State exam	71 (32.3)	55 (24.9)			
Master's degree, diploma, or similar	4 (1.8)	5 (2.3)			
Bachelor's degree or similar	0 (0)	1 (0.5)			
Other	3 (1.4)	4 (1.8)			
No data	1 (0.5)	0 (0)			
D009: Geographic location <sup>a</sup>			22.0	31	.89

<sup>a</sup>The data for the individual subcategories (on a federal state level for Germany, Austria, and Switzerland, as well as per country for other members of the European Union) is intentionally not shown here for brevity reasons. It is, however, available upon request from the authors.

**Table 7.** Interest in digital topics and app usage patterns, stratified by randomly assigned groups A and B.

Interest demographic (D0)	Group A (N=220), n (%)	Group B (N=221), n (%)	$\chi^2$	df	P value
<b>D010: Interest in digital topics</b>			<b>0.2</b>	<b>4</b>	<b>&gt;.99</b>
Highly interested	68 (30.9)	68 (30.8)			
Interested	95 (43.2)	98 (44.3)			
Partly interested	41 (18.6)	38 (17.2)			
Less interested	12 (5.5)	13 (5.9)			
Not interested	4 (1.8)	4 (1.8)			
<b>D011: Private use of apps</b>			<b>1.8</b>	<b>1</b>	<b>.18</b>
Yes	201 (91.4)	210 (95.0)			
No	19 (8.6)	11 (5.0)			
<b>D012: Use of apps at work</b>			<b>0</b>	<b>1</b>	<b>&gt;.99</b>
Yes	153 (69.5)	154 (69.7)			
No	67 (30.5)	67 (30.3)			
<b>D013: Have been asked about apps by patients</b>			<b>0.001</b>	<b>1</b>	<b>.97</b>
Yes	68 (30.9)	67 (30.3)			
No	152 (69.1)	154 (69.7)			
<b>A303<sup>a</sup>: App or its description was known</b>			<b>3.4</b>	<b>2</b>	<b>.18</b>
Yes	12 (5.5)	5 (2.3)			
No	206 (93.6)	215 (97.3)			
Do not know	2 (0.9)	1 (0.5)			

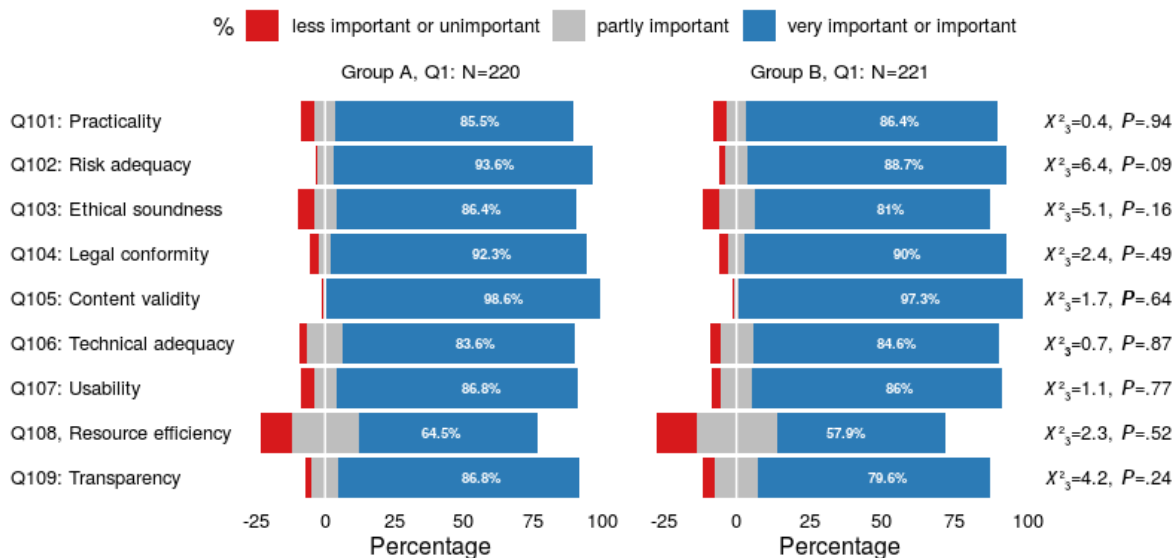
<sup>a</sup>This question was shown on page 7 of assessment 3 (A3) and was not part of the demographics part of the survey on page 8; therefore, it was encoded differently.

### Baseline Assessment of the Nine Predefined Quality Principles

In their initial assessment of the presented quality principles—see Table 2 for the definition of the nine quality principles—the majority of participants rated the criteria as *very important* or *important* (80%–99% of the participants). However, there was one notable exception: for resource efficiency, only

about two-thirds of the participants perceived this criterion as *important* or *very important*, while around 1 in 4 participants rated it as partially important. As for the demographic variables listed in Table 7, there were again no statistically significant differences between the groups regarding their assessment of the nine quality principles (see Figure 2; Multimedia Appendix 1 provides further information about the answers given by the participants).

**Figure 2.** Participants’ assessments, stratified by group, of the perceived importance of the nine quality principles, including corresponding results for the chi-square tests that only show statistically insignificant differences between groups. Q1: questionnaire 1.



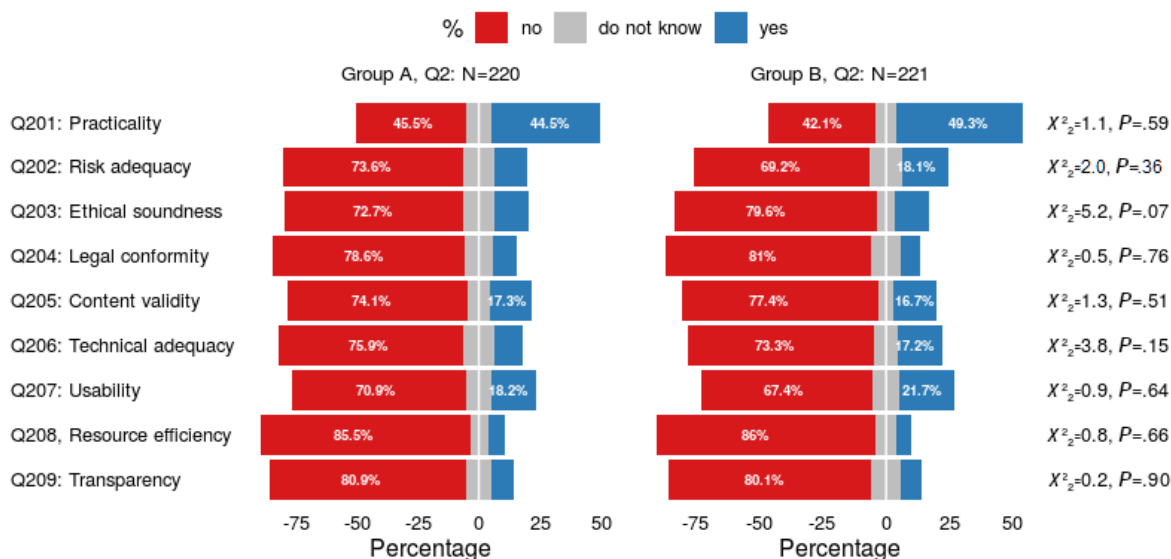
### Do the App Descriptions Provide Sufficient Information?

#### Availability of Information That Covers the Nine Quality Principles

With the exception of the practicality principle, based on the questions presented in Table 3, the participants were only rarely able to find sufficient information covering the nine quality

principles (see Figure 3). For practicality, there was almost an equilibrium between positive versus negative ratings in both groups; for all other aspects, negative ratings in both groups prevailed, with rates ranging from 70% to more than 80%. As shown by the results of the chi-square test shown in Figure 3, there were only negligible, statistically insignificant differences between test group A and validation group B (see Multimedia Appendix 2 for a more detailed listing of the results for Q2).

**Figure 3.** Graphical representation of the participants’ assessments of whether the provided app descriptions contain sufficient information with respect to the nine quality principles, stratified by the randomly assigned groups: group A (test) and group B (validation). Results for the corresponding chi-square tests are also shown. Q2: questionnaire 2.

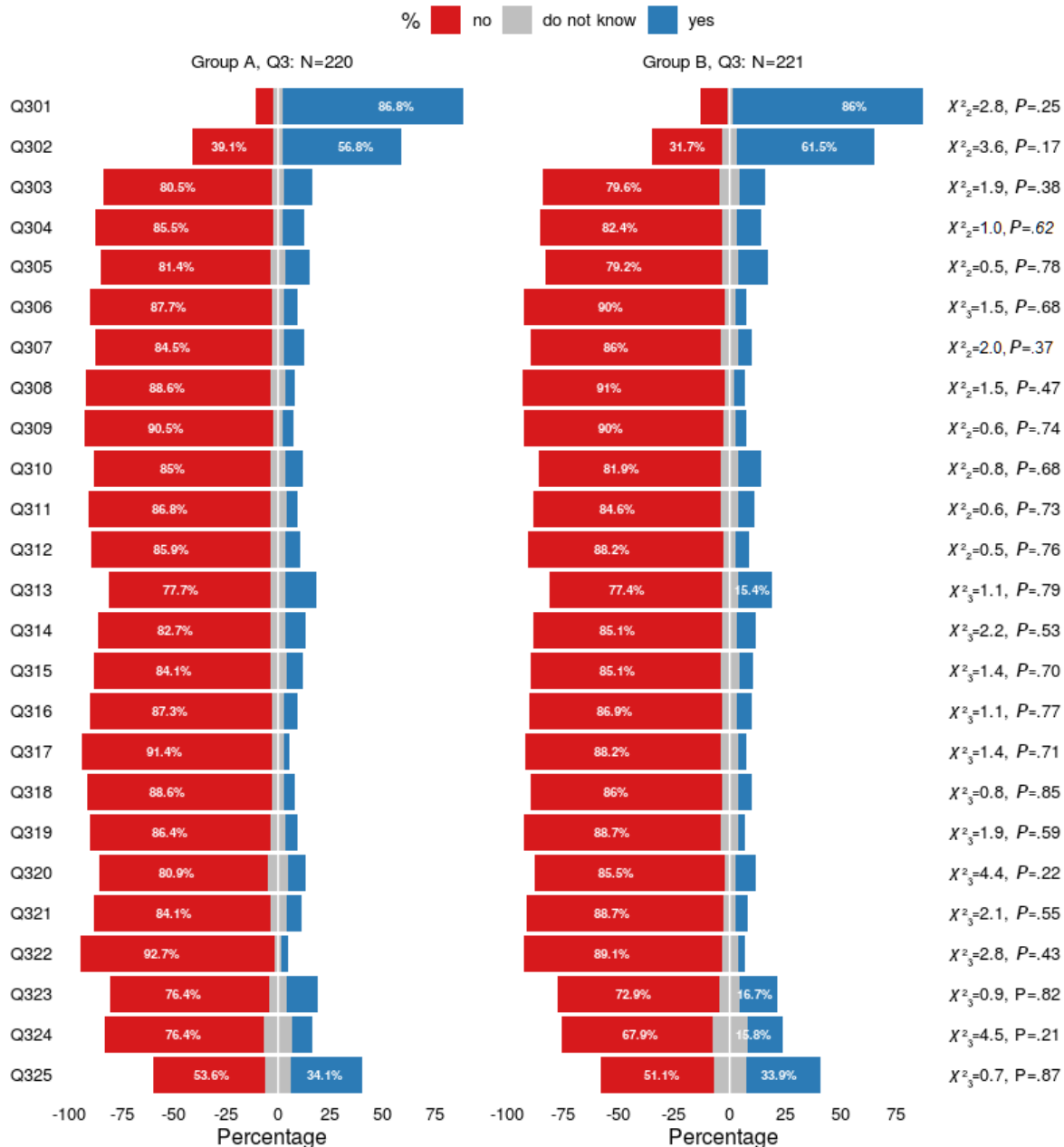


**Assessment of the Quality Principles Using 25 Key Questions**

Figure 4 is a graphical representation of the participants' assessments of whether the provided app descriptions contain sufficient information with respect to the nine quality principles.

Assessments were made using 25 key questions (Q301-Q325) and were stratified by the two randomly assigned groups: group A and group B. The figure also shows results of the chi-square tests. Detailed results for Q3 are listed in Multimedia Appendix 3.

**Figure 4.** Graphical representation of the participants' assessments of whether the provided app descriptions contain sufficient information with respect to the nine quality principles, this time assessed using 25 key questions (Q301-Q325) and stratified by the two randomly assigned groups: group A and group B. Results of the chi-square tests are also provided. Q3: questionnaire 3.



**Does the Confrontation With the Quality Principles Influence Participants' App Usage Decisions?**

**Differences Between the Initial Usage Decisions (Assessment 1) and the Intermediate Assessments (Assessment 2)**

As shown in Table 8, after having worked with the nine quality principles in steps involving Q1 and Q2 (see Figure 1, as well as Tables 2 and 3 for reference), based on the descriptions, the

participants perceived the apps they were confronted with in a more critical manner. For group A, assessments from 63 out of 220 participants (28.6%) changed in a negative direction, both in A1 (ie, from do not know to insufficient) and in A2 (ie, from sufficient to either do not know or insufficient). For group B, the amount of assessment change in a negative direction was similar (62/221, 28.1%). For both groups, based on the results of the McNemar-Bowker test (see Table 9), these changes were statistically significant ( $P<.001$  for both groups).

**Table 8.** Presentation of the contingency table (A1<sup>a</sup> vs A2<sup>b</sup>) before and after the clarification of quality principles and the targeted search for these quality principles.

A1 (before)	A2 (after): Group A (N=220), n (%)				A2 (after): Group B (N=221), n (%)			
	Insufficient	Do not know	Sufficient	Total	Insufficient	Do not know	Sufficient	Total
Insufficient	97 (44.1)	2 (0.9)	5 (2.3)	104 (47.3)	91 (41.2)	2 (0.9)	2 (0.9)	95 (43.0)
Do not know	10 (4.5)	4 (1.8)	1 (0.5)	15 (6.8)	10 (4.5)	3 (1.4)	1 (0.5)	14 (6.3)
Sufficient	48 (21.8)	5 (2.3)	48 (21.8)	101 (45.9)	42 (19.0)	10 (4.5)	60 (27.1)	112 (50.7)
Total	155 (70.5)	11 (5.0)	54 (24.5)	220 (100)	143 (64.7)	15 (6.8)	63 (28.5)	221 (100)

<sup>a</sup>A1: assessment 1.<sup>b</sup>A2: assessment 2.**Table 9.** McNemar chi-square, Cohen g, and odds ratio (OR): assessment 1 vs assessment 2.

Group	Statistic			Cohen g <sup>b</sup> (df=3)		
	McNemar $\chi^2_3$			OR <sup>b</sup>		
	DPRS <sup>a</sup>	Test statistic	P value	OR <sup>b</sup>	P value	g
Group A (N=220)	0.195	42.89	<.001	7.9	.89	0.39
Group B (N=221)	0.222	49.06	<.001	12.4	.93	0.42

<sup>a</sup>DPRS: discordant proportion ratio sum.<sup>b</sup>Calculated with the R package rcompanion.

### ***Differences Between the Intermediate Usage Decisions (Assessment 2) and the Final Assessments (Assessment 3)***

The subsequent application of the quality principles based on the more elaborate 25 key questions (see [Table 4](#) for reference) had considerably less influence on the final usage decision, as

depicted by the changes from decisions given in A2 to those given in A3. Here, an increase in critical ratings was only observed for 25 out of 220 participants (11.4%) from group A, and for 20 out of 221 participants (9.0%) from group B. However, these changes were statistically insignificant (see [Tables 10](#) and [11](#)).

**Table 10.** Presentation of the contingency table (A2<sup>a</sup> vs A3<sup>b</sup>) before and after the clarification of quality principles and the targeted search for these quality principles.

A2 (before)	A3 (after): Group A (N=220), n (%)				A3 (after): Group B (N=221), n (%)			
	Insufficient	Do not know	Sufficient	Total	Insufficient	Do not know	Sufficient	Total
Insufficient	145 (65.9)	3 (1.4)	7 (3.2)	155 (70.5)	136 (61.5)	1 (0.5)	6 (2.7)	143 (64.7)
Do not know	9 (4.1)	1 (0.5)	1 (0.5)	11 (5.0)	2 (0.9)	6 (2.7)	7 (3.2)	15 (6.8)
Sufficient	14 (6.4)	2 (0.9)	38 (17.3)	54 (24.5)	16 (7.2)	2 (0.9)	45 (20.4)	63 (28.5)
Total	168 (76.4)	6 (2.7)	46 (20.9)	220 (100)	154 (69.7)	9 (4.1)	58 (26.2)	221 (100)

<sup>a</sup>A2: assessment 2.<sup>b</sup>A3: assessment 3.

**Table 11.** McNemar chi-square, Cohen g, and odds ratio (OR): assessment 2 vs assessment 3.

Group	Statistic			Cohen g <sup>b</sup> (df=3)		
	McNemar $\chi^2_3$	Test statistic	P value	OR <sup>b</sup>	P value	g
Group A (N=220)	0.026	5.67	.13	2.3	.69	0.19
Group B (N=221)	0.035	7.66	.05	2.8	.73	0.23

<sup>a</sup>DPRS: discordant proportion ratio sum.

<sup>b</sup>Calculated with the R package rcompanion.

### Correlations and Regression

We were interested in determining whether there were certain factors exerting an influence on app usage decisions (eg, participants' demographics or the availability of information about the nine quality principles within the app descriptions). We defined A1→A2 as a binary variable, where 1 represented participants with a *negative* change (ie, *sufficient to do not know*, *sufficient to insufficient*, and *do not know to insufficient*) and 0 represented *no* change or *other* changes. As described in the previous two subsections, significant changes for the usage decisions were only noted between the initial (A1) and intermediate assessments (A2). Therefore, we applied the chi-square test for demographics versus change in usage decision (ie, demographics vs A1→A2; see [Table 12](#)) and for availability of information corresponding to the nine quality principles (ie, quality principle vs A1→A2; see [Table 13](#)). Regarding

demographics, with the exception of education in group B ( $P=.04$ ), there was no statistically significant influence (see [Table 12](#)).

For availability of information about the nine quality principles versus the change in usage decision, results differed between groups. For group A, there was a correlation between change in usage decision and the quality principles *practicality* ( $P=.01$ ), *risk adequacy* ( $P<.001$ ), and *ethical soundness* ( $P=.01$ ); however, there was only an insignificant influence of these principles on change in usage decision for group B. In contrast, *content validity* ( $P=.03$ ), *resource efficiency* ( $P=.02$ ), and *transparency* ( $P=.01$ ) correlated in a statistically significant manner with participants' changed usage decisions in group B; however, this was not the case for group A (see [Table 13](#)). Detailed results of the regression analysis can be found in [Multimedia Appendix 4](#).

**Table 12.** Demographic factors versus changes in assessment: initial to intermediate assessment (ie, changes from A1<sup>a</sup> to A2<sup>b</sup> toward a more critical assessment).

Demographic (D0) vs A1→A2	Group A			Group B		
	$\chi^2$	df <sup>c</sup>	P value	$\chi^2$	df <sup>c</sup>	P value
D001: Age	0.5	4	.97	6.7	4	.15
D002: Gender	5.2	2	.08	0.9	2	.63
D003: Years of work	4.3	6	.64	7.2	5	.21
D004: Specialty	26.0	21	.22	10.0	20	.96
D005: Function	2.7	5	.74	6.3	4	.18
D006: Sector	0.5	3	.92	3.6	3	.31
D007: Type of employment	1.2	2	.54	2.8	3	.43
D008: Education	3.1	5	.68	12.0	5	.04
D009: Location	31.0	26	.22	22.0	23	.53
D010: Interest in digital topics	5.8	4	.21	5.2	4	.27
D011: Private app use	1.2	1	.27	<.001	1	>.99
D012: App use at work	0	1	>.99	1.2	1	.28
D013: App recommended	2.6	1	.11	3.0	1	.08
A303 <sup>d</sup> : App known to participants	0.5	2	.77	2.4	2	.30

<sup>a</sup>A1: assessment 1.

<sup>b</sup>A2: assessment 2.

<sup>c</sup>Degrees of freedom may differ between both groups because, for either group, there may be categories with a row sum of zero.

<sup>d</sup>This question was shown on page 7 of assessment 3 (A3) and was not part of the demographics part of the survey on page 8; therefore, it was encoded differently.

**Table 13.** Assessment of whether there was sufficient information matching the nine quality principles versus changes in assessment: initial to intermediate assessment (ie, changes from A1<sup>a</sup> to A2<sup>b</sup> toward a more critical assessment).

Quality principle (as listed in questionnaire 2 [Q2]) vs A1→A2	Group A			Group B		
	$\chi^2$	df	<i>P</i> value	$\chi^2$	df	<i>P</i> value
Q201: Practicality	8.6	2	.01	1.1	2	.59
Q202: Risk adequacy	15.0	2	<.001	3.2	2	.21
Q203: Ethical soundness	9.0	2	.01	3.8	2	.15
Q204: Legal conformity	0.043	2	.98	5.9	2	.05
Q205: Content validity	4.3	2	.11	6.9	2	.03
Q206: Technical adequacy	1.1	2	.57	3.5	2	.17
Q207: Usability	3.8	2	.15	0.9	2	.65
Q208: Resource efficiency	0.1	2	.94	8.4	2	.02
Q209: Transparency	2.4	2	.30	9.7	2	.01

<sup>a</sup>A1: assessment 1.

<sup>b</sup>A2: assessment 2.

## Discussion

### Principal Findings

#### *Confrontation With the Quality Principles Influences Participants' App Usage Decisions*

For groups A and B, there were statistically significant changes in the physicians' assessments toward a more critical appraisal of the sufficiency of app descriptions for their individual app usage decisions after confronting the nine quality principles (see contingency Table 8 listing the participants' decisions for A1 vs A2). Cumulative changes from *do not know* to *insufficient* and changes from *sufficient* to either *do not know* or *insufficient* occurred for 28.6% (63/220) of participants in group A (McNemar-Bowker test [df=3]=42.89, *P*<.001, Cohen *g*=0.39). Cumulative changes from *do not know* to *insufficient* and changes from *sufficient* to either *do not know* or *insufficient* occurred for 28.1% (62/221) of participants from group B (McNemar-Bowker test [df=3]=49.06, *P*<.001, Cohen *g*=0.42). The effect was very strong in both groups. There was a less pronounced, but statistically insignificant, effect regarding changes in opinion toward a more critical appraisal after applying the 25 questions (Q3) to the app descriptions (see contingency Table 11 showing the participants' decisions for A2 vs A3). Opinions changed toward a more critical appraisal for 11.4% (25/220) of participants in group A (McNemar-Bowker test [df=3]=5.67, *P*=.13, Cohen *g*=0.19). Opinions changed toward a more critical appraisal for 9.0% (20/221) of participants in group B (McNemar-Bowker test [df=3]=7.66, *P*=.05, Cohen *g*=0.23). The similarly comprehensible change in assessment in both groups can be attributed to increased awareness about, and confrontation with, the quality principles and criteria, as well as their characteristics and requirements. Working with the nine generic criteria in Q1 and Q2 already led to changes in almost one-third of the cases. In contrast, the more detailed assessments performed in the subsequent step of the study (Q3), based on the 25 questions, only led to changes in assessment in just under 10% of cases

(see Table 10). This suggests that even contemplating one's own prioritization of quality principles has a major effect on decision making. This also suggests that the more elaborate and detailed evaluation based on the 25 filter questions provided in a checklist format might have a supplementary effect. Assuredly, the latter tool is also well suited for the assessment process, but its completion requires considerably more time and effort. Also, it is less flexible than the generic quality principles. The question of which of the two instruments, or whether the combination of both, is more promising will need to be addressed in future studies, especially when it comes to questions of acceptance of the tools presented here and their economy of use.

#### *Quality Principles Are Relevant*

The statements among the two investigated groups differed regarding which quality principles were associated with a conservative appraisal (see Table 13). In group A, the quality principles *practicality* (*P*=.01), *risk adequacy* (*P*<.001), and *ethical soundness* (*P*=.01) correlated with a conservative appraisal. However, in group B, there was a statistically significant association with a conservative appraisal for *content validity* (*P*=.03), *resource efficiency* (*P*=.02), and *transparency* (*P*=.01). With the exception of one characteristic in group B, groups A and B did not differ significantly when considering demographics versus changes in assessment from A1 to A2 (see Table 12). As well, the evaluation of the basic relevance for the usage decision did not differ in any characteristic (see Figure 2 and Multimedia Appendix 1). This suggests that either the app descriptions or the random allocation of highly heterogeneous app descriptions led to this result.

Regarding sufficiency of information for the nine quality principles, we wanted to know which of these principles accounted for the largest proportion of the changes in decisions. Based on the regression analysis, contributions of the nine principles to changes in usage decisions differed between both groups (see Multimedia Appendix 4). However, the validity of the models was relatively weak, with a Nagelkerke's *R*<sup>2</sup> ranging

from .051 to .085. Therefore, follow-up studies need to be carried out to further examine how much each variable may possibly contribute.

### ***Insufficient Information Quality of the App Descriptions***

It can be assumed that the lack of information available in the app descriptions was the most important factor influencing the results of the correlations and regression models. Descriptively, the poor information quality could be mapped to both the nine generic quality principles and the 25 filter questions. Again, groups A and B did not differ significantly in their distribution. Information on the nine quality principles could only be identified for the given principles in about 5%-20% of the cases (see [Multimedia Appendix 2](#)). The exception was for the quality principle *practicality*, for which there was information deemed sufficient for decision making, discernible from the app descriptions in 45.5% and 49.3% of the individual cases in group A and group B, respectively. The more detailed examination, performed by means of the 25 filter questions, confirmed this observation (see [Multimedia Appendix 3](#)). Five attributes were assigned to the quality principle *practicality* (see [Table 4](#)). While information on *specification of purpose* was found in about 86% of the cases (group A: 191/220, 86.8%; group B: 190/221, 86.0%), information on functionality in this context was only available in just 60% of the cases (see [Multimedia Appendix 3](#); group A: 125/220, 56.8%; group B: 136/221, 61.5%). The situation for *methods that are applied* (group A: 30/220, 13.6%; group B: 25/221, 11.3%), the *proofs regarding the assigned purpose* (group A: 22/220, 10.0%; group B: 25/221, 11.3%), and *suitability* (group A: 25/220, 11.4%; group B: 20/221, 9.0%) was considerably worse. Here, information was found for less than 15% of the cases. Apart from a few exceptions, less than 10% of the other filter questions for both groups lead to information on the respective aspects being identified. Exceptions to this were isolated aspects such as *target group* (group A: 75/220, 34.1%; group B: 75/221, 33.9%), *quality of content* (group A: 32/220, 14.5%; group B: 34/221, 15.4%), *scope of information* (group A: 33/220, 15.0%; group B: 37/221, 16.7%), and *validity of information* (group A: 22/220, 10.0%; group B: 35/221, 15.8%). Thus, the presentation of information stands in stark contrast to the assessment of the importance of the quality principles for the decision on use.

### ***Physicians Predominantly Perceive All Quality Principles as Important***

Initially, the physicians considered the nine criteria to be largely *important* or *very important* for their usage decision (see [Multimedia Appendix 1](#)), with *content validity* (group A: 217/220, 98.6%; group B: 215/220, 97.3%), *risk adequacy* (group A: 206/220, 93.6%; group B: 196/221, 88.7%), and *legal conformity* (group A: 203/220, 92.3%; group B: 199/221, 90.0%) receiving the greatest approval. *Resource efficiency* (group A: 142/220, 64.5%; group B: 128/221, 57.9%) ranked last. This contrasts strongly with the information actually found in the app descriptions. It was only for *practicality* (group A: 188/220, 85.5%; group B: 191/221, 86.4%) that the necessary information could be identified in about half of the apps. For all other criteria, the corresponding information was available in less than one-fifth of the provided app descriptions.

### **Comparison With the Previous Study**

The results presented here confirm those of our previous study conducted with a comparable design, although with medical students [9]. A more detailed analysis will be presented in a subsequent article, although the core aspects will be outlined here. In our previous study, it was possible to show that an exploration of quality principles already influences individual usage decisions that are made based on app descriptions. Based on the results presented in this study, however, it could be shown that this influence can largely be attributed to nine generic principles and that working with these nine principles resulted in the greater proportion of changes for the usage decisions. Due to the design of the predecessor study, it was impossible to show this effect previously. The information concerning the identification of specific information from the app descriptions in the context of both studies was comparably sobering. In 83.17% of the cases (3301/3969 ratings for all nine quality principles in Q2; cumulative answers for *no* and *do not know*), the physicians included in this study were unable to find sufficient information for the nine principles, while among the students, this was the case in 80% [9] of the submitted ratings. Descriptively, physicians and students seem to agree with each other in their requirements as to which quality principles are of relevance for the individual decision on use. Each of the nine quality principles was predominantly regarded as *very important* or *important*, with *resource efficiency* still receiving the lowest percentage of approval among both populations (group A: 64.5%; group B: 57.9%; medical students from previous study: 63.4%). One difference was found for the principle of *technical adequacy*, which physicians in both groups predominantly regarded as *very important* or *important*, while students viewed it more cautiously (group A: 83.6%; group B: 84.6%; medical students from previous study: 41.4%). It is possible that these rather technical aspects were too abstract for the target groups in the quality context.

As far as identifying information that matches the generic principles is concerned, there is agreement between physicians and students that if information is available, it is most often found for the quality principle of *practicality*. The physicians, however, were stricter in their assessment than the students from the previous study (group A: 44.5%; group B: 49.3%; medical students from previous study: 71.7%). In a follow-up study, the comparability of student and physician attitudes will be examined in more detail in order to further investigate the validity of the differences and their causes.

### **Limitations**

Due to the study's design, the selection of apps was limited to cardiological and pulmonological apps only. The identical set of apps that was employed in the previous study was used and no changes were made to the selection. Therefore, the limitations concerning the app selection correspond to those of the previous study [9]. The actuality of the apps does not matter, since archived app descriptions were used.

The study population included only members of the DGIM who were willing to participate and is, thus, certainly not representative of the entire medical profession and not even representative of the more than 21,000 members of the

association. However, the sole recruitment method via this society was chosen since control mechanisms, such as targeted but anonymous addressing of physicians, involvement of stakeholders, suitable communication channels, and recruitment instruments of this society, could be used to obtain the required number of cases. However, the authors assume that the examination of quality aspects relating to software is independent of the discipline. This is a conclusion based on our observations and experiences as well as the participation of one of the authors in numerous initiatives of different medical societies, where identical problems are likewise debated.

Despite the large number of invitations that were sent out, only a relatively small number of individuals (N=724) chose to start the survey. Of those, only 441 completed all pages of the survey, with varying numbers of dropouts at different stages (see [Table 5](#)); an additional 15 had not given their consent and stated so on the introductory page. For those who simply stopped answering the survey, there was no way to determine the reasons for not continuing; some of the free-text comments given by those who had persevered may still provide insight into possible reasons for not proceeding. For some participants, it may not have been clear that the different approaches presented for evaluation at different stages might influence their opinion; in their final comments after having already finished the survey, they therefore voiced their discontent about seemingly having to answer similar questions (eg, the usage decisions in A1-A3) multiple times. Others were highly frustrated by some of the presented app descriptions, specifically regarding content and language (ie, spelling and grammar), also stating that they were unable to see the relevance of the randomly assigned app description for their line of work. There was also a statement to the effect that the presented criteria and questions were too extensive and time-consuming. For future work, the latter point may possibly be addressed based on our results: while working with the nine criteria had a definite influence on the usage decision, the effect of the 25 detailed key questions was negligible; for those who want to quickly come to a decision, applying the nine base criteria may therefore suffice. We believe these criteria to be a valuable tool in order to not forget important aspects when weighing the pros and cons of using an app, although there were also some participants who stated that simply installing and testing an app would suffice.

All nine criteria were rated mostly *important* or *very important* by the physicians. Therefore, it is difficult to significantly identify a hierarchy of relevance. This aspect will be addressed in a follow-up study that will apply designs from requirements analysis, as prioritization of aspects with supposedly identical relevance is a known issue of requirements engineering.

Our sample size considerations have only partially been met. Detecting a difference between A1 and A2 was covered with our initial assumptions; however, the comparison of A2 and A3 would have required a larger sample size to detect a difference, a fact that we were not aware of at the planning stage.

The questionnaire design was adopted from the previous study and was revised in a detailed pretest (see subsection *Pretest of the Questionnaires* under Methods). In particular, the comprehensibility of the questions and tasks, as well as usability aspects related to the electronic system, were dealt with in order to create the greatest possible comfort for the participants. However, the 25 detailed questions could not be reworded with regard to their complexity, as they corresponded to the wording of a technical document [20]. The objective here was to avoid any falsification. Nevertheless, while in the pretest, the questions and tasks were all interpreted and paraphrased correctly, it cannot be ruled out that the questions might have been too complex for fast readers.

With the given setup, we could not assess test-retest reliability. It would have been an interesting aspect to determine whether participants would have given the same answer under similar conditions.

## Outlook

In line with the aim of this study, specifically the effect of sensitization to quality aspects of health software—health apps in particular—the following work is being carried out: (1) an evaluation of the existing dataset with respect to questions of acceptance, (2) a detailed comparison of the preceding study with medical students with the question of whether there are differences in attitude and evaluation between medical students and physicians, (3) a study to determine which combination of quality principles, particularly filter criteria, has the greatest acceptance among the target group, and (4) a study that examines the app selections that were used and tries to determine which information should be searched for beyond app descriptions in order to be able to assess the relevance of the descriptions in relation to other information sources.

## Conclusions

The study confirms that sensitizing physicians to the topic of quality principles via questions about attitudes toward established quality principles and their applications results in a more critical evaluation of the sufficiency of app descriptions. Even working with only nine generic criteria was sufficient to bring about the majority of decision changes, while the additional, checklist-like processing of 25 detailed requirements contributed only slightly to the overall share of decision changes. All physicians shared their views on the relevance of the nine quality principles, and the predominant opinion was that these principles are *important* or *very important*. *Content validity*, *risk adequacy*, and *legal conformity* received the greatest approval. However, in the app descriptions themselves, it was rarely possible to identify any specific information matching these principles. At best, for *practicality*, such data were detected in about 50% of the descriptions. In up to 85% of the other eight criteria, no meaningful information could be identified.



## Acknowledgments

The authors would like to thank the board members of DGIM eV for patronage of this project and kind support by their office. We are also grateful to the members of DGIM for their valuable participation. Lastly, the authors would like to express their appreciation to Professor Marie-Luise Dierks and Professor Stefan Engeli, both from Hannover Medical School, for the critical discussions. There was no external funding for the study.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Comparison of how the participants assessed the relevance of the nine quality principles (questionnaire 1 [Q1]) for their own usage decision (group A: N=220; group B: N=221).

[[PDF File \(Adobe PDF File\), 2425 KB - mhealth\\_v7i11e16442\\_app1.pdf](#)]

### Multimedia Appendix 2

Assessment as to whether information for the nine quality principles could be found within the available app descriptions: assessments for group A (N=220) and for group B (N=221).

[[PDF File \(Adobe PDF File\), 1992 KB - mhealth\\_v7i11e16442\\_app2.pdf](#)]

### Multimedia Appendix 3

Assessment of whether there was sufficient information to answer 25 detailed questions on the basis of the available app descriptions: assessments for group A (N=220) and group B (N=221).

[[PDF File \(Adobe PDF File\), 2060 KB - mhealth\\_v7i11e16442\\_app3.pdf](#)]

### Multimedia Appendix 4

Table D-1 (group A): Results of the regression analysis for group A for changes from assessment 1 (A1) to a more critical assessment in assessment 2 (A2) versus sufficiency of the information provided for assessing the nine quality principles; Table D-2 (group B): Results of the regression analysis for group B for changes from A1 to a more critical assessment in A2 versus sufficiency of the information provided for assessing the nine quality principles. Reference category for both tables was “no.” NR<sup>2</sup>: Nagelkerke’s R squared.

[[PDF File \(Adobe PDF File\), 2030 KB - mhealth\\_v7i11e16442\\_app4.pdf](#)]

## References

1. Albrecht UV, Höhn M, von Jan U. Gesundheits-Apps und markt. In: Albrecht UV, editor. Chancen und Risiken von Gesundheits-Apps (CHARISMHA) [Chances and Risks of Mobile Health Apps (CHARISMHA)]. Hannover, Germany: Medizinische Hochschule Hannover; 2016:62-82 URL: <http://www.digibib.tu-bs.de/?docid=60007>
2. Kebede M, Steenbock B, Helmer SM, Sill J, Möllers T, Pischke CR. Identifying evidence-informed physical activity apps: Content analysis. JMIR Mhealth Uhealth 2018 Dec 18;6(12):e10314 [[FREE Full text](#)] [doi: [10.2196/10314](https://doi.org/10.2196/10314)] [Medline: [30563810](https://pubmed.ncbi.nlm.nih.gov/30563810/)]
3. Yu L, Luo X, Qian C, Wang S, Leung HKN. Enhancing the description-to-behavior fidelity in Android apps with privacy policy. IEEE Trans Softw Eng 2018 Sep 1;44(9):834-854. [doi: [10.1109/tse.2017.2730198](https://doi.org/10.1109/tse.2017.2730198)]
4. Lee G, Raghu TS, Park S. Do app descriptions matter? Evidence from mobile app product descriptions. SSRN 2015 Mar 21. [doi: [10.2139/ssrn.2581279](https://doi.org/10.2139/ssrn.2581279)]
5. Han W, Wang W, Zhang X, Peng W, Fang Z. APP betting based on the consistency of description and APK. In: Proceedings of the International Conference on Trusted Systems 2014 (INTRUST 2014). Cham, Switzerland: Springer International Publishing; 2015 Presented at: International Conference on Trusted Systems 2014 (INTRUST 2014); December 16-17, 2014; Beijing, China p. 259-277. [doi: [10.1007/978-3-319-27998-5\\_17](https://doi.org/10.1007/978-3-319-27998-5_17)]
6. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: Lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: [10.1136/eb-2012-100885](https://doi.org/10.1136/eb-2012-100885)] [Medline: [22923708](https://pubmed.ncbi.nlm.nih.gov/22923708/)]
7. Hillebrand U, von Jan U, Albrecht UV. Concepts for quality assurance of health-related apps. Stud Health Technol Inform 2016;226:209-212. [Medline: [27350506](https://pubmed.ncbi.nlm.nih.gov/27350506/)]
8. Albrecht UV, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: Analysis. JMIR Mhealth Uhealth 2018 Apr 25;6(4):e10394 [[FREE Full text](#)] [doi: [10.2196/10394](https://doi.org/10.2196/10394)] [Medline: [29695374](https://pubmed.ncbi.nlm.nih.gov/29695374/)]
9. Albrecht U, Malinka C, Long S, Raupach T, Hasenfuß G, von Jan U. Quality principles of app description texts and their significance in deciding to use health apps as assessed by medical students: Survey study. JMIR Mhealth Uhealth 2019 Feb 27;7(2):e13375 [[FREE Full text](#)] [doi: [10.2196/13375](https://doi.org/10.2196/13375)] [Medline: [30810534](https://pubmed.ncbi.nlm.nih.gov/30810534/)]

10. Scott IA, Scuffham P, Gupta D, Harch TM, Borch J, Richards B. Going digital: A narrative overview of the effects, quality and utility of mobile apps in chronic disease self-management. *Aust Health Rev* 2018 Nov 13;1. [doi: [10.1071/AH18064](https://doi.org/10.1071/AH18064)] [Medline: [30419185](https://pubmed.ncbi.nlm.nih.gov/30419185/)]
11. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: A systematic review of prevalent and innovative methods. *Am J Prev Med* 2016 Dec;51(6):1051-1059. [doi: [10.1016/j.amepre.2016.07.009](https://doi.org/10.1016/j.amepre.2016.07.009)] [Medline: [27659122](https://pubmed.ncbi.nlm.nih.gov/27659122/)]
12. Ministry of Health – Manatū Hauora. 2017. Health applications assessment guidance. URL: <https://www.health.govt.nz/our-work/ehealth/digital-health-sector-architecture-standards-and-governance/health-applications-assessment-guidance> [accessed 2019-11-10]
13. Guiding Principles for Physicians Recommending Mobile Health Applications to Patients. Ottawa, Ontario: Canadian Medical Association; 2015 May 30. URL: <https://policybase.cma.ca/documents/Policypdf/PD15-13.pdf> [accessed 2019-11-11]
14. Digital Assessment Questions V2.2. Leeds, UK: NHS Digital; 2019. URL: <https://digital.nhs.uk/binaries/content/assets/website-assets/services/nhs-apps-library/daq.pdf> [accessed 2019-11-10]
15. Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS. 2018. APPKRI: Kriterien für Gesundheits-Apps. URL: <https://ehealth-services.fokus.fraunhofer.de/BMG-APPS/> [accessed 2019-11-10]
16. Albrecht UV. Transparenz schaffen und Orientierung bieten: Methoden und Werkzeuge als Entscheidungshilfe für die Nutzung von Gesundheits-Apps. Erstellung einer ersten Auslegeordnung zur Entwicklung eines Hilfsmittels für schweizerische Anwender. Bern, Switzerland: eHealth Suisse; 2019. [doi: [10.26068/mhhrpm/20190116-000](https://doi.org/10.26068/mhhrpm/20190116-000)]
17. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. *J Health Commun* 2017 Mar;22(3):274-284. [doi: [10.1080/10810730.2017.1284286](https://doi.org/10.1080/10810730.2017.1284286)] [Medline: [28248626](https://pubmed.ncbi.nlm.nih.gov/28248626/)]
18. Martínez-Pérez B, de la Torre-Díez I, Candelas-Plasencia S, López-Coronado M. Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications. *J Med Syst* 2013 Oct;37(5):9976. [doi: [10.1007/s10916-013-9976-x](https://doi.org/10.1007/s10916-013-9976-x)] [Medline: [24014255](https://pubmed.ncbi.nlm.nih.gov/24014255/)]
19. Tarricone R, Cucciniello M, Armeni P, Petracca F, Desouza KC, Hall LK, et al. Mobile health divide between clinicians and patients in cancer care: Results from a cross-sectional international survey. *JMIR Mhealth Uhealth* 2019 Sep 06;7(9):e13584 [FREE Full text] [doi: [10.2196/13584](https://doi.org/10.2196/13584)] [Medline: [31493318](https://pubmed.ncbi.nlm.nih.gov/31493318/)]
20. Albrecht UV. Einheitlicher Kriterienkatalog zur Selbstdeklaration der Qualität von Gesundheits-Apps. Bern, Switzerland: eHealth Suisse; 2019.
21. SoSci Survey. URL: <http://www.soscsurvey.com> [accessed 2019-08-13]
22. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947 Jun;12(2):153-157. [doi: [10.1007/BF02295996](https://doi.org/10.1007/BF02295996)]
23. Bowker AH. A test for symmetry in contingency tables. *J Am Stat Assoc* 1948 Dec;43(244):572. [doi: [10.2307/2280710](https://doi.org/10.2307/2280710)]
24. PASS Sample Size Software. Kaysville, UT: NCSS; 2018. Tests for multiple correlated proportions (McNemar-Bowker test of symmetry). URL: [https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Tests\\_for\\_Multiple\\_Correlated\\_Proportions-McNemar-Bowker\\_Test\\_of\\_Symmetry.pdf](https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Tests_for_Multiple_Correlated_Proportions-McNemar-Bowker_Test_of_Symmetry.pdf) [accessed 2019-11-10]
25. Chow SC, Shao J, Wang H. Sample Size Calculations in Clinical Research. 2nd edition. Boca Raton, FL: Chapman and Hall/CRC; 2008.
26. Wickham H, François R, Henry L, Müller K. The Comprehensive R Archive Network. 2019 Jul 04. dplyr: A grammar of data manipulation. URL: <https://CRAN.R-project.org/package=dplyr> [accessed 2019-11-10]
27. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer; 2016.
28. Neuwirth E. The Comprehensive R Archive Network. 2014 Dec 07. RColorBrewer: ColorBrewer palettes. URL: <https://CRAN.R-project.org/package=RColorBrewer> [accessed 2019-11-10]
29. Heinzen E, Sinnwell J, Atkinson E, Gunderson T, Dougherty G. The Comprehensive R Archive Network. 2019 Sep 07. arsenal: An arsenal of R functions for large-scale statistical summaries. URL: <https://CRAN.R-project.org/package=arsenal> [accessed 2019-11-10]
30. DeWitt P. The Comprehensive R Archive Network. 2019 Mar 15. qwraps2: Quick Wraps 2. URL: <https://CRAN.R-project.org/package=qwraps2> [accessed 2019-11-10]
31. Harrell FEJ, Dupont C. The Comprehensive R Archive Network. 2019 Nov 07. Hmisc: Harrell Miscellaneous. URL: <https://CRAN.R-project.org/package=Hmisc> [accessed 2019-11-10]
32. Signorell A. The Comprehensive R Archive Network. 2019 Oct 22. DescTools: Tools for descriptive statistics. URL: <https://cran.r-project.org/package=DescTools> [accessed 2019-11-10]
33. Mangiafico S. The Comprehensive R Archive Network. 2019 Sep 29. rcompanion: Functions to support extension education program evaluation. Version 2.3. URL: <https://CRAN.R-project.org/package=rcompanion> [accessed 2019-11-10]
34. The R Project for Statistical Computing. Vienna, Austria: The R Foundation; 2019. URL: <https://www.R-project.org/> [accessed 2019-11-10]

## Abbreviations

**A1:** assessment 1

**A2:** assessment 2  
**A3:** assessment 3  
**D0:** participant demographic  
**DGIM:** Deutsche Gesellschaft für Innere Medizin  
**DPRS:** discordant proportion ratio sum  
**OR:** odds ratio  
**PASS:** Power Analysis and Sample Size  
**Q1:** questionnaire 1  
**Q2:** questionnaire 2  
**Q3:** questionnaire 3

*Edited by G Eysenbach; submitted 30.09.19; peer-reviewed by B Breil, C Reis; comments to author 23.10.19; revised version received 25.10.19; accepted 31.10.19; published 18.11.19.*

*Please cite as:*

Albrecht UV, Framke T, von Jan U

*Quality Awareness and Its Influence on the Evaluation of App Meta-Information by Physicians: Validation Study*

*JMIR Mhealth Uhealth* 2019;7(11):e16442

URL: <http://mhealth.jmir.org/2019/11/e16442/>

doi: [10.2196/16442](https://doi.org/10.2196/16442)

PMID: [31738179](https://pubmed.ncbi.nlm.nih.gov/31738179/)

©Urs-Vito Albrecht, Theodor Framke, Ute von Jan. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 18.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Continuous Digital Monitoring of Walking Speed in Frail Elderly Patients: Noninterventional Validation Study and Longitudinal Clinical Trial

Arne Mueller<sup>1</sup>, PhD; Holger Alfons Hoefling<sup>1</sup>, PhD; Amir Muaremi<sup>1</sup>, PhD; Jens Praestgaard<sup>2</sup>, PhD; Lorcan C Walsh<sup>3</sup>, BEng, PhD; Ola Bunte<sup>1</sup>, MA; Roland Martin Huber<sup>1</sup>, PhD; Julian Fürmetz<sup>4</sup>, Dr med; Alexander Martin Keppler<sup>4</sup>, Dr med; Matthias Schieker<sup>1,4</sup>, Dr med, Prof Dr; Wolfgang Böcker<sup>4</sup>, Dr med, Prof Dr; Ronenn Roubenoff<sup>1</sup>, MD, MHS; Sophie Brachat<sup>1</sup>, PhD; Daniel S Rooks<sup>5</sup>, MD, PhD; Ieuan Clay<sup>1</sup>, PhD

<sup>1</sup>Novartis Institutes for BioMedical Research, Basel, Switzerland

<sup>2</sup>Biostatistics and Pharmacometrics, Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States

<sup>3</sup>Novartis Business Services, Novartis Ireland Ltd, Dublin, Ireland

<sup>4</sup>University Hospital, Ludwigs-Maximilians Universität, Munich, Germany

<sup>5</sup>Novartis Institutes for BioMedical Research, Cambridge, MA, United States

**Corresponding Author:**

Ieuan Clay, PhD

Novartis Institutes for BioMedical Research

Campus St Johan

Basel

Switzerland

Phone: 41 61 324 1111

Email: [ieuan.clay@novartis.com](mailto:ieuan.clay@novartis.com)

## Abstract

**Background:** Digital technologies and advanced analytics have drastically improved our ability to capture and interpret health-relevant data from patients. However, only limited data and results have been published that demonstrate accuracy in target indications, real-world feasibility, or the validity and value of these novel approaches.

**Objective:** This study aimed to establish accuracy, feasibility, and validity of continuous digital monitoring of walking speed in frail, elderly patients with sarcopenia and to create an open source repository of raw, derived, and reference data as a resource for the community.

**Methods:** Data described here were collected as a part of 2 clinical studies: an independent, noninterventional validation study and a phase 2b interventional clinical trial in older adults with sarcopenia. In both studies, participants were monitored by using a waist-worn inertial sensor. The cross-sectional, independent validation study collected data at a single site from 26 naturally slow-walking elderly subjects during a parcours course through the clinic, designed to simulate a real-world environment. In the phase 2b interventional clinical trial, 217 patients with sarcopenia were recruited across 32 sites globally, where patients were monitored over 25 weeks, both during and between visits.

**Results:** We have demonstrated that our approach can capture in-clinic gait speed in frail slow-walking adults with a residual standard error of 0.08 m per second in the independent validation study and 0.08, 0.09, and 0.07 m per second for the 4 m walk test (4mWT), 6-min walk test (6MWT), and 400 m walk test (400mWT) standard gait speed assessments, respectively, in the interventional clinical trial. We demonstrated the feasibility of our approach by capturing 9668 patient-days of real-world data from 192 patients and 32 sites, as part of the interventional clinical trial. We derived inferred contextual information describing the length of a given walking bout and uncovered positive associations between the short 4mWT gait speed assessment and gait speed in bouts between 5 and 20 steps (correlation of 0.23) and longer 6MWT and 400mWT assessments with bouts of 80 to 640 steps (correlations of 0.48 and 0.59, respectively).

**Conclusions:** This study showed, for the first time, accurate capture of real-world gait speed in slow-walking older adults with sarcopenia. We demonstrated the feasibility of long-term digital monitoring of mobility in geriatric populations, establishing that sufficient data can be collected to allow robust monitoring of gait behaviors outside the clinic, even in the absence of feedback or incentives. Using inferred context, we demonstrated the ecological validity of in-clinic gait assessments, describing positive

associations between in-clinic performance and real-world walking behavior. We make all data available as an open source resource for the community, providing a basis for further study of the relationship between standardized physical performance assessment and real-world behavior and independence.

(*JMIR Mhealth Uhealth* 2019;7(11):e15191) doi:[10.2196/15191](https://doi.org/10.2196/15191)

## KEYWORDS

gait; walking speed; mobility limitation; accelerometry; clinical trials; frailty; wearable electronic devices; algorithms; open source data; data collection; dataset

## Introduction

### Background

Gait speed is considered a key prognostic marker of survival [1] and adverse events [2] in older adults and has been shown to decline over time with healthy aging [3]. These observations have primarily been driven by data collected as part of controlled-environment performance tests, and we still understand very little in terms of how performance in the clinic relates to behavior in the real world, although it is clear that these represent distinct, but related, aspects of function as related to mobility [4].

Digital sensor technologies have drastically improved our ability to capture health-relevant data from patients, in particular, data describing real-world behaviors [5]. However, progress on interpretation via advanced analytics is undermined by a lack of algorithms with validated accuracy and performance in disease populations [6], and most clinical trials still focus exclusively on established performance tests.

### Objectives

Age-related muscle loss and weakness (ie, sarcopenia [7]) can result in an accelerated loss of patient mobility, progressive limitations in independence, and reduction in health-related quality of life (HRQoL). Developing therapies with the potential to maintain and improve real-world patient mobility is therefore a critical need. Previous early phase studies have demonstrated the therapeutic potential of increasing muscle mass in patients with sarcopenia [8]. To explore real-world functional consequences associated with changes in ability captured by clinical assessments, we incorporated continuous patient monitoring using a wearable inertial sensor, during and between their planned clinical site visits in the next clinical phase.

The work presented here describes progress on (1) ensuring that we are accurately capturing gait speed in our target population, (2) the feasibility of deployment in a global clinical trial, and (3) comparison with established gait speed performance measures to explore the validity of this novel, digital, continuous monitoring approach.

We hypothesized that this approach would allow us to examine real-world mobility on an individual patient basis and explore the relationship between currently accepted measures of mobility and real-world mobility behaviors, laying a foundation for measuring personalized response to therapy in a way that was not previously possible using in-clinic tests.

## Methods

Data described here were collected from 2 clinical studies: an independent, noninterventional validation study and a phase 2b interventional clinical trial in older adults with sarcopenia.

### Study Design: Independent Validation Study

The independent validation study was performed between May and August 2018 as a cross-sectional design at a single site, with no pharmaceutical treatment, where 26 naturally slow-walking elderly subjects were recruited in 4 cohorts based on their baseline self-selected gait speed over 4 m (4 m walk test, 4mWT): below 0.5 m per second, 0.5 to 0.6 m per second, 0.6 to 0.7 m per second, and 0.7 to 0.8 m per second. Subjects whose walking speed and natural movement were restricted by orthopedic or neurological complications or other relevant medical conditions were not eligible for the study. Data were recorded with a wearable inertial sensor (see section Accelerometry) while the subjects completed a *parcours* course, at least twice, through the clinic that was designed to simulate a real-world environment [9]. The *parcours* course included straight corridors, stairs, a ramp, and a flat outdoor section. Reference walking speeds were simultaneously measured using a novel combination of a standard distance measuring wheel and an inertial sensor, operated by an assistant, allowing correlation analyses between the patient's actual speed and the algorithm-derived gait speed estimation [10]. A standard smartphone video device on the measuring wheel recorded the subject's footfall in slow motion, allowing a more detailed analysis of different sections of the *parcours* course. A summary of the demographic data of the subjects in the validation study is given in Table 1. Raw and derived data as well as annotations and metadata captured during the study are made available as delimited text files (see section *How to Access*).

**Table 1.** Summaries of demographic data for subjects included in the independent validation study.

Demographic data	Values
Total subjects enrolled, n	26
<b>Gender, n</b>	
Male	11
Female	15
<b>Age (years), n (%)</b>	
60-65	2 (8)
66-75	4 (15)
76-89	18 (69)
>89	2 (8)
4 m walk test gait speed at enrolment in meter per second, mean (SD)	0.62 (0.12)

### Study Design: Interventional Clinical Trial

The phase 2b, interventional clinical trial (ClinicalTrials.gov identifier: NCT02333331 [11]) recruited 217 patients across 32 sites globally, based on criteria for their lean muscle mass, age, grip strength, and gait speed [12]. Between January 2015 and December 2018, patients were enrolled for a period of 25 weeks each, with clinical assessments occurring across 8 visits. Accelerometry data were collected from patients during clinical gait assessments at baseline, at weeks 9, 17, and 25, and during intervening periods at home. A summary of the demographic data of the patients and of accelerometry and clinical assessment

data collected in the interventional clinical trial is provided in Table 2.

At each clinical site visit, patients completed several clinical assessments including a 4mWT [13], 6-min walk test (6MWT [14]), and 400 m walk test (400mWT [15]) from which we derived average walking speed. For each of these assessments, accelerometry data were recorded, along with accurate timestamps delimiting when the patient started and finished each assessment.

Data captured during the trial and metadata describing the assessments are made available as delimited text files (see section *How to Access*).

**Table 2.** A summary of demographic data for patients included in the interventional clinical trial.

Demographic data	Values
Total patients enrolled, n	217
Total patients with accelerometry data, n	192
<b>Gender, n</b>	
Male	91
Female	126
Age (years), mean (SD)	79.0 (5.45)
4 m walk test gait speed at enrolment, mean (SD)	0.648 (0.1048)

### Accelerometry

For both studies, participants were monitored using a waist-worn inertial sensor (actibelt RCT2, Trium Analysis Online, Munich [3,6,16,17]), which recorded acceleration in 3 dimensions at a sampling frequency of 100 Hz (12-bit resolution) and a range of 6g. The devices did not require charging or other interaction from the patients and were otherwise self-managed by the patients.

For the independent validation study, subjects wore a single device for the duration of their single-visit assessment.

In the interventional clinical trial, each patient was instructed to wear the device continuously for a minimum period of 5 days before each planned clinical site visit, and in cases of self-reported noncompliance, for a further 5 days following the

visit. Multiple devices, each of the same model, design, and placement, were used over the observation period. Recorded data were stored locally on the device and downloaded periodically following exchange of devices at planned clinical site visits. Resulting recordings were then merged to form a single observation period. In total, 9668 patient-days of data were collected from 192 patients across 32 sites.

All files are made available in the HDF5 format (see section *How to Access*).

### Derived Gait Parameters and Annotations

Furthermore, algorithmically derived, aggregated gait parameters were calculated for periods when the patients or subjects were actively performing clinical assessments and for periods when patients were passively monitored in their real-world home

environment. Individual raw acceleration data files for a given patient were processed using an algorithm similar to that used in a study by Sabatini et al [18], where steps are first detected and parameterized before a Hilbert transform is used to calculate an analytical signal from which gait speed is projected per step using a linear model. In the first step, a short-time Fourier transform is used to extract dominant frequencies from the raw signal: a broad band (0.7-3 Hz) filter pass removes some noise from the signal before it is divided into overlapping windows of approximately 2.5 seconds; and a fast Fourier transform then calculates the frequency domain for each axis. For each window, these results are then combined to determine the dominant frequencies, removing windows where the angle toward gravity or overall activity is not plausible for upright walking or where there is no dominant frequency. This ensures that false-positive (ie, nonwalking) motions are removed. For windows that pass these checks, a Butterworth filter is applied, and a Hilbert transform is used to determine the frequency ( $F$ ), phase ( $\phi$ ), and amplitude ( $A$ ) for each axis (vertical,  $x$ ; lateral,  $y$ ; and longitudinal,  $z$ ).  $A_x$ ,  $A_y$ , and  $A_z$  give good indication of the force involved in a step, independent of the exact time point, whereas  $F$  indicates step frequency and ( $\phi$ ) yields the relative position within a step. Finally, to predict gait speed, a linear model is fit to the parameters  $A_x$ ,  $A_y$ , and  $A_z$  and their interaction terms. By combining information from all 3 axes, this approach allows for much improved step detection in slow, frail walkers, compared with previous methods [9].

To allow fair comparisons over time, periods of continuous walking (*bouts*) were grouped according to length, similar to our previous work [16]. To achieve this, contiguous windows were combined into a single bout, provided that the mean vertical frequency ( $F_x$ ) did not vary by more than an empirical factor of 1.6-fold between windows.

Derived data describing real-world and in-clinic gait parameters are made available as delimited text (daily summaries and in-clinic assessments) or HDF5 (step- and bout-level summaries) files.

### How to Access

Written informed consent was obtained from the patients for collection and use of the data, and a local ethics committee waiver was granted for collection and publication of the datasets. Anonymized, open source derived data and metadata for the independent validation study [19] and interventional clinical trial [20] are made available. Full datasets, including raw

accelerometry data, for both studies are also available for download [21].

### Statement of Ethics

The ethics committee of the Medical Faculty of the Ludwig Maximilian University of Munich (Ref. 17-798) approved the independent validation study. Individual local ethics approval was obtained for all sites in the interventional clinical trial (ClinicalTrials.gov identifier: NCT02333331).

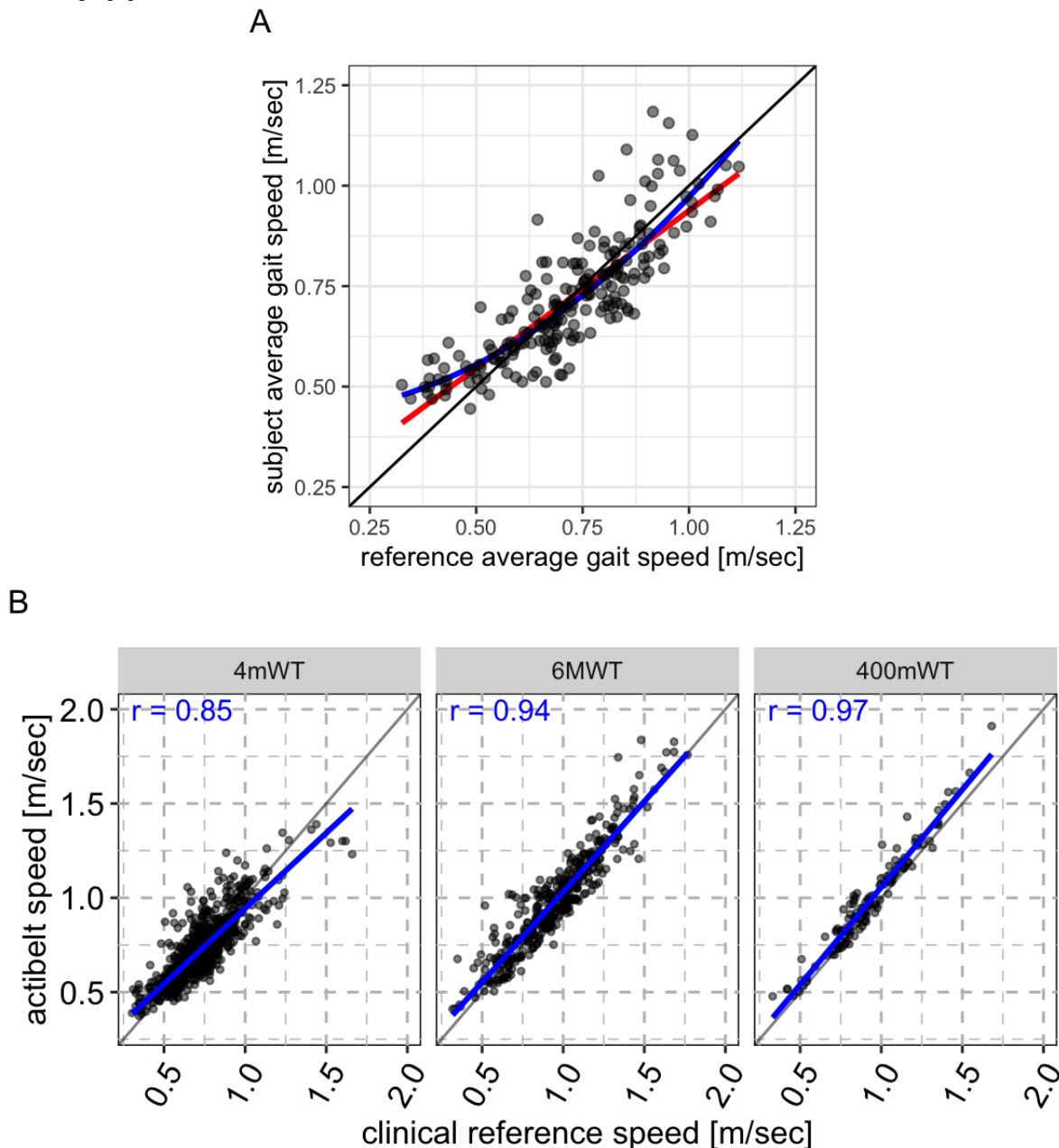
## Results

### Accurate Monitoring of Real-World Gait Speed in Slow-Walking, Older Adults in Controlled Settings

Previous work has shown effective gait speed monitoring in healthy populations by algorithms combining individual detection and parameterization of steps with estimation of gait speed for a given step. Step detection has been achieved with a range of methods including continuous wavelet transform [16,22], whereas gait speed estimation is typically done by using supervised methods such as support vector regression [3]. However, when applied to slow-walking populations, for example, multiple sclerosis (MS), these algorithms produce an observable overestimation of gait speed, particularly in the slowest walkers [6], presumably because of varying relevance of the feature set in patients with pathological gait relative to training sets. The step detector described in our study (see section Methods) is specifically tailored to achieve excellent step detection performance in slow and frail walkers [9]. To validate the accuracy of this new algorithm, in our target population, we performed an independent validation study comparing gait speed continuously captured via an assistant-operated device [10] with gait speed estimated from accelerometry data. The study collected data from 26 frail, elderly adults across a *parcours* course, which simulated a range of real-world situations including indoor corridors, stairs, and a ramp as well as a short outdoor section on uneven ground [9]. A summary of study enrolment is provided in Table 1.

As shown in Figure 1, we observed a strong association ( $r=0.84$ ; residual standard error=0.08 m per second) between the reference and sensor-estimated gait speed across all subjects and walking environments (see also Supplementary Figure 1 in Multimedia Appendix 1). This demonstrates, for the first time, accuracy of gait speed capture by a single system across a wide range of nonsimulated gait speeds (<0.5 to >1 m per second) in a frail population.

**Figure 1.** Accuracy of the algorithm in frail, slow walking adults. (A) Results from the independent validation study “parcours”. Reference gait speed continuously captured using an assistant-operated device is shown on the x axis, and accelerometer-derived patient gait speed is shown on the y-axis. Each datapoint represents the median speed for a given subject and parcours section. Derived gait speed is shown to strongly associate with reference gait speed in this parcours setting, the intercept for the linear fit (red line) is 0.15 and the slope is 0.78 the residual standard error is 0.08 m/sec. For comparison, a cubic fit is included (blue line). (B) Results from sarcopenic adults as captured during scheduled clinical walking test assessments in the interventional clinical trial. Reference gait speed (calculated as the distance traveled by the patient during the assessment divided by the time taken to complete the assessment) is shown on the x-axis, and accelerometer-derived gait speed is shown on the y-axis. Each datapoint represents the average speed for a given patient and assessment. The intercept for the linear fit is 0.15, 0.09, 0.09 from left to right and the slope is 0.79, 0.95 and 0.96, the residual standard error is 0.08, 0.09 and 0.07 m/sec for the 4mWT, 6MWT and 400mWT (panels, left to right), respectively. Note, the 400 meter walk test was collected for relatively few patients. A strong linear association is observed between derived and reference gait speed in all assessments indicating accuracy in our target population of frail, slow walkers.



The same system (hardware and algorithm) was deployed to monitor patients in a global, multisite phase 2b interventional clinical trial involving patients with age-related muscle loss and slow walking speed (see section Methods). To further evaluate the accuracy of the system, we conducted internal validation within the interventional clinical trial dataset by performing a head-to-head comparison of gait speed calculated from accelerometry data and gait speed captured using canonical, clinical standard assessments. The assessments included a

4mWT [13], a 6MWT [14], and a 400mWT [15] from which we derived walking speed. Using timestamps demarcating the specific periods ( $\pm 2$  seconds) in which the assessments were performed, we extracted those intervals from the aligned raw accelerometry data stream and calculated the average gait speed using the algorithm described in the Methods section. Comparing these values with the corresponding clinical reference values (Figure 1), we could confirm that the new algorithm performs with less than or equal to 0.1 m per second



residual standard error in our target population, with a correlation of 0.85, 0.94, and 0.97 for 4mWT, 6MWT, and 400mWT, respectively. The relatively high error seen in the 4mWT is because of the very short nature of the assessment, typically only taking a few seconds to complete, which, in turn, gives a greater weight to any intrinsic human error in recording the start and end of the assessment. The accuracy of our approach can overcome these errors and enable clinicians, which previously relied on the 4mWT, to draw more accurate conclusions.

### Feasibility and Patient Compliance in Long-Term, Real-World Settings

The above results demonstrated the accuracy of both the hardware and algorithm for active, controlled-environment data collection. The monitoring of real-world behavior presents a more complex challenge because of dependence on patient compliance outside of the controlled clinic environment. We therefore set out to evaluate compliance and its effect on our ability to capture real-world walking behavior.

In the interventional clinical trial, patients were requested to wear the device at home for 5 days during the week before every scheduled visit to the clinical site. No feedback was directly given to patients, but if a patient reported that they had not worn the device during the week preceding a visit, they were requested to wear it for the following week. Compliance was monitored through regular data reviews, and feedback provided to clinical sites. We avoided providing any feedback on an individual patient basis so as not to influence their patterns of mobility and confound the overall trial. Ultimately, 9668 patient-days of accelerometry data were collected from 192 patients and 32 sites (see Supplementary Figure 2 in [Multimedia Appendix 1](#) for patterns of observed compliance). A summary of enrolment and data collected is presented in [Table 2](#). This demonstrates the feasibility of continuous monitoring in global clinical trial settings and elderly populations and enabled us to collect an unprecedented volume of data describing real-world mobility. We make the full raw and derived datasets for both the independent validation study and interventional clinical trial, publicly available as part of this publication (see section Methods for instructions on how to access).

We observed that patient compliance patterns were highly variable, with some patients greatly exceeding the requested

wear time and others contributing far less. As previously reported [16], we sought to define a set of minimum thresholds for compliant wear time from which gait behaviors could be stably estimated. This 2-component threshold (hours per day and days around visits) is an attempt to maximize the number of patients included in the analysis and minimize variation because of the sampling of an unrepresentative, too short period of a patient's daily life.

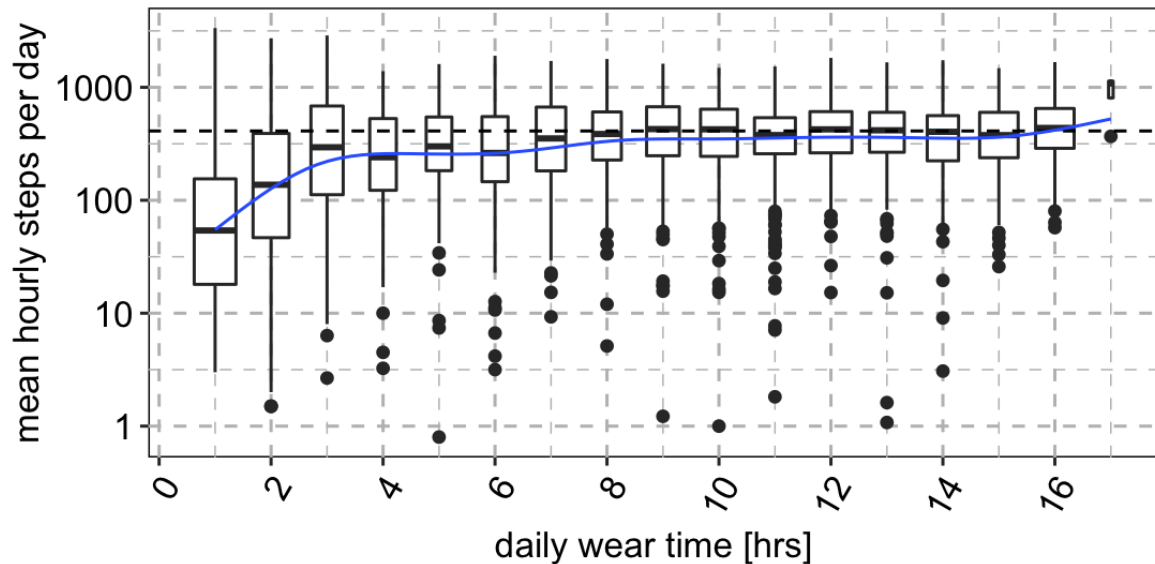
We extracted, for each patient, a 20-day period straddling each clinical site visit excluding the visit day itself. [Figure 2](#) shows the total amount of daily wear time and daily total step count normalized to the total daily wear time. We expect the wear time-normalized daily step count to distribute around the grand mean (410 steps per hour, horizontal line); however, at less than 3 hours of wear time per day, the normalized step count decreases strongly, suggesting that, in our patient population, at least 3 hours of wear time on a given day is required for stable and representative sampling of walking behavior. Additional hours of wear time do not affect the normalized step count but reduce between-day step count variability as indicated by data comparing 3 versus 9 hours of daily wear time ([Figure 2](#)).

The patients in our study were asked to wear their belt with the accelerometer for at least five days around each scheduled clinical site visit, and this is reflected in the distribution of wear time in days around visits ([Supplementary Figure 3](#) in [Multimedia Appendix 1](#)). However, some patients wore the belt for longer periods, and others for shorter periods. [Figure 2](#) shows the mean daily step count when sampling for different number of days around visits. The normalized step count estimate starts to stabilize at 3 days in our population. Where a patient is only compliant for 1 or 2 days in the week of their visit, the variation is relatively high compared with longer sampling periods.

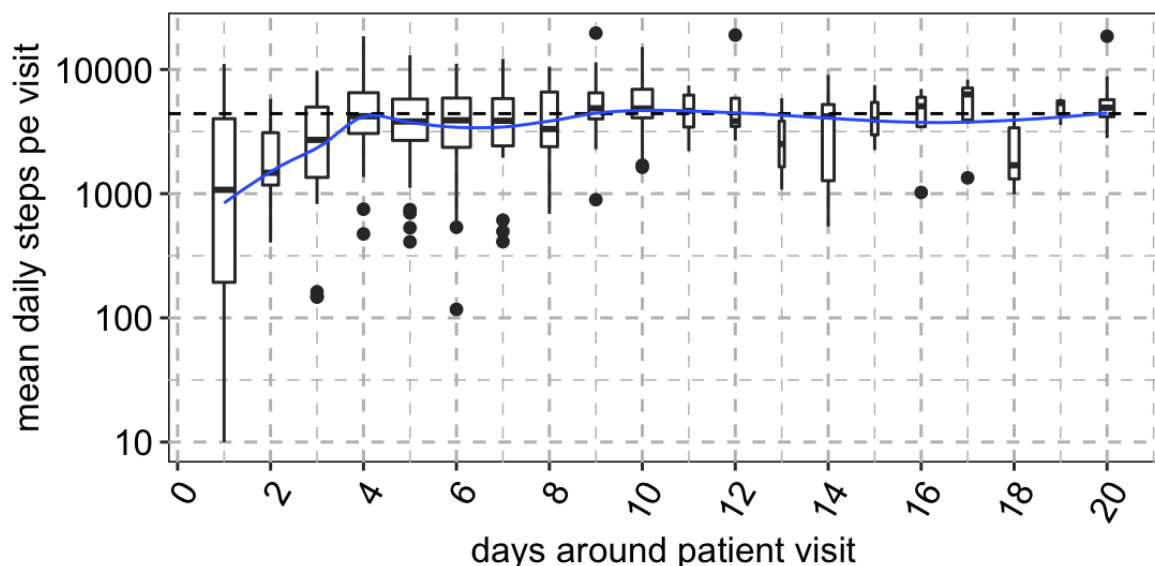
Combining these 2 observations, we chose a minimum criterion of at least three hours of wear time per day for at least three days per visit epoch for our real-world gait analysis. This enabled us to avoid biased estimates but also to minimize exclusion of patient visits and placed a very low burden on the patients. After applying these criteria, 398 visits from 160 patients remained from a complete dataset of 594 visits from 192 patients. This subset was used for all subsequent analyses and figures.

**Figure 2.** Effect of daily wear time and compliance around a visit on step count estimation. (A) The mean hourly steps per day is calculated for each day of patient observation as the total step count normalized to the detected wear time for that day. On the y-axis we show the distribution of mean hourly steps per day and patient, grouped by the total daily wear time in hours on the x-axis. The blue line is a smoothed Loess-fit. Mean daily steps are seen to drop sharply where less than 3 hours of wear time are detected. (B) After removing days of observation with less than 3 hours of wear time, we then calculated the distribution of average normalized daily step counts (“mean daily step count per visit”; total step count in a 20 day window straddling each planned visit divided by the number of compliant “days around patient visit” with a minimum 3 hours wear time). The “mean daily step count per visit” is plotted on the y-axis and the “days around patient visit” on the x-axis. The blue line is a smoothed Loess-fit. Mean daily steps per visit is observed to drop sharply where less than 3 compliant days are detected around a given visit. Combining the results of (A) and (B) we arrived at a two-component threshold of at least 3 days with at least 3 hours of compliance for robust capture of walking behavior in our population.

A



B



### Context Dependence of Real-World Gait Speed

Variability observed in real-world behaviors is heavily influenced by external factors [23]; for example, gait speeds observed in this study may be influenced by changing footwear or weather. Research is now beginning to show how overlaying external contextual information, for example, location [24], onto objectively captured sensor data and patient-reported outcome data can help explain some of the observed variations and enable even more meaningful comparisons. Although the wearable sensor that we deployed does not allow for the direct

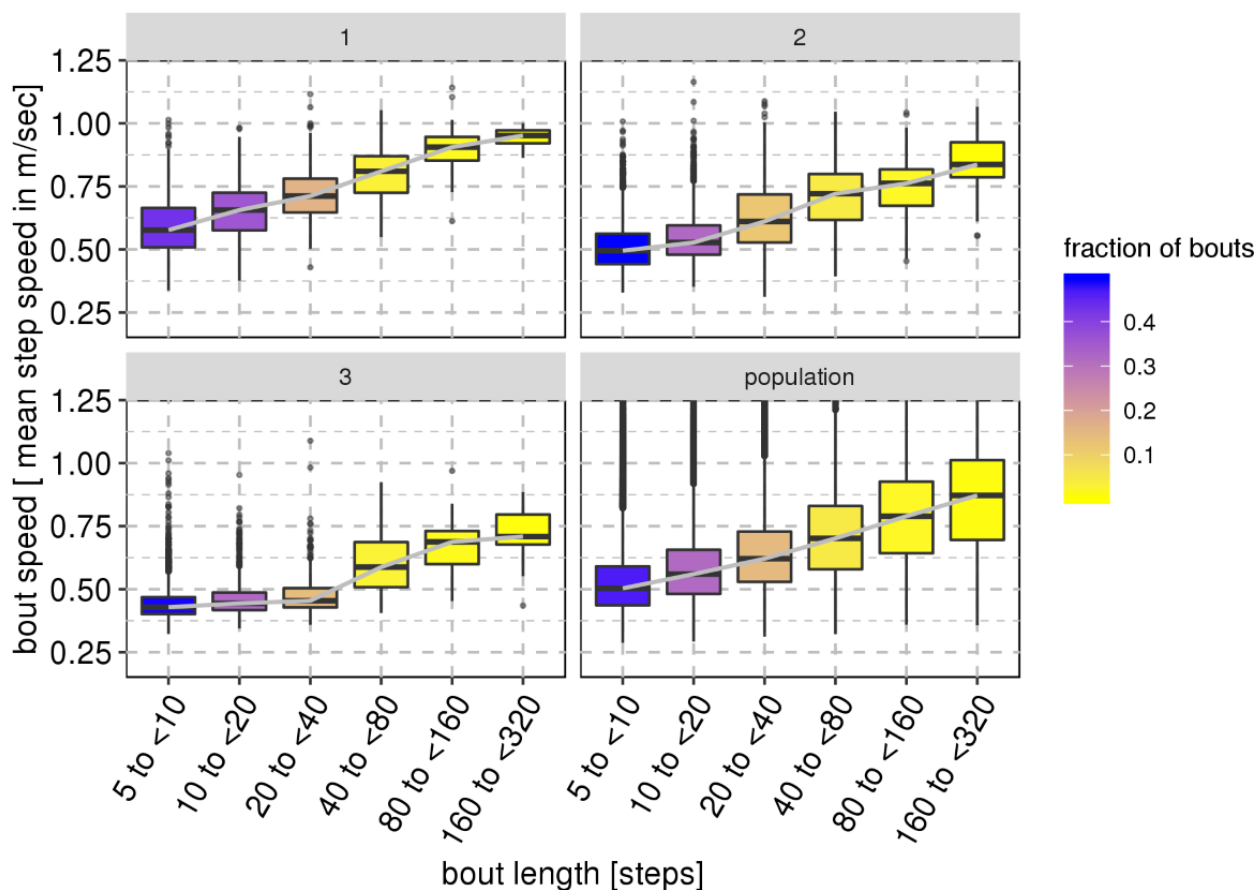
capture of contextual information, it has been shown that by grouping concerted periods of walking (*bouts*) by length, some context can be inferred and enable comparisons of distinct behaviors over time and between individuals [16,25,26].

We observed, on both the population and individual patient level, highly variable and skewed distributions in both bout length (Supplementary Figure 4 in [Multimedia Appendix 1](#)) and real-world gait speed (Supplementary Figure 5 in [Multimedia Appendix 1](#)).

The majority of observed real-world walking behavior comprised relatively short and slow bouts, although it is debatable whether some of these very short bouts of low acceleration intensity constitute true walking behavior. We

examined the relationship between bout length and real-world gait speed and found that real-world gait speed strongly increases with bout length on both the population and individual patient level (Figure 3).

**Figure 3.** Comparison of gait speed with bout length on a population-level (bottom right panel) and for three representative individual patients (other panels). For all panels, the x-axis shows bout length, divided into groups of increasing numbers of steps, from very short bouts (fewer than 10 steps) to very long bouts (<320 steps), and the y-axis shows the distribution of mean gait speed for each bout. Each boxplot is colored by the fraction of total bouts within that bout length range on a scale between dark blue (boxplots representing a large fraction of bouts) to light yellow (boxplots representing a small fraction of bouts). We observe that gait speed increases with bout length, and the majority of bouts are short in length (i.e. contain few steps).



### Comparison of Gait Speed in Real-World Behavior and In-Clinic Performance

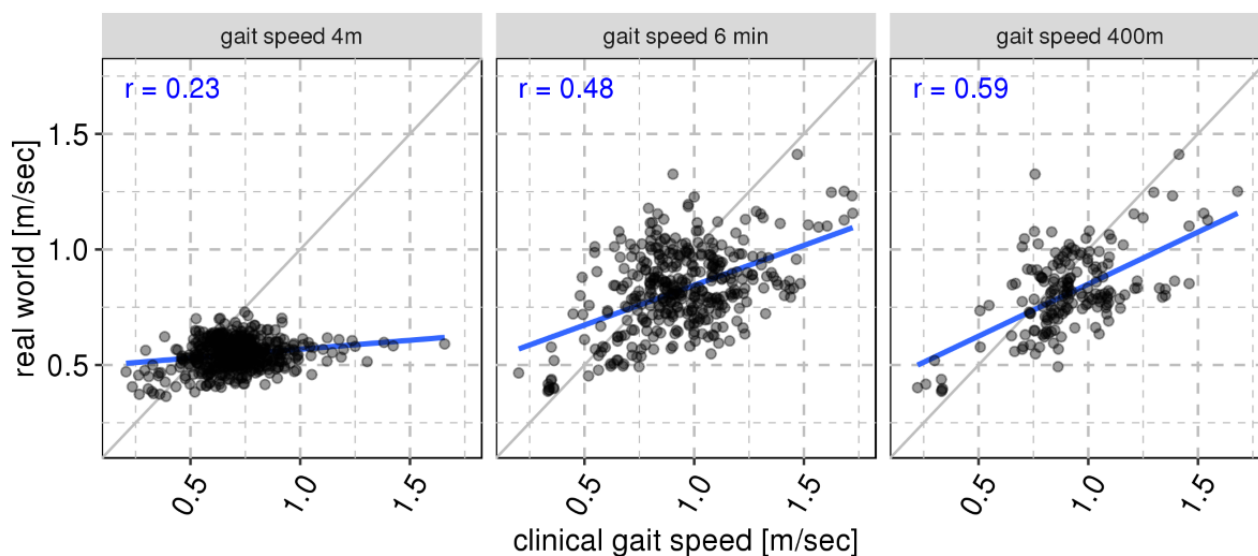
Understanding how real-world walking behavior is influenced by changing performance ability in clinical assessments is a key step in understanding how performance and subjective perceptions of independence are linked. Making comparisons between tests performed in the clinic and real-world behavior has proven more difficult, for example, in MS populations, and it was challenging to find sufficient *real-world* 6MWT events to make a meaningful comparison [4].

We compared real-world gait speed from short (1-20 steps) and long (160-320 steps) bouts with the corresponding short (4mWT) and long (6MWT and 400mWT) clinical gait assessments. Applying our threshold for compliance, each comparison was made between a given clinical assessment and real-world gait from the 20 days surrounding that assessment, excluding the days of the assessment. We found strong linear relationships between real-world and in-clinic gait, in particular, for the longer gait assessments when comparing similar length

bouts (Figure 4). Comparing the short 4mWT gait speed with gait speed in bouts between 5 and 20 steps, we saw a correlation of 0.23, and comparing the longer 6MWT and 400mWT results to bouts of 80 to 640 steps, we saw correlations of 0.48 and 0.59, respectively. Weaker associations were observed for comparison of nonsimilar bout lengths between the in-clinic assessments and real-world gait (Supplementary Figure 6 in Multimedia Appendix 1). Our results indicate that longer gait tests are the most reflective of real-world walking behavior.

We observed that real-world gait speed is consistently lower than what is observed in the corresponding clinic-based assessment. For example, a patient who records a 1 m per second 6MWT will, in their daily life, walk, on average, at 0.75 m per second for bouts of at least 80 steps. We found that this trend is reversed for the slowest walkers (below 0.5 m per second in clinical gait assessments), where the captured real-world gait speed is slightly higher. Overall, the positive associations between those measurements indicate that a higher mobility capacity is reflective of higher habitual gait speed in the real world.

**Figure 4.** Comparison of in-clinic gait speed performance measures and adjacent real-world gait speed behavior. Only patient visits with at least 3 days and 3 hours per day of wearing in a 20 day window around the visit are included. Gait speed in the 4mWT is compared to real-world gait speed in bouts of length between 5 and <20 steps, and 6MWT and 400mWT gait speeds are compared to real-world gait speed in bouts containing between 80 and <640 steps. Gait speed in the clinical assessment is plotted on the x-axis and real-world gait speed is plotted on the y-axis.



## Discussion

### Principal Findings

We show, for the first time, accurate capture of real-world gait speed in slow-walking older adults with sarcopenia. In an independent validation study, recruiting 26 subjects with a mean gait speed of 0.62 m per second, we demonstrate a mean residual error of 0.08 m per second when estimating gait speed through a *parcours* course. Furthermore, using data captured from patients with sarcopenia with a mean gait speed of 0.648 m per second, as part of clinical gait assessments during an interventional clinical trial, we demonstrate residual standard errors of 0.08, 0.09, and 0.07 m per second for estimating gait speed during the 4mWT, 6MWT, and 400mWT clinical gait speed assessments, respectively.

We demonstrate the feasibility of long-term, real-world monitoring of gait and mobility in geriatric populations with slow-walking speed, capturing 9668 patient-days of accelerometry data from 192 patients and 32 sites. Our results establish that even in the absence of feedback or other incentives, sufficient data can be collected to allow robust monitoring of gait behaviors outside the clinic.

We indirectly infer context, in this case, bout length, to partly explain some of the large variation we see in real-world behaviors. Using this contextual information, we demonstrate the ecological validity of in-clinic gait assessments and explore the relationship between in-clinic performance and real-world behaviors relating to gait. We show that a linear relationship exists between real-world gait speed and clinical assessments, but only between comparable bout lengths.

### Strengths and Limitations

Improving our understanding of how a patient's capacity, that is, what a patient *can* do, relates to their functional behavior, that is, what they *actually* do, is foundational to building a bridge

between physiological changes induced by a given therapy and changes in a patient's independence [27]. A more complete chain of evidence, incorporating physiological, functional, and subjective data, will enable development of interventions that genuinely improve patient HRQoL.

Our 2-component threshold for robust capture of real-world walking behavior, a minimum daily wear time of 3 hours and minimum period of 3 days, is well below our initial expectations and achievable in the vast majority of this population without any reminder system or other motivational tools. Moreover, this low threshold can still provide valuable insights into long-term mobility behaviors. Importantly, although relevant to this study population and protocol, this threshold may not be directly transferable to other settings without further evaluation. Other populations may differ in their wear time behavior and periodicity of activity; for example, in a working population, nonworking waking hours might be vastly different from working hours. We believe similar approaches could be modeled on what we present here to define these thresholds in the future populations and settings.

A recent study also compared distributions of real-world gait speed with 4mWT gait speed and found no association [28]. We find these results to be consistent with our own, as we see no relationship to overall distributions, but only find positive associations when comparing similar bouts between clinical assessments and real-world gait. In addition, we demonstrate that real-world gait was consistently slower than what was seen in the clinic, an observation also noted by Van Ancum et al [28], and reflecting other data showing that a patient's gait speed can be increased simply by the knowledge that they are being observed [29].

### Outlook and Conclusions

Future work will focus on the clinical relevance and value of this novel continuous monitoring approach by examining changes over time and response to therapeutic intervention. We

will also continue exploration of how context can enable interpretation of real-world behavioral data, for example, overlaying weather data from a patient's locality to model seasonal changes in mobility [30]. These efforts may further explain variation in the real-world data and enable more sensitive longitudinal comparisons. Further stratification of bouts may be possible by combining information on location from anonymized global positioning system data [24] or gyroscope data to distinguish linear and nonlinear bouts.

Direct, accurate measurement of performance capacity and behavior specifically relating to physical activity is still an emerging field but holds the promise of a better understanding of how mobility, independence, and HRQoL are interrelated on an individual patient level. The work presented here is based

on a relatively simple accelerometry-based system and still provides many advantages alongside traditional approaches for assessing mobility, yet capturing broader, multimodal data covering domains such as social interaction [31], stress [32], or vital signs could potentially increase our understanding of physical effort for a specific activity [33]. Directly capturing context, behavior, and activities of daily living using ambient [34] and smartphone technologies [24] will help us to better relate objective measurements to events and transitions in a patient's life. Ultimately, we aim to build on the progress presented here in terms of establishing accuracy, feasibility, and leveraging context to make meaningful comparisons to further build the link between performance, behavior, subjective perceptions of health, and clinical outcomes [35] and to predict long-term changes in health.

---

## Acknowledgments

The authors would firstly like to thank the subjects, patients and clinical sites for their contributions to this research, and acknowledge the support, critical review, and assistance of Jörg Goldhahn, Sarah Hemsley, Therese Swan, Nicholas Panchaud, Budhaditya Goswami, Sam Hariry, Jason Laramie, Scott Kennedy, Michaela Kneissel, and Evan Beckman. They would also like to thank their collaborators at Trium Analysis Online GmbH, especially Timur Nuritdinow, Astrid Sitte, Bettina Greese, Gerhard Aigner, and Martin Daumer, for resources and software provided and for their teamwork in the course of this work. Both the independent validation study and interventional clinical trial were funded by Novartis, with the former cosponsored by the Ludwig Maximilian University of Munich, Trium Analysis Online GmbH and Novartis, and the latter sponsored by Novartis.

---

## Authors' Contributions

RR, SB, DR, and IC were responsible for conceptualization of the work presented here. A Mueller, LW, OB, RMH, JF, AK, MS, WB, SB, DR, and IC contributed to design of the independent validation study. LW, OB, RMH, JF, AK, MS, and WB conducted the independent validation study. A Mueller, HH, A Muaremi, and IC carried out analysis of the independent validation study. RR and DR designed the interventional clinical trial. A Mueller, HH, JP, DR, and IC were involved in investigation and analysis of the interventional clinical trial. A Mueller, HH, A Muaremi, JP, LW, OB, RMH, JF, AK, RR, SB, DR, and IC drafted and revised the manuscript.

---

## Conflicts of Interest

A Mueller, HAH, A Muaremi, JTP, LCW, OB, RMH, MS, RR, SB, DSR, IC are employees of, and may hold stock in, Novartis. JF, AMK, and WB are employees of Ludwig-Maximilians University, Munich. MS also holds positions at Ludwig-Maximilians University Munich. Novartis is a partner in the Mobilise-D European Innovative Medicines Initiative 2 Joint Undertaking which aims to develop, validate, and ensure regulation of better mobility outcomes.

---

## Multimedia Appendix 1

Supplementary figures 1-6.

[[DOCX File, 1231 KB](#) - [mhealth\\_v7i11e15191\\_app1.docx](#) ]

---

## References

1. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. *J Am Med Assoc* 2011 Jan 5;305(1):50-58 [FREE Full text] [doi: [10.1001/jama.2010.1923](https://doi.org/10.1001/jama.2010.1923)] [Medline: [21205966](https://pubmed.ncbi.nlm.nih.gov/21205966/)]
2. Perera S, Patel KV, Rosano C, Rubin SM, Satterfield S, Harris T, et al. Gait speed predicts incident disability: a pooled analysis. *J Gerontol A Biol Sci Med Sci* 2016 Jan;71(1):63-71 [FREE Full text] [doi: [10.1093/gerona/glv126](https://doi.org/10.1093/gerona/glv126)] [Medline: [26297942](https://pubmed.ncbi.nlm.nih.gov/26297942/)]
3. Schimpl M, Moore C, Lederer C, Neuhaus A, Sambrook J, Danesh J, et al. Association between walking speed and age in healthy, free-living individuals using mobile accelerometry--a cross-sectional study. *PLoS One* 2011;6(8):e23299 [FREE Full text] [doi: [10.1371/journal.pone.0023299](https://doi.org/10.1371/journal.pone.0023299)] [Medline: [21853107](https://pubmed.ncbi.nlm.nih.gov/21853107/)]
4. Stellmann JP, Neuhaus A, Götze N, Briken S, Lederer C, Schimpl M, et al. Ecological validity of walking capacity tests in multiple sclerosis. *PLoS One* 2015;10(4):e0123822 [FREE Full text] [doi: [10.1371/journal.pone.0123822](https://doi.org/10.1371/journal.pone.0123822)] [Medline: [25879750](https://pubmed.ncbi.nlm.nih.gov/25879750/)]

5. Clay I. Impact of digital technologies on novel endpoint capture in clinical trials. *Clin Pharmacol Ther* 2017 Dec;102(6):912-913. [doi: [10.1002/cpt.866](https://doi.org/10.1002/cpt.866)] [Medline: [29027665](https://pubmed.ncbi.nlm.nih.gov/29027665/)]
6. Motl RW, Weikert M, Suh Y, Sosnoff JJ, Pula J, Soaz C, et al. Accuracy of the actibelt(®) accelerometer for measuring walking speed in a controlled environment among persons with multiple sclerosis. *Gait Posture* 2012 Feb;35(2):192-196. [doi: [10.1016/j.gaitpost.2011.09.005](https://doi.org/10.1016/j.gaitpost.2011.09.005)] [Medline: [21945386](https://pubmed.ncbi.nlm.nih.gov/21945386/)]
7. Fielding RA, Vellas B, Evans WJ, Bhasin S, Morley JE, Newman AB, et al. Sarcopenia: an undiagnosed condition in older adults. Current consensus definition: prevalence, etiology, and consequences. International working group on sarcopenia. *J Am Med Dir Assoc* 2011 May;12(4):249-256 [FREE Full text] [doi: [10.1016/j.jamda.2011.01.003](https://doi.org/10.1016/j.jamda.2011.01.003)] [Medline: [21527165](https://pubmed.ncbi.nlm.nih.gov/21527165/)]
8. Rooks D, Praetgaard J, Hariry S, Laurent D, Petricoul O, Perry RG, et al. Treatment of sarcopenia with bimagrumab: results from a phase II, randomized, controlled, proof-of-concept study. *J Am Geriatr Soc* 2017 Sep;65(9):1988-1995. [doi: [10.1111/jgs.14927](https://doi.org/10.1111/jgs.14927)] [Medline: [28653345](https://pubmed.ncbi.nlm.nih.gov/28653345/)]
9. Keppler AM, Nuritidinow T, Mueller A, Hoefling H, Schieker M, Clay I, et al. Validity of accelerometry in step detection and gait speed measurement in orthogeriatric patients. *PLoS One* 2019;14(8):e0221732 [FREE Full text] [doi: [10.1371/journal.pone.0221732](https://doi.org/10.1371/journal.pone.0221732)] [Medline: [31469864](https://pubmed.ncbi.nlm.nih.gov/31469864/)]
10. Aigner G, Grimm B, Lederer C, Daumer M. Method to collect ground truth data for walking speed in real-world environments: description and validation. *PeerJ Preprints* 2019;7:e27558v1 [FREE Full text] [doi: [10.7287/peerj.preprints.27558v1](https://doi.org/10.7287/peerj.preprints.27558v1)]
11. Clinical Trials. Dose Range Finding Study of Bimagrumab in Sarcopenia. URL: <https://clinicaltrials.gov/ct2/show/NCT02333331> [accessed 2019-10-01]
12. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, European Working Group on Sarcopenia in Older People. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on Sarcopenia in Older People. *Age Ageing* 2010 Jul;39(4):412-423 [FREE Full text] [doi: [10.1093/ageing/afq034](https://doi.org/10.1093/ageing/afq034)] [Medline: [20392703](https://pubmed.ncbi.nlm.nih.gov/20392703/)]
13. Lauretani F, Ticinesi A, Gionti L, Prati B, Nouvenne A, Tana C, et al. Short-Physical Performance Battery (SPPB) score is associated with falls in older outpatients. *Aging Clin Exp Res* 2019 Oct;31(10):1435-1442. [doi: [10.1007/s40520-018-1082-y](https://doi.org/10.1007/s40520-018-1082-y)] [Medline: [30515724](https://pubmed.ncbi.nlm.nih.gov/30515724/)]
14. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002 Jul 1;166(1):111-117. [doi: [10.1164/ajrccm.166.1.at1102](https://doi.org/10.1164/ajrccm.166.1.at1102)] [Medline: [12091180](https://pubmed.ncbi.nlm.nih.gov/12091180/)]
15. Chen H, Ambrosius WT, Murphy TE, Fielding R, Pahor M, Santanasto A, LIFE Study. Imputation of Gait speed for noncompleters in the 400-meter walk: application to the lifestyle interventions for elders study. *J Am Geriatr Soc* 2017 Dec;65(12):2566-2571 [FREE Full text] [doi: [10.1111/jgs.15078](https://doi.org/10.1111/jgs.15078)] [Medline: [28884789](https://pubmed.ncbi.nlm.nih.gov/28884789/)]
16. Mueller A, Hoefling H, Nuritidinow T, Holway N, Schieker M, Daumer M, et al. Continuous monitoring of patient mobility for 18 months using inertial sensors following traumatic knee injury: a case study. *Digit Biomark* 2018;2(2):79-89. [doi: [10.1159/000490919](https://doi.org/10.1159/000490919)]
17. actibelt. actibelt® - Experts in real-world monitoring of human motion with mobile accelerometry in clinical trials. URL: <http://www.actibelt.com/> [accessed 2019-10-01]
18. Sabatini AM, Ligorio G, Mannini A. Fourier-based integration of quasi-periodic gait accelerations for drift-free displacement estimation using inertial sensors. *Biomed Eng Online* 2015 Nov 23;14:106 [FREE Full text] [doi: [10.1186/s12938-015-0103-8](https://doi.org/10.1186/s12938-015-0103-8)] [Medline: [26597696](https://pubmed.ncbi.nlm.nih.gov/26597696/)]
19. Mueller A, Hoefling H, Praetgaard J, Walsh L, Bunte O, Huber R, et al. Zenodo - Research Shared. Continuous Digital Monitoring of Walking Speed in Frail Elderly Patients: Noninterventional Validation Study and Longitudinal Clinical Trial (Data for independent validation study). URL: <https://doi.org/10.5281/zenodo.2841297> [accessed 2019-10-03]
20. Mueller A, Rooks D, Brachet S, Roubenoff R, Hoefling H, Praetgaard J, et al. Zenodo. Continuous Digital Monitoring of Walking Speed in Frail Elderly Patients: Noninterventional Validation Study and Longitudinal Clinical Trial (Data for interventional clinical trial). URL: <https://doi.org/10.5281/zenodo.2846013> [accessed 2019-10-03]
21. Amazon Web Services. Index of bucket 'mueller-et-al-2019'. URL: <https://mueller-et-al-2019.s3.amazonaws.com/index.html> [accessed 2019-10-03]
22. Godfrey A, Din S, Barry G, Mathers JC, Rochester L. Instrumenting gait with an accelerometer: a system and algorithm examination. *Med Eng Phys* 2015 Apr;37(4):400-407 [FREE Full text] [doi: [10.1016/j.medengphy.2015.02.003](https://doi.org/10.1016/j.medengphy.2015.02.003)] [Medline: [25749552](https://pubmed.ncbi.nlm.nih.gov/25749552/)]
23. Patterson MR, Whelan D, Reginatto B, Caprani N, Walsh L, Smeaton AF, et al. Does external walking environment affect gait patterns? *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:2981-2984. [doi: [10.1109/EMBC.2014.6944249](https://doi.org/10.1109/EMBC.2014.6944249)] [Medline: [25570617](https://pubmed.ncbi.nlm.nih.gov/25570617/)]
24. Reichert M, Tost H, Reinhard I, Zipf A, Salize H, Meyer-Lindenberg A, et al. Within-subject associations between mood dimensions and non-exercise activity: an ambulatory assessment approach using repeated real-time and objective data. *Front Psychol* 2016;7:918 [FREE Full text] [doi: [10.3389/fpsyg.2016.00918](https://doi.org/10.3389/fpsyg.2016.00918)] [Medline: [27445891](https://pubmed.ncbi.nlm.nih.gov/27445891/)]
25. Hootman JM, Theis KA, Barbour KE, Paul P, Carlson SA. Leisure time and transportation walking among adults with and without arthritis in the United States, 2010. *Arthritis Care Res (Hoboken)* 2019 Feb;71(2):178-188. [doi: [10.1002/acr.23790](https://doi.org/10.1002/acr.23790)] [Medline: [30346654](https://pubmed.ncbi.nlm.nih.gov/30346654/)]

26. Hickey A, Del Din S, Rochester L, Godfrey A. Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol Meas* 2017 Jan;38(1):N1-15. [doi: [10.1088/1361-6579/38/1/N1](https://doi.org/10.1088/1361-6579/38/1/N1)] [Medline: [27941238](https://pubmed.ncbi.nlm.nih.gov/27941238/)]
27. Kaye J, Reynolds C, Bowman M, Sharma N, Riley T, Golonka O, et al. Methodology for establishing a community-wide life laboratory for capturing unobtrusive and continuous remote activity and health data. *J Vis Exp* 2018 Jul 27(137) [FREE Full text] [doi: [10.3791/56942](https://doi.org/10.3791/56942)] [Medline: [30102277](https://pubmed.ncbi.nlm.nih.gov/30102277/)]
28. van Ancum JM, van Schooten KS, Jonkman NH, Huijben B, van Lummel RC, Meskers CG, et al. Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults. *Maturitas* 2019 Mar;121:28-34. [doi: [10.1016/j.maturitas.2018.12.008](https://doi.org/10.1016/j.maturitas.2018.12.008)] [Medline: [30704562](https://pubmed.ncbi.nlm.nih.gov/30704562/)]
29. Hutchinson LA, Brown MJ, Deluzio KJ, de Asha AR. Self-Selected walking speed increases when individuals are aware of being recorded. *Gait Posture* 2019 Feb;68:78-80. [doi: [10.1016/j.gaitpost.2018.11.016](https://doi.org/10.1016/j.gaitpost.2018.11.016)] [Medline: [30465945](https://pubmed.ncbi.nlm.nih.gov/30465945/)]
30. Boutou AK, Raste Y, Demeyer H, Troosters T, Polkey MI, Vogiatzis I, et al. Progression of physical inactivity in COPD patients: the effect of time and climate conditions - a multicenter prospective cohort study. *Int J Chron Obstruct Pulmon Dis* 2019;14:1979-1992 [FREE Full text] [doi: [10.2147/COPD.S208826](https://doi.org/10.2147/COPD.S208826)] [Medline: [31564846](https://pubmed.ncbi.nlm.nih.gov/31564846/)]
31. Bilderbeck AC, Penninx BW, Arango C, van der Wee N, Kahn R, Winter-van Rossum I, et al. Overview of the clinical implementation of a study exploring social withdrawal in patients with schizophrenia and Alzheimer's disease. *Neurosci Biobehav Rev* 2019 Feb;97:87-93 [FREE Full text] [doi: [10.1016/j.neubiorev.2018.06.019](https://doi.org/10.1016/j.neubiorev.2018.06.019)] [Medline: [29940238](https://pubmed.ncbi.nlm.nih.gov/29940238/)]
32. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *J Med Internet Res* 2018 Jun 8;20(6):e210 [FREE Full text] [doi: [10.2196/jmir.9410](https://doi.org/10.2196/jmir.9410)] [Medline: [29884610](https://pubmed.ncbi.nlm.nih.gov/29884610/)]
33. Prahm KP, Witting N, Vissing J. Decreased variability of the 6-minute walk test by heart rate correction in patients with neuromuscular disease. *PLoS One* 2014;9(12):e114273 [FREE Full text] [doi: [10.1371/journal.pone.0114273](https://doi.org/10.1371/journal.pone.0114273)] [Medline: [25479403](https://pubmed.ncbi.nlm.nih.gov/25479403/)]
34. Uddin MZ, Khaksar W, Torresen J. Ambient sensors for elderly care and independent living: a survey. *Sensors (Basel)* 2018 Jun 25;18(7) [FREE Full text] [doi: [10.3390/s18072027](https://doi.org/10.3390/s18072027)] [Medline: [29941804](https://pubmed.ncbi.nlm.nih.gov/29941804/)]
35. Bahej I, Clay I, Jaggi M, de Luca V. Prediction of Patient-Reported Physical Activity Scores from Wearable Accelerometer Data: A Feasibility Study. In: 4th International Conference on NeuroRehabilitation. 2018 Presented at: ICNR'18; October 16-20, 2018; Pisa, Italy URL: [https://doi.org/10.1007/978-3-030-01845-0\\_133](https://doi.org/10.1007/978-3-030-01845-0_133) [doi: [10.1007/978-3-030-01845-0\\_133](https://doi.org/10.1007/978-3-030-01845-0_133)]

## Abbreviations

**400mWT:** 400 m walk test

**4mWT:** 4 m walk test

**6MWT:** 6-min walk test

**HRQoL:** health-related quality of life

**MS:** multiple sclerosis

*Edited by G Eysenbach; submitted 26.06.19; peer-reviewed by J Sasaki, R Dorsey; comments to author 31.08.19; revised version received 09.09.19; accepted 24.09.19; published 27.11.19.*

*Please cite as:*

Mueller A, Hoefling HA, Muaremi A, Praestgaard J, Walsh LC, Bunte O, Huber RM, Fürmetz J, Keppler AM, Schieker M, Böcker W, Roubenoff R, Brachat S, Rooks DS, Clay I

*Continuous Digital Monitoring of Walking Speed in Frail Elderly Patients: Noninterventional Validation Study and Longitudinal Clinical Trial*

*JMIR Mhealth Uhealth* 2019;7(11):e15191

URL: <http://mhealth.jmir.org/2019/11/e15191/>

doi:[10.2196/15191](https://doi.org/10.2196/15191)

PMID:[31774406](https://pubmed.ncbi.nlm.nih.gov/31774406/)

©Arne Mueller, Holger Alfons Hoefling, Amir Muaremi, Jens Praestgaard, Lorcan C. Walsh, Ola Bunte, Roland Martin Huber, Julian Fürmetz, Alexander Martin Keppler, Matthias Schieker, Wolfgang Böcker, Ronenn Roubenoff, Sophie Brachat, Daniel S. Rooks, Ieuan Clay. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 27.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Parental Perspectives of a Wearable Activity Tracker for Children Younger Than 13 Years: Acceptability and Usability Study

Kelly A Mackintosh<sup>1</sup>, PhD; Stephanie E Chappel<sup>2</sup>, BExSportSci (Hons); Jo Salmon<sup>2</sup>, PhD; Anna Timperio<sup>2</sup>, PhD; Kylie Ball<sup>2</sup>, PhD; Helen Brown<sup>2</sup>, PhD; Susie Macfarlane<sup>3</sup>, BSc (Hons), Grad Cert Higher Edn; Nicola D Ridgers<sup>2</sup>, PhD

<sup>1</sup>Applied Sports Science, Technology, Exercise and Medicine Research Centre, Swansea University, Swansea, United Kingdom

<sup>2</sup>Institute for Physical Activity and Nutrition, Deakin University, Burwood, Australia

<sup>3</sup>Learning Futures, Deakin University, Burwood, Australia

**Corresponding Author:**

Nicola D Ridgers, PhD

Institute for Physical Activity and Nutrition

Deakin University

221 Burwood Highway

Burwood, VIC 3125

Australia

Phone: 61 3 9244 6718

Email: [nicky.ridgers@deakin.edu.au](mailto:nicky.ridgers@deakin.edu.au)

## Abstract

**Background:** There is increasing availability of, and interest in, wearable activity trackers for children younger than 13 years. However, little is known about how children and parents use these activity trackers or perceive their acceptability.

**Objective:** This study primarily aimed to ascertain parental perspectives on the acceptability and usability of wearables designed to monitor children's physical activity levels. Secondary aims were to (1) identify practical considerations for future use in physical activity interventions and promotion initiatives; (2) determine use of different features and functions incorporated into the accompanying app; and (3) identify parents' awareness of their child's current physical activity levels.

**Methods:** In total, 36 children (18 boys and 18 girls) aged 7-12 years were asked to wear a wrist-worn activity tracker (KidFit) for 4 consecutive weeks and to use the accompanying app with parental assistance and guidance. Each week, one parent from each family (n=25; 21 mothers and 4 fathers) completed a Web-based survey to record their child's activity tracker use, app interaction, and overall experiences. At the end of the 4-week period, a subsample of 10 parents (all mothers) participated in face-to-face interviews exploring perceptions of the acceptability and usability of wearable activity trackers and accompanying apps. Quantitative and qualitative data were analyzed descriptively and thematically, respectively. Thematic data are presented using pen profiles, which were constructed from verbatim transcripts.

**Results:** Parents reported that they and their children typically found the associated app easy to use for activity tracking, though only step or distance information was generally accessed and some difficulties interpreting the data were reported. Children were frustrated with not being able to access real-time feedback, as the features and functions were only available through the app, which was typically accessed by, or in the presence of, parents. Parents identified that children wanted additional functions including a visual display to track and self-monitor activity, access to the app for goal setting, and the option of undertaking challenges against schools or significant others. Other barriers to the use of wearable activity trackers included discomfort of wearing the monitor because of the design and the inability to wear for water- or contact-based sports.

**Conclusions:** Most parents reported that the wearable activity tracker was easy for their child or children to use and a useful tool for tracking their children's daily activity. However, several barriers were identified, which may impact sustained use over time; both the functionality and wearability of the activity tracker should therefore be considered. Overall, wearable activity trackers for children have the potential to be integrated into targeted physical activity promotion initiatives.

(*JMIR Mhealth Uhealth* 2019;7(11):e13858) doi:[10.2196/13858](https://doi.org/10.2196/13858)

**KEYWORDS**

mobile applications; physical activity; child; monitoring, ambulatory; wearable electronic devices



## Introduction

### Background

Physical inactivity is identified as the fourth leading cause of death globally [1], with regular physical activity associated with numerous physiological and psychosocial health benefits in youth. Physical activity during childhood plays a critical role in reducing cardiometabolic risk factors, which are associated with long-term health [2]. Government guidelines across the world (eg, Australia, the United Kingdom, and the United States of America) recommend that children engage in at least 60 min of moderate- to vigorous-intensity physical activity (MVPA) every day [3-5]. Given that most children globally fail to meet these recommended levels [6-8], effective and sustainable interventions to increase children's physical activity levels are needed.

Children's lack of understanding regarding the amount and intensity of physical activity they do across the day is a frequently cited barrier to their overall physical activity levels [9,10]. This is further evidenced by research supporting the overestimation of self-reported data [11,12], suggesting that youth (children and adolescents) are not able to accurately self-monitor. Given that self-monitoring has been identified as a useful behavior change technique for enhancing awareness of, and potentially stimulating change in, physical activity levels [13], wearable activity trackers could play an important role in improving awareness of children's physical activity levels among parents and their children.

Over the past decade, there has been significant growth and interest in commercially available wearable activity trackers (eg, Fitbit, Garmin, and Xiaomi), becoming the number 1 worldwide fitness trend in 2017 [14]. Although research has examined the validity and reliability of such wearable devices for measuring key outcomes such as steps, distance traveled, active minutes, and energy expenditure, their potential impact as tools for promoting physical activity remains poorly understood [15,16]. Nonetheless, wearable activity trackers have shown good validity for measuring physical activity in preschool children [17,18], children [19], and adolescents [20], as well as promising acceptability for adolescents [21], adults [22], and older adults [23]; however, little research has been conducted with children younger than 13 years of age. This may be attributed to restricted data storage and access imposed by manufacturers in accordance with the American Children's Online Data Protection Act. Despite physical activity promotion interventions using wearable activity trackers for a range of populations [24-27], few have targeted youth [28] given that commercially available wearable activity trackers specifically designed for children (eg, KidFit, Fitbit Ace, and Garmin Vivofit Jr 2) have only more recently become available.

Mobile health (mHealth) apps are a promising approach for physical activity promotion, leveraging the ubiquity and pervasive nature of smart devices. When coupled with wearable activity trackers, these self-monitoring systems [29] can provide real-time feedback, as well as other behavior change techniques [30], such as goal setting, peer comparison, social support, and nudges [31]. Masteller et al [32] identified 36 well-established

behavior change techniques across a range of features and functions within 3 commercially available youth-oriented activity trackers and accompanying websites [16,29]. However, Riley et al [33] advocated that the capability of wearable activity trackers to enhance active behaviors relies on the device's ability to engage users to encourage and support positive behavior change in the first place. Owing to data restriction regulations for children younger than 13 years, parents are not only the gatekeepers of children's physical activity but also gatekeepers to mHealth apps. Therefore, parental perceptions are imperative. Understanding the users' experience, in this case parents, is essential to provide recommendations for future programs seeking to use activity trackers with children.

To date, research conducted using wearable activity trackers has typically used multiple devices simultaneously, prescribed app/website access and minimum durations, or used child-based questionnaires in isolation [32,34]. Moreover, the implications associated with the storage of children's data may limit app interaction, and thereby in-depth real-time feedback, an integral component of wearables [16,28], through the requirement to sync via parental/guardian smart devices. Further research is therefore required in free-living settings to understand an individual's natural use and interaction with such wearable activity trackers.

### Objectives

The aim of this study was to ascertain parental perceptions on the acceptability and usability of wearable activity trackers designed to monitor children's physical activity levels (<13 years). Secondary aims were to (1) identify practical considerations for future use in physical activity interventions and promotion initiatives; (2) determine and understand the use of different features and functions incorporated into the accompanying app; and (3) identify parents' awareness of their child's current physical activity levels.

## Methods

### Participants and Settings

Participants were recruited through an email advertisement circulated to staff and students working at or attending an Australian university. Parents with children aged 7-12 years who did not currently own or had not previously used the KidFit (a commercially available activity tracker) were eligible to participate in the study. In total, 37 families expressed interest in the study, with 25 parents (4 fathers and 21 mothers) providing written informed consent for themselves and their children (n=36; 18 boys and 18 girls) to participate in the study (25/37; 68% response rate). In total, 11 families had 2 children participate. Participation involved children wearing the KidFit activity tracker for 4 consecutive weeks, and parents completing a weekly Web-based survey. Following completion of the fourth week, a subsample of parents (n=10) provided consent to participate in a one-on-one interview. As children were not directly involved in data collection, no written assent for participation was required. Ethical approval was provided by the Deakin University Human Ethics Advisory Group (Health).

### **Wearable Activity Tracker**

This study investigated the acceptability and usability of wearable activity trackers specifically designed for children. Acceptability was defined as the perceived usefulness of the device for tracking activity behaviors, although usability was defined as the perceived ease of use of the device [21]. The KidFit (X-Doria International, Santa Monica, CA, USA; AUD \$70 per device) was selected as it was smaller and targeted a broader age range than the only other commercially available wearable activity trackers specifically designed for children younger than 13 years of age at the time of the study. However, it is important to highlight that the aim of the study was to assess the acceptability and usability of wearable activity trackers, rather than device-specific information, because of the rapid technological turnover. Nonetheless, the KidFit ([Multimedia Appendix 1](#)) is a splash-proof, rechargeable activity tracker (3.5 × 2.2 × 1.0 cm<sup>3</sup>) worn on the wrist using a snap band that collects physical activity and sleep data. No feedback is provided to the wearer about their activity levels; this information is only accessible via the accompanying app, which is free to download from the App Store or Google Play Store. Data are manually synced with the app via Bluetooth when pressing the button on the front of the device. Information provided in the app includes the number of steps, distance traveled (based on steps taken), a KidFit score (based on activity levels), and sleep time, which can be tracked over time (days/weeks) and facilitates goal setting. Challenges are also provided in the app (eg, hit your daily goal 3 times in a row). Similar to other wearable activity trackers for children, the KidFit requires charging approximately every 5 days. It is compliant with the American Children's Online Privacy Protection Act.

### **Protocol**

Each child was provided with a KidFit and asked to wear it for 4 consecutive weeks. Data collection took place between September and December 2015 (Spring-Summer). As participants had not previously used the KidFit, a research assistant met with each parent to help them set the device up for their child. This process involved downloading the KidFit app on to the parents' smartphone or tablet, showing them how

their child should wear the activity tracker, and familiarizing them with basic functions of the activity tracker and app. These included how to charge the KidFit, how to sync it with the app, what the different flashing lights on the KidFit meant, and how to navigate the app. A user manual that was developed by the research team was also provided to parents, which included troubleshooting information on these functions. This manual was based on a quick start manual that was provided with the KidFit and additional information available over Web at the time of the study. No instructions were provided on how to customize daily goals or how to interpret the information generated by the app. Parents were asked to sync their children's data each day to ensure no data were lost. Parents were not provided any specific directions regarding app interaction frequency or duration, engagement with app content (eg challenges), or data interpretation, with their children.

### **Measures**

#### **Survey**

Parents were asked to complete a short Web-based survey at the end of each week that their child wore the activity tracker. Where more than 1 child from each family was participating, parents completed 1 survey per child. Items developed for use in this study (see [Table 1](#)) assessed the use of the activity tracker and app in the previous week, daily wear compliance, reasons for nonwear or nonuse, perceptions of the wearable activity tracker and app (eg, ease of use, comfort, awareness of activity, and the ability to understand information provided), enjoyment of use, problems experienced using the device and app, and how information was used by the parents and their children (eg, rewarding goals achieved). All quantitative items were rated on a 4-point Likert scale from 0 (strongly disagree) to 3 (strongly agree). Open-ended items asked about the parents' perspectives (eg, how did you use the information provided?) and their children's perspectives (eg, what does your child like about the wearable activity tracker?). Demographic data (eg, age and sex of parent and child) were collected in the survey completed at the end of week 1. Parents' ownership and use of activity trackers was also assessed in week 1.

**Table 1.** Descriptive survey data on wearable activity tracker use and perceptions among parents (parents were asked to complete 1 survey per child).

Activity tracker use <sup>a</sup>	Week 1	Week 2	Week 3	Week 4
Worn for 7 days, n/N (%)	19/35 (54)	14/35 (40)	19/34 (56)	14/31 (45)
Not worn at all, n/N (%)	1/35 (3)	1/35 (3)	2/34 (6)	4/31 (13)
<b>Acceptability (% agreeing), n/N (%)</b>				
My child likes wearing the activity tracker	30/35 (86)	30/35 (86)	27/34 (79)	28/31 (90)
My child is embarrassed to wear the activity tracker	1/34 (3)	7/35 (20)	5/34 (15)	6/31 (19)
The activity tracker has made me think about how I can help my child to be more active	16/34 (47)	17/35 (49)	11/34 (32)	15/31 (48)
My child is more active than normal <i>because</i> they are wearing the activity tracker	16/34 (46)	14/35 (40)	10/34 (29)	7/31 (23)
<b>Usability (% agreeing), n/N (%)</b>				
The activity tracker is easy to use	30/34 (88)	29/35 (83)	25/34 (74)	24/31 (77)
My child finds the activity tracker uncomfortable to wear	17/34 (50)	20/35 (57)	19/34 (56)	15/31 (48)
I (or my child) experienced problems in the last week using or wearing the activity tracker	27/35 (77)	20/35 (57)	19/34 (56)	10/31 (32)
My child had trouble remembering to put on their activity tracker	9/34 (26)	13/35 (37)	13/34 (38)	13/31 (42)
The activity tracker activity score reflects how active my child is	26/34 (76)	25/35 (71)	23/34 (68)	22/31 (71)
The activity tracker data are easy to understand	27/34 (79)	30/35 (86)	27/34 (79)	25/31 (81)

<sup>a</sup>The N values differ due to some questions not being completed in the survey.

## Interviews

The semistructured interviews explored parents' views about the acceptability and usability of the wearable activity tracker and accompanying app as well as the potential use of commercially available wearable activity trackers as a tool for promoting physical activity in real-world settings. Example questions included:

*What things did you (and/or your child) like about the KidFit?*

*Did you experience any issues using the KidFit?*

*Can you explain what features you/your child used and why?*

*Did the KidFit increase your child's activity and/or your awareness of their activity levels?*

*How could such devices be used within physical activity programs?*

Interviews were conducted by the same research assistant and continued until data saturation was reached. In total, 10 interviews (all mothers) were conducted and digitally recorded. Interviews (mean duration 19.5 [SD 4.4] min) were then transcribed verbatim, producing 81 pages (Times New Roman, size 12) of raw transcription data for analysis.

## Data Analyses

### Survey Data

All quantitative survey data were collapsed to form binary percentage agreement data. Specifically, *strongly agree* and *agree*, and *disagree* and *strongly disagree* categories were merged to those who agreed and disagreed with each of the

statements, respectively. The percentage of parents who agreed was then calculated.

### Qualitative Data

After reading and undergoing familiarization with the transcripts, 2 researchers (KAM and SEC), independent of the project team and each other, thematically analyzed [35] the transcripts and developed an initial list of codes in NVivo 12 (QSR, Southport, the United Kingdom). Consistent with recommended approaches [36], these authors then discussed their independent coding, as well as any discrepancies and deviant cases, until a consensus was reached. This led to a small number of new codes being developed and existing codes being deleted or refined. Once the transcripts had been fully coded, data under specific codes were retrieved and overarching or central themes identified. Pen profiles—an increasingly used technique to present analysis outcomes via diagrams of composite key emergent themes [37]—were then constructed, with example verbatim quotations provided to further illustrate the theme. To provide an indication of the prevalence of the themes, the number of parents who mentioned a specific theme is also presented [37].

A third author (NDR) with expertise in qualitative analyses [38] then analyzed the data in reverse from the pen profiles back to the transcripts, critically questioning the presented thematic analyses [39]. This process allowed authors to offer alternative interpretations and interrogate the data until a consensus was reached, assuring credibility [37]. The adopted inductive analytic approach enabled within- and between-transcript comparisons, which ensured findings stayed grounded in these data. Overall, methodological rigor (ie, dependability, credibility, and transferability) was demonstrated through the comparison of

pen profiles with verbatim transcriptions, verbatim citations, and the bidirectional triangular consensus process, and the likelihood of bias was minimized through the independent coding of transcripts and exploration of deviant cases. To provide a detailed analysis of parental perceptions, both quantitative descriptive data from the surveys and qualitative data from the follow-up interviews are presented.

## Results

### Survey Data

Daily wear compliance and activity tracker perceptions are presented in [Table 1](#). Although the percentage of children wearing the activity trackers for 7 days across the 4-week period fluctuated, there was little variation in the reasons provided for nonwear across the weeks. Reasons included forgetting to wear the activity tracker, needing to charge it or it having a flat battery, or being asked to remove it for sports. Other reasons provided were device malfunction, being ill, losing the device, or staying somewhere other than their primary home. Although 1 child chose not to wear the wearable activity tracker at all beyond the initial set up, another chose not to wear it toward the end of the study because of losing interest as a consequence of the perceived inaccuracy of the device to capture their activity. Parents reported that the reasons some children did not wear the activity tracker at all included the following: *not being into it*; too bulky and restrictive for sports; it malfunctioning; *hating it*; or having lost it. The percentage of parents reporting that their child found the activity tracker embarrassing to wear increased from 2.9% (1/35) in week 1 to 17.1%-20.0% (5/34 to 7/34) across the rest of the weeks, but parents did not report that this affected wear-time.

Although the percentage of parents stating that their child liked wearing the activity tracker decreased during the 4-week monitoring period (see [Table 1](#)), the reasons for their child either liking or disliking the device remained consistent. Parents perceived that their child liked wearing the activity tracker because it facilitated goal setting, was colorful, tracked their physical activity, was *fun* and *cool*, enabled peer comparison and competition, and they liked the snap band. Conversely, those whose children did not like the activity tracker said this was because of the device lacking functions and prohibiting water-based activities, as well as being too bulky, *not cool enough*, uncomfortable/sore, and/or a burden.

Parents generally reported that the activity tracker and associated data were easy to use and understand. There was little change over time in how much parents perceived the trackers made them more aware of, or wanted to increase, their child's physical activity levels. Although parents initially perceived that the activity tracker increased their child's physical activity levels, the percentage progressively decreased from week 1 to week 4. Approximately half of the parents reported that their children found the device uncomfortable over the duration of the study.

Across the 4 weeks, the reporting of problems associated with the activity tracker decreased. The key problems cited included the poor battery life, not being able to sync the data, and perceived poor tracking of their child's physical activity data.

Parents generally accessed the app with their child, though this decreased from 88.6% (31/35) at week 1 to 60.0% (18/30) following week 4. During the first week, most parents accessed the app with their children at least once per day (16/35; 45.7%), though (9/35; 25.7%) accessed it at least twice per day. App access was mainly to track activity (17/35; 48.6%) and sync data (17/35; 48.6%), with very few children and parents reporting accessing sleep data and setting goals. Over the 4 weeks of the study, there was a shift in the way parents accessed activity data or synced the activity tracker. Whilst the majority used the app with their child (22/34; 64.7%) at the start, the majority tended to access data on their own, without their child (15/28; 53.8%), at the end of this period. The only other times parents accessed the app without their child was to try to resolve problems, such as syncing.

### Qualitative Data

Data were themed and are presented in 6 pen profiles ([Figures 1-6](#)), focusing on the positive and negative aspects of the wearable activity tracker, use of information, perceptions, feedback, and future considerations.

#### Wearable Activity Tracker (Positive and Negative)

The activity tracker itself was viewed as both a facilitator ([Figure 1](#)) and barrier ([Figure 2](#)) to use. Parents reported that the key enablers for their child to wear the activity tracker were the design and ease of use, yet barriers included the burden (ie, syncing and charging the device) and wearability, and its design, limitations and malfunctions. Although some reported the snap band grew tighter across the duration of wear, others highlighted that the band was easy to wear and that their child preferred it. The band was more frequently described as uncomfortable (n=7; 70%) than comfortable (n=1; 10%) among this subsample of parents, with the majority of reasons surrounding the nature of the wristband being tight, hot, and sweaty:

*So it slowly would get tighter and tighter, and especially when she's out running around, it would get really hot, and kind of sweaty underneath the band, because they don't breathe either.* [Parent 3]

Furthermore, several parents reported that their children found the activity tracker *bulky* on the wrist. The snap band fastener (ie, not secure), comfort, and lack of device functionality (ie, no digital time display or real-time feedback) were frequently cited as barriers to wearability, as well as to potential long-term use. Several other factors impacted on the use and wearability of the activity tracker, including forgetting to wear it (4/10; 40%) following removal for sport, sleeping or charging, and being unable to wear it while sleeping (5/10; 50%) because of it being painful and waking up with marks. Parents commonly described their child's frustration that the trackers were not waterproof and had to be removed for certain sports, such as swimming, basketball, and ballet. However, 1 parent described how their child moved the snap band to their ankle to allow their activity still to be captured:

*So then she put it on her ankle, and put her sock over the top so that she could still wear it.* [Parent 7]

The activity tracker itself was identified as easy to use by 30% (3/10) of parents and with others mentioning the ease of syncing

(n=1; 10%) and using the app (n=3; 30%). However, several inaccuracies and malfunctions were noted in the sample.

Most parents reported that the device occasionally did not track their child’s activity (n=6; 60%) or reset to a new day resulting in data accumulation and false results (n=5; 50%). There were also issues with syncing (n=5; 50%) and charging (n=2; 20%). These factors collectively led to the children feeling disappointed and demotivated to continue to wear the activity tracker:

*Well then she just didn’t want to wear it, or she was kind of like, oh it’s a silly thing.* [Parent 8]

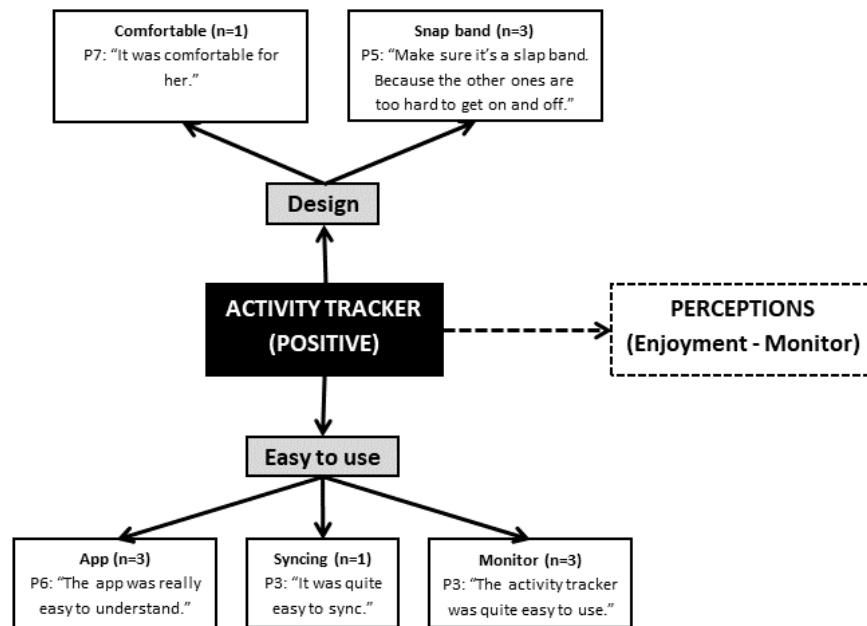
Parents reported that they had difficulties syncing the activity tracker and the app (n=5; 50%), and that it was a burden to do every night (n=4; 40%):

*...with the syncing, it did take me a while to understand first of all how to sync.* [Parent 1]

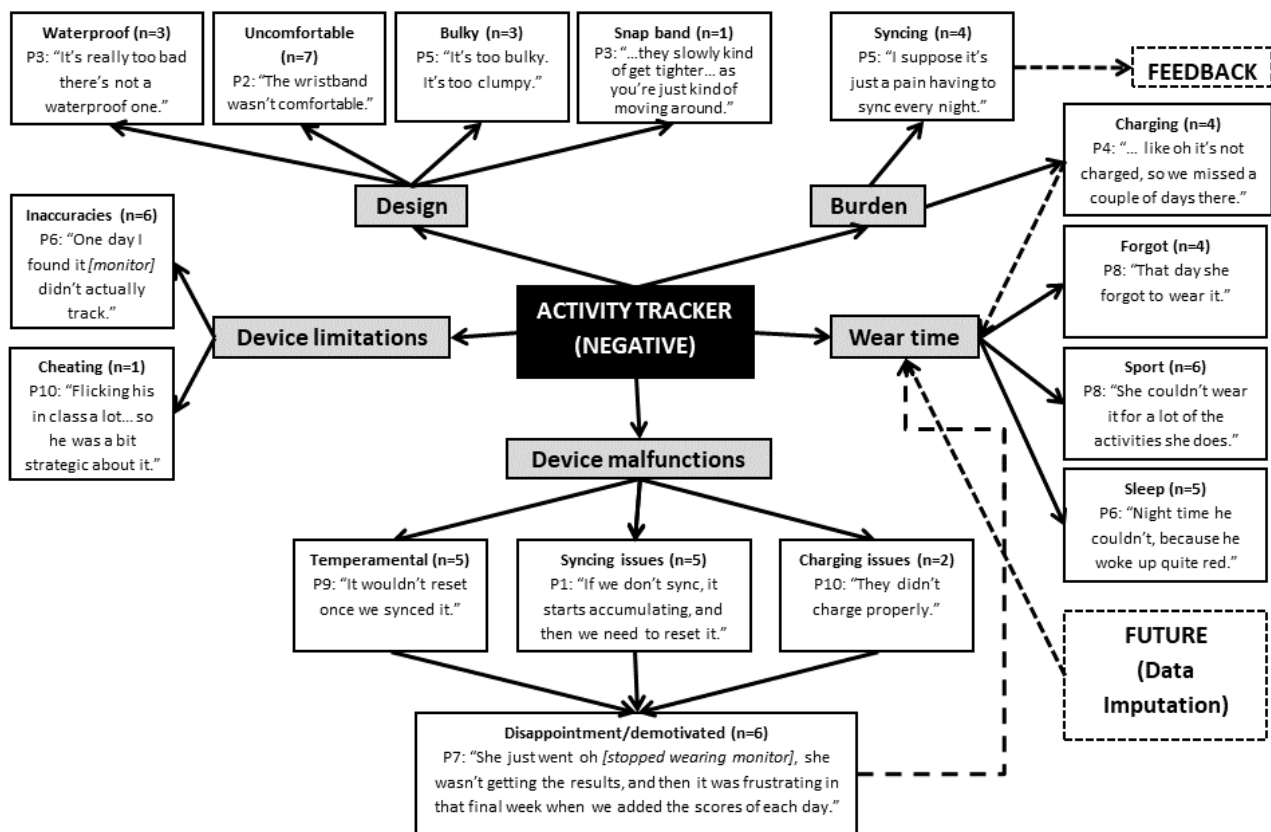
In line with this, parents also found that charging the activity tracker was a burden (n=4; 40%), thereby impacting their child’s ability to wear it:

*I would sync it, and then I’d put it on sleep mode, if I remembered. Or I’d have to take it off to charge the battery.* [Parent 5]

**Figure 1.** Positive perceptions of the wearable activity tracker. The dashed line indicates link made between different themes noted by the researchers from the points discussed, rather than directly mentioned by the parents (P).



**Figure 2.** Negative perceptions of the wearable activity tracker. The dashed line indicates link made between different themes noted by the researchers from the points discussed, rather than directly mentioned by the parents (P).

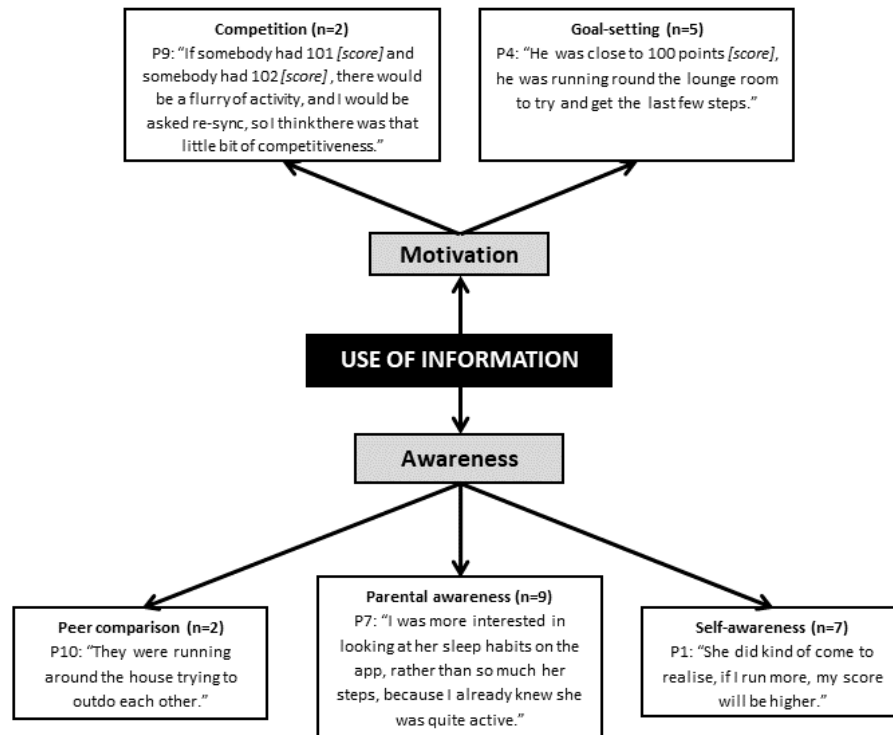


**Use of Information**

Parents reported that children predominantly used the data provided by the app for motivation to be more physically active or aware of their physical activity levels (see Figure 3). There was consensus among parents that the data provided information

on, or verification of, their child’s physical activity levels. Parents (7/10; 70%) mentioned that the information provided within the app also enhanced their child’s awareness of their own physical activity levels, with 2 parents specifically highlighting peer comparison and enhanced motivation, through either competition (n=2; 20%) or goal setting (n=5; 50%).

Figure 3. Use of information collected by the wearable activity tracker. P: parent.



**Parents’ Perceptions**

Most children enjoyed and were enthusiastic about the activity tracker (see Figure 4). The majority of the enjoyment stemmed from the tracker itself being cool to wear. However, 1 parent reported that their children initially felt self-conscious wearing the tracker because of its size and bright color. Most parents found that the activity tracker had a positive effect on their child’s enthusiasm for being active. However, several parents (n=4; 40%) identified that the wearable tracker could result in a lack of enthusiasm:

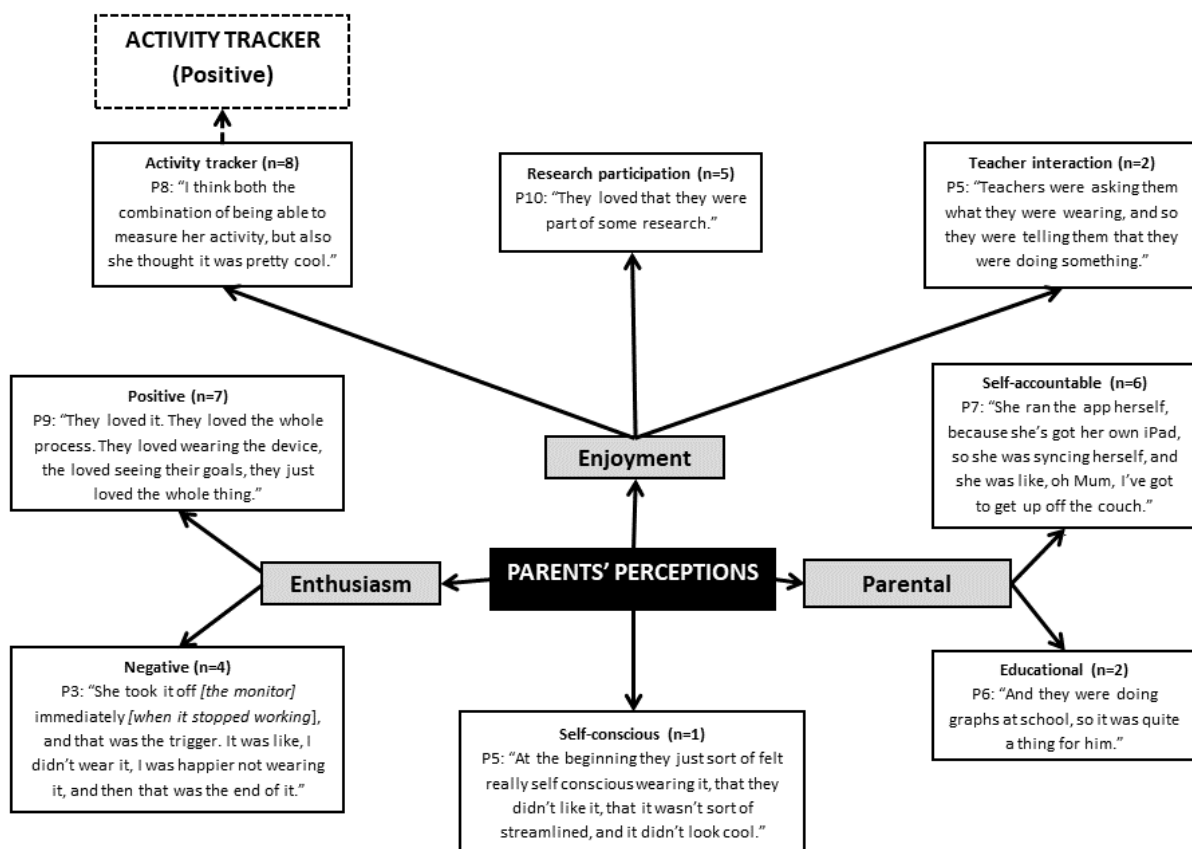
*She was uncomfortable wearing it [the wearable activity tracker]. It wasn’t even providing her with*

*immediate feedback to sort of provide her with intrinsic motivation...[she was like] wow this is annoying, I don’t like it, so I’m turning it off. [Parent 3]*

When asked about their own perceptions, most parents (6/10; 60%) liked that the activity tracker made their children more accountable regarding their own activity and took responsibility for syncing and charging it. Furthermore, 2 parents reported that the use of the tracker and app linked into their children’s current curriculum:

*...they were doing graphs at school, so it was quite a thing for him. [Parent 6]*

**Figure 4.** Parents' (P) perceptions on children's use of the wearable activity tracker. The dashed line indicates link made between different themes noted by the researchers from the points discussed, rather than directly mentioned by the parents.



**Feedback**

The feedback provided through the app was important to the parents and their children (see Figure 5). A common theme was their children's desire to be provided with real-time feedback to gain further information about their physical activity levels (8/10; 80%). For example, some parents reported that their children expressed a desire to have feedback through a visual display on the wearable activity tracker, rather than just the app, to facilitate continuous monitoring. Although 2 parents enjoyed the nudges the activity tracker provided when *they were being lazy, and [to] get off the couch* (Parent 10), several others perceived the lack of feedback did not, and was unlikely to, help increase their child's physical activity levels in the long term. Indeed, the ability to access the app only through the parents' smartphone or tablet was commonly cited as an inhibiting factor to receiving feedback, particularly on weekdays. Conversely, weekend data access was much easier:

*On weekends we might check it 3 or 4 times. Because they'd want to see what they'd scored.* [Parent 10]

However, 1 child had their own smart device, which allowed her to have more control over increasing their activity levels:

*...really she ran the app herself, because she's got her own iPad, so she was syncing herself, and she*

*was like, oh Mum, I've got to get up off the couch.* [Parent 7]

Although some parents really liked and enjoyed the app (4/10; 40%), others found it was acceptable but not of interest for themselves or their children. The most commonly accessed app features included the score, steps, and distance. Parents also used the graphs as a visualization aid to explore how the output from the activity tracker related to their children's activity across the day:

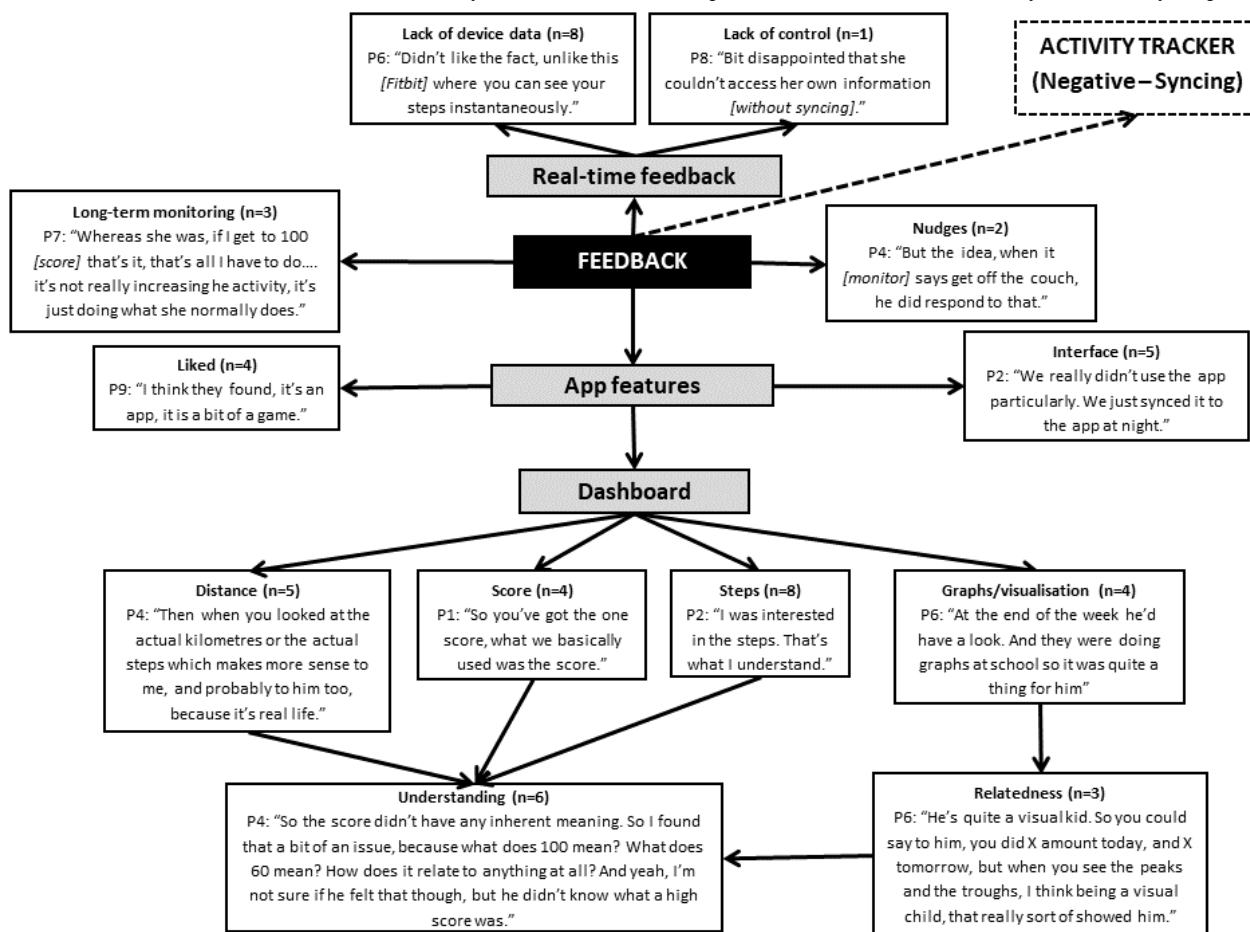
*...obviously it was sitting in the classroom, so you could see, and then when she went out for lunch, so we talked about that's when it [physical activity] spiked up.* [Parent 7]

However, the use of these features differed across all the parents. Some parents demonstrated a lack of understanding regarding various app components, such as the score and its link to total steps and distance covered, the interpretation of the graphs, and, how to meet and set goals. Moreover, some parents did not access the various features available to them. Specifically, 1 parent felt that it was not appropriate to share the app data with their children:

*because it then becomes a competition* [Parent 5]



**Figure 5.** Parents’ perspectives on feedback provided by the wearable activity tracker and how this was used by them and their children. The dashed line indicates link made between different themes noted by the researchers from the points discussed, rather than directly mentioned by the parents (P).



**Future Considerations**

A number of future considerations and suggestions were raised by the parents (see Figure 6). These include the app and tracker features, the use of competition, and whether children younger than 13 years would continue to wear the tracker in the long run. There was a feeling that the app needed to allow children to enter in times when they did physical activity that the wearable could not detect, similar to adult devices. In addition, they felt that some improvements could be made to the goal-setting function to allow more meaningful individual (n=2; 20%) and group (n=1; 10%) goals between family, friends, and teams. Parents conveyed that their child also felt that the activity tracker itself could be enhanced through adding other functions such as heart rate and a time and date display.

One parent suggested the addition of kinetic charging to eliminate the inconvenience of plugging in the activity tracker to recharge. Indeed, parents’ thought that children would capitalize on the opportunity future activity trackers’ could offer, to not only provide real-time feedback on their daily goal but also to challenge themselves or others, either competitively or cooperatively. This was evident from the suggestion of 2 key environments: school and family. From a family perspective,

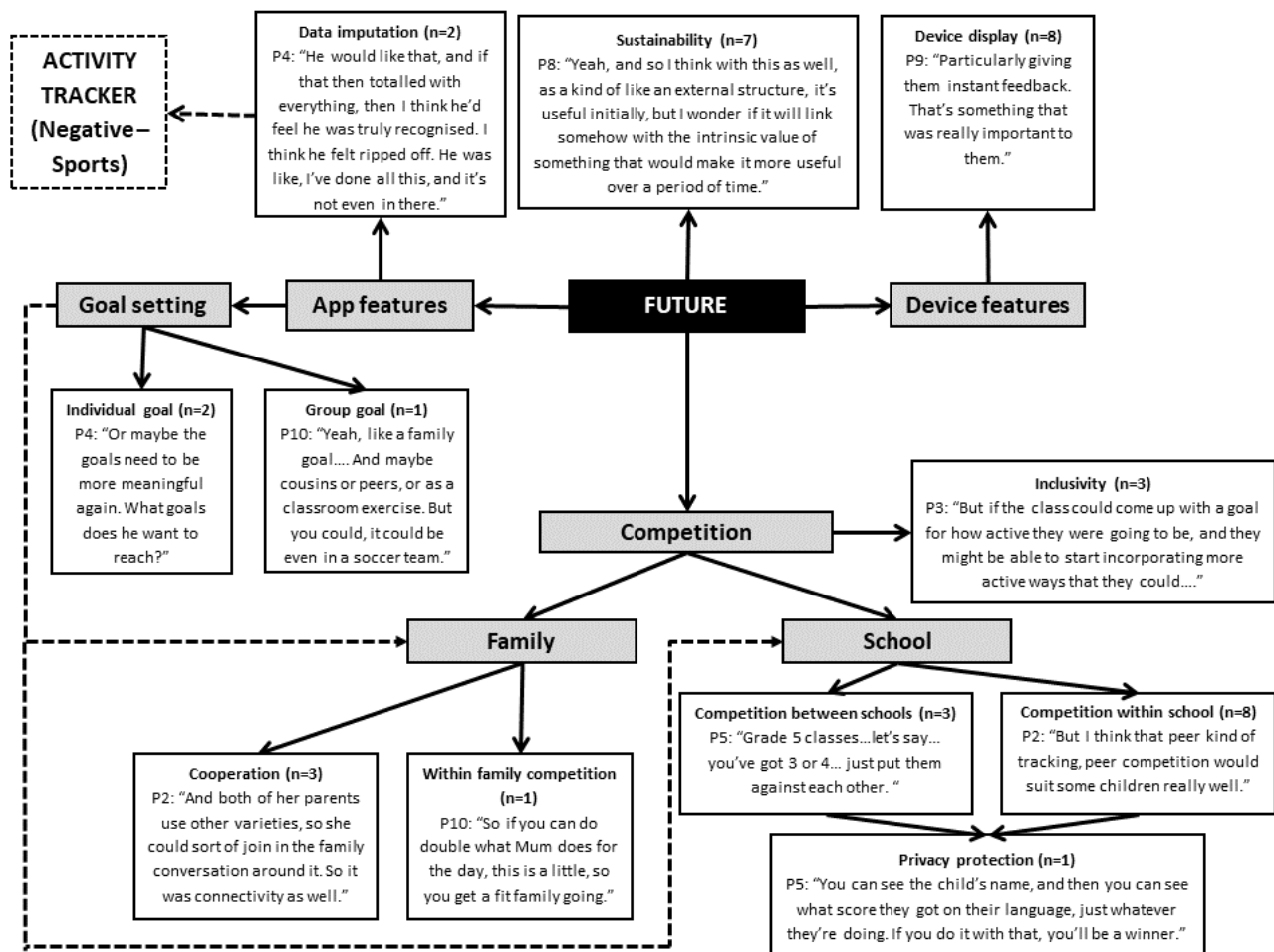
parents suggested connecting family members through the wearable devices, which could also facilitate healthy competition between family members. Indeed, most parents proposed developing a more competitive nature, particularly within school (8/10; 80%) whereby students know what their peers are doing. However, the issue of privacy protection if children’s names are being shown was also raised (see Figure 6). This inclusive group-based competition, regardless of environment, was thought to be a method to increase the inclusiveness of wearing activity trackers beyond just those children who are active:

*It would be great to have a competition between the classes, rather than amongst each individual kid, because then they’re helping each other along. [Parent 7]*

The long-term wear compliance and engagement of an activity tracker was highlighted as an important consideration for future use because of differences between individual children’s motivation. Indeed, some children wanted to purchase their own wearable tracker, whereas there was a novelty effect for others:

*I said, oh would you want to wear them, like would you want one of your own? They said no, we’ve kind of used it now. [Parent 10]*

**Figure 6.** Parents' considerations and recommendations for using wearable activity trackers in the future. The dashed line indicates link made between different themes was noted by the researchers from the points discussed, rather than directly mentioned by the parents (P).



## Discussion

### Principal Findings

Although self-monitoring systems have the potential to be used to promote physical activity levels in children, few studies have evaluated the acceptability and usability of wearable activity trackers in this age group. The main aim of this study was to examine parents' perceptions regarding the acceptability and usability of a wearable activity tracker (KidFit) for children over a 4-week period. Approximately half of parents reported that children wore the activity tracker every day and found it easy to use for activity tracking. However, parents noted that their children's main frustration was not being able to access real-time feedback, with features and functions only accessible through smart devices typically owned by their parents. Indeed, although newer generations of commercially available wearable activity trackers for children have overcome some of the barriers reported in this study, especially the real-time feedback, it is important to consider critical recommendations for all targeted wearable activity trackers as a potential intervention tool, rather than specific to the KidFit per se.

### Comparisons With Prior Work

Although there is a lack of research examining the usability and acceptability of wearable activity trackers in youth [28], research conducted in young children [34], adolescents [21], and adults

[22] has indicated that such devices are often viewed favorably. In agreement with previous research, the ease of use of the wearable activity tracker for children was identified as an important factor [21,24,40-44], with parents highlighting that the accompanying app, on the whole, was easy to navigate. Despite this, many visual app features were not accessed, particularly in relation to tracking sleep. This is contrary to Ridgers et al [21], where despite emphasizing a desire to access sleep data, adolescents felt difficulties and discomfort associated with wearing the activity tracker prevented this. Promoting good sleep habits is becoming increasingly important given the introduction of 24-hour movement guidelines in some countries [45], though there are concerns regarding the accuracy of wearables for tracking sleep [46]. The lack of interest in sleep data in this study could be because of the lack of comfort, which led to 50% of the children not wearing the devices overnight. Newer generation wearable trackers should therefore primarily focus on comfort and 24-hour *wearability*, identified as 1 of the 9 key factors that companies should consider when designing activity trackers [47]. Future research should seek to monitor the time children wear the activity trackers daily.

Of interest, the number of perceived issues with the activity tracker reported by the parents substantially decreased over the 4-week period (77.1%-28.6%). Although this could be attributed to parents identifying problems and knowing how to overcome them, this finding may also be explained by the reduction in

children wearing the activity trackers over time (ie, those who had problems simply stopped wearing them). This suggests that future research seeking to use wearable activity trackers in children should integrate a familiarization component for key features and functions to ensure ability to access and understand feedback, which includes problem-solving monitoring issues, to try to ensure that the tracker is used as intended [21]. More specifically, consideration needs to be given to parents, who may require training in the wearable activity tracker and app use [24], and children, where personalization of the device and app in relation to daily goals could be more important [19]. However, ceased wear could also be a result of children understanding what a sufficiently active day encompasses and, therefore, more accurately interpreting their own behaviors, negating wearable use.

Wearable activity trackers have been increasingly used in physical activity and sedentary behavior interventions, using self-monitoring systems that utilize theory-driven behavior change techniques [16,32]. Although several functions and features are incorporated into the KitFit app, such as a daily score and step counts, as strategies to target key behavior change techniques, these are limited compared with the wearables available to adolescents and adults [16] and the new-generation models for children. Nonetheless, 36 well-established behavior change techniques were identified in a recent content analysis across a range of features and functions within 3 commercially available youth-oriented activity trackers [32]. However, contrary to similar research in adolescents [21], parents in this study reported that they and their children used few features and functions, and therefore, the integration of more behavior change techniques may not be beneficial to the child when accessed through a parental app. This could be explored in evaluations for more recently available wearables for children that incorporate more behavior change techniques.

Those children who did use the associated wearable activity tracker app with their parents mainly used it for goal setting, self-monitoring, and individualized behavioral feedback. These behavior change techniques not only are integral to wearable activity trackers [12] but also are the most important components to influence behavior [13,43]. Despite the relatively short study duration, some children started to implement goal-setting strategies, with qualitative data revealing that most parents interviewed felt that the feedback enhanced their awareness of their children's current physical activity levels. Parents also noted such feedback also increased their children's awareness of their own activity levels. It could be argued that this increased awareness via the feedback provided was the most important outcome, enabling some children to understand the need to enhance, and others maintain, their activity behaviors [48]. Indeed, research has suggested that the development and implementation of self-regulation is a key objective for utilizing wearable activity trackers [49]. However, the accuracy issues inherent to wearable activity trackers raise concerns about the interpretation of associated data, particularly if the output is not directly relevant to government guidelines (ie, minutes of MVPA).

Irrespective of the different self-monitoring systems used across studies, there were numerous similarities in their acceptability

by young children [32] and adolescents [21]. Only the tangible activity data (ie, step or distance), on the whole, was accessed in this study, with the KidFit score rarely being used because of a lack of understanding of how it related to actual activity levels. Newer generation activity trackers, such as the Fitbit Ace, overcome this issue by providing time spent being active in accord with government guidelines. Parents emphasized that children wanted to visually track and self-monitor their activity. Indeed, Masteller et al [32] reported that children liked creating and changing the avatars associated with 2 child-specific wearable activity trackers, as a form of visual feedback. Therefore, *how* the activity tracker itself (eg, visual display) and associated app are designed may be fundamental for long-term use. As such, further research is warranted to explore how these constructs are presented on wearable activity trackers and associated feedback mechanisms, and *whether* and indeed *how* they need to change with long-term use.

Congruent with the previous research, the manner in which the wearable activity tracker fastened around the wrist was the biggest issue [21]. However, Müller et al [34] found that the Garmin Vivofit, which has now evolved from a stretchy bracelet to a watch strap, was a feasible wearable bracelet for assessment in children aged 4-10 years, though it is important to acknowledge that the devices were only worn for 7 days and removed for sleep. In this study, discomfort, coupled with device malfunction or maintenance (ie, syncing and charging), often led to frustration and consequently poor wear-time compliance. Congruent with the studies by Müller et al [34] and Shih et al [50], in children and adults, respectively, aesthetic and practical concerns, such as the activity tracker being *too bright* or *bulky*, suggest that the appealing design of the device is a crucial factor for long-term use [47].

Parents cited peer comparison, relating to awareness or competition, as a desired motivational component. During interviews, only 2 parents reported their children to use the feedback for peer comparison or competition, though this is unknown for the broader sample and could be because of the lack of siblings included in the subsample study, thereby limiting such interaction. Nonetheless, in accord with the previous research [32], group features, such as those incorporated into newer wearable activity trackers, may encourage more regular use by children. These findings are similar to those of Masteller et al [32] who found that 75% of children aged 6-11 years would wear the activity trackers more if their peers were wearing them, though it was the social aspects of the associated activity tracker websites that children liked more than the activity tracking features per se [32]. This concurs with the previous research in adolescents, whereby a wearable activity tracker was combined with social media facilitating social comparison and support [26,44,51]. Given the influence of peers on youth physical activity [52], future research should consider how to integrate and promote peer comparison into wearable activity tracker interventions, while simultaneously enhancing autonomy and minimizing peer pressure, or indeed dropout [53,54].

There are numerous strengths associated with this study. Specifically, this is the first to determine the usability and acceptability of wearable activity trackers in children, from a parents' perspective, using a mixed-methods approach.

Moreover, the survey and qualitative data were consistent, with the latter reaching data saturation. Although short in duration, this study was an ecologically valid acceptability and feasibility study, likely precluding the impact of a novelty effect. However, several limitations are noteworthy of discussion. The recruitment strategy (ie, university staff circular) may have resulted in a nonrepresentative sample incorporating participants from a higher socioeconomic background. However, contrary to findings among adults, evidence of socioeconomic gradients in children's physical activity is equivocal [55], and the generalizability of our findings across socioeconomic strata is unknown. Nonetheless, the recruitment strategy may have resulted in a sample with higher physical activity levels, which may explain the decreased interest across the study duration. Future research should target families from wider socioeconomic backgrounds and lower physical activity levels. It should be noted that although this was only 1 of 2 activity trackers marketed at children at the time, a number of parents owned activity trackers, which may have influenced perceptions of the acceptability and usability of the wearable activity tracker. Moreover, the KidFit is no longer in production and has since been superseded by newer devices, including the release of the Fitbit Ace, which has addressed issues such as real-time feedback and integrated familial and peer competition. Finally, although this study provides insights into children and parents'

initial experiences, via parental report, of using a wearable activity tracker for children, it is not known *whether or how* this would impact longer-term use.

## Conclusions

Taken together, parents reported that the wearable activity tracker was easy to use and a useful tool for tracking their children's daily activity. However, compliance in wearing the activity tracker was modest and declined across the study. Although parents acknowledged that the wearable activity tracker was acceptable for use in children, there were numerous barriers highlighted, which may preclude long-term wear and monitoring. Further research is needed to determine whether these findings are reflective of generic wrist-worn commercially available wearable activity trackers. Although wearable activity trackers for children hold promise as relatively inexpensive, scalable methods for the delivery of evidence-based behavior change techniques, it is important to address potential barriers and behavior change efficacy. These findings provide insights into how children and parents engage with wearable activity trackers designed for use by younger children (<13 years), which in turn may help to identify how to integrate wearable activity trackers into physical activity interventions or initiatives for children and what factors may need to be addressed to try to facilitate longer-term sustainability.

## Acknowledgments

SEC was supported by a Deakin University Postgraduate Research Scholarship. JS was supported by a National Health and Medical Research Council (NHMRC) Principal Research Fellowship (APP1026216) during this project. KB was supported by an NHMRC Principal Research Fellowship (APP1042442); AT was supported by a Future Leader Fellowship from the National Heart Foundation of Australia (Award ID 100046) during this project. NDR was supported by a Future Leader Fellowship from the National Heart Foundation of Australia (Award ID 101895). The authors gratefully acknowledge the contribution of Kara Richards to the project.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

KidFit monitor.

[PNG File, 16 KB - [mhealth\\_v7i11e13858\\_app1.png](#)]

## References

1. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, Lancet Physical Activity Series Working Group. The pandemic of physical inactivity: global action for public health. *Lancet* 2012 Jul 21;380(9838):294-305. [doi: [10.1016/S0140-6736\(12\)60898-8](https://doi.org/10.1016/S0140-6736(12)60898-8)] [Medline: [22818941](https://pubmed.ncbi.nlm.nih.gov/22818941/)]
2. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010 May 11;7:40 [FREE Full text] [doi: [10.1186/1479-5868-7-40](https://doi.org/10.1186/1479-5868-7-40)] [Medline: [20459784](https://pubmed.ncbi.nlm.nih.gov/20459784/)]
3. Health.Gov. 2008. Physical Activity Guidelines. URL: <http://www.health.gov/paguidelines/> [accessed 2019-07-01] [WebCite Cache ID 5gebKFYpM]
4. Australian Government Department of Health. Australian 24-Hour Movement Guidelines for Children and Young People (5-17 years) – An Integration of Physical Activity, Sedentary Behaviour and Sleep URL: <https://www.health.gov.au/internet/main/publishing.nsf/Content/health-24-hours-phys-act-guidelines> [accessed 2019-07-01]
5. Department of Health and Social Care. Physical Activity Guidelines for Children and Young People (5–18 Years). URL: [http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/documents/digitalasset/dh\\_128144.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_128144.pdf) [accessed 2018-02-11] [WebCite Cache ID 636UIoSkV]

6. Schranz NK, Olds T, Boyd R, Evans J, Gomersall SR, Hardy L, et al. Results from Australia's 2016 report card on physical activity for children and youth. *J Phys Act Health* 2016 Nov;13(11 Suppl 2):S87-S94. [doi: [10.1123/jpah.2016-0345](https://doi.org/10.1123/jpah.2016-0345)] [Medline: [27848737](https://pubmed.ncbi.nlm.nih.gov/27848737/)]
7. Katzmarzyk PT, Denstel KD, Beals K, Bolling C, Wright C, Crouter SE, et al. Results from the United States of America's 2016 report card on physical activity for children and youth. *J Phys Act Health* 2016 Nov;13(11 Suppl 2):S307-S313 [FREE Full text] [doi: [10.1123/jpah.2016-0321](https://doi.org/10.1123/jpah.2016-0321)] [Medline: [27848726](https://pubmed.ncbi.nlm.nih.gov/27848726/)]
8. Tyler R, Mannello M, Mattingley R, Roberts C, Sage R, Taylor SR, et al. Results from Wales' 2016 report card on physical activity for children and youth: is Wales turning the tide on children's inactivity? *J Phys Act Health* 2016 Nov;13(11 Suppl 2):S330-S336. [doi: [10.1123/jpah.2016-0309](https://doi.org/10.1123/jpah.2016-0309)] [Medline: [27848724](https://pubmed.ncbi.nlm.nih.gov/27848724/)]
9. Grewal TS. Western University Electronic Thesis and Dissertation Repository. 2013. Awareness of Physical Activity Levels and Sedentary Behaviour: An Assessment of Awareness of Physical Activity Levels and Sedentary Behaviour Among Parents and Children. URL: <https://ir.lib.uwo.ca/etd/1155/> [accessed 2018-06-06] [WebCite Cache ID 76US3XXAH]
10. Knox EC, Esliger DW, Biddle SJ, Sherar LB. Lack of knowledge of physical activity guidelines: can physical activity promotion campaigns do better? *BMJ Open* 2013 Dec 5;3(12):e003633 [FREE Full text] [doi: [10.1136/bmjopen-2013-003633](https://doi.org/10.1136/bmjopen-2013-003633)] [Medline: [24319271](https://pubmed.ncbi.nlm.nih.gov/24319271/)]
11. Hussey J, Bell C, Gormley J. The measurement of physical activity in children. *Phys Ther Rev* 2013 Jul 19;12(1):52-58. [doi: [10.1179/108331907X174989](https://doi.org/10.1179/108331907X174989)]
12. 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](https://doi.org/10.1186/1479-5868-8-115)] [Medline: [22018588](https://pubmed.ncbi.nlm.nih.gov/22018588/)]
13. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
14. Thompson WR. Worldwide survey of fitness trends for 2017. *ACSMS Health Fit J* 2017;20(6):8-17. [doi: [10.1249/01.FIT.0000422568.47859.35](https://doi.org/10.1249/01.FIT.0000422568.47859.35)]
15. Salmon J, Ridgers ND. Is wearable technology an activity motivator, or a fad that wears thin? *Med J Aust* 2017 Feb 20;206(3):119-120. [doi: [10.5694/mja16.01242](https://doi.org/10.5694/mja16.01242)] [Medline: [28208042](https://pubmed.ncbi.nlm.nih.gov/28208042/)]
16. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res* 2014 Aug 15;16(8):e192 [FREE Full text] [doi: [10.2196/jmir.3469](https://doi.org/10.2196/jmir.3469)] [Medline: [25131661](https://pubmed.ncbi.nlm.nih.gov/25131661/)]
17. Sharp CA, Mackintosh KA, Erjavec M, Pascoe DM, Horne PJ. Validity and reliability of the Fitbit Zip as a measure of preschool children's step count. *BMJ Open Sport Exerc Med* 2017;3(1):e000272 [FREE Full text] [doi: [10.1136/bmjsem-2017-000272](https://doi.org/10.1136/bmjsem-2017-000272)] [Medline: [29081984](https://pubmed.ncbi.nlm.nih.gov/29081984/)]
18. Byun W, Lee JM, Kim Y, Brusseau TA. Classification accuracy of a wearable activity tracker for assessing sedentary behavior and physical activity in 3-5-year-old children. *Int J Environ Res Public Health* 2018 Mar 26;15(4):pii: E594 [FREE Full text] [doi: [10.3390/ijerph15040594](https://doi.org/10.3390/ijerph15040594)] [Medline: [29587432](https://pubmed.ncbi.nlm.nih.gov/29587432/)]
19. Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. *Sleep* 2015 Aug 1;38(8):1323-1330 [FREE Full text] [doi: [10.5665/sleep.4918](https://doi.org/10.5665/sleep.4918)] [Medline: [26118555](https://pubmed.ncbi.nlm.nih.gov/26118555/)]
20. Schneider M, Chau L. Validation of the Fitbit Zip for monitoring physical activity among free-living adolescents. *BMC Res Notes* 2016 Sep 21;9(1):448 [FREE Full text] [doi: [10.1186/s13104-016-2253-6](https://doi.org/10.1186/s13104-016-2253-6)] [Medline: [27655477](https://pubmed.ncbi.nlm.nih.gov/27655477/)]
21. Ridgers ND, Timperio A, Brown H, Ball K, Macfarlane S, Lai SK, et al. Wearable activity tracker use among Australian adolescents: usability and acceptability study. *JMIR Mhealth Uhealth* 2018 Apr 11;6(4):e86 [FREE Full text] [doi: [10.2196/mhealth.9199](https://doi.org/10.2196/mhealth.9199)] [Medline: [29643054](https://pubmed.ncbi.nlm.nih.gov/29643054/)]
22. Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. *BMC Public Health* 2017 Nov 15;17(1):880 [FREE Full text] [doi: [10.1186/s12889-017-4888-1](https://doi.org/10.1186/s12889-017-4888-1)] [Medline: [29141607](https://pubmed.ncbi.nlm.nih.gov/29141607/)]
23. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. *JMIR Mhealth Uhealth* 2017 Mar 6;5(3):e28 [FREE Full text] [doi: [10.2196/mhealth.6967](https://doi.org/10.2196/mhealth.6967)] [Medline: [28264796](https://pubmed.ncbi.nlm.nih.gov/28264796/)]
24. Sloatmaker SM, Chinapaw MJ, Schuit AJ, Seidell JC, van Mechelen W. Feasibility and effectiveness of online physical activity advice based on a personal activity monitor: randomized controlled trial. *J Med Internet Res* 2009 Jul 29;11(3):e27 [FREE Full text] [doi: [10.2196/jmir.1139](https://doi.org/10.2196/jmir.1139)] [Medline: [19674956](https://pubmed.ncbi.nlm.nih.gov/19674956/)]
25. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. *Am J Prev Med* 2015 Sep;49(3):414-418 [FREE Full text] [doi: [10.1016/j.amepre.2015.01.020](https://doi.org/10.1016/j.amepre.2015.01.020)] [Medline: [26071863](https://pubmed.ncbi.nlm.nih.gov/26071863/)]
26. Pumper MA, Mendoza JA, Arseniev-Koehler A, Holm M, Waite A, Moreno MA. Using a facebook group as an adjunct to a pilot mhealth physical activity intervention: a mixed methods approach. *Stud Health Technol Inform* 2015;219:97-101. [doi: [10.3233/978-1-61499-595-1-97](https://doi.org/10.3233/978-1-61499-595-1-97)] [Medline: [26799887](https://pubmed.ncbi.nlm.nih.gov/26799887/)]

27. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable sensor/device (Fitbit One) and SMS text-messaging prompts to increase physical activity in overweight and obese adults: a randomized controlled trial. *Telemed J E Health* 2015 Oct;21(10):782-792 [FREE Full text] [doi: [10.1089/tmj.2014.0176](https://doi.org/10.1089/tmj.2014.0176)] [Medline: [26431257](https://pubmed.ncbi.nlm.nih.gov/26431257/)]
28. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. *JMIR Mhealth Uhealth* 2016 Nov 23;4(4):e129 [FREE Full text] [doi: [10.2196/mhealth.6540](https://doi.org/10.2196/mhealth.6540)] [Medline: [27881359](https://pubmed.ncbi.nlm.nih.gov/27881359/)]
29. Duncan M, Murawski B, Short CE, Rebar AL, Schoeppe S, Alley S, et al. Activity trackers implement different behavior change techniques for activity, sleep, and sedentary behaviors. *Interact J Med Res* 2017 Aug 14;6(2):e13 [FREE Full text] [doi: [10.2196/ijmr.6685](https://doi.org/10.2196/ijmr.6685)] [Medline: [28807889](https://pubmed.ncbi.nlm.nih.gov/28807889/)]
30. Rich E, Miah A. Mobile, wearable and ingestible health technologies: towards a critical research agenda. *Health Sociol Rev* 2016 Jul 28;26(1):84-97. [doi: [10.1080/14461242.2016.1211486](https://doi.org/10.1080/14461242.2016.1211486)]
31. Atienza AA, Patrick K. Mobile health: the killer app for cyberinfrastructure and consumer health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S151-S153. [doi: [10.1016/j.amepre.2011.01.008](https://doi.org/10.1016/j.amepre.2011.01.008)] [Medline: [21521588](https://pubmed.ncbi.nlm.nih.gov/21521588/)]
32. Masteller B, Sirard J, Freedson P. The physical activity tracker testing in youth (PATTY) study: content analysis and children's perceptions. *JMIR Mhealth Uhealth* 2017 Apr 28;5(4):e55 [FREE Full text] [doi: [10.2196/mhealth.6347](https://doi.org/10.2196/mhealth.6347)] [Medline: [28455278](https://pubmed.ncbi.nlm.nih.gov/28455278/)]
33. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med* 2011 Mar;1(1):53-71 [FREE Full text] [doi: [10.1007/s13142-011-0021-7](https://doi.org/10.1007/s13142-011-0021-7)] [Medline: [21796270](https://pubmed.ncbi.nlm.nih.gov/21796270/)]
34. Müller J, Hoch AM, Zoller V, Oberhoffer R. Feasibility of physical activity assessment with wearable devices in children aged 4-10 years-a pilot study. *Front Pediatr* 2018;6:5 [FREE Full text] [doi: [10.3389/fped.2018.00005](https://doi.org/10.3389/fped.2018.00005)] [Medline: [29435438](https://pubmed.ncbi.nlm.nih.gov/29435438/)]
35. 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](https://doi.org/10.3402/qhw.v9.26152)] [Medline: [25326092](https://pubmed.ncbi.nlm.nih.gov/25326092/)]
36. MacPhail C, Khoza N, Abler L, Ranganathan M. Process guidelines for establishing intercoder reliability in qualitative studies. *Qual Res* 2015 Apr 20;16(2):198-212. [doi: [10.1177/1468794115577012](https://doi.org/10.1177/1468794115577012)]
37. Mackintosh KA, Knowles ZR, Ridgers ND, Fairclough SJ. Using formative research to develop CHANGE!: a curriculum-based physical activity promoting intervention. *BMC Public Health* 2011 Oct 27;11:831 [FREE Full text] [doi: [10.1186/1471-2458-11-831](https://doi.org/10.1186/1471-2458-11-831)] [Medline: [22032540](https://pubmed.ncbi.nlm.nih.gov/22032540/)]
38. Ridgers ND, Knowles ZR, Sayers J. Encouraging play in the natural environment: a child-focused case study of forest school. *Child Geogr* 2012 Feb;10(1):49-65 [FREE Full text] [doi: [10.1080/14733285.2011.638176](https://doi.org/10.1080/14733285.2011.638176)]
39. Miles MB, Huberman AM. *Qualitative Data Analysis: A Expanded Sourcebook*. Second Edition. Thousand Oaks, California: Sage Publications; 1994.
40. Hooke MC, Gilchrist L, Tanner L, Hart N, Withycombe JS. Use of a fitness tracker to promote physical activity in children with acute lymphoblastic leukemia. *Pediatr Blood Cancer* 2016 Apr;63(4):684-689. [doi: [10.1002/pbc.25860](https://doi.org/10.1002/pbc.25860)] [Medline: [26756736](https://pubmed.ncbi.nlm.nih.gov/26756736/)]
41. Schaefer SE, van Loan M, German BJ. A feasibility study of wearable activity monitors for pre-adolescent school-age children. *Prev Chronic Dis* 2014 May 22;11:E85 [FREE Full text] [doi: [10.5888/pcd11.130262](https://doi.org/10.5888/pcd11.130262)] [Medline: [24854236](https://pubmed.ncbi.nlm.nih.gov/24854236/)]
42. Schaefer SE, Ching CC, Breen H, German BJ. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. *Am J Health Educ* 2016 Jan 8;47(1):8-16. [doi: [10.1080/19325037.2015.1111174](https://doi.org/10.1080/19325037.2015.1111174)]
43. Ledger D, McCaffrey D. Medium. 2017. Inside Wearables: How the Science of Human Behavior Change Offers the Secret to Long-Term Engagement. URL: <https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3> [accessed 2016-11-18]
44. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mhealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: a pilot study. *Pediatr Blood Cancer* 2017 Dec;64(12). [doi: [10.1002/pbc.26660](https://doi.org/10.1002/pbc.26660)] [Medline: [28618158](https://pubmed.ncbi.nlm.nih.gov/28618158/)]
45. Tremblay MS, Carson V, Chaput JP. Introduction to the Canadian 24-hour movement guidelines for children and youth: an integration of physical activity, sedentary behaviour, and sleep. *Appl Physiol Nutr Metab* 2016 Jun;41(6 Suppl 3):iii-iv. [doi: [10.1139/apnm-2016-0203](https://doi.org/10.1139/apnm-2016-0203)] [Medline: [27306430](https://pubmed.ncbi.nlm.nih.gov/27306430/)]
46. Sargent C, Zhou X, Matthews RW, Darwent D, Roach GD. Daily rhythms of hunger and satiety in healthy men during one week of sleep restriction and circadian misalignment. *Int J Environ Res Public Health* 2016 Jan 29;13(2):170 [FREE Full text] [doi: [10.3390/ijerph13020170](https://doi.org/10.3390/ijerph13020170)] [Medline: [26840322](https://pubmed.ncbi.nlm.nih.gov/26840322/)]
47. Ledger D. Medium. 2014. Inside Wearables Part 2: A Look at the Uncertain Future of Smart Wearable Devices, and Five Industry Developments That Will Be Necessary for Meaningful Mass Market Adoption and Sustained Engagement. URL: <https://blog.endeavour.partners/inside-wearables-part-2-july-2014-ef301d425cdd?gi=1492a1eb659c> [accessed 2018-09-27]
48. Corder K, van Sluijs EM, Goodyer I, Ridgway CL, Steele RM, Bamber D, et al. Physical activity awareness of British adolescents. *Arch Pediatr Adolesc Med* 2011 Jul;165(7):603-609. [doi: [10.1001/archpediatrics.2011.94](https://doi.org/10.1001/archpediatrics.2011.94)] [Medline: [21727272](https://pubmed.ncbi.nlm.nih.gov/21727272/)]
49. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. *JMIR Mhealth Uhealth* 2016 Apr 27;4(2):e40 [FREE Full text] [doi: [10.2196/mhealth.4461](https://doi.org/10.2196/mhealth.4461)] [Medline: [27122452](https://pubmed.ncbi.nlm.nih.gov/27122452/)]

50. Shih PC, Han K, Poole ES, Rosson MB, Carroll JM. Use and Adoption Challenges of Wearable Activity Trackers. In: Proceedings of the iConference. 2015 Presented at: iConference'15; March 24-27, 2015; Newport Beach, CA, USA. URL: <http://hdl.handle.net/2142/73649>
51. Ridgers ND, Timperio A, Brown H, Ball K, Macfarlane S, Lai SK, et al. A cluster-randomised controlled trial to promote physical activity in adolescents: the raising awareness of physical activity (RAW-PA) study. BMC Public Health 2017 Jan 4;17(1):6 [FREE Full text] [doi: [10.1186/s12889-016-3945-5](https://doi.org/10.1186/s12889-016-3945-5)] [Medline: [28052773](https://pubmed.ncbi.nlm.nih.gov/28052773/)]
52. Hohepa M, Scragg R, Schofield G, Kolt GS, Schaaf D. Social support for youth physical activity: importance of siblings, parents, friends and school support across a segmented school day. Int J Behav Nutr Phys Act 2007 Nov 8;4:54 [FREE Full text] [doi: [10.1186/1479-5868-4-54](https://doi.org/10.1186/1479-5868-4-54)] [Medline: [17996070](https://pubmed.ncbi.nlm.nih.gov/17996070/)]
53. Goodyear VA, Kerner C, Quennerstedt M. Young people's uses of wearable healthy lifestyle technologies; surveillance, self-surveillance and resistance. Sport Educ Soc 2017 Sep 22;24(3):212-225. [doi: [10.1080/13573322.2017.1375907](https://doi.org/10.1080/13573322.2017.1375907)]
54. Kerner C, Goodyear VA. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on Fitbits with adolescents. Am J Health Educ 2017 Jul 28;48(5):287-297. [doi: [10.1080/19325037.2017.1343161](https://doi.org/10.1080/19325037.2017.1343161)]
55. O'Donoghue G, Kennedy A, Puggina A, Aleksovska A, Buck C, Burns C, et al. Socio-economic determinants of physical activity across the life course: a 'DEterminants of DIet and Physical ACTivity' (DEDIPAC) umbrella literature review. PLoS One 2018;13(1):e0190737 [FREE Full text] [doi: [10.1371/journal.pone.0190737](https://doi.org/10.1371/journal.pone.0190737)] [Medline: [29351286](https://pubmed.ncbi.nlm.nih.gov/29351286/)]

## Abbreviations

**mHealth:** mobile health

**MVPA:** moderate- to vigorous-intensity physical activity

**NHMRC:** National Health and Medical Research Council

*Edited by G Eysenbach; submitted 27.02.19; peer-reviewed by S Edney, J Sasaki, B Worm; comments to author 11.05.19; revised version received 07.07.19; accepted 19.07.19; published 04.11.19.*

*Please cite as:*

*Mackintosh KA, Chappel SE, Salmon J, Timperio A, Ball K, Brown H, Macfarlane S, Ridgers ND*

*Parental Perspectives of a Wearable Activity Tracker for Children Younger Than 13 Years: Acceptability and Usability Study*

*JMIR Mhealth Uhealth 2019;7(11):e13858*

*URL: <https://mhealth.jmir.org/2019/11/e13858>*

*doi: [10.2196/13858](https://doi.org/10.2196/13858)*

*PMID: [31682585](https://pubmed.ncbi.nlm.nih.gov/31682585/)*

©Kelly A D Mackintosh, Stephanie E Chappel, Jo Salmon, Anna Timperio, Kylie Ball, Helen Brown, Susie Macfarlane, Nicola D Ridgers. Originally published in JMIR Mhealth and Uhealth (<http://mhealth.jmir.org>), 04.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Opportunities and Pitfalls in Applying Emotion Recognition Software for Persons With a Visual Impairment: Simulated Real Life Conversations

Hendrik Buimer<sup>1,2\*</sup>, MSc; Renske Schellens<sup>1\*</sup>, MSc; Tjerk Kostelijk<sup>3</sup>, MSc; Abdellatif Nemri<sup>2</sup>, PhD; Yan Zhao<sup>2</sup>, PhD; Thea Van der Geest<sup>4</sup>, PhD; Richard Van Wezel<sup>1,2</sup>, PhD

<sup>1</sup>Department of Biophysics, Radboud University, Nijmegen, Netherlands

<sup>2</sup>Department of Biomedical Signals and Systems, University of Twente, Enschede, Netherlands

<sup>3</sup>VicarVision, Amsterdam, Netherlands

<sup>4</sup>Center IT + Media, Hogeschool van Arnhem en Nijmegen University of Applied Sciences, Arnhem, Netherlands

\*these authors contributed equally

**Corresponding Author:**

Hendrik Buimer, MSc

Department of Biophysics

Radboud University

Heijendaalseweg 135

Nijmegen

Netherlands

Phone: 31 24 3652428

Email: [h.buimer@donders.ru.nl](mailto:h.buimer@donders.ru.nl)

## Abstract

**Background:** A large part of the communication cues exchanged between persons is nonverbal. Persons with a visual impairment are often unable to perceive these cues, such as gestures or facial expression of emotions. In a previous study, we have determined that visually impaired persons can increase their ability to recognize facial expressions of emotions from validated pictures and videos by using an emotion recognition system that signals vibrotactile cues associated with one of the six basic emotions.

**Objective:** The aim of this study was to determine whether the previously tested emotion recognition system worked equally well in realistic situations and under controlled laboratory conditions.

**Methods:** The emotion recognition system consists of a camera mounted on spectacles, a tablet running facial emotion recognition software, and a waist belt with vibrotactile stimulators to provide haptic feedback representing Ekman's six universal emotions. A total of 8 visually impaired persons (4 females and 4 males; mean age 46.75 years, age range 28-66 years) participated in two training sessions followed by one experimental session. During the experiment, participants engaged in two 15 minute conversations, in one of which they wore the emotion recognition system. To conclude the study, exit interviews were conducted to assess the experiences of the participants. Due to technical issues with the registration of the emotion recognition software, only 6 participants were included in the video analysis.

**Results:** We found that participants were quickly able to learn, distinguish, and remember vibrotactile signals associated with the six emotions. A total of 4 participants felt that they were able to use the vibrotactile signals in the conversation. Moreover, 5 out of the 6 participants had no difficulties in keeping the camera focused on the conversation partner. The emotion recognition was very accurate in detecting happiness but performed unsatisfactorily in recognizing the other five universal emotions.

**Conclusions:** The system requires some essential improvements in performance and wearability before it is ready to support visually impaired persons in their daily life interactions. Nevertheless, the participants saw potential in the system as an assistive technology, assuming their user requirements can be met.

(*JMIR Mhealth Uhealth* 2019;7(11):e13722) doi:[10.2196/13722](https://doi.org/10.2196/13722)

**KEYWORDS**

visual impairment; emotion recognition; tactile; social interaction



## Introduction

### Background

A large number of communication cues exchanged between persons are nonverbal (eg, gestures, facial expressions, and gaze direction). As a result, the inability to perceive these cues leads to a loss of communication effectiveness. Not perceiving these nonverbal cues is particularly present in persons who are blind or visually impaired (hereafter referred to as visually impaired persons) and can lead to feelings of exclusion [1,2]. Previous inventory studies of assistive technology needs [3,4] found that a need among the community of visually impaired persons still exists for a solution that makes nonverbal signals accessible.

To convey such information and make it accessible for visually impaired persons, visual information can be *translated* into auditory or tactile cues, which is the foundation of sensory substitution devices (SSDs). The SSD developed by Bach-y-Rita [5], often considered as the first SSD, translated a video feed into a grid of vibration motors attached to the back of a chair. Visually impaired persons trained to use this system were able to pick up an object. More recently, the vOICE system (which translated a video with very high contrast directly into auditory cues) and a Tongue Display Unit were developed (which translated video into electric signals on the tongue) [6-8]. However, these systems rely largely on senses and body parts that are crucial for interpersonal communication, that is, mouth and ears. In the last years, various researchers have acknowledged the issue of nonverbal information and worked toward a range of sensory-substitution designs: information such as interpersonal distance, location of others, facial expressions of emotions, and gestures was conveyed using various vibrotactile devices, including the back of a chair [9,10], a glove [11], and a vibrotactile belt [12,13].

In addition, recent advances in machine learning and computer vision technologies, which improved the ability of computers to recognize patterns from images, allow for very specific recognition tasks. One of the opportunities that arose from these advancements is the possibility to train a computer to recognize faces and facial expressions from a video. This idea was pioneered by Bartlett et al in a collaboration with Paul Ekman [14], leading to effective emotion recognition software. Among its numerous applications, such software is particularly interesting for people who are unable to detect and interpret facial expressions of emotions, such as visually impaired persons. Given the need from the community and the potential of emotion recognition software in meeting these needs, we investigated to what extent such emotion recognition software can support visually impaired persons in their daily lives.

Previously, we presented a wearable system for facial emotion recognition consisting of a head-mounted camera and a vibrating belt [15]. It was designed to help visually impaired persons perceive facial expressions of the 6 universal emotions [16]. Via a tiny camera attached to standard spectacles, pointing in the direction where the wearer is looking, a continuous video stream was available for processing. To determine which emotions were expressed, FaceReader facial expression recognition software was applied. This software can be used to

detect Ekman's universal emotions from pictures and videos [16-18]. In this study, FaceReader was applied to analyze the video stream, originating from the video camera, in real time. Once a face is detected, the software attempts to create a map of the face including almost 500 facial landmarks. If a face muscle movement is detected, the deviation from a baseline neutral face is recognized as a facial expression and then classified as an emotion. Depending on the occlusion of the faces, either a standard model (called 3D Active Appearance Model) or proprietary deep neural networks are used to assign a score to the intensity of the facial expression (emotion). Once an expression was detected as being 1 of the 6 universal emotions, vibrotactile signals on a waist belt conveyed this information to the wearer of the system. Previous research has shown that vibrotactile signals were interpretable and usable in cognitively and physically demanding tasks such as flying a helicopter or steering a boat [19]. By signaling emotions through haptic feedback at the waist during social interactions, the ears of the users remained free to engage in a conversation.

This system was previously tested with visually impaired persons in a controlled laboratory setting [15], with pictures from the Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) [20] and videos from the Amsterdam Dynamic Facial Expression Set (ADFES) [21]. Such validated sets of images and videos of actors showing strong emotional expressions in controlled laboratory conditions play an important role in the training and development of emotion recognition software but do not entirely resemble realistic facial expressions [22]. The results of our first study were promising, as the FaceReader software reached high recognition scores and users were able to easily learn how to interpret and use the vibrotactile signals from the system [15]. That the system worked well under controlled laboratory lighting in classifying very expressive facial expressions was expected, as emotion recognition software often achieved high recognition scores when tested with validated sets [18,23]. On validated sets such as the ADFES and WSEFEP, FaceReader was known to achieve recognition rates of 89% and 88%, respectively [17]. However, these sets are often acquired during photoshoots with actors who express strong emotional states under optimal lighting conditions. Facial expressions in real life are often much subtler than those found in the validated sets, and lighting conditions also differ between real life and controlled conditions [22,24,25]. As a result, we cannot be certain that the recognition scores achieved from these validated sets can be replicated once such software is faced with genuine facial expressions of emotions, expressed in natural conversations. Thus, to determine whether the technology applied is useful for this purpose, its performance must be further tested under real-world conditions [26].

### Objectives

To investigate whether the system is usable and useful during realistic conversations, we aimed to examine the following questions: (1) were visually impaired persons able to keep the camera focused on the face of the conversation partner?; (2) did the software reach a satisfactory recognition rate when confronted with an unstable video stream and suboptimal lighting conditions?; (3) were visually impaired persons able to interpret the emotion cues while engaging in a meaningful

conversation?; and (4) did the system live up to prospective users' expectations, and did they believe they could use such technologies in their daily lives?

## Methods

### Ethical Approval

The study was designed in accordance with the Declaration of Helsinki and approved by the ethical committee of the Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente, Enschede.

**Table 1.** Participants.

ID	Age (years)	Gender	Visual impairment	Sight	10 or more years of vision	Emotion logs available <sup>a</sup>
1	28	Male	Neurological damage to eye nerves	Tunnel	No	Yes
2 <sup>b</sup>	28	Male	Persistent fetal vasculature	None	No	No
3 <sup>b</sup>	66	Male	Retinitis pigmentosa	None	No	No
4	59	Female	Retinitis pigmentosa, glaucoma, and cataract	None	Yes	Yes
5	47	Male	Cone dystrophy	Peripheral	Yes	Yes
6	64	Female	Congenital rubella syndrome	None	No	Yes
7	50	Female	Retinitis pigmentosa	None	Yes	Yes
8	32	Female	Aniridia	<5%	Yes	Yes

<sup>a</sup>See the section Data Collection and Analysis.

<sup>b</sup>No emotion logs were collected during the experiment.

### Apparatus

The experiments were conducted at an observation room at the HAN University of Applied Sciences (Arnhem, The Netherlands), where it was possible to discreetly observe the experiment from an adjacent viewing room and record with 3 different cameras and 2 microphones for postanalysis.

The emotion recognition system used in the study consisted of a small USB camera clipped on spectacles; a Microsoft Surface Pro 4 tablet running the FaceReader 6 SDK emotion recognition software (running the general face model); and a chain of 9 vibrotactile stimulators attached to a Velcro belt, 6 of which were used (see Figure 1). To ensure steady connections between the components, cable connections were used.

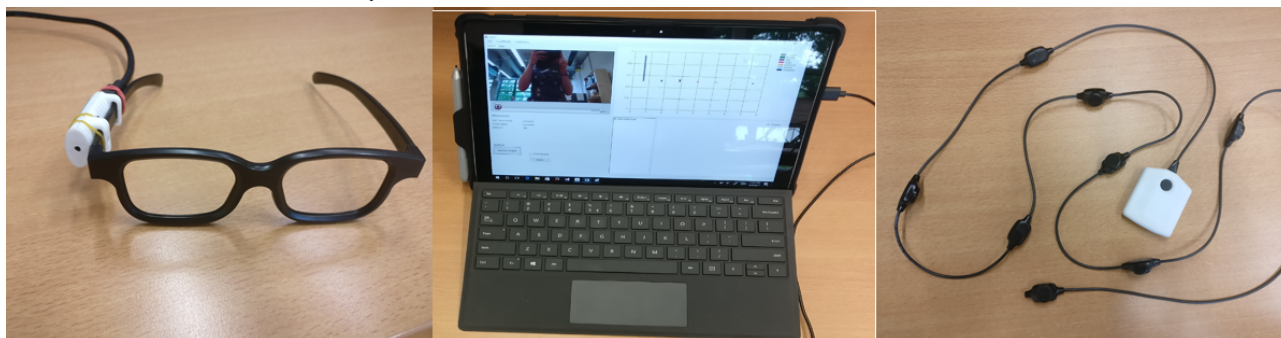
The FaceReader system classified facial expressions into 1 of the 6 of Ekman's universal emotions: happiness, sadness, anger,

### Participants

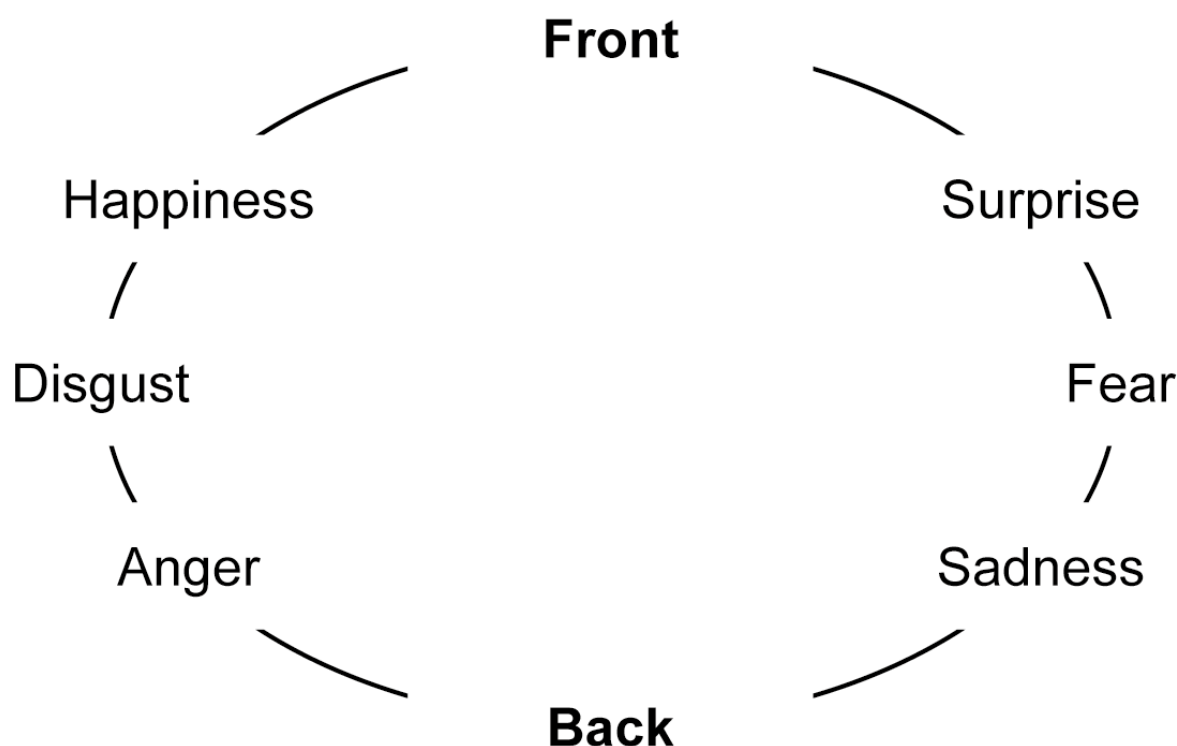
A total of 8 visually impaired persons (4 females and 4 males; mean age 46.75 years, age range 28-66 years) were included in the study. Moreover, 4 of the participants were visually impaired from birth, whereas the others became visually impaired later in life but had experience with vision for at least 10 years. All participants reported difficulties with the recognition of facial expressions and did not suffer from any other cognitive or sensory impairments (Table 1).

fear, surprise, and disgust [16]. When an emotion is recognized, feedback was given to the participant via the vibrotactile stimulators worn around the waist. Each emotion was assigned to a single vibrotactile stimulator (Figure 2). The decision to use a belt with vibrotactile stimulators and let it be worn on the waist had several reasons. First, the tactile modality was chosen because the use of the auditory modality is undesirable during social interactions, in which auditory information is very important for visually impaired persons and should, therefore, not be interfered with. Second, the tactile resolution of the waist is enough to place 6 vibrotactile stimulators without the risk of users being unable to distinguish them. A belt can be worn unobtrusively underneath clothing and wearing it at the waist means that the system does not interfere with the conversations in the sense that it keeps your eyes, ears, and hands free.

**Figure 1.** Schematic overview of the used system.



**Figure 2.** Emotion mapping. The mapping of Ekman's universal emotions on the waist band.



Vibrotactile stimulators were activated when the associated emotion was detected above a certain threshold of certainty, after which the vibration strength indicated the certainty that the software had detected an emotion. The standard threshold to initiate a vibrating pulse was 0.3 (on the FaceReader scale from 0 to 1), with 2 exceptions. In pretests with an actor, the FaceReader software was more sensitive to some emotions than to others. For example, when the actor was concentrating on listening and looking interested, his facial expression was at times classified as angry. Therefore, the threshold for the emotion anger was raised to 0.6. A threshold of 0.7 was chosen for the emotion happiness as it was rapidly detected even if the expression was barely visible.

### Procedure

The study consisted of 4 different phases: initial training, refresher training, the experiment, and an exit interview.

First, the participants were trained to interpret the vibrotactile signals during 2 separate training sessions (Table 2). The first session took place about 2 weeks before the experiment. After a brief introduction to the tactile mapping around the waist, the training included application of 4 sets of 36 vibrotactile stimuli. After each stimulus, the participant was asked to indicate which emotion was signaled. These stimuli were initially accompanied with matching auditory cues that matched the vibrotactile feedback. The number of audio-accompanied stimuli steadily decreased (100%, 50%, 33%, and 0%), resulting in a full dependency on vibrotactile signals in the fourth set. In these sets, each emotion was equally represented. The fifth and last set included only vibrotactile feedback. The minimal number

of stimuli presented in this set was 20, whereas the maximum number of stimuli was set at 96. It was ended when participants achieved a 95% success rate over the last 20 stimuli that were presented to them. The initial training phase lasted approximately 65 min.

The refresher training session took place just before the experiment and consisted of 3 sets. The first 2 sets included 24 stimuli in which the emotions were equally represented. The first set was fully accompanied by audio, whereas the second set included only vibrotactile feedback. The third set was the same as the final set of the initial training. It included 96 vibrotactile-only stimuli and was ended when participant achieved a 95% correct rate in the last 20 stimuli presented to them. The refresher training had an average duration of approximately 30 min.

After training was completed, the participants engaged in 2 conversations with an actor for approximately 15 min. The conversations were set up as a mock job interview in which the actor played the role of the director of a fictional company. The actor and the participant were seated at a table approximately 1.5 m opposite of each other.

The first conversation was colloquial, whereas the second was more formal. This was done to put the participants at ease and because we expected that the quantity and type of facial expressions of emotions might differ between the two.

The actor followed a similar structure for all conversations with the 8 participants. He was instructed not to exaggerate any facial expressions to keep the conversations as realistic as possible. A total of 4 participants wore the system in the first

conversation, whereas the remaining 4 wore it during the second conversation.

After the experimental session, a semistructured exit interview was conducted to allow participants to share their thoughts and

ideas about the system. It addressed various themes associated with system functionality, user experience, and technology acceptance (see [Table 3](#)).

**Table 2.** Stimuli distribution. Overview of the number of stimuli that were presented to the participants in each training session and the percentage of audio-accompanied stimuli.

Session	Total number of stimuli (percentage audio accompanied stimuli), n (%)				
1	36 (100)	36 (50)	36 (33)	36 (0)	20-96 (0) <sup>a</sup>
2	24 (100)	24 (0)	20-96 (0) <sup>a</sup>	— <sup>b</sup>	—

<sup>a</sup>Ended after a 95% correct score in the last 20 stimuli.

<sup>b</sup>The second session included only three sets of stimuli.

**Table 3.** Exit interview topics. An overview of the topics discussed in the exit interview, the associated questions, and the subthemes of interest.

Topic, exit interview question	Subthemes
<b>Experience</b>	
How was your experience using the system?	Accuracy, usefulness, feedback, fun of use, and ease of use
What was the value of the system during the conversation?	Additional value, distraction from the conversation, and interpretation of the system
<b>Potential utility</b>	
Would you like to use the system in your daily life?	System usage, situations of use, and recommendation to others
Imagine you have a job interview next week and we will lend you the system. Would you use it during the job interview?	Usage in job interview
<b>Adjustments to the system</b>	
If you are the manager of the team that develops this system, what do you think they should tackle first?	Improvements of the system and additions to the system
<b>Social acceptance</b>	
How do you expect people in your surroundings (eg, family, friends, and colleagues) will react to you using a system like this?	Reaction of surroundings and introduction of the system

## Data Collection and Analysis

After the experimental sessions, 5 sets of data were acquired. First, the training data were collected and analyzed by tallying the number of correct and wrong answers for each set of stimuli. Second, to determine the performance of the emotion recognition software, video recordings and emotion logs were saved and reviewed. The actor and the participant were filmed from 3 different camera positions (1 ceiling camera aimed at the participant, and 1 stationary camera aimed at the participant,

and 1 ceiling camera focused on the actor). The emotion logs kept a record of the emotion recognition software by saving timestamps together with the strength of a detected emotion (a value between 0 and 1; an example of such a log can be found in [Table 4](#)). Each time the software detected an emotion above threshold for at least 5 timestamps within a time span of 1 second, a fragment was labeled as an emotion signal. Due to technical issues, the emotion logs of only 6 out of the 8 experiments were analyzable.

**Table 4.** Example of an emotion log. The emotion log provides a value between 0 and 1 for each emotion at each timestamp (hours [hh]:minutes [mm]:seconds[ss]:milliseconds[ms]).

hh	mm	ss	ms	Neutral	Happiness	Sadness	Anger	Surprise	Fear	Disgust
14	1	24	226	0.848577	0.261171	0.045551	0.037921	0.066328	0.101152	0.034363
14	1	24	279	0.853517	0.246223	0.052067	0.039493	0.064576	0.089338	0.034737
14	1	24	354	0.860102	0.227184	0.061851	0.045071	0.060259	0.075662	0.033203
14	1	24	439	0.868132	0.206724	0.068306	0.054500	0.054876	0.063745	0.032779
14	1	24	550	0.875653	0.181231	0.069503	0.060954	0.048684	0.050882	0.034567
14	1	24	617	0.866770	0.157769	0.065179	0.063756	0.041683	0.042549	0.050491
14	1	24	673	0.850679	0.149551	0.064955	0.064503	0.038547	0.039288	0.058762
14	1	24	742	0.834428	0.144484	0.065088	0.065431	0.036603	0.038934	0.061535
14	1	24	798	0.807851	0.137719	0.065118	0.067514	0.033447	0.040757	0.061940

The third dataset consisted of the relationship between the emotion log fragments and the video recordings. For every instance FaceReader recognized an emotion, as was derived from the emotion logs, a video fragment was created, resulting in 166 video fragments that required annotation across all videos. A total of 2 independent coders then analyzed each fragment to determine if they recognized a facial expression of any of the 6 universal emotions (happiness, sadness, anger, fear, surprise, and disgust) and annotated the video fragments accordingly. In cases where the coders did not identify 1 of these emotions (be it seeing another emotion or seeing no emotion at all), the fragment was coded as “none/other.” Eventual disagreements were discussed in person to reach consensus.

To analyze how well the software detected emotions, a fourth set was used. This set consisted of a sample of every third minute that was extracted from the videos. This resulted in a total of 29 min of video that were analyzed for facial expressions of emotions, independent of the emotion logs. A first coder analyzed the videos and identified fragments in which facial expressions of emotions were thought to be visible. A second coder analyzed the same sample of randomly selected videos while seeing the selected fragments by the first coder. The second coder checked whether there was an agreement with the expressions identified by the first coder or if new expression were to be added. Again, any disagreements were discussed to reach consensus. Finally, the fragments in which facial expressions were detected by the 2 coders were compared with

the emotion logs to analyze how well the system detected facial expressions of emotions during the conversations.

The fifth and final set were the exit interview transcriptions. The exit interviews were recorded and transcribed verbatim, which resulted in 70 pages of text. A content analysis was performed to extract the most important information fragments from the transcriptions. Codes were assigned to all fragments, based on the question to which the fragments provided an answer. To structure the answers, the subthemes mentioned in [Table 3](#) were assigned to the answers.

## Results

### Learn to Interpret Vibrotactile Signals

Both training sessions successfully taught participants how to interpret the vibrotactile signals. Participants were quickly able to correctly identify most stimuli. The average performance gradually improved over each set of stimuli. Moreover, 6 out of the 8 participants achieved perfect scores in the final set ([Table 5](#)).

The refresher session, right before the experiment, showed even better results, ending with a perfect recognition rate for all participants in the third and final set ([Table 6](#)).

It can be concluded that during the training session, most participants were able to learn how to identify the correct emotions from the vibrotactile signals.

**Table 5.** Number of errors during the initial training. The table shows the numbers of errors of all participants during the initial training session.

ID	Set 1	Set 2	Set 3		Set 4	Set 5	Total	
	A <sup>a</sup> (n=36)	A (n=18)	T <sup>b</sup> (n=18)	A (n=12)	T (n=24)	T (n=36)		T (n=20-96)
1	3	0	4	0	0	1	0	8
2 <sup>c</sup>	0	0	0	0	0	0	0	0
3 <sup>c</sup>	10	3	0	1	0	1	7	22
4	0	0	0	0	1	0	0	1
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	1	0	2	0	0	0	0	3
8	0	0	1	1	0	1	2	5
Total	14	3	7	2	1	3	9	39

<sup>a</sup>A: auditory accompanied vibrotactile stimuli.

<sup>b</sup>T: tactile only stimuli.

<sup>c</sup>No emotion logs were available for this participant in the experiment.

**Table 6.** Number of errors during the refresher training session. The table shows the number of errors of all participants during the refresher training session.

ID	Set 1, A <sup>a</sup> (n=24)	Set 2, T <sup>b</sup> (n=24)	Set 3, T (n=20-96)	Total
1	1	0	0	1
2 <sup>c</sup>	0	1	0	1
3 <sup>c</sup>	6	0	0	6
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	2	0	0	2
Total	9	1	0	10

<sup>a</sup>A: auditory accompanied vibrotactile stimuli.

<sup>b</sup>T: tactile only stimuli.

<sup>c</sup>No emotion logs were available for this participant in the experiment.

## Camera Pointing

A prerequisite for the emotion recognition system is that participants can focus the camera on the face of the conversation partner. Generally, the participants were able to focus the camera on the actor who sat on the other side of the table. In 5 out of the 6 videos analyzed, the face of the actor was out of sight for only a couple of seconds out of the approximately 15 min that the conversation lasted. Moreover, 1 participant was an exception to this rule. This participant tended to slightly lift her head upward. As a result, the face of the actor was out of the field of view of the camera many times, rendering the emotion recognition software useless in our context.

## Emotion Recognition Performance

### Agreement With Software-Detected Emotions

This analysis was based on the video fragments in which the software had detected an emotion. A total of 2 coders annotated 166 fragments for which the emotion log values exceeded threshold levels for both emotion value and duration. Initially, the coders disagreed 30 times, but mutual consensus was achieved after reviewing the fragments together.

The agreement between the codes assigned by the coders and emotion recognition software was poor. In only 54 out of the 166 fragments, the coders and FaceReader classified the same universal emotions (Table 7).

In 13 cases the coders agreed upon seeing 1 of the 6 universal emotions but differed from the emotion detected by the software. For most fragments (99), however, the coders disagreed with

the software in the sense that the software detected a universal emotion, whereas the coders did not. In this analysis, the coders and the software showed most agreement on the emotion happiness, whereas the performance for other emotions is far

worse. It can be concluded that the coders disagreed with the signals that FaceReader conveyed for 5 out of the 6 emotions. Only when it came to the signaling of happiness, the coders largely agreed with the software.

**Table 7.** Crosstabs of agreement between coders and software. The table shows a tally of the number of time the coders and FaceReader classified a fragment as a particular emotion. The diagonal shows the number of times that the coders and FaceReader classified a fragment as the same emotion.

Facereader	Coders							Total
	Anger	Disgust	Happiness	Sadness	Fear	Surprise	None/other	
Anger	2	0	0	0	0	0	8	10
Disgust	1	0	0	0	0	0	12	13
Happiness	0	0	42	0	0	1	3	46
Sadness	1	0	2	4	0	1	41	49
Fear	0	0	2	0	0	1	4	7
Surprise	0	0	4	0	0	6	31	41
Total	4	0	50	4	0	9	99	166

### System Performance to Detect Facial Expressions of Emotions

To determine how well the system derived facial expressions from the actor, we checked the random sample of every third minute. In this time, the 2 coders detected 72 facial expressions of basic emotions and another 12 expressions not necessarily being associated with 1 of Ekman's emotions. The software detected only 44 in the same time frame. In 20 cases there was an agreement between the human coders and the system detection. A total of 24 times the human coders and FaceReader disagreed about emotions being expressed. The remaining 51

facial expressions of universal emotions that were detected by the coders (31 instances of happiness and 20 instances of surprise) were not recognized by FaceReader (Table 8).

These results show that in the random sample of videos, mostly happiness and surprise were expressed according to the coders. In the 44 cases where the software did detect an emotion, the coders disagreed in almost half of the cases. The only exception seemed to be happiness, which, if it was detected, was agreed upon in 17 of 19 cases. However, the coders detected many more instances of happiness that the software was unable to detect. This was similar for facial expressions of surprise.

**Table 8.** Overview of the agreement between human coders and the software for the facial expressions detected. The table shows the classification of detected emotions.

Emotions	Agreement	Disagreement	Not detected by FaceReader <sup>a</sup>
Anger	0	5	0
Happiness	17	2	31
Disgust	0	2	0
Sadness	0	10	0
Fear	0	2	0
Surprise	3	3	20
Total	20	24	51

<sup>a</sup>The coders identified 11 other facial expressions that did not classify as 1 of the universal emotions.

### User Experience

Besides the objective measures to analyze the technical performance of the system in a realistic condition, we asked the participants how was their experience using the system and whether they saw potential utility of the system in their daily lives.

How did persons perceive the vibrotactile cues during the conversation?:

*There was no doubt about the transferred information. Once you know the location of emotions, the usage*

*of the system is very convenient.* [P2, male, aged 8 years, fully blind]

A total of 4 participants even believed they were able to subconsciously interpret the vibrotactile signals:

*I am consciously participating in the conversation while I am taking the feedback subconsciously into account, so the vibrations did not distract me, absolutely not.* [P4, female, aged 59 years, fully blind]

Moreover, 3 others found it difficult to focus their attention on the signals and struggled to interpret them. Participant 2 (male, aged 28 years, fully blind) felt that the prolonged stimuli really

disrupted his ability to engage in the conversation with the actor. Participant 5 (male, aged 47 years, partially sighted) found it particularly difficult to interpret the signal while he was speaking himself, although he felt that the interpretation of “happiness” and “surprise” was easier than the other emotions as these occurred more frequently than other emotions. Indeed, during his conversation, happiness and surprise were conveyed quite often (7 and 5 times, respectively), although sadness was conveyed as often as surprise.

In addition, 3 participants did not realize that vibration strength was linked to the intensity of the expressed emotion and explicitly requested for such a functionality. Participant 1 (male, aged 28 years, partially sighted) suggested to use pulsating vibrations instead to convey intensity. Thus, to convey such information other methods are wished for.

A total of 7 participants were skeptical about the accuracy of the conveyed emotions, as these did not always correspond to the vocal cues and the atmosphere during the conversation. Participant 4 (female, aged 59 years, fully blind), for example, reported to have received signals of a disgusted facial expression, whereas she felt there was absolutely no reason to believe the actor looked disgusted. In the video analysis of her interview with the actor, no facial expressions of disgust were identified by the 2 coders, whereas there were 7 instances where the software conveyed disgust. Participant 1 (male, aged 28 years, partially sighted) felt that the system did consistently provide information at the moments it was most necessary. Moreover, 2 persons reported receiving only limited signals.

Despite the inaccuracy, 6 participants were convinced about the usefulness of the concept of an emotion recognition system. Participant 7 (female, aged 50 years, fully blind) believed such a system would especially be useful in the absence of vocal cues, where it would be the only source of information. In addition, confirmation of the emotions derived from vocal cues led to more confidence. Furthermore, 1 participant (P5, male, aged 47 years, partially sighted) mentioned that the system could be valuable while the user himself is talking, as feedback on emotional expressions of the conversation partner in such situations can make the user more confident about the story he or she is telling. On the downside, 2 persons reported that because of inaccuracies, the system caused more confusion than clarity. Participant 4 (female, aged 59 years, fully blind) mentioned that she believed she could detect faulty signals and therefore was not so much deterred by the errors. It could be that some participants are more willing to accept errors as these did not weigh up against the experienced usefulness of the system.

A total of 4 participants wanted to use the system in their daily life, whereas 3 others wished for more time with the system before they could decide. Moreover, 2 persons were convinced that the system, as it is, would be an added value to their life. In addition, 3 others believed that the system would have an added value after some adjustments.

Besides the obvious accuracy improvements, participants wished for improvements in the wearability of the system and requested additional functionalities. A total of 7 participants said they would like a more portable system, which could be achieved

by going fully wireless. All but 1 participant requested for a camera that is integrated in the spectacles. The 1 participant who did not request for a camera integrated in spectacles (P5, male, aged 47 years, partially sighted) suggested that the camera should be clearly visible to make conversation partners aware they are being filmed. Moreover, 2 participants suggested to get rid of the spectacles completely and allow users to clip the camera anywhere on their clothing.

A total of 7 out of the 8 participants believed that a person recognition function would increase the usefulness of the system. Participant 1 (male, aged 28 years, partially sighted) even stated that emotion recognition and face recognition are inseparable, as in a busy environment you need to know who you look at before emotion recognition becomes relevant. Such technologies could also be applied to locate a specific person that you are looking for. Related to this, participant 2 (male, aged 28 years, fully blind) wished for a function to request for feedback from the system at any time to check whether the system is still on and focused after long periods of inactivity. Such a trigger was also suggested by participant 1 (male, aged 28 years, partially sighted), who wished for a trigger to switch the feedback on or off to avoid an information overload from the system.

## Discussion

### Principal Findings

The main objective of this study was to determine whether the system presented in an earlier study [15], which was successfully tested in laboratory conditions, would work in a realistic conversation situation. The study affirmed that a positive laboratory test does not necessarily mean that a system is ready for real-life usage [26]. Although the concept of our system seems to have potential, it requires essential improvements before it is ready to support visually impaired persons in real life.

We have confirmed findings from a previous study [5] and showed that it is very easy to learn how to interpret vibrotactile cues signaled with a haptic waist belt. That in itself is not a surprise, both older [5,27] and more recent [28,29] studies have shown that much more complex tactile patterns can be successfully learned. Furthermore, it was already known that healthy persons can interpret vibrotactile cues in demanding environments [30]. However, what is interesting is that we now have indications that users of the system with a visual impairment were able to use information conveyed through vibrotactile signals while they engaged in a conversation.

Most of the participants were able to keep the camera focused in the direction of the actor for most, if not all the, time. That means that the principle of putting a camera on spectacles will provide a sufficient camera feed for such computer vision analysis. Moreover, 1 person, however, aimed the camera above the actor for most of the time. Before using such a system, this person would require training on how to aim the camera to benefit from the system. Another option would be to redesign the camera and spectacles in such a way that the camera angle can be altered to match the wearer's preference. This is a feature that some mobile eye trackers already have.



The accuracy of the software was not sufficient to convey all of Ekman's [16] 6 universal emotions under our experimental conditions. Earlier studies on the performance of the software were conducted using validated sets composed of pictures including high-quality full-face pictures in ideal lighting conditions. In such conditions, the software used in the study achieved accuracy scores as high as 88% [17]. In our experiment, however, the software was confronted with suboptimal lighting conditions and unstable shots because of a moving camera and actor. In these conditions, the only emotion that was accurately recognized was happiness, which was also reported by the participants. This is if the emotion was recognized, as the coders detected 31 more instances of a facial expression of happiness that the software did not detect. The other 5 emotions were poorly recognized. As a result, the performance achieved in the experiment was too poor to be a support for visually impaired persons in real life.

It is unlikely that the system presented in the study would be used by the participants in their daily lives in its current form. Too many adjustments were suggested to believe that implementation would be easy (eg, emotion recognition accuracy, portability, person recognition, and on/off switch). Nevertheless, there is reason to believe that such a system might be used, once the required adjustments are adapted.

### Limitations

The emotions were not uniformly represented in the conversations. It is clear that the emotion happiness was far

more often detected and conveyed via vibrotactile feedback than the other emotions. Therefore, it is difficult to draw definitive conclusions on the accuracy of the other emotions.

During the experiment, it was impossible to objectively measure if and how well participants were able to perceive the vibration signals. Therefore, we had to rely on the reports by the participants when it comes to the interpretation of signals.

We chose not to encode any other emotions than the 6 universal emotions. As a result, we cannot provide an overview of all the facial expressions of emotions that were expressed in the conversations. Thus, we cannot state which emotions are more often present in realistic conversations and could possibly be added to the emotion recognition software.

### Conclusions

For systems such as the one presented in this study, the main improvements to be made should be sought in the direction of both the wearability of the system and improvement of the emotion recognition software. The solution for wearability would be to make components of the system smaller (from tablet to smartphone) and by making all connections between system components wireless. To improve emotion recognition software, we suggest training with more ambiguous and subtle facial expressions under different lighting conditions. We do believe that if these concerns are taken care of, systems such as the one presented here might be of great benefit for visually impaired persons in their future daily social interactions.

### Acknowledgments

The authors would like to thank the participants for their enthusiasm and feedback; the professional actor, Wim Bouwens; VicarVision for the use of the FaceReader software and continuous support of Marian Bittner; and Jan Henk Annema and Jurre Brienne for their support and hospitality during the use of the HAN ICA Experience Lab.

### Conflicts of Interest

TK was employed at VicarVision at the time of the study. VicarVision is the company which developed the FaceReader software used in this study. TK did not have a say in the study design, execution, analyses, or outcomes.

### References

1. Vision 2020 NL. 2005. Vermijdbare Blindheid En Slechtziendheid in Nederland URL: <http://www.vision2020.nl/contents/V2020NLrapport.pdf> [accessed 2019-09-10]
2. Schölvinc AM, Pittens CA, Broerse JE. The research priorities of people with visual impairments in the Netherlands. *J Vis Impair Blind* 2018 Dec 13;111(3):201-217 [FREE Full text] [doi: [10.1177/0145482X1711100302](https://doi.org/10.1177/0145482X1711100302)]
3. van der Gees TM, Buimer HP. User-centered priority setting for accessible devices and applications. In: Weisbecker A, Burmester M, Schmidt A, editors. *Mensch und Computer 2015 – Workshopband*. Berlin: de Gruyter Oldenbourg; 2015:383-389.
4. Krishna S, Colbry D, Black J, Balasubramanian V, Panchanathan S. A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. In: *Workshop on Computer Vision Applications for the Visually Impaired*. 2008 Presented at: CVPR'08; October, 2008; Marseille, France URL: <http://hal.inria.fr/inria-00325432> [doi: [10.1177/0145482X1711100302](https://doi.org/10.1177/0145482X1711100302)]
5. Bach-y-Rita P. Sensory plasticity. Applications to a vision substitution system. *Acta Neurol Scand* 1967;43(4):417-426. [doi: [10.1111/j.1600-0404.1967.tb05747.x](https://doi.org/10.1111/j.1600-0404.1967.tb05747.x)] [Medline: [5583047](https://pubmed.ncbi.nlm.nih.gov/5583047/)]
6. Ward J, Meijer P. Visual experiences in the blind induced by an auditory sensory substitution device. *Conscious Cogn* 2010 Mar;19(1):492-500. [doi: [10.1016/j.concog.2009.10.006](https://doi.org/10.1016/j.concog.2009.10.006)] [Medline: [19955003](https://pubmed.ncbi.nlm.nih.gov/19955003/)]
7. Meijer PB. An experimental system for auditory image representations. *IEEE Trans Biomed Eng* 1992 Feb;39(2):112-121. [doi: [10.1109/10.121642](https://doi.org/10.1109/10.121642)] [Medline: [1612614](https://pubmed.ncbi.nlm.nih.gov/1612614/)]

8. Kaczmarek KA. The tongue display unit (TDU) for electrotactile spatiotemporal pattern presentation. *Sci Iran D Comput Sci Eng Electr Eng* 2011 Dec;18(6):1476-1485 [FREE Full text] [doi: [10.1016/j.scient.2011.08.020](https://doi.org/10.1016/j.scient.2011.08.020)] [Medline: [28748231](https://pubmed.ncbi.nlm.nih.gov/28748231/)]
9. McDaniel R, Bala S, Rosenthal J, Tadayon R, Tadayon A, Panchanathan S. Affective Haptics for Enhancing Access to Social Interactions for Individuals Who are Blind. In: *Proceedings of the International Conference on Universal Access in Human-Computer Interaction*. 2014 Presented at: UAHCI'14; June 22-27, 2014; Heraklion, Crete, Greece p. 419-429. [doi: [10.1007/978-3-319-07437-5\\_40](https://doi.org/10.1007/978-3-319-07437-5_40)]
10. McDaniel T, Tran D, Chowdhury A, Fakhri B, Panchanathan S. Recognition of tactile facial action units by individuals who are blind and sighted: a comparative study. *Multimodal Technol Interact* 2019 May 8;3(2):32 [FREE Full text] [doi: [10.3390/mti3020032](https://doi.org/10.3390/mti3020032)]
11. Krishna S, Bala S, McDaniel T, McGuire S, Panchanathan S. Vibroglove: An Assistive Technology Aid for Conveying Facial Expressions. In: *Extended Abstracts on Human Factors in Computing Systems*. 2010 Aug Presented at: CHI EA'10; April 10-15, 2010; Atlanta, Georgia, USA p. 3637-3642. [doi: [10.1145/1753846.1754031](https://doi.org/10.1145/1753846.1754031)]
12. McDaniel TL, Krishna S, Colbry D, Panchanathan S. Using Tactile Rhythm to Convey Interpersonal Distances to Individuals Who Are Blind. In: *Extended Abstracts on Human Factors in Computing Systems*. 2009 Presented at: CHI EA'09; April 4-9, 2009; Boston, MA, USA p. 4669-4674. [doi: [10.1145/1520340.1520718](https://doi.org/10.1145/1520340.1520718)]
13. McDaniel T, Krishna S, Balasubramanian V, Colbry D, Panchanathan S. Using a Haptic Belt to Convey Nonverbal Communication Cues During Social Interactions to Individuals Who Are Blind. In: *Proceedings of the International Workshop on Haptic Audio visual Environments and Games*. 2008 Presented at: HAVE'08; October 18-19, 2008; Ottawa, ON, Canada p. 13-18. [doi: [10.1109/HAVE.2008.4685291](https://doi.org/10.1109/HAVE.2008.4685291)]
14. Bartlett MS, Hager JC, Ekman P, Sejnowski TJ. Measuring facial expressions by computer image analysis. *Psychophysiology* 1999 Mar;36(2):253-263. [doi: [10.1017/s0048577299971664](https://doi.org/10.1017/s0048577299971664)] [Medline: [10194972](https://pubmed.ncbi.nlm.nih.gov/10194972/)]
15. Buimer HP, Bittner M, Kostelijik T, van der Geest TM, Nemri A, van Wezel RJ, et al. Conveying facial expressions to blind and visually impaired persons through a wearable vibrotactile device. *PLoS One* 2018;13(3):e0194737 [FREE Full text] [doi: [10.1371/journal.pone.0194737](https://doi.org/10.1371/journal.pone.0194737)] [Medline: [29584738](https://pubmed.ncbi.nlm.nih.gov/29584738/)]
16. Ekman P. Universal facial expressions of emotion. *Calif Ment Heal Res Dig* 1970;8(4):151-158 [FREE Full text]
17. Lewinski P, den Uyl TM, Butler C. Automated facial coding: validation of basic emotions and FACS AUs in FaceReader. *J Neurosci Psychol Econ* 2014 Dec;7(4):227-236. [doi: [10.1037/npe0000028](https://doi.org/10.1037/npe0000028)]
18. den Uyl MJ, van Kuilenburg J, Lebert E. The FaceReader : online facial expression recognition. In: Noldus LP, Grieco F, Loijens LW, Zimmerman PH, editors. *Proceedings of Measuring Behavior*. Netherlands: Wageningen Publishers; 2005:589-590.
19. van Erp JB, van Veen HA, Jansen C, Dobbins T. Waypoint navigation with a vibrotactile waist belt. *ACM Trans Appl Percept* 2005 Apr 1;2(2):106-117. [doi: [10.1145/1060581.1060585](https://doi.org/10.1145/1060581.1060585)]
20. Olszanowski M, Pochwatko G, Kuklinski K, Scibor-Rylski M, Lewinski P, Ohme RK. Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. *Front Psychol* 2014;5:1516 [FREE Full text] [doi: [10.3389/fpsyg.2014.01516](https://doi.org/10.3389/fpsyg.2014.01516)] [Medline: [25601846](https://pubmed.ncbi.nlm.nih.gov/25601846/)]
21. van der Schalk J, Hawk ST, Fischer AH, Doosje B. Moving faces, looking places: validation of the Amsterdam dynamic facial expression set (ADFES). *Emotion* 2011 Aug;11(4):907-920. [doi: [10.1037/a0023853](https://doi.org/10.1037/a0023853)] [Medline: [21859206](https://pubmed.ncbi.nlm.nih.gov/21859206/)]
22. Dhall A, Goecke R, Lucey S, Gedeon T. Static Facial Expression Analysis in Tough Conditions: Data, Evaluation Protocol and Benchmark. In: *Proceedings of the International Conference on Computer Vision Workshops*. 2011 Presented at: ICCV'11; November 6-13, 2011; Barcelona, Spain p. 2106-2112. [doi: [10.1109/ICCVW.2011.6130508](https://doi.org/10.1109/ICCVW.2011.6130508)]
23. Kharat GU, Dudul SV. Emotion Recognition from Facial Expression Using Neural Networks. In: *Proceedings of the Human-Computer Systems Interaction*. 2008 Presented at: HCI'2008; May 25-27, 2008; Krakow, Poland p. 207-219. [doi: [10.1109/HSI.2008.4581476](https://doi.org/10.1109/HSI.2008.4581476)]
24. Mehta D, Siddiqui MF, Javaid AY. Facial emotion recognition: a survey and real-world user experiences in mixed reality. *Sensors (Basel)* 2018 Feb 1;18(2):E416 [FREE Full text] [doi: [10.3390/s18020416](https://doi.org/10.3390/s18020416)] [Medline: [29389845](https://pubmed.ncbi.nlm.nih.gov/29389845/)]
25. Dhall A, Goecke R, Gedeon T, Sebe N. Emotion recognition in the wild. *J Multimodal User In* 2016 Mar 3;10(2):95-97. [doi: [10.1007/s12193-016-0213-z](https://doi.org/10.1007/s12193-016-0213-z)]
26. Pawluk DT, Adams RJ, Kitada R. Designing haptic assistive technology for individuals who are blind or visually impaired. *IEEE Trans Haptics* 2015;8(3):258-278. [doi: [10.1109/TOH.2015.2471300](https://doi.org/10.1109/TOH.2015.2471300)] [Medline: [26336151](https://pubmed.ncbi.nlm.nih.gov/26336151/)]
27. White B. Perceptual findings with the vision-substitution system. *IEEE Trans Man Mach Syst* 1970 Mar;11(1):54-58. [doi: [10.1109/TMMS.1970.299962](https://doi.org/10.1109/TMMS.1970.299962)]
28. McDaniel T, Tran D, Devkota S, DiLorenzo K, Fakhri B, Panchanathan S. Tactile Facial Expressions and Associated Emotions toward Accessible Social Interactions for Individuals Who Are Blind. In: *Proceedings of the 2018 Workshop on Multimedia for Accessible Human Computer Interface*. 2018 Presented at: MAHCI'18; October 22, 2018; Seoul, Republic of Korea p. 25-32. [doi: [10.1145/3264856.3264860](https://doi.org/10.1145/3264856.3264860)]
29. Bala S, McDaniel T, Panchanathan S. Visual-to-Tactile Mapping of Facial Movements for Enriched Social Interactions. In: *Proceedings of the International Symposium on Haptic, Audio and Visual Environments and Games*. 2014 Presented at: HAVE'14; October 10-11, 2014; Richardson, TX, USA p. 82-87. [doi: [10.1109/HAVE.2014.6954336](https://doi.org/10.1109/HAVE.2014.6954336)]

30. van Erp J. Presenting directions with a vibrotactile torso display. *Ergonomics* 2005 Feb 22;48(3):302-313. [doi: [10.1080/0014013042000327670](https://doi.org/10.1080/0014013042000327670)] [Medline: [15764328](https://pubmed.ncbi.nlm.nih.gov/15764328/)]

## Abbreviations

**ADFES:** Amsterdam Dynamic Facial Expression Set

**SSD:** sensory substitution device

**WSEFEP:** Warsaw Set of Emotional Facial Expression Pictures

*Edited by G Eysenbach; submitted 15.02.19; peer-reviewed by N Ahmadpour, R Dewey; comments to author 17.05.19; revised version received 12.07.19; accepted 19.07.19; published 21.11.19.*

*Please cite as:*

*Buimer H, Schellens R, Kostelijk T, Nemri A, Zhao Y, Van der Geest T, Van Wezel R*

*Opportunities and Pitfalls in Applying Emotion Recognition Software for Persons With a Visual Impairment: Simulated Real Life Conversations*

*JMIR Mhealth Uhealth* 2019;7(11):e13722

URL: <https://mhealth.jmir.org/2019/11/e13722>

doi: [10.2196/13722](https://doi.org/10.2196/13722)

PMID: [31750838](https://pubmed.ncbi.nlm.nih.gov/31750838/)

©Hendrik Buimer, Renske Schellens, Tjerk Kostelijk, Abdellatif Nemri, Yan Zhao, Thea Van der Geest, Richard Van Wezel. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 21.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mhealth and uhealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

# Real-Time Mobile Monitoring of Drinking Episodes in Young Adult Heavy Drinkers: Development and Comparative Survey Study

Daniel J Fridberg<sup>1</sup>, PhD; James Faria<sup>1</sup>, BA; Dingcai Cao<sup>2</sup>, PhD; Andrea C King<sup>1</sup>, PhD

<sup>1</sup>Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, IL, United States

<sup>2</sup>Department of Ophthalmology, The University of Illinois at Chicago, Chicago, IL, United States

**Corresponding Author:**

Daniel J Fridberg, PhD

Department of Psychiatry and Behavioral Neuroscience

The University of Chicago

5841 S Maryland Ave

MC 3077

Chicago, IL, 60637

United States

Phone: 1 7738343598

Email: [dfridberg@bsd.uchicago.edu](mailto:dfridberg@bsd.uchicago.edu)

## Abstract

**Background:** Binge drinking, defined as consuming five or more standard alcoholic drinks for men (four for women) within a 2-hour period, is common among young adults and is associated with significant alcohol-related morbidity and mortality. To date, most research on this problem in young adults has relied upon retrospective questionnaires or costly laboratory-based procedures. Smartphone-based ecological momentary assessment (EMA) may address these limitations by allowing researchers to measure alcohol use and related consequences in real time and in drinkers' natural environments. To date, however, relatively less research has systematically examined the utility of this approach in a sample of young adults targeting real-world heavy drinking episodes specifically.

**Objective:** This study aimed to evaluate the feasibility, acceptability, and safety of a smartphone-based EMA method targeting binge drinking and related outcomes in heavy drinking young adults during real-world drinking occasions.

**Methods:** Young adult binge drinkers in the smartphone group (N=83; mean 25.4 (SD 2.6) years; 58% (48/83) male; bingeing on 23.2% (6.5/28) days in the past month) completed baseline measures of alcohol use and drinking-related consequences, followed by up to two smartphone-based EMA sessions of typical drinking behavior and related outcomes in their natural environments. They also completed next-day and two-week follow-up surveys further assessing alcohol use and related consequences during the EMA sessions and two weeks after study participation, respectively. A separate demographic- and drinking-matched safety comparison group (N=25) completed the baseline and two-week follow-up surveys but did not complete EMA of real-world drinking behavior.

**Results:** Most participants (71%, 59/83) in the smartphone group engaged in binge drinking during at least one 3-hour EMA session, consuming 7.3 (SD 3.0) standard alcoholic drinks. They completed 87.2% (507/581) system-initiated EMA prompts during the real-world drinking episode, supporting the feasibility of this approach. The procedure was acceptable, as evidenced by high participant ratings for overall satisfaction with the EMA software and study procedures and low ratings for intrusiveness of the mobile surveys. Regarding safety, participants endorsed few drinking-related consequences during or after the real-world drinking episode, with no adverse or serious adverse events reported. There were no differences between the groups in terms of changes in drinking behavior or consequences from baseline to two-week follow-up.

**Conclusions:** This study provided preliminary support for the feasibility, acceptability, and safety of a smartphone-based EMA of real-time alcohol use and related outcomes in young adult heavy drinkers. The results suggest that young adults can use smartphones to safely monitor drinking even during very heavy drinking episodes. Smartphone-based EMA has strong potential to inform future research on the epidemiology of and intervention for alcohol use disorder by providing researchers with an efficient and inexpensive way to capture large amounts of data on real-world drinking behavior and consequences.

(*JMIR Mhealth Uhealth* 2019;7(11):e13765) doi:[10.2196/13765](https://doi.org/10.2196/13765)

**KEYWORDS**

ecological momentary assessment; young adults; binge drinking

## Introduction

### Background

Binge drinking, also known as heavy episodic drinking, is a significant public health problem in the United States and accounts for more than half of the 88,000 alcohol-related deaths in the United States each year [1,2]. Approximately 25% of young adults aged 18 to 34 years report past-month binge drinking, defined as consuming 5 or more standard alcoholic drinks for men (4 for women) within a 2-hour period [3], which is the highest prevalence rate across the life span [1]. These individuals are at significant risk for developing alcohol use disorder (AUD) and morbidity and mortality related to excessive alcohol consumption [4]. Thus, improving our knowledge of the factors that maintain binge drinking behavior in young adults is critical to developing new interventions for AUD. To date, most research on experiences of young adults during binge drinking has either used retrospective surveys (eg, [5]) or laboratory-based alcohol challenge paradigms (eg, [6-8]); however, these approaches are limited by recall bias and unclear ecological validity, respectively. One potential solution to these limitations is the use of ecological momentary assessment (EMA), which allows researchers to use technology such as mobile phones to measure participants' real-world behaviors in real time and in their natural environments [9]. In this study, we report on the feasibility, acceptability, and safety of smartphone-based EMA targeting binge drinking and related outcomes in young adult heavy drinkers' natural environments.

Prior research has used EMA to measure real-world light-to-moderate alcohol use (ie, 2-4 standard drinks consumed on average) with good compliance across studies (ie, 78-90% of survey prompts completed) [10-15]. EMA methodologies in drinkers have varied from random prompts of mood and behaviors throughout the day to prompts specifically targeting drinking episodes. Most of this work has been conducted in general samples of young adult drinkers and not specifically in hazardous drinkers during binge drinking episodes specifically, although prior studies have captured EMA data on heavy drinkers and binge episodes through recruitment of broader samples of young adult drinkers generally (refer to the studies by, eg, Kuntsche and Labhart, Piasecki et al, and Groefsema et al [12,14,16]). Although EMA methods are acceptable in young social drinkers [13], with little [13] or no effect [11] on subsequent drinking behavior (ie, reactivity to the assessment method), it is unclear whether such acute monitoring of real-world drinking will also be feasible and safe in young adult chronic heavy drinkers who regularly drink to intoxication. To advance the field and increase the clinical relevance of EMA methods of monitoring drinking in hazardous users, additional work is needed on the acceptability of the procedure and reactivity to EMA in young adult heavy drinkers and others at risk for AUD [17].

Given our current era of rapid advances in mobile technology and changing user preferences, the modality of alcohol

monitoring is important. Most prior EMA studies of drinking have employed either palmtop computers [14] or older cellular phone technology such as texting (eg, [12]). However, recent advancements in smartphone technology and the popularity of these devices (94% of young adults aged 18-29 years own a smartphone; [18]) have rendered these older methods obsolete and ushered in a new era for user-friendly, real-time monitoring of participants' real-world experiences. To our knowledge, however, no previous studies have focused specifically on the ability of smartphone-based EMA to assess real-time drinking behavior and related outcomes in young adult heavy drinkers during binge drinking episodes. Although modern smartphones have several advantages over prior EMA technology in terms of ease of use and familiarity, they also offer numerous distractions competing for users' attention. As such, it is possible that they may not be a useful platform for assessing alcohol use and other outcomes during a heavy drinking episode.

### Objectives

Therefore, the purpose of this study was to evaluate the feasibility, acceptability, and safety of a real-time smartphone-based EMA method to obtain self-reported alcohol use and related outcomes (drinking context and subjective alcohol responses) during real-world drinking events in young adult heavy drinkers. We targeted the EMA procedure in this study toward binge drinking episodes specifically because previous work in this field has focused mainly on capturing drinking behavior generally over a period, for example, several days or weeks [14,16,19], rather than binge drinking outcomes per se, and because more data are needed to establish the acceptability, feasibility, and safety of EMA of binge drinking episodes specifically. We hypothesized that the method would be feasible, as evidenced by good compliance (ie,  $\geq 80\%$  response rate of survey prompts), and acceptable, as evidenced by high reported satisfaction with study participation and low perceived intrusiveness of the mobile assessments. We also evaluated the safety of the mobile method by examining participants' drinking behaviors and related consequences before and several weeks after completing the assessments relative to a comparison group who did not undergo monitoring of binge drinking.

## Methods

### Design

The Mobile Alcohol Response Study was conducted between April 2017 and June 2018 and employed a within-subject design with each participant undergoing 1 or 2 separate real-time smartphone assessments of typical drinking episodes in their natural environment. As stated above, a separate safety comparison sample completed baseline and follow-up surveys of drinking behavior and related outcomes but did not undergo real-time mobile monitoring. As part of a larger study, participants also completed a laboratory session with alcohol administration to validate the smartphone-based EMA method; those data will be reported elsewhere. All study procedures were

approved by the University of Chicago Institutional Review Board.

## Participants

Candidates were recruited from Web-based advertisements and word-of-mouth referrals. Inclusion criteria were age 21 to 29 years, generally healthy young adults with weekly heavy alcohol consumption ( $\geq 5$  drinks in 1-4 weekly occasions for men and  $\geq 4$  drinks for women; [3]) for at least the past year, and consumption of  $\geq 14$  drinks per week for men or  $\geq 7$  drinks per week for women, similar to previous studies by our group [6,20-22]. Candidates meeting the basic study criteria from telephone screening were invited to an in-person visit to confirm their eligibility for the study. They were instructed to abstain from alcohol and recreational drugs for 24 hours before screening, and breath alcohol concentration (BrAC) and urine drug toxicology tests (cocaine, amphetamines, methamphetamines, opioids, and benzodiazepines) corroborated recent abstinence from these substances, with the exception of 1 candidate who tested positive for benzodiazepine use and was excluded from further participation. As approximately one-third of young adult binge drinkers report using cannabis [23], current cannabis use was not an exclusion criterion, as long as the reported frequency did not exceed 3 times per week, the candidate agreed to refrain from using it during the real-world mobile drinking monitoring sessions, and they did not meet Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria for cannabis use disorder.

The screening visit included informed consent, an explanation of study procedures, surveys, and interviews conducted by a trained research assistant. The surveys included the Alcohol Use Disorder Identification Test [24], the Beck Depression Inventory [25] and Spielberger Trait Anxiety Inventory [26] to assess current symptoms of depression and anxiety, and a demographic and health history questionnaire. Interviews included the Timeline Followback (TLFB) calendar [27] for past-month alcohol drinking, the Alcohol Quantity-Frequency Interview [28] for typical and maximum drinking over the past 6 months, and the Structured Clinical Interview for the DSM-5, Research Version [29]. Candidates were excluded if they met criteria for severe AUD or a major psychiatric disorder. Identical recruitment methods were employed to enroll 25 participants for the safety comparison group.

Of the 103 candidates screened, 93 (90%) were deemed eligible to participate, and the majority of them (94%, 87/93) completed at least one real-time assessment for 3 hours when drinking.

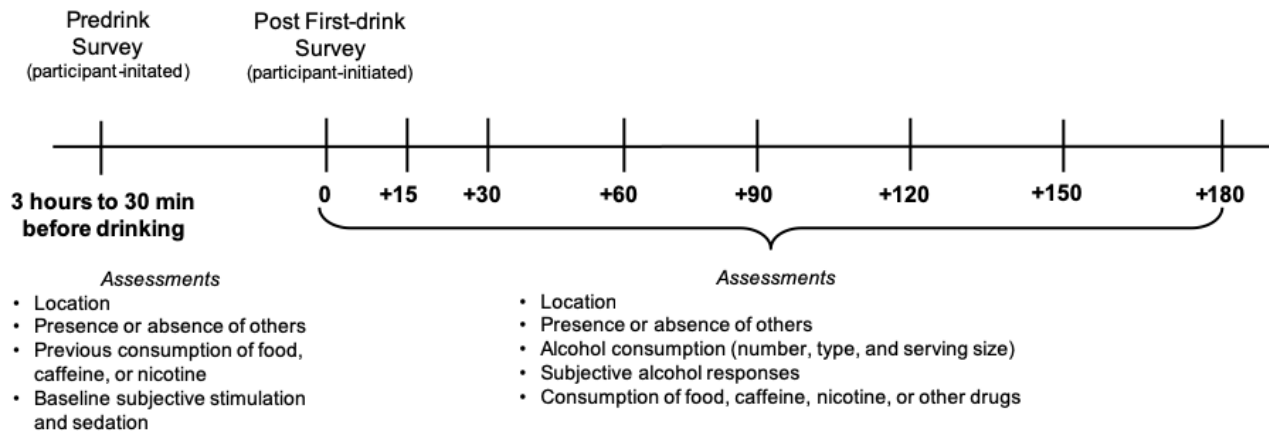
## Procedure

After eligibility determination at study screening, the research assistant installed the mobile EMA app (MetricWire, Inc) on each participants' personal Android or iPhone operating system smartphone. During study orientation, each participant was trained on using the app, setting alert notifications for survey

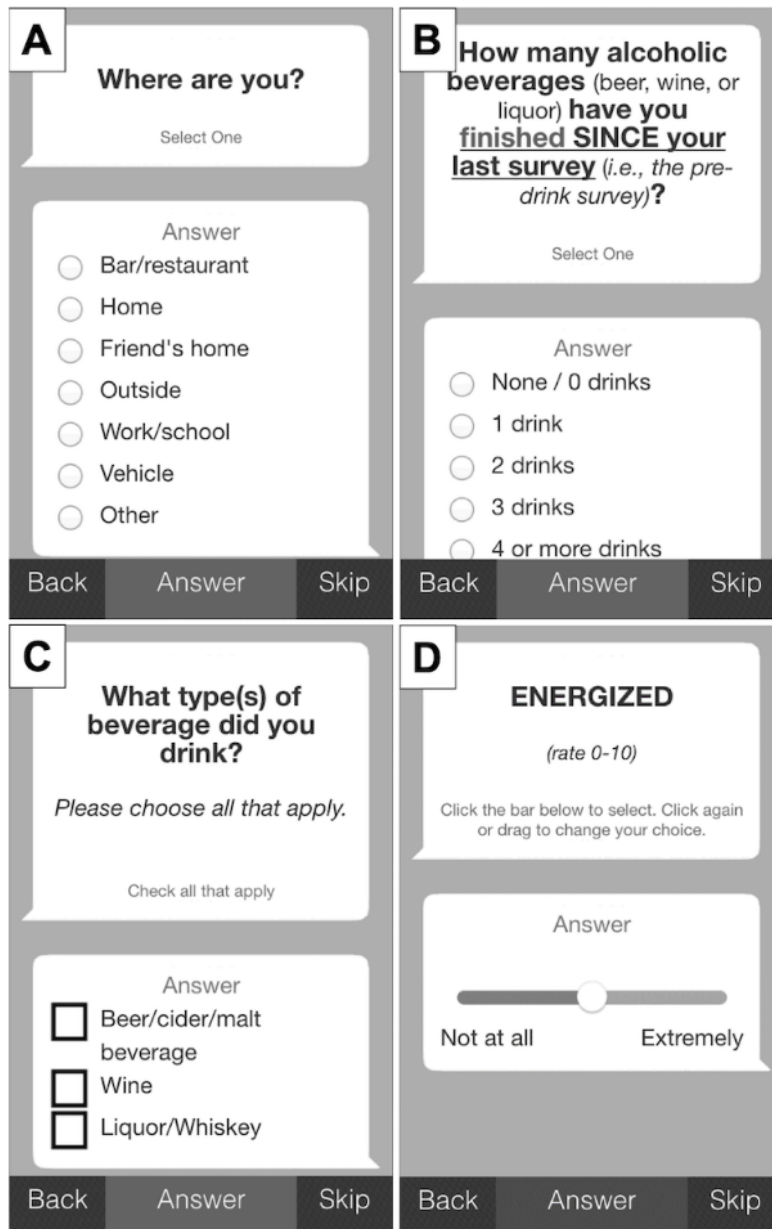
prompts, and employing methods to preserve confidentiality. The purpose of the assessments was explained as assessing one's real-time drinking behavior, subjective responses, and contextual factors (eg, location and presence of others.) for 1 or 2 drinking episodes. The participant was advised to engage in the real-time mobile monitoring during a typical drinking occasion and refrain from smoking or recreational drugs (including cannabis) during that episode. He/she was also trained on properly classifying the types of beverages consumed during mobile monitoring as "beer," "wine," or "liquor" using the EMA app. The TLFB calendar was used as a guide for days of the week each participant was most likely to engage in binge drinking, and he or she was encouraged to complete the EMA session(s) during those days. However, for safety and ethical reasons, the participant was not instructed or encouraged to consume a specific amount of alcohol during their future drinking episode(s). He/she was also reminded not to drive or operate machinery during or after the drinking episode and, as a safety measure, was provided financial compensation for taxi/rideshare at the end of the real-time assessment, if needed.

Figure 1 depicts the timeline for the smartphone assessments. The participant was instructed to self-initiate the mobile assessment procedure by completing a predrinking baseline survey on his or her mobile device via the smartphone app 30 to 180 min before drinking alcohol. The purpose of this survey was to capture predrinking subjective responses; these outcomes will be the subject of a separate report. The participant then self-initiated the first survey after finishing their first drink. All subsequent surveys were delivered automatically to the participant's smartphone via the EMA app at 15, 30, 60, 90, 120, 150, and 180 min thereafter. Although some previous studies have used longer EMA sessions to capture drinking behavior in young adults (eg, [14]), we opted to limit EMA session length to 3 hours in this study for 2 main reasons. First, the relatively short duration was intended to limit participant burden related to completing multiple surveys over 2 drinking episodes. Second, based on the typical pattern of alcohol consumption reported by young adult heavy drinkers in our previous studies [6,20], we expected that this assessment period would be sufficiently long to capture drinking behavior and subjective effects corresponding to ascending, peak, and early descending blood alcohol concentration (BAC) for a large proportion of participants. Of note, studies using EMA to track real-world drinking behavior and related subjective outcomes in young adults have varied in terms of the duration and frequency of monitoring [13,14] and, to date, no empirical studies have established best practice guidelines for EMA of alcohol use and subjective responses in that population [17]. The mobile surveys administered during the smartphone-based EMA sessions were designed to be brief (approximately 1 min in length) and user-friendly. As shown in Figure 2, participants used simple touch screen controls (eg, radio buttons and sliders) to enter and submit their responses to each item.

**Figure 1.** Timeline of the real-time monitoring of a drinking episode.



**Figure 2.** Screenshots of the smartphone interface and sample items from the real-time mobile assessments, including menu-based choices to record contextual factors (eg, (A), drinking location; (B), alcohol quantity; (C), beverage type; and (D), slider-based input to record alcohol responses).



After the final survey at 180 min, the participant received a message via the smartphone app thanking them for their time and reminding them that they would be prompted to complete a next-day survey the following day. This survey was prompted automatically at 11:00 am the next day by the smartphone app. After completing the next-day survey following the first mobile drinking assessment, the participant was offered the option to complete a second mobile assessment on a separate drinking occasion for additional compensation. The second mobile assessment was identical to the first assessment and was completed at least 24 hours after the first drinking occasion. Most participants (61/83, 73%) completed a second mobile assessment, approximately 1 to 2 weeks after the first assessment (mean 11.9, SD 11.4 days). All data from the mobile assessments were uploaded wirelessly from participants' smartphones to a secure server for analysis.

Moreover, 2 weeks after the final EMA of a drinking episode, study staff emailed the participant a link to a follow-up survey assessing drinking behavior and related consequences for the interval since completing the last real-time assessment (for details, refer to the Measures section). The participant was then compensated for his/her participation and debriefed.

## Measures

### Smartphone Drinking Episode Assessments

The first item of the predrinking baseline survey asked whether alcohol had been consumed that day. If the response was yes, then the assessment did not proceed and the participant was reminded that the baseline survey needed to be completed before drinking was initiated, and to try back again in  $\geq 24$  hours. If the response was no, that is, no alcohol consumed yet, then the 1-min baseline survey proceeded to assess context (current location; presence/absence of others; and consumption of food, nonalcoholic beverages, caffeine, nicotine, and other drugs up to the time of the survey) and baseline subjective measures.

For the post first-drink survey and 7 subsequent follow-up surveys, the participant was asked to select the alcohol type (beer, wine, or liquor), number of drinks (ie, "0," "1," "2," "3," or "4 or more") finished since the last survey, and the size of drinks consumed. For beer, drink size options were "less than 12 oz.," "12 oz. (bottle or can)," "16 oz. (pint/pounder)," "24 oz. (tall boy)," "32 oz.," or "40 oz. or more"; for wine, drink size options were "5 oz. (standard glass)," "10 oz.," or "more than 10 oz.," and for liquor, drink size options were "1.5 oz. (1 shot)," "3 oz. (2 shots)," "4.5 oz. (3 shots)," "6 oz. (4 shots)," or "more than 6 oz." Participants were also asked about current contextual variables as in the baseline survey (see prior paragraph). They also completed 10 adjective-based alcohol-subjective effects items [30,31] that will be the subject of a separate report and are not presented here.

### Next-Day Survey

The EMA app issued a survey at 11:00 am following the conclusion of each real-world drinking assessment to capture additional information regarding that drinking episode. This *next day* survey assessed the type/brand of alcohol consumed, duration of continued drinking after the 3-hour EMA period, participants' activities during the assessment (time spent

working, socializing, standing, sitting, dancing or engaging in other physical activity, and playing games; rated from 0="none" to 8="a lot"), and consequences of drinking via a modified 24-item Brief Young Adult Alcohol Consequences Questionnaire (BYAACQ; [32]) assessing drinking-related consequences over the past 24 hours. Finally, an item asked the participant to indicate if the prior days' drinking episode reflected a *typical* drinking occasion (answered "yes" or "no").

### Follow-Up Survey

The follow-up survey included TLFB calendar and modified BYAACQ assessing alcohol use and related consequences during the past 2 weeks, respectively, and 5 items assessing the acceptability of the real-time monitoring method. The acceptability items were as follows: (1) "Overall, the mobile app was easy to use"; (2) "The mobile assessments (surveys) were intrusive"; (3) "The mobile assessments (surveys) were too long"; (4) "I would recommend the study to other potential participants"; and (5) "Overall, I was satisfied with the study experience." Each acceptability item was rated on a 1 to 5 scale, with 1="strongly disagree" and 5="strongly agree." Participants in the safety comparison group completed an identical follow-up survey 2 weeks after their screening, excluding the acceptability items related to the smartphone assessments.

### Statistical Analyses

Feasibility was examined by (1) the percentage of nonparticipant-initiated prompts completed and (2) the estimated number of standard drinks reported at each time point. The number of estimated standard drinks consumed was used to determine the percentage of episodes that included heavy drinking. In addition, this information was used to estimate BAC (estimated BAC, eBAC) levels throughout the drinking episode according to the equation of Matthews and Miller [33]:  $eBAC = [(c/2) \times (CG/w)] - (\beta_{00} \times t)$ , where  $c$  is the number of standard drinks consumed to that point in the drinking episode,  $GC$  is a gender constant (7.5 for men and 9.0 for women),  $w$  is weight in pounds,  $\beta_{00}$  is a constant representing the average population alcohol metabolism rate (0.017 g/dl per hour), and  $t$  is the time in hours since drinking began. We assumed that participants consumed the first drink in the episode over 20 min when calculating  $t$ , as in previous studies [10,14]. This equation approximates actual BrAC and is most accurate at  $BAC \leq 0.08$  g/dL [34].

Acceptability was determined from responses to the satisfaction items from the follow-up survey. Finally, safety of the procedure was examined by comparing drinking behavior and alcohol consequences at baseline versus the 2-week follow-up for the experimental and control group with generalized estimating equation (GEE) analyses examining group, time, and their interaction. Skewed data were log-transformed before analysis, as appropriate.



## Results

### Mobile Assessment Compliance and Study Demographics

Examination of data from the mobile assessments revealed that a high majority of participants correctly recorded their drinking during the drinking episode (83/87, 95%). Despite training participants during screening to indicate only the number of drinks finished since the prior EMA prompt, 4 participants incorrectly reported their cumulative number of drinks at each prompt instead. Thus, their data could not be interpreted, resulting in a final study sample of 83 participants. eBAC calculated for a small number of survey responses was  $\geq 0.30$

g/dl, a level of intoxication associated with loss of consciousness or death. We assumed that those time points reflected errors in reporting [14] and opted to treat reported alcohol consumption and calculated eBAC for those specific responses as missing when calculating mean number of standard drinks consumed and eBAC per time point. This resulted in dropping reported alcohol use and eBAC data from 16 survey responses (16/673, 2.4%) across all participants who completed at least one survey (n=83) and 25 responses (25/996, 2.5%) across all those who completed both surveys (n=61). Demographic and drinking information for the sample and the safety comparison group is presented in Table 1. The groups did not differ on any demographic or drinking-related variables.

**Table 1.** Demographic characteristics, baseline, next-day, and 2-week follow-up drinking and safety outcomes for the smartphone (n=83) and safety comparison (n=25) groups.

Outcome	Smartphone group	Safety comparison group	P value
<b>Demographics</b>			
Age (years), mean (SD)	25.3 (2.6)	24.7 (2.5)	.26 <sup>a</sup>
Male, n (%)	48 (58)	15 (60)	.85 <sup>b</sup>
White, n (%)	54 (65)	20 (80)	.16 <sup>b</sup>
Education (years), mean (SD)	15.9 (1.8)	15.6 (1.9)	.42 <sup>a</sup>
<b>Baseline drinking and consequences, mean (SD)</b>			
% Drinking days (past month)	47.4 (18.9)	45.3 (16.7)	.61 <sup>a</sup>
% Binge drinking days (past month)	23.1 (10.3)	19.7 (7.7)	.13 <sup>a</sup>
Drinks/drinking day (past month)	4.7 (1.5)	4.1 (1.2)	.11 <sup>a</sup>
BYAACQ <sup>c</sup> (past 2 weeks)	2.9 (2.6)	3.6 (3.0)	.46 <sup>a,d</sup>
Alcohol Use Disorders Identification Test total	11.1 (4.3)	11.4 (4.9)	.74 <sup>a</sup>
<b>Next-day survey (smartphone group only), mean (SD)</b>			
Duration of continued drinking after the 3-hour EMA <sup>e</sup> period (hours)	2.7 (1.9)	— <sup>f</sup>	—
Additional standard drinks consumed following the 3-hour EMA period	2.9 (1.9)	—	—
BYAACQ (past 24 hours)	2.1 (2.1)	—	—
<b>Two-week follow-up drinking and consequences, mean (SD)</b>			
% Drinking days (past 2 weeks)	42.8 (17.8)	40.0 (12.0)	.47 <sup>a</sup>
% Binge drinking days (past 2 weeks)	23.2 (11.5)	19.7 (11.7)	.19 <sup>a</sup>
Drinks/drinking day (past 2 weeks)	4.6 (1.6)	3.9 (1.5)	.06 <sup>a</sup>
BYAACQ (past 2 weeks)	4.1 (3.4)	3.4 (2.8)	.44 <sup>a,d</sup>

<sup>a</sup>t test.

<sup>b</sup>Chi-square test.

<sup>c</sup>BYAACQ: Brief Young Adult Alcohol Consequences Questionnaire.

<sup>d</sup>Data log-transformed before analysis.

<sup>e</sup>EMA: ecological momentary assessment.

<sup>f</sup>Participants in the safety comparison group that did not complete the next-day survey.

### Feasibility

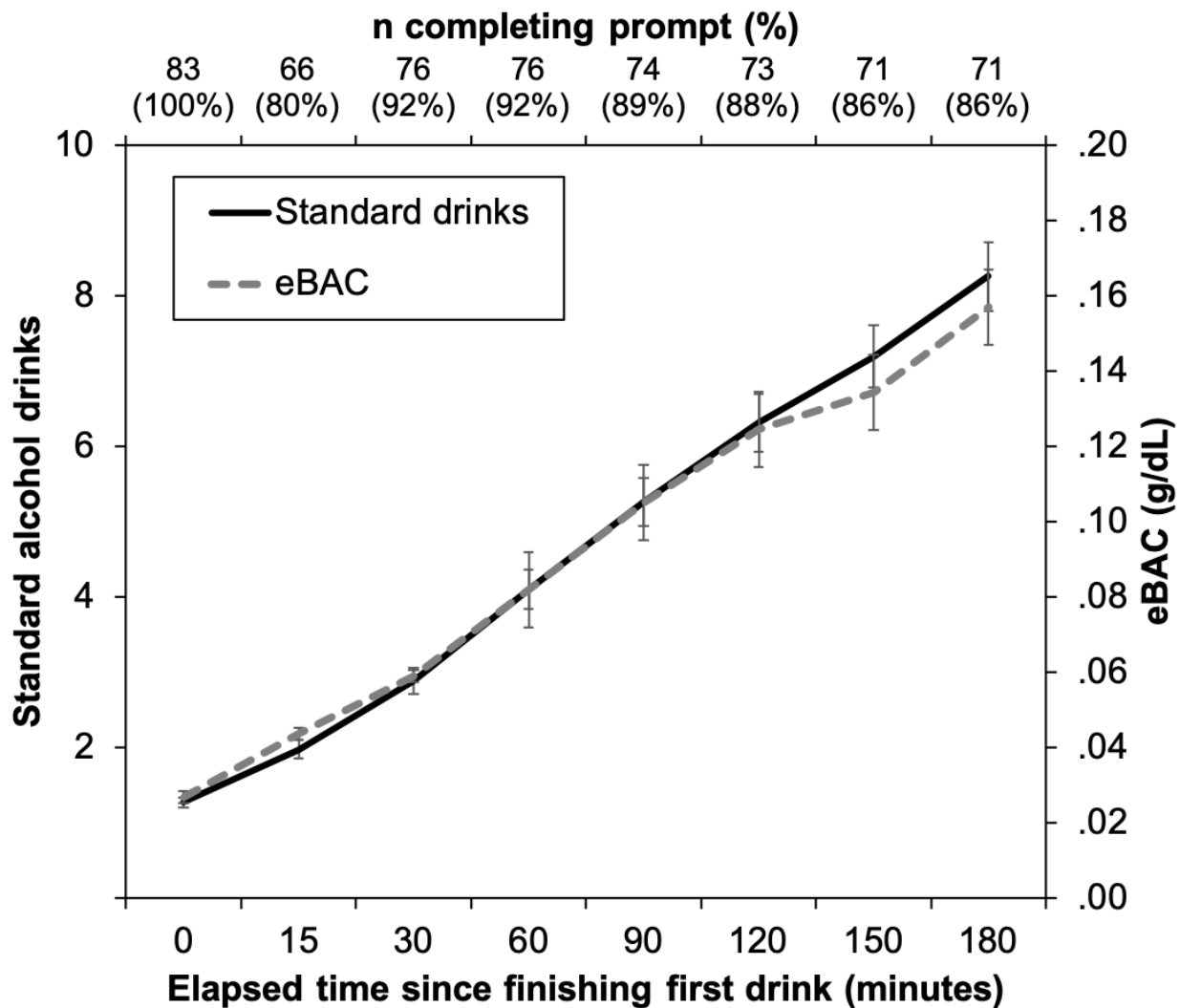
All participants completed the self-initiated baseline and post first-drink survey prompts and 87.2% (507/581) of the remaining

7 system-initiated prompts over the ensuing 3 hours of the EMA session (Figure 3). Most participants (64/83, 77%) started drinking during the evening hours (ie, 5:00 pm-12:00 am) on a Thursday, Friday, or Saturday (58/83, 70%). The majority of

prompts were completed in the presence of others (556/657 prompts, 84.6%) in various locations (bar/restaurant: 33.7% [221/656], one's own home: 32.0% [221/656], or friend's home: 24.2% [221/656]). Food consumption was not common during drinking, with participants reporting eating during only 12.2%

(70/574) of prompts. Cigarette smoking or recreational drug use were also not common as they were reported during only 3.1% (18/574) and 0.2% (1/574) of prompts, respectively. For the latter, only 1 participant reported using drugs (cocaine/stimulants) during only a single prompt (0.2%).

**Figure 3.** Mean (SE) of the mean standard alcoholic drinks (left y-axis) and estimated BAC (eBAC; right y-axis) at each survey time point (bottom x-axis) during the real-world drinking episode. Time 0 represents the survey that participants completed immediately after finishing their first drink in the drinking episode. The number of participants completing the survey prompt at each time point and corresponding percentage of the total study sample (n=83) is presented for each time point on the top x-axis. Drinking and eBAC data for time points where eBAC was  $\geq 0.30$  g/dl were excluded from calculations of the means for those outcomes; see the Results section for details. BAC: blood alcohol concentration; eBAC: estimated blood alcohol concentration.



### Real-World Drinking Behavior

Participants reported consuming any alcohol on 75.4% (445/590) of 590 completed prompts from the self-initiated post first-drink survey onward (ie, not including the predrinking baseline survey). They consumed 7.3 (SD 3.0) standard drinks during the 3-hour drinking assessment, with eBAC reaching 0.13 (SD 0.07) g/dL at the final time point (+180 min; Figure 3). Most participants (71%, 59/83) engaged in heavy (binge) drinking during the first 2 hours of the assessment, consuming 7.4 (SD 2.8) standard drinks during that time (9.1 [SD 2.9] drinks over the entire 3-hour drinking episode), with eBAC reaching 0.14 (SD 0.06) g/dL (eBAC 0.17 [SD 0.07] g/dL at the conclusion

of the EMA period). In contrast, participants who did not binge consumed 2.9 (SD 0.9) drinks during the first 2 hours of the EMA period (4.5 [SD 1.8] drinks over the entire 3-hour drinking episode), with eBAC reaching 0.04 (SD 0.02) g/dL during that time (eBAC 0.07 [SD 0.03] g/dL at the conclusion of the 3-hour monitoring period). For all participants, beer was the most common alcoholic drink consumed at each prompt (48.1%, 214/445), followed by liquor (40.2%, 179/445) and wine (16.9%, 75/445; note that percentages sum to greater than 100% because participants could report consuming multiple types of alcohol at each time point). Participants averaged 1.1 (SD 1.0) standard drinks reported at each of the 7 post baseline survey prompts.

### Participants Completing Both Mobile Sessions

Background characteristics and alcohol consumption during drinking episodes did not differ between participants who completed 2 smartphone drinking assessments (73%, 61/83) and those who completed 1 assessment ( $n=22$ ), all  $P_s>.60$ . Among participants who completed 2 drinking episode assessments, most (>60%) had high response rates and completed all the study prompts (9/9) with an average of 8.1 and 8.3 prompts completed in the first and second episodes, respectively. There was more drinking alone during the second episode (109/504, 21.6%) than the first (78/491, 15.8%;  $\chi^2_1=5.4$ ;  $P=.02$ ). Participants' reported location when completing the prompts or consumption of food, nicotine, or other drugs did not differ across assessment sessions, all  $P_s>.07$ .

### Acceptability

Participants rated the smartphone survey software as easy to use overall (mean 4.6, SD 0.7 out of 5 points) and endorsed high satisfaction with study participation (mean 4.6, SD 0.6). Similarly, participants did not consider the surveys to be time consuming (mean 2.1, SD 1.0) or intrusive (mean 2.0, SD 0.9). Overall, participants agreed that they would recommend participation in the study to others (mean 4.6, SD 0.6). Intrusiveness ratings were slightly higher among participants who completed only 1 smartphone assessment than among those who completed 2 assessments (mean 2.5, SD 1.1 vs mean 1.9, SD 0.9;  $P=.02$ ), but other acceptability ratings did not differ based upon the number of smartphone sessions completed ( $P_s>.16$ ).

### Safety

#### Next-Day Survey

A total of 93% (77/83) of participants in the smartphone group completed the next-day survey after the real-world drinking episode. Most participants (60%, 46/77) reported continued drinking after the final survey prompt for the episode (see Table 1). Participants reported few drinking-related consequences during or after the real-world drinking episode on the modified past 24-hour BYAACQ (Table 1). The most frequently reported consequences (endorsed by  $\geq 10\%$  of participants) were less energy after drinking (61%, 47/77), hangover (40%, 31/77), drinking larger amounts than anticipated (14%, 11/77), and saying or doing something embarrassing as a result of drinking (13%, 10/77). None of the participants reported any serious adverse events (eg, arrest or injury) related to their drinking during the assessment period.

#### Two-Week Follow-Up Survey

GEE revealed that self-reported drinking frequency decreased at follow-up for both the smartphone and control groups (time: beta [SE]= $-4.7$  [1.8];  $P=.008$ ), but binge drinking frequency and drinks consumed per drinking day did not change for either group ( $P_s>.12$ ). There was a significant increase in self-reported alcohol consequences from baseline to 2-week follow-up (time: beta [SE]= $0.4$  [0.2];  $P=.02$  [data log-transformed]; see Table 1) but no main effect or interaction of group for that outcome ( $P_s>.22$ ).

## Discussion

### Principal Findings

This study demonstrated the feasibility, acceptability, and safety of using smartphone-based EMA to measure real-time alcohol use and related outcomes in young adult heavy drinkers, most of whom (71%, 59/83) completed the EMA protocol in the setting of a drinking binge. The smartphone-based approach was feasible, as evidenced by a high response rate (87.2%, 507/581) to system-initiated study prompts during the 3-hour assessment period, and acceptable, with participants reporting high satisfaction study procedures and low ratings on intrusiveness for the brief (1 min each) EMA surveys. Safety was evidenced by few drinking-related consequences and no increases in drinking quantity or frequency over the 2-week follow-up. The results of this study show that young adult heavy drinkers can complete mobile assessments of their drinking behavior and related outcomes even during very heavy drinking episodes at BAC well over the threshold for intoxication. The method described here is user-friendly and easy to administer on participants' own smartphones and could be customized to facilitate future research on a variety of topics related to risky drinking. Furthermore, it may be less susceptible than retrospective or laboratory-based studies of binge drinking outcomes to limitations posed by recall bias and unclear ecological validity, respectively.

Relative to prior EMA studies in social drinkers, our 87% response rate to survey prompts is similar to, or higher than, response rates reported in previous studies [10,12-15]. Nearly three-quarters (71%, 59/83) of participants engaged in binge drinking during the first 2 hours and averaged just over 7 standard alcoholic drinks (eBAC 0.13g/dL) through 3 hours (see Figure 2), which is considerably higher than the 2 to 4 drinks reported in prior EMA studies [10,11,14,35]. In contrast to those studies, however, this study recruited heavy social drinkers specifically (ie, light drinkers and those with severe AUD were excluded from participation) and limited measurement of alcohol use and outcomes for up to 2 self-selected drinking events. Previous EMA studies have recorded drinking behavior and outcomes over longer durations (eg, 3 weeks [14]), which can provide data on heavier- and lighter-drinking events. Of note, the drinking episode was not an isolated event for the study participants, with the vast majority (85%, 66/77 completing the next day survey) reporting that it was indicative of their typical drinking.

Regarding feasibility and acceptability of the EMA approach, participants reported good compliance with study directions to avoid alcohol or other drug use during the assessments and high overall satisfaction with study procedures. The data from this study suggest that heavy drinkers can comply with instructions to avoid those substances even during binge episodes with high BAC and that this approach can produce data on drinking-related outcomes that are contaminated only minimally by other substance use. The high acceptability ratings reported in this study echo previous findings from EMA studies in general young adult drinker samples [13,36]. The brief duration of the mobile surveys (approximately 1 min each) may have facilitated

the overall high acceptability of the mobile procedure despite the repeated survey prompts throughout the 3-hour drinking episode. Furthermore, the high rates of smartphone use among young adults and participants' high degree of familiarity with these devices may have contributed to the high overall satisfaction and usability ratings reported in this study.

Importantly, the results of this study suggest that EMA of a real-world binge drinking episode in heavy drinkers is safe. Participants reported few alcohol-related problems during the drinking episode, and those that were reported tended to be minor (eg, hangover and less energy after drinking). Both groups reported a small decrease in their frequency of alcohol use from baseline to 2-week follow-up, with no changes in binge drinking frequency. This finding is in line with the results of previous studies indicating that EMA of alcohol use is associated with minimal or no changes in reported drinking behavior [11,37-39]. However, studies of reactivity to multiple assessments of alcohol use via retrospective self-report or interview (eg, TLFB) spaced several weeks/months apart have shown that, in general, self-reported alcohol use and consequences decrease over time with those repeated assessments (for a review, refer to the study by Schrimsher and Filtz [40]). In contrast, we observed a small, but significant, increase in self-reported alcohol consequences over the 2-week follow-up period across all participants, but all reported consequences tended to be mild, with no serious adverse events reported. This effect appeared to be driven mainly by the small (approximately 1 point) increase in mean BYAACQ scores from baseline to 2-week follow-up among the smartphone group (Table 1), although our GEE model did not show a significant effect or interaction of group for that outcome. We speculate that completing the next-day assessments of alcohol-related consequences related to the real-world drinking episodes increased awareness of those consequences among the smartphone group, resulting in them reporting slightly higher scores on the 2-week follow-up BYAACQ relative to baseline. However, this study was not designed to test this hypothesis directly, and to our knowledge, no previous studies have examined the effect of monitoring drinking on perceptions of alcohol-related consequences over time. Whether EMA of alcohol use affects individuals' awareness of drinking-related risks may be a promising area for future research.

### Strengths and Limitations

This study featured several strengths, including a well-characterized sample of young adult heavy drinkers, application of existing mobile phone technology to support EMA of real-world heavy drinking and related outcomes, and inclusion of an independent heavy drinking sample as a safety comparison group. However, there are also some limitations worth noting. First, although we recruited current binge drinkers, light drinkers and those meeting criteria for severe AUD were excluded from participation. Thus, based on the current data, we cannot infer the feasibility, acceptability, or safety of our smartphone-based EMA approach in those groups. Second, the duration of our monitoring period (3 hours per drinking occasion, over a maximum of 2 occasions) was brief and fixed, in contrast to previous studies that have assessed alcohol use outcomes over several weeks [11] and scaled the duration of

the EMA sessions according to participant-reported drinking patterns [14]. Participants in this study reported continued drinking for nearly 3 hours after the EMA period ended, consuming approximately 3 additional drinks during that period. Additional research is needed to evaluate the feasibility, acceptability, and safety of using this assessment method over longer drinking episodes, including any potential effect of a longer monitoring duration on compliance with the survey prompts. Third, the protocol used in this study relied upon participant self-report to estimate alcohol consumption and BAC, and the equation we used to compute eBAC is less accurate at  $BAC \geq 0.08$  g/dL [34]. Future research on EMA of drinking behavior may benefit from incorporating recently developed wearable alcohol biosensors, which can measure consumption passively at the skin, to provide an objective measure of alcohol use [41]. Fourth, for the purposes of this study, "safety" was conceptualized in terms of scores on the next-day BYAACQ (for the smartphone group) and changes in drinking behavior or self-reported alcohol-related consequences on the BYAACQ at 2-week follow-up, compared across the smartphone and safety comparison groups. We did not examine other safety-related outcomes, such as the potential for the smartphone prompts to serve as a distraction while walking or driving (although we note that participants were instructed not to drive after consuming alcohol during the study), nor did we collect next-day report data from members of the safety comparison group, which prevented us from examining possible group-related differences on consequences related to real-world drinking events. Future studies could implement a broader definition of "safety" to evaluate the potential effects of EMA of alcohol use on a wider range of safety-related outcomes. Finally, this study does not speak to the reliability and validity of using smartphones to measure alcohol use and related outcomes during real-world (binge) drinking events. We will present data on the reliability of the smartphone-based EMA method across multiple real-world drinking sessions and validity of the approach relative to the laboratory alcohol challenge session in a separate report.

### Conclusions

In sum, the results of this study provide preliminary support for the feasibility, acceptability, and safety of using smartphone-based EMA to assess alcohol use and related outcomes (eg, subjective responses and drinking context) in young adult heavy drinkers. These data suggest that young adults can use smartphones to monitor their drinking even during very heavy drinking episodes, with average consumption approximately double (or more) that of previous work in this field [10,11,14,35]. Future research could use this technology to further study the dynamics of binge drinking behavior, refine EMA-based treatment approaches targeting risky drinking specifically [42,43], or develop "just-in-time" interventions to reduce heavy drinking and associated risks in young adults [44]. In sum, smartphone-based EMA has great potential to provide a practical, inexpensive, and efficient way to capture a large amount of data on real-world drinking behavior and associated consequences, which may inform future research on the epidemiology of and intervention for AUD.

## Acknowledgments

This research was supported by grant 1R21AA024901 to DJF and ACK.

## Conflicts of Interest

None declared.

## References

1. Kanny D, Naimi TS, Liu Y, Lu H, Brewer RD. Annual total binge drinks consumed by US Adults, 2015. *Am J Prev Med* 2018 Apr;54(4):486-496 [FREE Full text] [doi: [10.1016/j.amepre.2017.12.021](https://doi.org/10.1016/j.amepre.2017.12.021)] [Medline: [29555021](https://pubmed.ncbi.nlm.nih.gov/29555021/)]
2. Stahre M, Roeber J, Kanny D, Brewer RD, Zhang X. Contribution of excessive alcohol consumption to deaths and years of potential life lost in the United States. *Prev Chronic Dis* 2014 Jun 26;11:E109 [FREE Full text] [doi: [10.5888/pcd11.130293](https://doi.org/10.5888/pcd11.130293)] [Medline: [24967831](https://pubmed.ncbi.nlm.nih.gov/24967831/)]
3. Rethinking Drinking: Alcohol and Your Health. 2016. URL: <https://www.rethinkingdrinking.niaaa.nih.gov> [accessed 2019-09-17]
4. White A, Hingson R. The burden of alcohol use: excessive alcohol consumption and related consequences among college students. *Alcohol Res* 2013;35(2):201-218 [FREE Full text] [Medline: [24881329](https://pubmed.ncbi.nlm.nih.gov/24881329/)]
5. Courtney KE, Polich J. Binge drinking in young adults: data, definitions, and determinants. *Psychol Bull* 2009 Jan;135(1):142-156 [FREE Full text] [doi: [10.1037/a0014414](https://doi.org/10.1037/a0014414)] [Medline: [19210057](https://pubmed.ncbi.nlm.nih.gov/19210057/)]
6. King AC, de Wit H, McNamara PJ, Cao D. Rewarding, stimulant, and sedative alcohol responses and relationship to future binge drinking. *Arch Gen Psychiatry* 2011 Apr;68(4):389-399 [FREE Full text] [doi: [10.1001/archgenpsychiatry.2011.26](https://doi.org/10.1001/archgenpsychiatry.2011.26)] [Medline: [21464363](https://pubmed.ncbi.nlm.nih.gov/21464363/)]
7. Pedersen SL, McCarthy DM. An examination of subjective response to alcohol in African Americans. *J Stud Alcohol Drugs* 2009 Mar;70(2):288-295 [FREE Full text] [doi: [10.15288/jsad.2009.70.288](https://doi.org/10.15288/jsad.2009.70.288)] [Medline: [19261241](https://pubmed.ncbi.nlm.nih.gov/19261241/)]
8. Schuckit MA. Subjective responses to alcohol in sons of alcoholics and control subjects. *Arch Gen Psychiatry* 1984 Sep;41(9):879-884. [doi: [10.1001/archpsyc.1984.01790200061008](https://doi.org/10.1001/archpsyc.1984.01790200061008)] [Medline: [6466047](https://pubmed.ncbi.nlm.nih.gov/6466047/)]
9. 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](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
10. Carpenter RW, Trela CJ, Lane SP, Wood PK, Piasecki TM, Trull TJ. Elevated rate of alcohol consumption in borderline personality disorder patients in daily life. *Psychopharmacology (Berl)* 2017 Nov;234(22):3395-3406 [FREE Full text] [doi: [10.1007/s00213-017-4727-1](https://doi.org/10.1007/s00213-017-4727-1)] [Medline: [28884321](https://pubmed.ncbi.nlm.nih.gov/28884321/)]
11. Hufford MR, Shields AL, Shiffman S, Paty J, Balabanis M. Reactivity to ecological momentary assessment: an example using undergraduate problem drinkers. *Psychol Addict Behav* 2002 Sep;16(3):205-211. [doi: [10.1037//0893-164x.16.3.205](https://doi.org/10.1037//0893-164x.16.3.205)] [Medline: [12236455](https://pubmed.ncbi.nlm.nih.gov/12236455/)]
12. Kuntsche E, Labhart F. Investigating the drinking patterns of young people over the course of the evening at weekends. *Drug Alcohol Depend* 2012 Aug 1;124(3):319-324. [doi: [10.1016/j.drugalcdep.2012.02.001](https://doi.org/10.1016/j.drugalcdep.2012.02.001)] [Medline: [22377089](https://pubmed.ncbi.nlm.nih.gov/22377089/)]
13. Luczak SE, Rosen IG, Wall TL. Development of a real-time repeated-measures assessment protocol to capture change over the course of a drinking episode. *Alcohol Alcohol* 2015 Mar;50(2):180-187 [FREE Full text] [doi: [10.1093/alcald/agu100](https://doi.org/10.1093/alcald/agu100)] [Medline: [25568142](https://pubmed.ncbi.nlm.nih.gov/25568142/)]
14. Piasecki TM, Wood PK, Shiffman S, Sher KJ, Heath AC. Responses to alcohol and cigarette use during ecologically assessed drinking episodes. *Psychopharmacology (Berl)* 2012 Oct;223(3):331-344 [FREE Full text] [doi: [10.1007/s00213-012-2721-1](https://doi.org/10.1007/s00213-012-2721-1)] [Medline: [22538731](https://pubmed.ncbi.nlm.nih.gov/22538731/)]
15. Smit K, Groefsema M, Luijten M, Engels R, Kuntsche E. Drinking motives moderate the effect of the social environment on alcohol use: an event-level study among young adults. *J Stud Alcohol Drugs* 2015 Nov;76(6):971-980. [doi: [10.15288/jsad.2015.76.971](https://doi.org/10.15288/jsad.2015.76.971)] [Medline: [26562607](https://pubmed.ncbi.nlm.nih.gov/26562607/)]
16. Groefsema M, Luijten M, Engels R, Kuntsche E. Young adults do not catch up missed drinks when starting later at night-an ecological momentary assessment study. *Exp Clin Psychopharmacol* 2019 Apr;27(2):160-165. [doi: [10.1037/pha0000236](https://doi.org/10.1037/pha0000236)] [Medline: [30382729](https://pubmed.ncbi.nlm.nih.gov/30382729/)]
17. Piasecki TM. Assessment of alcohol use in the natural environment. *Alcohol Clin Exp Res* 2019 Apr;43(4):564-577. [doi: [10.1111/acer.13975](https://doi.org/10.1111/acer.13975)] [Medline: [30748019](https://pubmed.ncbi.nlm.nih.gov/30748019/)]
18. Pew Research Center. 2018. Mobile Fact Sheet URL: <http://www.pewinternet.org/fact-sheet/mobile/> [accessed 2018-09-24]
19. Dvorak RD, Pearson MR, Day AM. Ecological momentary assessment of acute alcohol use disorder symptoms: associations with mood, motives, and use on planned drinking days. *Exp Clin Psychopharmacol* 2014 Aug;22(4):285-297 [FREE Full text] [doi: [10.1037/a0037157](https://doi.org/10.1037/a0037157)] [Medline: [24932896](https://pubmed.ncbi.nlm.nih.gov/24932896/)]
20. Fridberg DJ, Cao D, King AC. Integrating alcohol response feedback in a brief intervention for young adult heavy drinkers who smoke: a pilot study. *Drug Alcohol Depend* 2015 Oct 1;155:293-297 [FREE Full text] [doi: [10.1016/j.drugalcdep.2015.08.017](https://doi.org/10.1016/j.drugalcdep.2015.08.017)] [Medline: [26341847](https://pubmed.ncbi.nlm.nih.gov/26341847/)]

21. King AC, Hasin D, O'Connor SJ, McNamara PJ, Cao D. A prospective 5-year re-examination of alcohol response in heavy drinkers progressing in alcohol use disorder. *Biol Psychiatry* 2016 Mar 15;79(6):489-498 [FREE Full text] [doi: [10.1016/j.biopsych.2015.05.007](https://doi.org/10.1016/j.biopsych.2015.05.007)] [Medline: [26117308](https://pubmed.ncbi.nlm.nih.gov/26117308/)]
22. King AC, McNamara PJ, Hasin DS, Cao D. Alcohol challenge responses predict future alcohol use disorder symptoms: a 6-year prospective study. *Biol Psychiatry* 2014 May 15;75(10):798-806 [FREE Full text] [doi: [10.1016/j.biopsych.2013.08.001](https://doi.org/10.1016/j.biopsych.2013.08.001)] [Medline: [24094754](https://pubmed.ncbi.nlm.nih.gov/24094754/)]
23. Substance Abuse and Mental Health Services Administration. 2018. 2017 National Survey on Drug Use and Health: Detailed Tables URL: <https://www.samhsa.gov/data/report/2017-nsduh-detailed-tables> [accessed 2018-09-24]
24. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. World Health Organization. 2002. The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care URL: [https://apps.who.int/iris/bitstream/handle/10665/67205/WHO\\_MSD\\_MSB\\_01.6a.pdf?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/67205/WHO_MSD_MSB_01.6a.pdf?sequence=1) [accessed 2019-09-17]
25. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961 Jun;4:561-571. [doi: [10.1001/archpsyc.1961.01710120031004](https://doi.org/10.1001/archpsyc.1961.01710120031004)] [Medline: [13688369](https://pubmed.ncbi.nlm.nih.gov/13688369/)]
26. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
27. Sobell LC, Sobell MB. National Institute on Alcohol Abuse and Alcoholism. 1995. Alcohol Timeline Followback (TLFB) URL: [https://pubs.niaaa.nih.gov/publications/assessingalcohol/instrumentpdfs/13\\_tlfb.pdf](https://pubs.niaaa.nih.gov/publications/assessingalcohol/instrumentpdfs/13_tlfb.pdf) [accessed 2019-09-17]
28. Cahalan D, Cisin IH, Crossley HM. American Drinking Practices: A National Study of Drinking Behavior Patterns. New Brunswick, NJ: Rutgers Center of Alcohol Studies; 1969.
29. First MB, Williams JB, Benjamin LS, Spitzer RL. Structured Clinical Interview for DSM-5 (R) Personality Disorders (SCID-5-PD). Arlington, VA: American Psychiatric Association; 2015.
30. Morean ME, de Wit H, King AC, Sofuoglu M, Rueger SY, O'Malley SS. The drug effects questionnaire: psychometric support across three drug types. *Psychopharmacology (Berl)* 2013 May;227(1):177-192 [FREE Full text] [doi: [10.1007/s00213-012-2954-z](https://doi.org/10.1007/s00213-012-2954-z)] [Medline: [23271193](https://pubmed.ncbi.nlm.nih.gov/23271193/)]
31. Rueger SY, McNamara PJ, King AC. Expanding the utility of the biphasic alcohol effects scale (BAES) and initial psychometric support for the brief-BAES (B-BAES). *Alcohol Clin Exp Res* 2009 May;33(5):916-924 [FREE Full text] [doi: [10.1111/j.1530-0277.2009.00914.x](https://doi.org/10.1111/j.1530-0277.2009.00914.x)] [Medline: [19320625](https://pubmed.ncbi.nlm.nih.gov/19320625/)]
32. Kahler CW, Strong DR, Read JP. Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: the brief young adult alcohol consequences questionnaire. *Alcohol Clin Exp Res* 2005 Jul;29(7):1180-1189. [doi: [10.1097/01.alc.0000171940.95813.a5](https://doi.org/10.1097/01.alc.0000171940.95813.a5)] [Medline: [16046873](https://pubmed.ncbi.nlm.nih.gov/16046873/)]
33. Matthews DB, Miller WR. Estimating blood alcohol concentration: two computer programs and their applications in therapy and research. *Addict Behav* 1979;4(1):55-60. [doi: [10.1016/0306-4603\(79\)90021-2](https://doi.org/10.1016/0306-4603(79)90021-2)] [Medline: [420046](https://pubmed.ncbi.nlm.nih.gov/420046/)]
34. Hustad JT, Carey KB. Using calculations to estimate blood alcohol concentrations for naturally occurring drinking episodes: a validity study. *J Stud Alcohol* 2005 Jan;66(1):130-138. [doi: [10.15288/jsa.2005.66.130](https://doi.org/10.15288/jsa.2005.66.130)] [Medline: [15830913](https://pubmed.ncbi.nlm.nih.gov/15830913/)]
35. Ray LA, Miranda R, Tidey JW, McGeary JE, MacKillop J, Gwaltney CJ, et al. Polymorphisms of the mu-opioid receptor and dopamine D4 receptor genes and subjective responses to alcohol in the natural environment. *J Abnorm Psychol* 2010 Feb;119(1):115-125 [FREE Full text] [doi: [10.1037/a0017550](https://doi.org/10.1037/a0017550)] [Medline: [20141248](https://pubmed.ncbi.nlm.nih.gov/20141248/)]
36. Kuntsche E, Labhart F. ICAT: development of an internet-based data collection method for ecological momentary assessment using personal cell phones. *Eur J Psychol Assess* 2013;29(2):140-148 [FREE Full text] [doi: [10.1027/1015-5759/a000137](https://doi.org/10.1027/1015-5759/a000137)] [Medline: [24285917](https://pubmed.ncbi.nlm.nih.gov/24285917/)]
37. Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papanonatos GD. Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* 1998 Aug;6(3):306-315. [doi: [10.1037/1064-1297.6.3.306](https://doi.org/10.1037/1064-1297.6.3.306)] [Medline: [9725114](https://pubmed.ncbi.nlm.nih.gov/9725114/)]
38. Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* 1998 Jan;17(1):48-52. [doi: [10.1037//0278-6133.17.1.48](https://doi.org/10.1037//0278-6133.17.1.48)] [Medline: [9459069](https://pubmed.ncbi.nlm.nih.gov/9459069/)]
39. Simpson TL, Kivlahan DR, Bush KR, McFall ME. Telephone self-monitoring among alcohol use disorder patients in early recovery: a randomized study of feasibility and measurement reactivity. *Drug Alcohol Depend* 2005 Aug 1;79(2):241-250. [doi: [10.1016/j.drugalcdep.2005.02.001](https://doi.org/10.1016/j.drugalcdep.2005.02.001)] [Medline: [16002033](https://pubmed.ncbi.nlm.nih.gov/16002033/)]
40. Schrimsher GW, Filtz K. Assessment reactivity: can assessment of alcohol use during research be an active treatment? *Alcohol Treat Q* 2011 Apr 19;29(2):108-115. [doi: [10.1080/07347324.2011.557983](https://doi.org/10.1080/07347324.2011.557983)]
41. Wang Y, Fridberg DJ, Leeman RF, Cook RL, Porges EC. Wrist-worn alcohol biosensors: strengths, limitations, and future directions. *Alcohol* 2018 Sep 1:- (epub ahead of print)(forthcoming)(forthcoming). [doi: [10.1016/j.alcohol.2018.08.013](https://doi.org/10.1016/j.alcohol.2018.08.013)] [Medline: [30179709](https://pubmed.ncbi.nlm.nih.gov/30179709/)]
42. Suffoletto B, Chung T, Muench F, Monti P, Clark DB. A text message intervention with adaptive goal support to reduce alcohol consumption among non-treatment-seeking young adults: non-randomized clinical trial with voluntary length of enrollment. *JMIR Mhealth Uhealth* 2018 Feb 16;6(2):e35 [FREE Full text] [doi: [10.2196/mhealth.8530](https://doi.org/10.2196/mhealth.8530)] [Medline: [29453191](https://pubmed.ncbi.nlm.nih.gov/29453191/)]
43. Voogt C, Kuntsche E, Kleinjan M, Poelen E, Engels R. Using ecological momentary assessment to test the effectiveness of a web-based brief alcohol intervention over time among heavy-drinking students: randomized controlled trial. *J Med Internet Res* 2014 Jan 8;16(1):e5 [FREE Full text] [doi: [10.2196/jmir.2817](https://doi.org/10.2196/jmir.2817)] [Medline: [24401555](https://pubmed.ncbi.nlm.nih.gov/24401555/)]

44. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAs) 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](https://doi.org/10.1007/s12160-016-9830-8)] [Medline: [27663578](https://pubmed.ncbi.nlm.nih.gov/27663578/)]

## Abbreviations

**AUD:** alcohol use disorder

**BAC:** blood alcohol concentration

**BrAC:** breath alcohol concentration

**BYAACQ:** Brief Young Adult Alcohol Consequences Questionnaire

**DSM-5:** Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

**eBAC:** estimated blood alcohol concentration

**EMA:** ecological momentary assessment

**GEE:** generalized estimating equation

**TLFB:** Timeline Followback

*Edited by G Eysenbach; submitted 20.02.19; peer-reviewed by T Piasecki, R O'Donnell; comments to author 18.05.19; revised version received 10.07.19; accepted 19.08.19; published 20.11.19.*

*Please cite as:*

*Fridberg DJ, Faria J, Cao D, King AC*

*Real-Time Mobile Monitoring of Drinking Episodes in Young Adult Heavy Drinkers: Development and Comparative Survey Study*  
*JMIR Mhealth Uhealth* 2019;7(11):e13765

*URL:* <https://mhealth.jmir.org/2019/11/e13765>

*doi:* [10.2196/13765](https://doi.org/10.2196/13765)

*PMID:* [31746774](https://pubmed.ncbi.nlm.nih.gov/31746774/)

©Daniel J Fridberg, James Faria, Dingcai Cao, Andrea C King. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 20.11.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.

---

Publisher:  
JMIR Publications  
130 Queens Quay East.  
Toronto, ON, M5A 3Y5  
Phone: (+1) 416-583-2040  
Email: [support@jmir.org](mailto:support@jmir.org)

---

<https://www.jmirpublications.com/>