JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 7 (2019), Issue 6 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Short Paper

Viewpoint

Using Mobile Apps for Health Management: A New Health Care Mode in China

Qing Lv[1*], MD; Yutong Jiang[1*], MD; Jun Qi[1], MD; Yanli Zhang[1], MD; Xi Zhang[1], MIM; Linkai Fang[1], MIM; Liudan Tu[1], MD; Mingcan Yang[1], MD; Zetao Liao[1], MD; Minjing Zhao[1], MD; Xinghua Guo[1], MIM; Minli Qiu[1], MIM; Jieruo Gu[1], MD; Zhiming Lin[1], MD

Department of Rheumatology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Zhiming Lin, MD
Department of Rheumatology
The Third Affiliated Hospital of Sun Yat-sen University
600 Tianhe Road
Guangzhou, 510930
China
Phone: 86 20 85252205
Fax: 86 20 85253396
Email: lzm-zj99@163.com

Abstract

Background: China has a large population; however, medical resources are unevenly distributed and extremely limited, and more medical services are needed. With the development and ever-increasing popularity of mobile internet communication, China has created a mode of mobile health (mHealth) care to resolve this problem.

Objective: The aim of this study was (1) to describe the problems associated with China's medical care practice, (2) explore the need for and the feasibility of internet-based medical care in China, and (3) analyze the functionality of and services offered by internet-based health care platforms for the management of chronic diseases.

Methods: Data search was performed by searching national websites, the popular search engine Baidu, the App Store, and websites of internet medical care institutions, using search terms like "mobile health," "Internet health," "mobile medical," "Internet medical," "digital medical," "digital health," and "online doctor." A total of 6 mobile apps and websites with the biggest enrollment targeting doctors and end users with chronic diseases in China were selected.

Results: We recognized the limitations of medical and health care providers and unequal distribution of medical resources in China. An mHealth care platform is a novel and efficient way for doctors and patients to follow up and manage chronic diseases. Services offered by these platforms include reservation and payment, medical consultation, medical education assessment, pharmaceutical and medical instruments sales, electronic medical records, and chronic disease management. China's health policies are now strongly promoting the implementation of mHealth solutions, particularly in response to the increasing burden of chronic diseases and aging in the population.

Conclusions: China's internet-based medical and health care mode can benefit the populace by providing people with high-quality medical resources. This can help other countries and regions with high population density and unevenly distributed medical resources manage their health care concerns.

(JMIR Mhealth Uhealth 2019;7(6):e10299) doi:10.2196/10299

KEYWORDS

mHealth; internet; health care; medical informatics

Introduction

Background

China has the largest population and the second largest economy in the world (2018 gross domestic product: 90 trillion Chinese Yuan [CNY]), but medical care resources in China are relatively limited. There are only 2.21 licensed (assistant) physicians for every 1000 people in China. There exists an uneven geographical distribution of medical resources as well. These two problems have led to countless transprovincial patients, resulting in numerous extra economic and time costs.

An understanding of the Chinese health care system and mobile phone usage provides a framework for understanding mobile health (mHealth) apps in China. Acknowledging the increasing pressure exerted by an aging population, behavioral changes, and rapid urbanization [1], the government's Healthy China 2030 plan [2] envisions the primary health care system as a means of addressing the emerging dual burden of chronic noncommunicable diseases [3-5] and increasing health expenditures. However, primary health care providers in China are usually inadequately trained and some are not even certified [6]. Professionals equipped with up-to-date medical knowledge play an important role in efficient health care, especially among those in poorly developed regions, and the lack of this has become a major issue in China.

With the rapid development of mobile internet communication and the increasing needs of medical services, a new mode of mHealth care is being implemented in China to effectively overcome some traditional barriers. Mobile phones (ie, mobile phones with advanced computing and internet access) and tablet computers have become the most popular and widespread types of mobile devices used. Mobile apps allow patients to access information, assessments, and treatments in a timely manner, and improve the health of those living with chronic diseases [7]. Therefore, extra costs, such those associated with travel, time, and doctor consultations, can be dramatically reduced. In addition, doctors could build another way to connect with their patients, practice without geographical limitation, enhance their reputation by providing remote chronic disease management services to patients, and gain extra income (since they are paid comparatively low wages and minimal benefits due to China's new health care reform [8]). It can also promote a balanced distribution of medical resources via regional medical information interconnection, which increases the accessibilty of high-quality medical resources in the rural areas. Furthermore, some internet hospitals have been established in China, although many of them are still underdeveloped and face various issues (eg, a scarcity of online doctors) [9]. Many mobile medical platforms, such as Doctor 7LK, Doctor Xingren, Micro-doctor, Doctor Hao, and Doctor Chunyu, usually rely on experts from hospitals. Patients can be reached and followed up via these internet-based platforms. For instance, Doctor 7LK is an comprehensive internet-based medical service enterprise. It can serve as an online practice site for doctors across the country and link hospital resources all over China to provide services such as appointments, remote diagnosis, examinations, and electronic prescriptions to patients.

Objectives

The aim of this study was (1) to describe the problems associated with China's medical care practice, (2) explore the need for and the feasibility of internet-based medical care in China, and (3) analyze the functionality of and services offered by internet-based health care platforms for the management of chronic diseases.

Methods

Data Collection

Data were collected by searching national websites, the popular search engine Baidu, and websites of internet-based medical care institutions (ie, the National Bureau of Statistics of the People's Republic of China, the National Health and Family Planning Commission of the People's Republic of China, the Hospital Management Institute of Fudan University, and the Ministry of Industry and Information Technology of the People's Republic of China), using search terms like "mobile health," "Internet health," "mobile medical," "Internet medical," "digital medical," "digital health," and "doctor," until June 2018. In addition, we used Baidu and the App Store to identify medical care platforms and mHealth apps. We found more than 200 medical care platforms and mHealth apps.

Choosing Medical Care Platforms

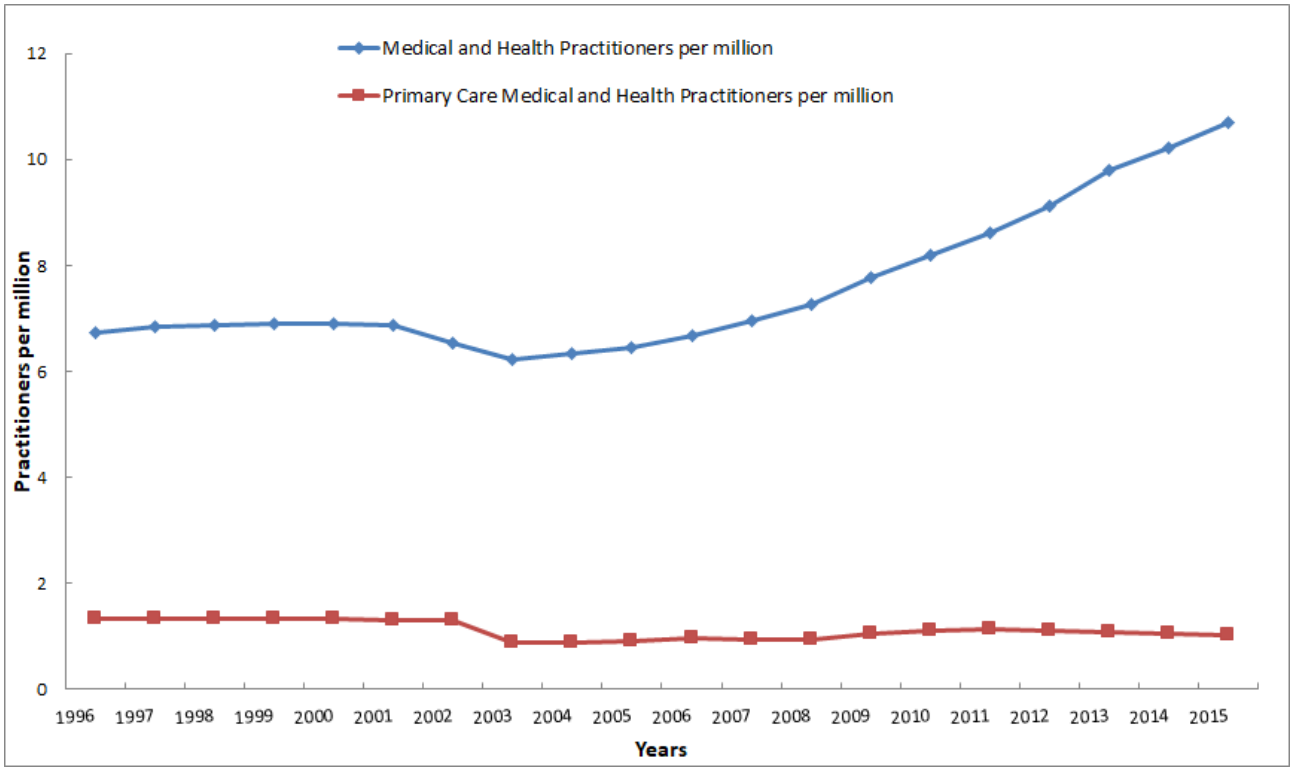
We chose the most frequently used mobile medical care platforms, such as Doctor 7LK, Doctor Xingren, Micro-doctor, Doctor Hao, Doctor Chunyu, and Doctor Shu, to analyze the number of registered doctors. Overall, there were 81,604 hospital-based specialists and 400,344 patients involved in these platforms. Subsequently, we compared registered doctors in several provinces on the Doctor 7LK platform, the permanent resident population of these areas, and the distribution of patients and doctors on the Doctor 7LK platform. Data management and analysis were performed using Microsoft Excel 2017. Nominal and ordinal data were presented by using frequencies, percentages, and bar charts.

Results

The Overwhelming Health Care Situation in China

China has limited health care resources and an overwhelming population, but the growth of primary care medical and health practitioners is relatively slow (Figure 1). From 1996 to 2015, the number of health practitioners in China rose by 58.6% from 6.74 million to 10.69 million. In contrast, primary care medical and health practitioners decreased from 1.32 million to 1.03 million, accounting for less than 10% of national medical and health practitioners.

Figure 1. The growth of total medical and health practitioners and primary care medical and health practitioners in the past 20 years in China.



The geographical distribution of medical resources is imbalanced in China. Among the top 100 comprehensive hospitals [1], more than half are located in Beijing (City), Shanghai (City), and Guangdong province, accounting for 23%, 19%, and 9%, respectively. In the 12 western provinces and autonomous regions, only Chongqing, Sichuan, and Shanxi provinces have some of the top 100 hospitals, accounting for 6%, 5%, and 3%, respectively.

The Current Status of Internet-Based Medical Platforms

In 2018, the coverage rate of mobile broadband network had reached 57%, which is 12 times that of 2002 [10]. By the end of 2016, statistics from the Ministry of Industry and Information Technology of the People's Republic of China revealed that there were 1.32 billion mobile phone users, 192 times of that observed in late 1996. Furthermore, the mobile phone penetration rate had reached 96.2% [10]. By 2016, the size of China's mobile medical market had reached 10.56 billion CNY, with a growth rate of 116.4% per year [11].

There are more than 100,000 registered doctors on multiple digital medical platforms [12-17]. For example, the Doctor Hao platform had attracted about 480,000 doctors to register online as of November 2017, followed by Doctor 7LK (389,000 doctors), Doctor Xingren (380,000 doctors), and Micro-doctor (260,000 doctors).

The unbalanced distribution of medical resource leads to the exploration and explosion of internet-based medical practice. Although the number of registered doctors in the southeastern area is higher, many doctors are more accepting of an internet-based medical practice in the relatively poor western regions (Figure 2). Gansu province in northwestern China is a good example. The population of Gansu province is 26 million, whereas the registered number of physicians is 8480 on the Doctor 7LK platform, the equivalent of 3.26 online registered doctors per 10 million people. In contrast, 2.08 registered doctors are shared by every 10 million people in Guangdong province, which is much more economically developed and has the largest number of registered doctors. The number of registered patients in the southeastern region is the highest in Guangdong province with 2.36 per 1000 registered patients, far more than that of the central and western regions.

Within the departments of chronic diseases, the largest number of registered users (patients and doctors) came from the cardiovascular division, which is related to the high incidence of cardiovascular diseases in the nation (Figure 3). In terms of the number of patients per registered doctor, rheumatologists were associated with the largest number of patients, with an average of 29.25 patients per rheumatologist, which reflects a relative shortage of rheumatologists at present. Dermatology is the most interactive department and enjoys the largest visit numbers per day, which is linked to special manifestations, diagnosis, and treatment features of skin diseases.

Figure 2. Registered physicians of Doctor 7LK in several provinces and the permanent resident populations of these areas.

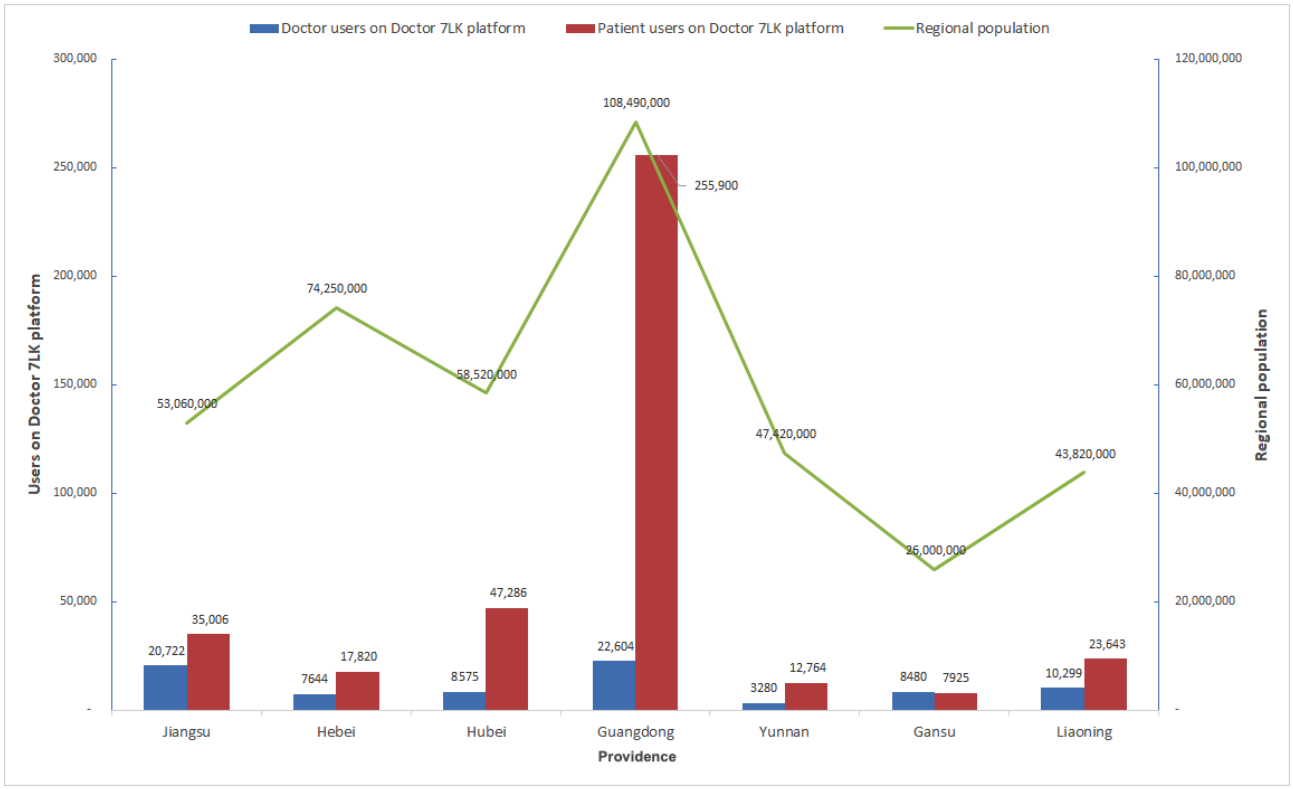
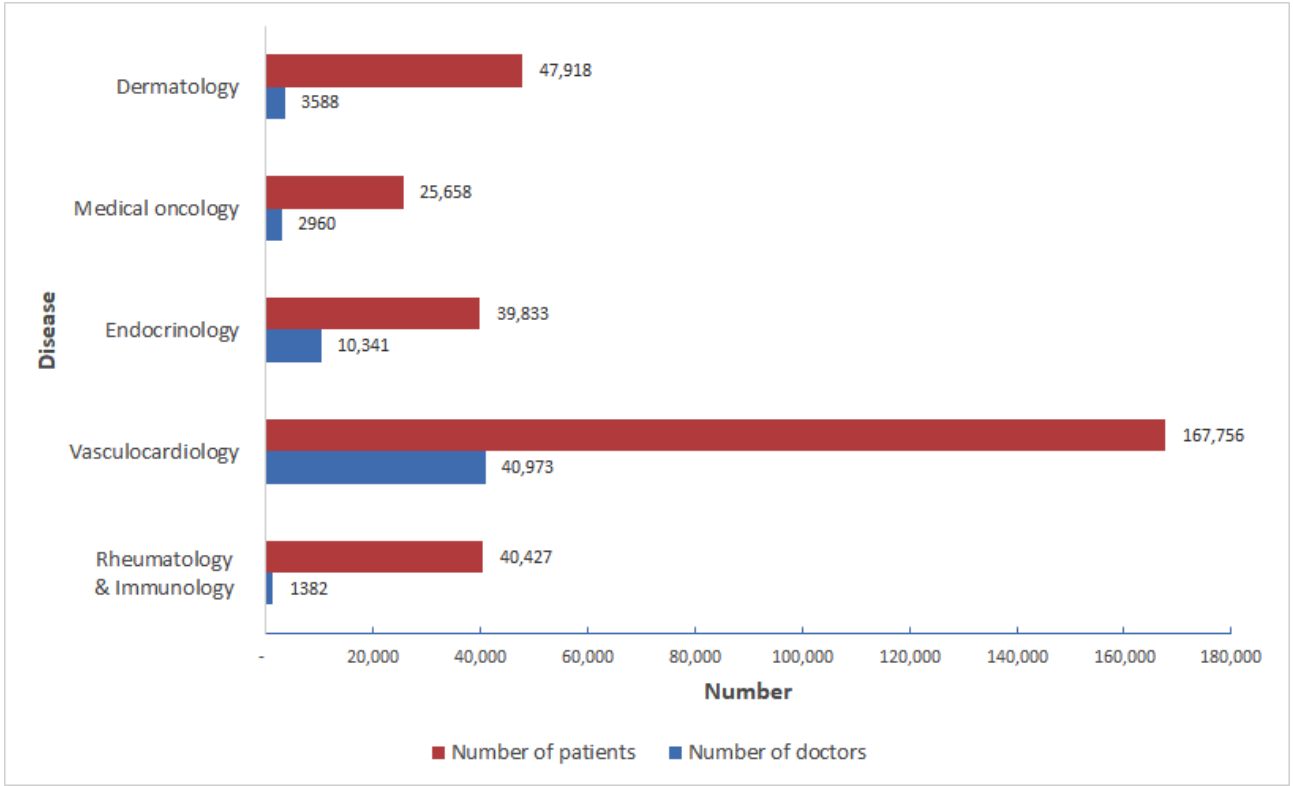


Figure 3. Numbers of patients and the distribution of doctors on the Doctor 7LK platform.



Discussion

Principal Findings

We provided an overview of China's current internet usage and mHealth care mode. The unbalanced distribution of medical resource leads to the exploration and explosion of internet-based medical practice. Internet-based mHealth care platforms could enable people to overcome geographical health care barriers and help reduce the uneven distribution of medical resources. A large number of venture capital funds have flowed into internet-based health care platforms, and the number of users of medical apps has increased gradually each year. Patients in

well-developed southeastern China have a higher recognition of internet medical behaviors. Within the departments of chronic diseases, the largest number of registered users (patients and doctors) came from the cardiovascular division, which is related to the high incidence of cardiovascular diseases in the nation.

By 2016, the total population of mainland China had reached 1.38 billion, but there were only 2.21 licensed (assistant) physicians per 1000 people and unlicensed doctors [6]. In the primary health care system, the number of licensed (assistant) doctors per 1000 people in each province ranged from 0.52 to 1.13, and the difference in rural doctors was 16 times (from 0.24 to 3.90). [6]. Similar to the huge differences in the distribution of doctors throughout the entire health care system [18], among the top 100 comprehensive hospitals [19], more than half are located in Beijing, Shanghai, and Guangdong province. On the contrary, the western region of China, which accounts for 72% of the territory and 29% of the whole population, is home to only 14% of the top hospitals. Most of the best resources (doctors) are distributed in large cities such as Beijing, Shanghai, and Guangzhou, although the growth of primary care medical and health practitioners is relatively slow. Due to the huge gap in human resources and clinical skills between general hospitals and primary medical institutions, 3.08 billion patients are accepted by 28,000 general hospitals annually; however, only 4.34 billion patients choose to go to the 921,000 primary medical institutions. This indicates that less than 3% of medical institutions account for more than 40% of medical and health service [2]. One of the main reasons for the shortage of high-quality medical services is that Chinese patients are free to choose their medical institutions and doctors. Even with mild symptoms, there is a tendency to visit high-end hospitals, which results in overcrowding [20].

Why do Chinese patients need the mHealth care platform? First, primary health care providers in China are usually inadequately trained and sometimes uncertified. According to the China Health and Family Planning Statistical Yearbook, only 24% of doctors had a regular license in village clinics in 2015 [6]. The continuing education for primary health care doctors is also insufficient. In addition, primary health care professionals are poorly paid, have the lowest benefits, and the payment policy does not reward high-quality medical services [8]. Skilled doctors are unwilling to work in communities or remote rural areas for financial and professional reasons [21]. In addition, many patients are reluctant to go to primary health care institutions because of a lack of confidence in the health professionals' skills and the quality of health care provided [22]. The desire to access high-quality care is a trigger for the development of the mHealth care platform.

The good news is that mobile internet communication is booming in China. Statistical data from the National Bureau of Statistics of the People's Republic of China showed that the broadband network covered 94.5% of the total villages and 78% of impoverished villages by the end of 2015 [1]. The mobile phone penetration rate has been increasing gradually each year. With the development of technology and people's attention to health, mHealth care platforms, including mHealth apps, have the potential to assist the self-management of many health conditions [7,23] and improve the health of those living with chronic diseases. In the Chinese health app market, the number of technology and health app users is growing exponentially. Although the quality of Chinese mHealth apps appears to be inferior to those in developed countries [24,25], they could help reduce the negative impact of the shortage of professional health care providers and meet the needs for high-quality information and care access [26,27]. We found that a large number of venture capital funds had flowed into internet-based medical platforms, and the number of internet users of medical apps have been growing gradually each year [11]. By the end of 2016, there were over 2000 mobile medical apps in China, involving hundreds of thousands of medical practitioners across the country [12-17].

According to the different stages of medical interventions, operation modes of digital medical services can be divided into different types based on the analysis of these mHealth apps. These include (1) reservation and payment: patients can make appointments with doctors through the platform and pay the consultation fee through the platform; (2) medical consultation and assessment: registered patients can consult doctors on the platform, and the registered doctors can answer the patient's questions. To ensure the standardization of diagnosis and treatment, these consultations usually involve common medical questions, and the doctors cannot provide diagnostic opinions or prescriptions; (3) medical education: these websites or apps provide patients with disease-related popular science and education articles, such as daily life considerations, self-monitoring, self-management, and emergency response to certain acute diseases; (4) pharmaceutical and medical instrument sales: patients can purchase over-the-counter medications or purchase prescription medicine after verifying a valid prescription on these websites; (5) electronic medical records: the users of such website or apps are usually doctors. It facilitates doctors to make medical electronic records; (6) chronic disease management: such websites or apps usually have doctor and patient versions. These types of platforms usually have multiple functions. For example, doctors can send text messages to patients to remind them about their follow-up visits, and the platforms would automatically make an appointment for a follow-up visit after receiving a reply from the patients. They could also send related medical education articles to the patients under the patients' request. Doctors can write their patients' medical records on the platform. Patients can purchase prescription medicine on these websites according to the doctor's prescription.

Other countries already use mobile phones to assess and treat certain diseases [28-31]. With the widespread availability of the internet and decreasing cost of distant medical services over time, the internet health care mode has had an impact on many developing countries, where medical services and resources are concentrated in large cities. Moreover, the system could improve health education and raise public health awareness, as patients can easily gain access to professional health consultancy.

Chronic disease management platforms aim to facilitate subsequent follow-ups for patients with chronic diseases. It enables remote medical treatment, which otherwise relies on the offline primary medical and health care institutions, clinical laboratories, and medical image centers. The platforms facilitate

setting up a medical closed loop of care for doctors, patients, and medicine, thus helping doctors to build up private clinics to better manage patients with chronic diseases. The proportion of doctors and patients associated with chronic diseases is very high; the top 5 departments of registered users are related to chronic diseases. This correlates to the nature of the disease and the need for follow-ups. Using internet-based health care platforms is an innovative and effective way to help doctors follow up and manage their patients with chronic diseases in China.

Nevertheless, in terms of the distribution of registered users, patients in the southeast have a higher recognition of these novel internet-based medical advances. One of the reasons is that it is a relatively economically developed region and the gross regional domestic product is related to higher medical input of the government. Moreover, people in this area are highly educated and they take their health very seriously. However, statistical bias may be caused by the differences among several internet-based medical companies in various areas. The widespread existence of internet-based health care platform has weakened the uneven distribution of medical resources. Many doctors are more likely to accept internet-based medical practices, especially in poor areas as the internet can make it easier for doctors to connect with patients who are geographically distant. Both doctors' and patients' needs are fundamental for the survival and development of these platforms.

Future Prospects

China's health care system needs to be considerably strengthened to manage both the rising burden of chronic diseases and increasing health-related expenditures [32]. On April 12, 2018, Premier Li Keqiang presided over an executive meeting of the central committee and the state council, which defined measures to accelerate the development of *internet plus medical and health care*, and allowed medical institutions to provide online services and carry out internet-based medical services to benefit more people with high-quality medical resources and improve the overall health of residents. The government will also expand telemedicine coverage to all medical federations, support high-speed broadband networks, and explore the sharing of prescription and drug retail information in medical institutions, thus improving the *internet plus medical care* system [33]. As China takes steps to accelerate the development of *internet plus medical and health care*, it has the opportunity to build an integrated, cooperative internet-based medical and health care system and gain valuable experiences through the whole process. How China's internet-based health care will evolve under existing policies and socioeconomic environment is worthy of attention.

Conclusions

China's internet health care mode can provide possible solutions to address the lack and uneven distribution of experienced health care providers in some countries or regions. Although it is still being explored, we highlight the important role of the mHealth care mode in China, which could also be helpful for other developing countries facing similar challenges.

Acknowledgments

A part of the data source was provided by Doctor 7LK and Jiangfengyuan chronic disease management platform.

This research was supported by the National Natural Science Foundation for the Youth of China (grant #81302583), the Guangdong Natural Science Funds for Distinguished Young Scholar (grant #2014A030306039), the high-level personnel of special support program for Technology Innovative Talents and the Top Young of Guangdong Province (grant #2015TQ01R516), the Fundamental Research Funds for the Central Universities (grant #16ykzd05), the Distinguished Young Scholar Candidates Program for the Third Affiliated Hospital of Sun Yat-sen University, and the Pearl River Nova Program of Guangzhou (grant #201610010005).

Authors' Contributions

QL, YJ, and ZL designed the research. JQ, YZ, and XZ searched relevant national and regional information. LF, LT, and MY gathered data. ZL, MZ, XG, and MQ did the statistical analyses. QL, YJ, and JQ wrote the first draft of the paper, with revisions from JG and ZL. All authors contributed to revisions and approved the final version.

Conflicts of Interest

None declared.

References

1. Gong P, Liang S, Carlton EJ, Jiang Q, Wu J, Wang L, et al. Urbanisation and health in China. Lancet 2012 Mar 3;379(9818):843-852 [FREE Full text] [doi: 10.1016/S0140-6736(11)61878-3] [Medline: 22386037]
2. [The State Council of the People's Republic of China]. 2016. [Government Information Disclosure Column - China Government Network] URL: http://www.gov.cn/zhengce/xxgkzl.htm [accessed 2018-12-12] [WebCite Cache ID 1544602587845851]
3. Yang G, Wang Y, Zeng Y, Gao GF, Liang X, Zhou M, et al. Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 2013 Jun 8;381(9882):1987-2015. [doi: 10.1016/S0140-6736(13)61097-1] [Medline: 23746901]

4. [Ministry of Health of the People's Republic of China]. 2002. [China Nutrition and Health Survey] URL: https://www.cpc.unc.edu/projects/china/ [WebCite Cache ID 1544602772801484]
5. [Chinese Center for Disease Control and Prevention]. [China National Plan for NCD Prevention and Treatment (2012-2015)] URL: http://www.chinacdc.cn/en/ne/ [accessed 2018-12-12] [WebCite Cache ID 1544603484158068]
6. National Health and Family Planning Commission of the People's Republic of China. China YearBooks. 2017. China health and family planning statistical yearbook 2016 URL: https://www.chinayearbooks.com/chinas-health-and-family-planning-statistical-yearbook-2016.html [WebCite Cache ID 1544601702565069]
7. 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;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
8. Li X, Lu J, Hu S, Cheng KK, De Maeseneer J, Meng Q, et al. The primary health-care system in China. Lancet 2017 Dec 9;390(10112):2584-2594. [doi: 10.1016/S0140-6736(17)33109-4] [Medline: 29231837]
9. Xie X, Zhou W, Lin L, Fan S, Lin F, Wang L, et al. Internet hospitals in China: cross-sectional survey. J Med Internet Res 2017 Jul 4;19(7):e239 [FREE Full text] [doi: 10.2196/jmir.7854] [Medline: 28676472]
10. [Ministry of Industry and Information Technology of the People's Republic of China]. [Work dynamics] URL: http://www.miit.gov.cn/n1146290/n1146402/index.html [accessed 2018-12-12] [WebCite Cache ID 1544604009399639]
11. [Easy View - Data Driven Lean Growth]. URL: https://www.analysys.cn/ [accessed 2018-12-12] [WebCite Cache ID 1544604248588093]
12. [Almond Doctor - Professional Development Partner of Chinese Excellent Doctors]. [China's leading doctor business platform] URL: https://xingren.com/index/doctors [accessed 2018-12-13] [WebCite Cache ID 1544681569544338]
13. [Micromedicine (registered network) - Internet hospital online diagnosis and treatment platform]. URL: https://www.guahao.com/ [WebCite Cache ID 1544511023254994]
14. [Good doctor online _ can find the medical platform of 200,000 public hospital doctors]. [All expert teams] URL: https://www.haodf.com/doctorteam/ [accessed 2018-12-11] [WebCite Cache ID 74ZzpsRci]
15. [Dr. Chun Yu - Consult a doctor online]. [Regional doctor list] URL: https://www.chunyuyisheng.com/pc/doctors/ [accessed 2018-12-13] [WebCite Cache ID 1544680357909063]
16. [Home_36]. URL: https://36kr.com/ [accessed 2018-12-12] [WebCite Cache ID 1544605584610445]
17. [Qi Lekang - Internet Medical Comprehensive Service Platform]. URL: http://wap.7lk.com [accessed 2018-12-13] [WebCite Cache ID 1544680555255218]
18. Anand S, Fan V, Zhang J. China's human resources for health: quantity, quality, and distribution. Lancet 2008;372:1774-1781. [doi: 10.1016/S0140-6736(08)61363-X]
19. [Hospital Management Institute, Fudan University]. [In-depth interpretation of Fudan's 2013 hospital rankings] URL: http://www.fudanmed.com/institute/NewsDetail.aspx?cid=24 [accessed 2018-12-12] [WebCite Cache ID 1544601884895736]
20. Wagstaff A, Lindelow M. Can insurance increase financial risk? The curious case of health insurance in China. J Health Econ 2008 Jul;27(4):990-1005. [doi: 10.1016/j.jhealeco.2008.02.002] [Medline: 18342963]
21. Tu J, Wang C, Wu S. The internet hospital: an emerging innovation in China. Lancet Glob Health 2015 Aug;3(8):e445-e446 [FREE Full text] [doi: 10.1016/S2214-109X(15)00042-X] [Medline: 26187488]
22. Yip W, Hsiao W. Harnessing the privatisation of China's fragmented health-care delivery. Lancet 2014 Aug 30;384(9945):805-818. [doi: 10.1016/S0140-6736(14)61120-X] [Medline: 25176551]
23. 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] [Medline: 27861583]
24. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The top Chinese mobile health apps: a systematic investigation. J Med Internet Res 2016 Aug 29;18(8):e222 [FREE Full text] [doi: 10.2196/jmir.5955] [Medline: 27573724]
25. Liu Y, Chang P, Li N, Lamb KV, Wang J, Huang X, et al. Quality of Chinese CVD medication apps. Stud Health Technol Inform 2016;225:1036-1037. [Medline: 27332470]
26. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
27. McMillan B, Hickey E, Mitchell C, Patel M. The need for quality assurance of health apps. Br Med J 2015 Nov 4;351:h5915. [doi: 10.1136/bmj.h5915] [Medline: 26536895]
28. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using mobile apps to assess and treat depression in Hispanic and Latino populations: fully remote randomized clinical trial. J Med Internet Res 2018 Aug 9;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]
29. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
30. Cheong IY, An SY, Cha WC, Rha MY, Kim ST, Chang DK, et al. Efficacy of mobile health care application and wearable device in improvement of physical performance in colorectal cancer patients undergoing chemotherapy. Clin Colorectal Cancer 2018 Jun;17(2):e353-e362. [doi: 10.1016/j.clcc.2018.02.002] [Medline: 29551558]

31. Sjöström M, Lindholm L, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a cost-effectiveness analysis. J Med Internet Res 2017 May 8;19(5):e154 [FREE Full text] [doi: 10.2196/jmir.7383] [Medline: 28483745]
32. Beaglehole R, Epping-Jordan J, Patel V, Chopra M, Ebrahim S, Kidd M, et al. Improving the prevention and management of chronic disease in low-income and middle-income countries: a priority for primary health care. Lancet 2008 Sep 13;372(9642):940-949. [doi: 10.1016/S0140-6736(08)61404-X] [Medline: 18790317]
33. [China Government Network_Central People's Government Portal]. [Speech] URL: http://www.gov.cn/premier/lkq_jh.htm [accessed 2018-12-12] [WebCite Cache ID 1544602296462608]

Abbreviations

CNY: Chinese Yuan
mHealth: mobile health

Edited by G Eysenbach; submitted 06.03.18; peer-reviewed by K Blondon, B Celler; comments to author 05.08.18; revised version received 30.09.18; accepted 30.03.19; published 03.06.19.

Please cite as:
Lv Q, Jiang Y, Qi J, Zhang Y, Zhang X, Fang L, Tu L, Yang M, Liao Z, Zhao M, Guo X, Qiu M, Gu J, Lin Z
Using Mobile Apps for Health Management: A New Health Care Mode in China
JMIR Mhealth Uhealth 2019;7(6):e10299
URL: https://mhealth.jmir.org/2019/6/e10299/
doi:10.2196/10299
PMID:31162131

Review

Accuracy of mHealth Devices for Atrial Fibrillation Screening: Systematic Review

Godwin Denk Giebel[1], MSc; Christian Gissel[1], PhD

Health Economics, Department of Economics and Business, Justus Liebig University, Giessen, Germany

Corresponding Author:
Christian Gissel, PhD
Health Economics
Department of Economics and Business
Justus Liebig University
Licher Strasse 62
Giessen, 35394
Germany
Phone: 49 641 99 22070
Email: christian.gissel@wirtschaft.uni-giessen.de

Abstract

Background: Mobile health (mHealth) devices can be used for the diagnosis of atrial fibrillation. Early diagnosis allows better treatment and prevention of secondary diseases like stroke. Although there are many different mHealth devices to screen for atrial fibrillation, their accuracy varies due to different technological approaches.

Objective: We aimed to systematically review available studies that assessed the accuracy of mHealth devices in screening for atrial fibrillation. The goal of this review was to provide a comprehensive overview of available technologies, specific characteristics, and accuracy of all relevant studies.

Methods: PubMed and Web of Science databases were searched from January 2014 until January 2019. Our systematic review was performed according to the Preferred Reporting Items for Systematic Review and Meta-Analyses. We restricted the search by year of publication, language, noninvasive methods, and focus on diagnosis of atrial fibrillation. Articles not including information about the accuracy of devices were excluded.

Results: We found 467 relevant studies. After removing duplicates and excluding ineligible records, 22 studies were included. The accuracy of mHealth devices varied among different technologies, their application settings, and study populations. We described and summarized the eligible studies.

Conclusions: Our systematic review identifies different technologies for screening for atrial fibrillation with mHealth devices. A specific technology's suitability depends on the underlying form of atrial fibrillation to be diagnosed. With the suitable use of mHealth, early diagnosis and treatment of atrial fibrillation are possible. Successful application of mHealth technologies could contribute to significantly reducing the cost of illness of atrial fibrillation.

(JMIR Mhealth Uhealth 2019;7(6):e13641) doi:10.2196/13641

KEYWORDS

mHealth; atrial fibrillation; wearable; app

Introduction

Background

Atrial fibrillation is a cardiac arrhythmia appearing in different forms. Globally, 33.5 million people are affected by atrial fibrillation [1]. This disease leads to a significantly increased risk of all-cause mortality, cardiovascular mortality, major cardiovascular events, stroke, ischemic stroke, ischemic heart disease, sudden cardiac death, heart failure, chronic kidney disease, and peripheral arterial disease [2].

Atrial fibrillation can occur in five different forms: first diagnosed, paroxysmal, persistent, long-standing persistent, and permanent. For patients above 65 years of age, opportunistic screening for atrial fibrillation is recommended by pulse taking or using an electrocardiogram (ECG) rhythm strip. The gold standard for atrial fibrillation detection is the 12-lead ECG [3].

Because atrial fibrillation is a serious risk factor for stroke and mortality, its treatment is inevitable for patients. Through medication with oral anticoagulation such as vitamin K antagonists or nonvitamin K antagonists, the risk for stroke and mortality in patients with atrial fibrillation can be markedly reduced [4,5].

In addition to high health risks for patients with atrial fibrillation, the economic burden of the disease is vast. An investigation carried out by Johnson et al estimated an average of €20,403-€26,544 for the cost of illness caused by atrial fibrillation over 3 years in the Danish health care system [6]. For Sweden and Germany, the cost of illness amounts to €7,241 and €5,586 per year, respectively [7]. In this context, secondary diseases like stroke cause the majority of costs. The difference in costs between treated and untreated atrial fibrillation is significant. A stroke survivor with atrial fibrillation receiving oral anticoagulation costs €17,518, and the cost for a stroke survivor with atrial fibrillation not receiving oral anticoagulation is €19,143 [8]. Furthermore, there are several studies confirming the cost-effectiveness of screening for atrial fibrillation [9-12].

One of the main challenges is detection of irregular forms of atrial fibrillation in an accurate way in order to start treatment as soon as possible. Even an ECG taken over a longer period (>24 hours) using a Holter monitor does not always lead to a reliable diagnosis of existing atrial fibrillation. In the case of paroxysmal atrial fibrillation, the occurrence of the disease often cannot be detected within the first 48 hours of ambulatory ECG monitoring [13].

To manage the increasing number of patients with atrial fibrillation and to cope with the consequences of this disease, an early diagnosis is fundamental. In this context, mobile health (mHealth) has often been suggested as a possible solution.

Among the reviews of the use of mHealth for the diagnosis, treatment, and prevalence estimation of arrhythmias [14-26], only one systematic review focused on outpatient cardiac rhythm monitoring in cryptogenic stroke [16]. Therefore, we conducted a systematic review focusing especially on the most recent and relevant noninvasive mHealth devices for the detection of atrial fibrillation. The aim of this article was to provide a systematic overview about the possible and real application of mHealth as well as to show its potentials and limitations by assessing the measurement quality.

Mobile Health

Smartphones, tablets, and mobile apps are widely used in many parts of the world. With an increasing rate of usage, two-thirds of the population in Europe and North America own at least one mobile device. Hence, there is already a basis for an mHealth approach in the context of atrial fibrillation, and the incremental costs for its use are relatively low.

There are two possible stages for mHealth use in the context of atrial fibrillation. First, the treatment of atrial fibrillation should start even before the occurrence of arrhythmia, in the form of prevention. Obesity, physical inactivity, and hypertension are preventable risk factors [3]. Despite the fact that behavior does not change by purchasing a wearable device or smartphone, these devices can contribute to a healthier and more active lifestyle [15].

Second, when atrial fibrillation has occurred, there are four possibilities to support the diagnosis and treatment: ECG or rhythm monitoring, heart rate monitoring, symptom and environmental annotation, and medication adherence [26].

Diverse propositions exist in the field of medication support. One approach is to support patients through communication of general knowledge about the disease, the mechanism of medication, and medication adherence [27-29]. Other applications provide guidelines and risk scores to support decision making for treatments [30,31].

In this review, the specific focus is on the diagnosis of atrial fibrillation by monitoring the heart rate and detection of arrhythmia by mHealth devices. For this purpose, event monitors or Holter devices are used. Monitoring can be done by either loop recorders or postevent recorders. The former is used over a long period, wherein electrodes are attached to the skin in order to monitor the heart activity when triggered by patients or an embedded algorithm. The patient-activated postevent ECG is not worn continuously, but used regularly or immediately after symptoms have occurred [32]. Nowadays, especially through the development in the field of mHealth and its simple use outside of health care, both approaches can record cardiac activities in an extensive way and thereby support the diagnosis of atrial fibrillation.

Despite the high cost of illness of atrial fibrillation, there are few economic assessments for mHealth solutions [33,34].

Methods

Our systematic review is performed according to the guidelines for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [35].

Article Retrieval

We ran literature searches in PubMed MEDLINE and Web of Science databases in January 2019. Regarding time of publication, we considered the year 2014 as the baseline, because there are some reviews covering previous years [19,24,26]. As search keywords, we used the terms “mHealth,” “telemedicine,” “wearable,” “mobile health,” “app,” and “digital treatment” in combination with the term “atrial fibrillation.”

Study Selection

Eligible studies had to meet the following predefined criteria: original research, focus on the diagnosis of atrial fibrillation, interventions using mHealth devices, noninvasive, and published in English language. The following were used as exclusion criteria: focus on technical descriptions or algorithms and lack of information about the accuracy of the investigated device.

Table 1. Overview of all studies included in the review.

Study and app/device	FDA[a] approval	Study population	Reference method	Recording duration (mHealth[b] device)	Sensitivity, %	Specificity, %	PPV[c], %	NPV[d], %
Bonomi et al [36]								
CM3 Generation-3, Connected Sensing	No	AF[e] patients before and after elective electrical cardioversion (n=20)	Actiwave Cardio (single-lead ECG[f])	60 min	97	100	99	98
CM3 Generation-3, Connected Sensing	No	Patients prescribed to undergo 24-/48-hour ECG Holter with either paroxysmal or persistent AF (n=40)	12-lead Holter ECG	Duration of the Holter monitoring period	93	100	N/A[g]	N/A
Brasier et al [37]								
Heartbeats app + iPhone 4S	No	In-house patients with presumed AF and matched controls in sinus rhythm (n=592)	Single-lead iECG[h] (AliveCor)	1 min	89.9	99.1	N/A	N/A
Heartbeats app + iPhone 4S	No	In-house patients with presumed AF and matched controls in sinus rhythm (n=592)	Single-lead iECG (AliveCor)	5 min	91.5	99.6	N/A	N/A
Bumgarner et al [38]								
AliveCor KardiaBand + Apple Watch + smartphone	Yes	Patients with a diagnosis of AF who presented for scheduled elective cardio version, aged 18-90 years (n=100)	12-lead ECG	N/A	93	84	N/A	N/A
Chan et al [39]								
Cardiio Rhythm + iPhone 4S	No	Patients with either hypertension or diabetes mellitus or aged ≥65 years (n=1013)	12-lead ECG (15 min)	17.1 s	92.9	97.7	53.1	99.8
AliveCor Heart Monitor	Yes	Patients with either hypertension or diabetes mellitus or aged ≥65 years (n=1013)	12-lead ECG (15 min)	30 s	71.4	99.4	76.9	99.2
Desteghe [33]								
AliveCor KardiaMobile	Yes	Patients at a cardiology ward (n=320)	12-lead ECG (10 sec)	30 s	36.8	96.1	56	91.1
AliveCor KardiaMobile	Yes	Patients at a geriatric ward (n=125)	6-lead ECG (30 sec)	30 s	72.7	98.1	88.9	94.4
MyDiagnostick	No	Patients at a cardiology ward (n=320)	12-lead ECG (10 sec)	30 s	60.5	93.3	54.8	94.6
MyDiagnostick	No	Patients at a geriatric ward (n=125)	6-lead ECG (30 sec)	30 s	81.8	96.1	81.8	96.1
Eerikäinen et al [40]								
CM3 Generation-3, Connected Sensing	No	Patients before and after an electrical cardioversion procedure in the hospital (n=18)	Single-lead (Actiwave Cardio) ECG and 24-hour Holter	2 h	92.3	60.7	N/A	N/A
CM3 Generation-3, Connected Sensing	No	24-hour measurements in normal everyday conditions (n=16)	Single-lead (Actiwave Cardio) ECG and 24-hour Holter	24 h	71.6	84.9	N/A	N/A

Study and app/device	FDA[a] approval	Study population	Reference method	Recording duration (mHealth[b] device)	Sensitivity, %	Specificity, %	PPV[c], %	NPV[d], %
Fan et al [41]								
HUAWEI Mate 9	No	Patients aged ≥18 years excluding patients with ICD[i] or pacemaker (n=108)	12-lead ECG (3 min)	3 min	94.41	100	100	95.43
HUAWEI Honor 7x	No	Patients aged ≥18 years excluding patients with ICD or pacemaker (n=108)	12-lead ECG (3 min)	3 min	95.56	99.4	99.23	96.49
HUAWEI Band 2	No	Patients aged ≥18 years excluding patients with ICD or pacemaker (n=108)	12-lead ECG (3 min)	3 min	95.36	99.7	99.63	96.24
Gropler et al [42]								
AliveCor KardiaMobile	Yes	Patients aged <18 years with standard 12-lead ECG ordered as part of routine visit testing (n=30)	Standard 12-lead ECG	30 s	N/A	87	N/A	N/A
Haberman et al [43]								
AliveCor KardiaMobile (iPhone case or iPad)	Yes	Division I athletes (n=123)	Standard 12-lead ECG	30 s	N/A	99.2	N/A	100
AliveCor KardiaMobile (iPhone case or iPad)	Yes	Healthy young adults (n=128)	Standard 12-lead ECG	30 s	N/A	100	N/A	100
AliveCor KardiaMobile (iPhone case or iPad)	Yes	Cardiology clinic patients (n=130)	Standard 12-lead ECG	30 s	94.4	99.1	94.4	99.1
Hochstadt et al [44]								
CardiacSense	In process	Patients aged ≥18 years excluding patients with ICD or pacemaker (n=108)	Simultaneously recorded ECG	30 min	100	N/A	N/A	N/A
Kang et al [45][j]								
CPstethoscope + Samsung Galaxy S5	No	Selected study participants (n=46)	Cardiologists using an electronic stethoscope	2 min	94	86	88	92
CPstethoscope + Samsung Galaxy S6	No	Selected study participants (n=46)	Cardiologists using an electronic stethoscope	2 min	94	79	83	92
CPstethoscope + LG G3	No	Selected study participants (n=46)	Cardiologists using an electronic stethoscope	2 min	81	100	100	82
Koltowski et al [46]								
AliveCor KardiaMobile	Yes	Patients of an academic cardiology care center (n=100)	12-lead ECG	N/A	92.8	100	N/A	N/A
Koshy et al [47][k]								
FitBit (Blaze) + Apple Watch (Series 1)	No	Patients in sinus rhythm or with arrhythmias, aged ≥18 years from a coronary care unit, an intensive care unit, and an emergency room (n=102)	12-lead ECG	30 min	N/A	N/A	N/A	N/A

Study and app/device	FDA[a] approval	Study population	Reference method	Recording duration (mHealth[b] device)	Sensitivity, %	Specificity, %	PPV[c], %	NPV[d], %
Krivoshei et al [48]								
iPhone 4S	No	Patients with AF or patients in sinus rhythm (n=80)	Heart rate monitor chest belt	5 min	95	95	N/A	N/A
Lahdenoja et al [49]								
Different smartphones, mostly Sony Xperia Z-Series	No	Patients with AF and healthy individuals as the control group (n=39)	Previous diagnosed AF	A few minutes (typically less than 5 min)	93.8	100	N/A	N/A
Lown et al [50]								
Polar-H7	No	Individuals from three general practices aged >65 years with and without AF (n=418)	12-lead ECG	N/A	96.34	98.21	N/A	N/A
AliveCor KardiaMobile	Yes	Individuals from three general practices aged >65 years with and without AF (n=418)	12-lead ECG	N/A	87.8	98.81	N/A	N/A
Firstbeat Bodyguard 2	No	Individuals from three general practices aged >65 years with and without AF (n=418)	12-lead ECG	N/A	96.34	98.51	N/A	N/A
WatchBP	Yes	Individuals from three general practices aged >65 years with and without AF	12-lead ECG	N/A	96.34	93.45	N/A	N/A
Lowres et al [51]								
AliveCor KardiaMobile	Yes	Persons aged ≥65 years entering a participating pharmacy (n=1000)	General practitioner review/12-lead ECG	30-60 s	98.5	91.4	N/A	N/A
Mena et al [52]								
Loop recorder ECG sensor device, Classifier, and a smartphone as central unit	No	Older adults (mean age 73.5, SD 11.8 years; n=100)	ECG by expert cardiologist	N/A	100	96.6	N/A	N/A
Rozen et al [53]								
Cardiio Rhythm + iPhone	No	Patients aged >18 years, scheduled for elective cardioversion (n=98)	Standard 12-lead ECG	3 times 20 s before and 3 times 20 s after cardioversion	93.1	90.9	92.2	92
Selder et al [54]								
AliveCor KardiaMobile	Yes	Population participating in the Hartwacht Arrhythmia program (n=233)	ECG interpreting team led by a cardiologist	30 s	92	95	80	98
Tison et al [55]								
Cardiogram application + Apple Watch	N/A[l]	Sedentary participants undergoing cardioversion (n=51)	Standard 12-lead ECG	≥20 min	98	90.2	90.9	97.8
Cardiogram application + Apple Watch	N/A[l]	Ambulatory participants (n=1617)	Standard 12-lead ECG	≥20 min	67.7	67.6	7.9	98.1

Study and app/device	FDA[a] approval	Study population	Reference method	Recording duration (mHealth[b] device)	Sensitivity, %	Specificity, %	PPV[c], %	NPV[d], %
William [56]								
AliveCor KardiaMobile + iPod	Yes/No	Patients aged 35-85 years with a history of paroxysmal or persistent AF (n=52)	12-lead ECG	30 s	96.6	94.1	N/A	N/A

[a]FDA: Food and Drug Administration.

[b]mHealth: mobile health.

[c]PPV: positive predictive value.

[d]NPV: negative predictive value.

[e]AF: atrial fibrillation.

[f]ECG: electrocardiogram.

[g]N/A: not applicable.

[h]iECG: internet-enabled electrocardiography.

[i]ICD: implantable cardioverter defibrillator.

[j]Accuracy of classification of the heart sounds into a correct category. Atrial fibrillation led to significantly fewer interpretable heart sounds. The app needs further improvement to diagnose atrial fibrillation.

[k]No data available on sensitivity, specificity, PPV, and NPV, but there was a significant correlation between device use and ECG in atrial arrhythmias (Apple Watch: r_s=0.83, FitBit: r_s=0.56; both P<.01)

[l]No information available about the series used in the study.

Data Extraction and Analysis

In the first step, we assessed the studies' eligibility by focusing on the inclusion and exclusion criteria mentioned above. After searching for the reference method used in the study, we searched for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) in relevant studies as indicators to evaluate the accuracy of the underlying mHealth device. In addition, we extracted characteristics about the study population and the size of the study population as well as recording duration and Food and Drug Administration (FDA) approval (Table 1).

Results

Literature Search

We identified 461 articles through database searching. We added seven relevant studies either known by the authors, found by searching the reference lists of key studies, or found through a manual search in the Journal of Medical Internet Research and its sister journals. After removing duplicates, there were 352 articles. Of this pool, 22 studies were relevant for our review and included in our work (Figure 1).

To present the results, we categorized the mHealth devices into three groups: apps ("app"), only smartphones or tablets used as a medium for diagnosis, and "wrist worn wearables" and "other devices."

Apps

Smartphone or tablet apps are characterized by their usability and the fact that no additional device is needed for atrial fibrillation screening. In this field, a general distinction between direct and indirect photoplethysmography (PPG) can be made. Direct PPGs require direct contact between the user and device. Thus, it is possible to measure the pulse by putting a finger above the camera and flashlight while running the app. Indirect PPGs do not require direct contact; they measure the pulse by scanning a body part over a distance.

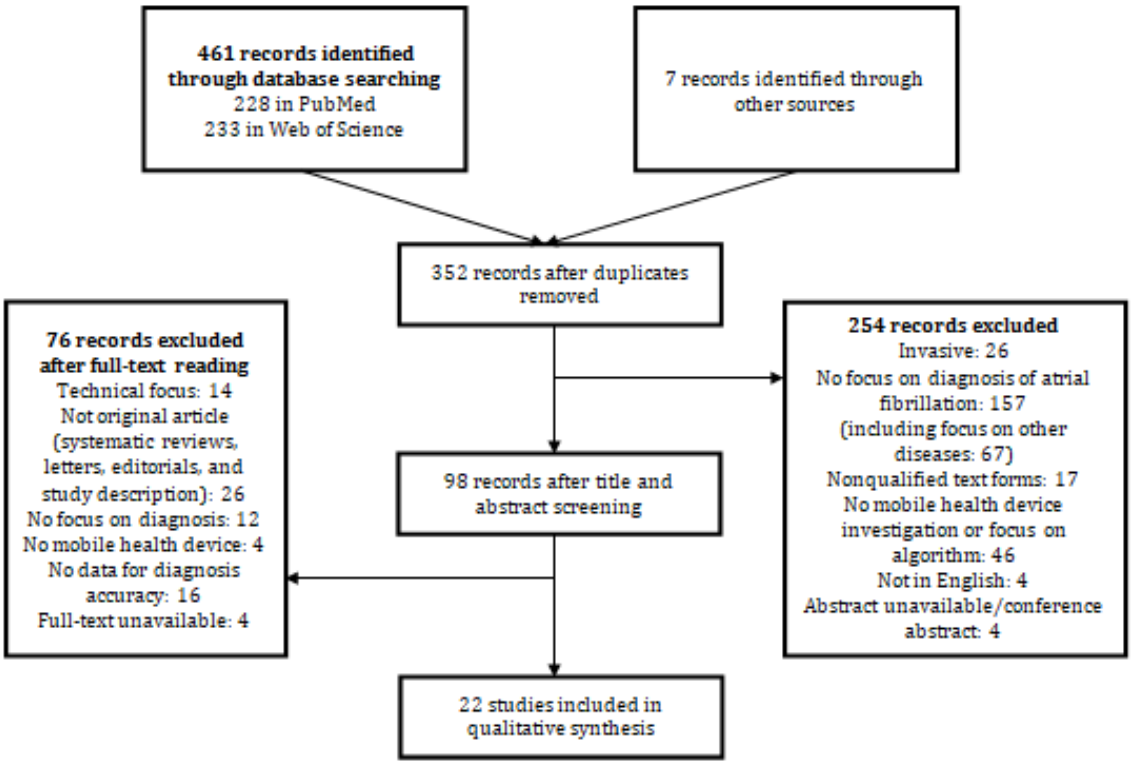
Several apps use the direct method. One of the most common smartphone apps in this context is "Cardiio Rhythm," which can be used either as a direct or an indirect heart rate monitor. Chan et al [39] and Rozen et al [53] investigated the direct use of this app and found a high accuracy in comparison with a single-lead ECG and a 12-lead ECG, respectively.

Krivoshei et al [48] proposed an unnamed app using the direct PPG method. Comparison of the diagnostic results of the app with a heart rate monitor chest belt as a reference method showed high sensitivity and specificity.

Fan et al [41] investigated atrial fibrillation detection through PPG with the aid of either one of two different smartphones or a smart band. Compared to 12-lead ECG, they found high accuracy in both smart phones but concluded that the final diagnosis should be based on ECG. Another study on atrial fibrillation screening with the aid of PPG showed that PPG-based algorithms can reach high accuracy; the authors recommended further investigation using population-based, large-scale atrial fibrillation screening studies [37].

In addition to apps using PPG, there are two fundamentally different approaches. The first one, proposed by Lahdenoja et al [49], is the diagnosis of atrial fibrillation with a smartphone app using the integrated inertial measurement unit. The device is placed on the chest of the patient to measure movement triggered by the heart. Second is the app "CPstethoscope" presented by Kang et al [45] to auscultate the heart. Using this method, they found vast differences in sensitivity, specificity, PPV, and NPV depending on the smartphone model running the app.

Figure 1. Flowchart of the systematic literature review process.



Wrist-Worn Wearables

The most popular wearables used to measure heart frequency and heart rhythm originate from the field of fitness. They appear either as simple bracelets or smartwatches. Besides their potential in supporting basic and clinical research by providing data [57], they have the capability to detect arrhythmias like atrial fibrillation.

In the context of atrial fibrillation, use of such devices can be made in three different ways. First, they can promote healthy behaviors like an active lifestyle. Second, they can support the diagnosis of atrial fibrillation through permanent tracking of heart frequency and rhythm. Lastly, they are able to facilitate coping with the disease [15]. In this review, we focused on diagnosis of atrial fibrillation and the accuracy of mHealth devices.

To detect pulse and heart rhythm, wrist-worn wearables use either PPGs or electrodes. In diagnosis of arrhythmias, the Apple Watch Series 4 is known as the most popular wrist-worn device. It uses both PPG and a two-lead ECG for detection of atrial fibrillation. For the ECG, the first electrode is installed in the digital crown, and the other one is installed on the back of the watch. Thus, this device allows both long-term surveillance of the heart rate through PPG and user-triggered ECG recording with a one-lead ECG.

Koshy et al [47] investigated Apple Watch Series 1 and FitBit Blaze. They showed a high correlation between use of the devices and ECG in patients with sinus rhythm or atrial flutter, but the heart rate in patients with atrial fibrillation tended to be underestimated.

While examining on the Cardiogram app using Apple Watch, Tison et al [55] noted a high accuracy in sedentary patients undergoing cardioversion, but lower accuracy in ambulatory participants.

Another study showed a high accuracy of Apple Watch [38] in combination with the AliveCor KardiaBand, which is evaluated in the AliveCor section below.

In patients with atrial fibrillation, Hochstadt et al [44] found a high correlation between use of such a PPG sensor and simultaneously recorded ECGs in the smartwatch CardiacSense.

Bonomi et al [36] conducted a study on a wrist-worn device that includes a PPG sensor and an accelerometer. When comparing the measurements with either a Holter monitor or a single-lead ECG, the accuracy was high.

Furthermore, while investigating the influence of various conditions on PPGs, Eerikäinen et al [40] found significant differences in sensitivity and specificity between the use in a hospital compared to the use under normal everyday conditions.

Another wrist-worn device, HUAWEI Band 2, was compared to a 12-lead ECG by Fan et al [41]. They found that the PPG smart band is a convenient tool to detect AF at high accuracy.

Other Devices

Just like wrist-worn wearables, other wearable devices have the capability to measure either the pulse or heart rhythm to detect arrhythmias by using loop or event recording. Most recent articles about atrial fibrillation–diagnosing wearables are either about AliveCor or ECG devices integrated in patches like ZioPatch. Despite its FDA approval, we did not find eligible studies focusing on the accuracy of ZioPatch compared to a reference method. Therefore, it was not part of our review. Due

to the high number of studies focusing on the accuracy of AliveCor, we first focused on AliveCor before analyzing studies about other devices.

AliveCor

AliveCor KardiaMobile is an event recorder that has been subject to various studies focusing on its accuracy. The device is an FDA-certified medical product that can record heart beat and rhythm by using a single-lead electrode. To measure heart rate and heart rhythm, the user has to put two fingers on the electrodes fixed to a small plastic plate, following which AliveCor KardiaMobile starts to write an ECG and transmits it to either a mobile phone or a tablet computer. It features a very high sensitivity and specificity. Koltowski et al [46] and Lowres et al [51] found high accuracy of this device compared to the standard 12-lead ECG. Furthermore, Selder et al [54] evaluated an arrhythmia program using AliveCor and reported its high accuracy compared to the reference method, which was a team assessing the device-recorded ECGs.

In a study assessing the accuracy of AliveCor in patients with a history of either paroxysmal or persistent atrial fibrillation, William et al [56] found a very high accuracy in the form of sensitivity and specificity.

Nevertheless, analyses with special populations like children [42] and elite athletes and cardiology clinic patients [43], or patients in either a cardiology or geriatric ward [33] showed slightly to significantly modified accuracy compared to the majority of studies focusing on AliveCor.

In a study by Lown et al [50], AliveCore yielded high accuracy but was not superior to inexpensive consumer devices.

A device related to KardiaMobile is KardiaBand. It is a watchband, but its function is similar to that of KardiaMobile; therefore, we deemed evaluation in combination with KardiaMobile appropriate, even though it is a wrist-worn device. A study by Bumgarner et al [38] found very high sensitivity of the KardiaBand used in combination with an Apple Watch (both FDA approved) compared to a 12-lead ECG.

Other Devices

Besides the devices mentioned above, there are some, less widespread forms of mHealth for the diagnosis of atrial fibrillation. Few of them have been assessed for accuracy.

MyDiagnostick is similar to AliveCor in its functionality. It is a rod-like device with two electrodes on the endings. Desteghe et al [33] compared the device with either 6-lead or 12-lead ECG for its sensitivity, specificity, PPV, and NPV on patients in a cardiology ward and a geriatric ward. Compared to the algorithm, they found that manual interpretation of the device-recorded data led to increased sensitivity, but decreased specificity.

To detect cardiac abnormalities in the home environment of elderly people residing in low and middle-income countries, Mena et al [52] designed and developed a loop recorder ECG sensor device. Two electrodes are attached to the chest and one to the right leg of the patient. The captured data are directly processed by a machine learning algorithm, and the patient receives feedback through his/her smartphone immediately. Furthermore, the data can be transmitted to health care providers. Tested on 100 older adults, the mobile ECG and the corresponding algorithm reached a very high accuracy (97%), sensitivity (100%), and specificity (96,6%). Thus, further development of the device seems useful.

Discussion

Overview

In addition to the devices included in our review, there are many other kinds of mHealth devices to screen for atrial fibrillation, for example, ECG patches like the ZioPatch. Despite its positive evaluation in a multitude of studies [13,58-60], there is no eligible study about its accuracy compared to a reference method. Most of the studies about the ZioPatch compare the detection rate over a given period to the reference method.

To provide an even more accurate diagnosis of atrial fibrillation through mHealth devices, Steijlen et al [61] presented a first approach to allow patients to record an accurate 12-lead ECG at home. They developed a device that can be worn within 8 minutes of first-time use. This device should be studied further.

Another study focusing on the benefit of Apple Watch in the context of irregular heart rhythm detection is the Apple Heart Study [62]. Data about the heart rhythm are received from the Apple Watch and automatically evaluated. If there are irregularities, an app notifies the study participant. Furthermore, there is the possibility for some participants to receive an ePatch and to wear it up to 7 days. After returning the ePatch, the experts offer feedback and recommend further medical care from the participant's own health care provider. The Apple Heart Study enrolls over 400,000 participants and is thus the largest ever study of its kind. The study results are not yet published.

Despite the overall good evaluation of mHealth devices in the context of atrial fibrillation, there are some possible limitations. Shcherbina et al found that exogeneous factors like dark skin color, higher body mass index, and male gender as well as mechanical separating or shifting of PPG during physical activities led to higher device errors [63]. Furthermore, part of the recorded atrial fibrillation screenings were noninterpretable by algorithms [38,56].

Principal Results

In this review, we presented various possibilities to screen for atrial fibrillation. mHealth devices appearing in different forms like smartphone apps, wrist-worn devices, small plates such as the AliveCor, or rod-like devices were investigated for their accuracy. These devices mostly use either ECG or PPG technology to detect atrial fibrillation.

Mobile apps provide a convenient way to screen for atrial fibrillation. Most common are apps using PPG, which allows detection of atrial fibrillation with a high accuracy compared to the gold standard. Furthermore, it is possible to develop apps that use the inertial measurement unit or can be used to auscultate the heart.

Wrist-worn wearables appearing as bracelets or smart watches provide the possibility to measure heart rhythm in an unobtrusive way. The most effective way to guarantee atrial fibrillation detection is to combine PPG and ECG in a wrist worn-wearable device in order to screen over a long-term period and record an accurate user-triggered ECG. This is the case for Apple Watch or Apple Watch in combination with KardiaBand.

The ECG-based AliveCor is one of the few FDA-approved devices. It reaches a very high overall accuracy and benefits from its ease of use. Overall, the use of mHealth devices is convenient [34,61,64,65]. Nevertheless, after atrial fibrillation detection through mHealth devices, the diagnosis should always be confirmed by standard 12-lead ECG Holter monitoring.

Economic Aspects

From an economic point of view, mHealth devices seem to be an eligible possibility to prevent expensive secondary diseases like stroke. Therefore, mobile apps have a high economic potential in screening for atrial fibrillation. Given the fact that smartphones are already widespread in many countries, the economic burden is low. Even if app accuracy does not reach the gold standard, mobile apps can provide a first approach to detect atrial fibrillation.

The integration of atrial fibrillation screening methods in smartwatches and bracelets could be valuable. Smart watches, in particular, have gained popularity during the last few years. Further investigation on the economic effect of subsidizing wrist-worn wearables, which are able to screen for atrial fibrillation, should be performed. A special focus should be placed on the accuracy of these devices to avoid costs due to misdiagnosis.

With a fundamentally different approach, AliveCor benefits from its ease of use. This device seems suitable to integrate in health care as already implemented in the Dutch Hartwacht program [54]. Orchard et al implemented a study to examine the cost-effectiveness of screening with AliveCor in a rural primary care setting. The aim was to screen 2000 patients aged ≥65 years for atrial fibrillation during 3-4 months and to evaluate the process through qualitative interviews as well as cost-effectiveness [56]. Especially for low-income countries, mHealth is a possible approach to screen for atrial fibrillation, which will reduce the economic burden [66,67]. Nevertheless, to assess the real economic potential of mHealth devices in the context of atrial fibrillation screening, further studies for all types of mHealth devices are needed.

Conclusions

The main advantage of mHealth in atrial fibrillation detection is its use in addition to standard care. Even if its accuracy is not yet as high as expected, it is an additional possibility to diagnose atrial fibrillation, especially in its silent, paroxysmal form. Economic assessment of mHealth devices should be further explored.

Conflicts of Interest

None declared.

References

1. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation 2014 Feb 25;129(8):837-847 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.113.005119] [Medline: 24345399]
2. Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG, Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. BMJ 2016 Sep 06;354:i4482 [FREE Full text] [doi: 10.1136/bmj.i4482] [Medline: 27599725]
3. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, ESC Scientific Document Group. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016 Dec 07;37(38):2893-2962. [doi: 10.1093/eurheartj/ehw210] [Medline: 27567408]
4. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007 Jun 19;146(12):857-867. [Medline: 17577005]
5. Ruff CT, Giugliano RP, Braunwald E, Hoffman EB, Deenadayalu N, Ezekowitz MD, et al. Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: a meta-analysis of randomised trials. Lancet 2014 Mar 15;383(9921):955-962. [doi: 10.1016/S0140-6736(13)62343-0] [Medline: 24315724]
6. Johnsen SP, Dalby LW, Täckström T, Olsen J, Fraschke A. Cost of illness of atrial fibrillation: a nationwide study of societal impact. BMC Health Serv Res 2017 Nov 10;17(1):714 [FREE Full text] [doi: 10.1186/s12913-017-2652-y] [Medline: 29126415]
7. Jönsson L, Eliasson A, Kindblom J, Almgren O, Edvardsson N. Cost of illness and drivers of cost in atrial fibrillation in Sweden and Germany. Appl Health Econ Health Policy 2010;8(5):317-325. [doi: 10.2165/11319880-000000000-00000] [Medline: 20804224]
8. Reinhold T, Belke R, Hauser T, Grebmer C, Lennerz C, Semmler V, et al. Cost Saving Potential of an Early Detection of Atrial Fibrillation in Patients after ICD Implantation. Biomed Res Int 2018;2018:3417643 [FREE Full text] [doi: 10.1155/2018/3417643] [Medline: 30186856]

9. Jacobs MS, Kaasenbrood F, Postma MJ, van Hulst M, Tieleman RG. Cost-effectiveness of screening for atrial fibrillation in primary care with a handheld, single-lead electrocardiogram device in the Netherlands. Europace 2016 Oct 12. [doi: 10.1093/europace/euw285] [Medline: 27733465]
10. Levin L, Husberg M, Sobocinski PD, Kull VF, Friberg L, Rosenqvist M, et al. A cost-effectiveness analysis of screening for silent atrial fibrillation after ischaemic stroke. Europace 2015 Feb;17(2):207-214. [doi: 10.1093/europace/euu213] [Medline: 25349228]
11. Aronsson M, Svennberg E, Rosenqvist M, Engdahl J, Al-Khalili F, Friberg L, et al. Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. Europace 2015 Jul;17(7):1023-1029. [doi: 10.1093/europace/euv083] [Medline: 25868469]
12. Mayer F, Stahrenberg R, Gröschel K, Mostardt S, Biermann J, Edelmann F, et al. Cost-effectiveness of 7-day-Holter monitoring alone or in combination with transthoracic echocardiography in patients with cerebral ischemia. Clin Res Cardiol 2013 Dec;102(12):875-884 [FREE Full text] [doi: 10.1007/s00392-013-0601-2] [Medline: 23904073]
13. Solomon MD, Yang J, Sung SH, Livingston ML, Sarlas G, Lenane JC, et al. Incidence and timing of potentially high-risk arrhythmias detected through long term continuous ambulatory electrocardiographic monitoring. BMC Cardiovasc Disord 2016 Feb 17;16:35 [FREE Full text] [doi: 10.1186/s12872-016-0210-x] [Medline: 26883019]
14. Bansal A, Joshi R. Portable out-of-hospital electrocardiography: A review of current technologies. J Arrhythm 2018 Dec;34(2):129-138 [FREE Full text] [doi: 10.1002/joa3.12035] [Medline: 29657588]
15. McConnell MV, Turakhia MP, Harrington RA, King AC, Ashley EA. Mobile Health Advances in Physical Activity, Fitness, and Atrial Fibrillation: Moving Hearts. J Am Coll Cardiol 2018 Jun 12;71(23):2691-2701. [doi: 10.1016/j.jacc.2018.04.030] [Medline: 29880130]
16. Afzal MR, Gunda S, Waheed S, Sehar N, Maybrook RJ, Dawn B, et al. Role of Outpatient Cardiac Rhythm Monitoring in Cryptogenic Stroke: A Systematic Review and Meta-Analysis. Pacing Clin Electrophysiol 2015 Oct;38(10):1236-1245. [doi: 10.1111/pace.12688] [Medline: 26172621]
17. Guzik P, Malik M. ECG by mobile technologies. J Electrocardiol 2016;49(6):894-901. [doi: 10.1016/j.jelectrocard.2016.07.030] [Medline: 27524474]
18. Singh N, Chun S, Hadley D, Froelicher V. Clinical Implications of Technological Advances in Screening for Atrial Fibrillation. Prog Cardiovasc Dis 2018;60(4-5):550-559. [doi: 10.1016/j.pcad.2018.01.007] [Medline: 29341881]
19. Bruining N, Caiani E, Chronaki C, Guzik P, van der Velde E, Task Force of the e-Cardiology Working. Acquisition and analysis of cardiovascular signals on smartphones: potential, pitfalls and perspectives: by the Task Force of the e-Cardiology Working Group of European Society of Cardiology. Eur J Prev Cardiol 2014 Nov;21(2 Suppl):4-13. [doi: 10.1177/2047487314552604] [Medline: 25354948]
20. Fung E, Järvelin M, Doshi RN, Shinbane JS, Carlson SK, Grazette LP, et al. Electrocardiographic patch devices and contemporary wireless cardiac monitoring. Front Physiol 2015;6:149 [FREE Full text] [doi: 10.3389/fphys.2015.00149] [Medline: 26074823]
21. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Per Med 2018 Sep;15(5):429-448 [FREE Full text] [doi: 10.2217/pme-2018-0044] [Medline: 30259801]
22. Shrivastav M, Padte S, Arora V, Biffi M. Pilot evaluation of an integrated monitor-adhesive patch for long-term cardiac arrhythmia detection in India. Expert Rev Cardiovasc Ther 2014 Jan;12(1):25-35. [doi: 10.1586/14779072.2013.867807] [Medline: 24325206]
23. Ajijola OA, Boyle NG, Shivkumar K. Detecting and monitoring arrhythmia recurrence following catheter ablation of atrial fibrillation. Front Physiol 2015;6:90 [FREE Full text] [doi: 10.3389/fphys.2015.00090] [Medline: 25870562]
24. Garabelli P, Stavrakis S, Po S. Smartphone-based arrhythmia monitoring. Curr Opin Cardiol 2017 Jan;32(1):53-57. [doi: 10.1097/HCO.0000000000000350] [Medline: 27875477]
25. Cheung CC, Krahn AD, Andrade JG. The Emerging Role of Wearable Technologies in Detection of Arrhythmia. Can J Cardiol 2018 Aug;34(8):1083-1087 [FREE Full text] [doi: 10.1016/j.cjca.2018.05.003] [Medline: 30049358]
26. Turakhia MP, Kaiser DW. Transforming the care of atrial fibrillation with mobile health. J Interv Card Electrophysiol 2016 Oct;47(1):45-50. [doi: 10.1007/s10840-016-0136-3] [Medline: 27306552]
27. Stephan LS, Almeida ED, Guimarães RB, Ley AG, Mathias RG, Assis MV, et al. Oral Anticoagulation in Atrial Fibrillation: Development and Evaluation of a Mobile Health Application to Support Shared Decision-Making. Arq Bras Cardiol 2018 Jan;110(1):7-15 [FREE Full text] [doi: 10.5935/abc.20170181] [Medline: 29412241]
28. Hirschey J, Bane S, Mansour M, Sperber J, Agboola S, Kvedar J, et al. Evaluating the Usability and Usefulness of a Mobile App for Atrial Fibrillation Using Qualitative Methods: Exploratory Pilot Study. JMIR Hum Factors 2018 Mar 15;5(1):e13 [FREE Full text] [doi: 10.2196/humanfactors.8004] [Medline: 29549073]
29. Desteghe L, Kluts K, Vijgen J, Koopman P, Dilling-Boer D, Schurmans J, et al. The Health Buddies App as a Novel Tool to Improve Adherence and Knowledge in Atrial Fibrillation Patients: A Pilot Study. JMIR Mhealth Uhealth 2017 Jul 19;5(7):e98 [FREE Full text] [doi: 10.2196/mhealth.7420] [Medline: 28724512]
30. Kotecha D, Chua WWL, Fabritz L, Hendriks J, Casadei B, Schotten U, European Society of Cardiology (ESC) Atrial Fibrillation Guidelines Taskforce, the CATCH ME consortiumthe European Heart Rhythm Association (EHRA). European

Society of Cardiology smartphone and tablet applications for patients with atrial fibrillation and their health care providers. Europace 2018 Dec 01;20(2):225-233 [FREE Full text] [doi: 10.1093/europace/eux299] [Medline: 29040548]
31. Guo Y, Chen Y, Lane DA, Liu L, Wang Y, Lip GYH. Mobile Health Technology for Atrial Fibrillation Management Integrating Decision Support, Education, and Patient Involvement: mAF App Trial. Am J Med 2017 Dec;130(12):1388-1396.e6 [FREE Full text] [doi: 10.1016/j.amjmed.2017.07.003] [Medline: 28847546]
32. Marouf M, Vukomanovic G, Saranovac L, Bozic M. Multi-purpose ECG telemetry system. Biomed Eng Online 2017 Jun 19;16(1):80 [FREE Full text] [doi: 10.1186/s12938-017-0371-6] [Medline: 28629374]
33. Desteghe L, Raymaekers Z, Lutin M, Vijgen J, Dilling-Boer D, Koopman P, et al. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. Europace 2017 Jan;19(1):29-39. [doi: 10.1093/europace/euw025] [Medline: 26893496]
34. Halcox JPJ, Wareham K, Cardew A, Gilmore M, Barry JP, Phillips C, et al. Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation: The REHEARSE-AF Study. Circulation 2017 Nov 07;136(19):1784-1794. [doi: 10.1161/CIRCULATIONAHA.117.030583] [Medline: 28851729]
35. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700 [FREE Full text] [Medline: 19622552]
36. Bonomi AG, Schipper F, Eerikäinen LM, Margarito J, van Dinther R, Muesch G, et al. Atrial Fibrillation Detection Using a Novel Cardiac Ambulatory Monitor Based on Photo-Plethysmography at the Wrist. J Am Heart Assoc 2018 Aug 07;7(15):e009351. [doi: 10.1161/JAHA.118.009351] [Medline: 30371247]
37. Brasier N, Raichle CJ, Dörr M, Becke A, Nohturfft V, Weber S, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace 2018 Jul 31. [doi: 10.1093/europace/euy176] [Medline: 30085018]
38. Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, et al. Smartwatch Algorithm for Automated Detection of Atrial Fibrillation. J Am Coll Cardiol 2018 May 29;71(21):2381-2388. [doi: 10.1016/j.jacc.2018.03.003] [Medline: 29535065]
39. Chan P, Wong C, Poh YC, Pun L, Leung WW, Wong Y, et al. Diagnostic Performance of a Smartphone-Based Photoplethysmographic Application for Atrial Fibrillation Screening in a Primary Care Setting. J Am Heart Assoc 2016 Dec 21;5(7) [FREE Full text] [doi: 10.1161/JAHA.116.003428] [Medline: 27444506]
40. Eerikäinen L, Dekker L, Bonomi A, Vullings R, Schipper F, Margarito J. Validating Feature for Atrial Fibrillation Detection from Photoplethysmogram under Hospital and Free-living Conditions. In: Computing in Cardiology. 2017 Presented at: 44th Computing in Cardiology Conference, CinC 2017; Sept 24-27, 2017; Rennes, France p. 1-4 URL: https://research.tue.nl/en/publications/validating-features-for-atrial-fibrillation-detection-from-photop-2 [doi: 10.22489/CinC.2017.050-177]
41. Fan Y, Li Y, Li J, Cheng W, Shan Z, Wang Y, et al. Diagnostic Performance of a Smart Device With Photoplethysmography Technology for Atrial Fibrillation Detection: Pilot Study (Pre-mAFA II Registry). JMIR Mhealth Uhealth 2019 Mar 05;7(3):e11437 [FREE Full text] [doi: 10.2196/11437] [Medline: 30835243]
42. Gropler MRF, Dalal AS, Van Hare GF, Silva JNA. Can smartphone wireless ECGs be used to accurately assess ECG intervals in pediatrics? A comparison of mobile health monitoring to standard 12-lead ECG. PLoS One 2018;13(9):e0204403 [FREE Full text] [doi: 10.1371/journal.pone.0204403] [Medline: 30260996]
43. Haberman ZC, Jahn RT, Bose R, Tun H, Shinbane JS, Doshi RN, et al. Wireless Smartphone ECG Enables Large-Scale Screening in Diverse Populations. J Cardiovasc Electrophysiol 2015 May;26(5):520-526. [doi: 10.1111/jce.12634] [Medline: 25651872]
44. Hochstadt A, Chorin E, Viskin S, Schwartz AL, Lubman N, Rosso R. Continuous heart rate monitoring for automatic detection of atrial fibrillation with novel bio-sensing technology. J Electrocardiol 2019;52:23-27. [doi: 10.1016/j.jelectrocard.2018.10.096] [Medline: 30476634]
45. Kang S, Joe B, Yoon Y, Cho G, Shin I, Suh J. Cardiac Auscultation Using Smartphones: Pilot Study. JMIR Mhealth Uhealth 2018 Feb 28;6(2):e49 [FREE Full text] [doi: 10.2196/mhealth.8946] [Medline: 29490899]
46. Koltowski L, Balsam P, Głowczynska R, Rokicki JK, Peller M, Maksym J, et al. Kardia Mobile applicability in clinical practice: A comparison of Kardia Mobile and standard 12-lead electrocardiogram records in 100 consecutive patients of a tertiary cardiovascular care center. Cardiol J 2019 Jan 15 [FREE Full text] [doi: 10.5603/CJ.a2019.0001] [Medline: 30644079]
47. Koshy AN, Sajeev JK, Nerlekar N, Brown AJ, Rajakariar K, Zureik M, et al. Smart watches for heart rate assessment in atrial arrhythmias. Int J Cardiol 2018 Sep 01;266:124-127. [doi: 10.1016/j.ijcard.2018.02.073] [Medline: 29887428]
48. Krivoshei L, Weber S, Burkard T, Maseli A, Brasier N, Kühne M, et al. Smart detection of atrial fibrillation†. Europace 2017 May 01;19(5):753-757 [FREE Full text] [doi: 10.1093/europace/euw125] [Medline: 27371660]
49. Lahdenoja O, Hurnanen T, Iftikhar Z, Nieminen S, Knuutila T, Saraste A, et al. Atrial Fibrillation Detection via Accelerometer and Gyroscope of a Smartphone. IEEE J Biomed Health Inform 2018 Jan;22(1):108-118. [doi: 10.1109/JBHI.2017.2688473] [Medline: 28391210]

50. Lown M, Yue AM, Shah BN, Corbett SJ, Lewith G, Stuart B, et al. Screening for Atrial Fibrillation Using Economical and Accurate Technology (From the SAFETY Study). Am J Cardiol 2018 Oct 15;122(8):1339-1344. [doi: 10.1016/j.amjcard.2018.07.003] [Medline: 30131106]
51. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. The SEARCH-AF study. Thromb Haemost 2014 Jun;111(6):1167-1176. [doi: 10.1160/TH14-03-0231] [Medline: 24687081]
52. Mena LJ, Félix VG, Ochoa A, Ostos R, González E, Aspuru J, et al. Mobile Personal Health Monitoring for Automated Classification of Electrocardiogram Signals in Elderly. Comput Math Methods Med 2018;2018:9128054 [FREE Full text] [doi: 10.1155/2018/9128054] [Medline: 30002725]
53. Rozen G, Vaid J, Hosseini SM, Kaadan MI, Rafael A, Roka A, et al. Diagnostic Accuracy of a Novel Mobile Phone Application for the Detection and Monitoring of Atrial Fibrillation. Am J Cardiol 2018 May 15;121(10):1187-1191. [doi: 10.1016/j.amjcard.2018.01.035] [Medline: 29525063]
54. Selder JL, Breukel L, Blok S, van Rossum AC, Tulevski II, Allaart CP. A mobile one-lead ECG device incorporated in a symptom-driven remote arrhythmia monitoring program. The first 5,982 Hartwacht ECGs. Neth Heart J 2019 Jan;27(1):38-45 [FREE Full text] [doi: 10.1007/s12471-018-1203-4] [Medline: 30523617]
55. Tison GH, Sanchez JM, Ballinger B, Singh A, Olgin JE, Pletcher MJ, et al. Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch. JAMA Cardiol 2018 May 01;3(5):409-416. [doi: 10.1001/jamacardio.2018.0136] [Medline: 29562087]
56. William AD, Kanbour M, Callahan T, Bhargava M, Varma N, Rickard J, et al. Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: The iREAD Study. Heart Rhythm 2018 Dec;15(10):1561-1565. [doi: 10.1016/j.hrthm.2018.06.037] [Medline: 30143448]
57. Lim WK, Davila S, Teo JX, Yang C, Pua CJ, Blöcker C, et al. Beyond fitness tracking: The use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. PLoS Biol 2018 Dec;16(2):e2004285 [FREE Full text] [doi: 10.1371/journal.pbio.2004285] [Medline: 29485983]
58. Turakhia MP, Ullal AJ, Hoang DD, Than CT, Miller JD, Friday KJ, et al. Feasibility of extended ambulatory electrocardiogram monitoring to identify silent atrial fibrillation in high-risk patients: the Screening Study for Undiagnosed Atrial Fibrillation (STUDY-AF). Clin Cardiol 2015 May;38(5):285-292 [FREE Full text] [doi: 10.1002/clc.22387] [Medline: 25873476]
59. Steinhubl SR, Waalen J, Edwards AM, Ariniello LM, Mehta RR, Ebner GS, et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA 2018 Dec 10;320(2):146-155. [doi: 10.1001/jama.2018.8102] [Medline: 29998336]
60. Barrett PM, Komatireddy R, Haaser S, Topol S, Sheard J, Encinas J, et al. Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. Am J Med 2014 Jan;127(1):95.e11-95.e17 [FREE Full text] [doi: 10.1016/j.amjmed.2013.10.003] [Medline: 24384108]
61. Steijlen AS, Jansen KM, Albayrak A, Verschure DO, Van Wijk DF. A Novel 12-Lead Electrocardiographic System for Home Use: Development and Usability Testing. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e10126 [FREE Full text] [doi: 10.2196/10126] [Medline: 30061094]
62. Turakhia MP, Desai M, Hedlin H, Rajmane A, Talati N, Ferris T, et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. Am Heart J 2019 Jan;207:66-75 [FREE Full text] [doi: 10.1016/j.ahj.2018.09.002] [Medline: 30392584]
63. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. J Pers Med 2017 May 24;7(2) [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
64. Kropp C, Ellis J, Nekkanti R, Sears S. Monitoring Patients With Implantable Cardioverter Defibrillators Using Mobile Phone Electrocardiogram: Case Study. JMIR Cardio 2018 Feb 21;2(1):e5. [doi: 10.2196/cardio.8710]
65. Veale EL, Stewart AJ, Mathie A, Lall SK, Rees-Roberts M, Savickas V, et al. Pharmacists detecting atrial fibrillation (PDAF) in primary care during the influenza vaccination season: a multisite, cross-sectional screening protocol. BMJ Open 2018 Dec 14;8(3):e021121 [FREE Full text] [doi: 10.1136/bmjopen-2017-021121] [Medline: 29540425]
66. Soni A, Earon A, Handorf A, Fahey N, Talati K, Bostrom J, et al. High Burden of Unrecognized Atrial Fibrillation in Rural India: An Innovative Community-Based Cross-Sectional Screening Program. JMIR Public Health Surveill 2016 Oct 13;2(2):e159 [FREE Full text] [doi: 10.2196/publichealth.6517] [Medline: 27737818]
67. Evans GF, Shirk A, Muturi P, Soliman EZ. Feasibility of Using Mobile ECG Recording Technology to Detect Atrial Fibrillation in Low-Resource Settings. Glob Heart 2017 Dec;12(4):285-289. [doi: 10.1016/j.gheart.2016.12.003] [Medline: 28302547]

Abbreviations

AF: atrial fibrillation
ECG: electrocardiogram

FDA: Food and Drug Administration
mHealth: mobile health
NPV: negative predictive value
PPG: photoplethysmography
PPV: positive predictive value

Edited by G Eysenbach; submitted 01.03.19; peer-reviewed by A Hernández, J Kors, L Saxon, R Casado; comments to author 20.03.19; revised version received 14.05.19; accepted 14.05.19; published 16.06.19.

Please cite as:
Giebel GD, Gissel C
Accuracy of mHealth Devices for Atrial Fibrillation Screening: Systematic Review
JMIR Mhealth Uhealth 2019;7(6):e13641
URL: http://mhealth.jmir.org/2019/6/e13641/
doi:10.2196/13641
PMID:31199337

Original Paper

Perceptions of Health Care Providers Regarding a Mobile Health Intervention to Manage Chronic Obstructive Pulmonary Disease: Qualitative Study

Meshari F Alwashmi[1], BN, MSc; Beverly Fitzpatrick[1], PhD; Erin Davis[1], PharmD; John-Michael Gamble[2], PhD; Jamie Farrell[1], MD, FRCPC; John Hawboldt[1], PharmD

[1]Memorial University of Newfoundland, St John's, NL, Canada

[2]School of Pharmacy, Faculty of Science, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Meshari F Alwashmi, BN, MSc
Memorial University of Newfoundland
Health Sciences Centre
300 Prince Philip Drive
St John's, NL, A1B 3V6
Canada
Phone: 1 7096910728
Email: mfa720@mun.ca

Abstract

Background: Using a mobile health (mHealth) intervention, consisting of a smartphone and compatible medical device, has the potential to enhance chronic obstructive pulmonary disease (COPD) treatment outcomes while mitigating health care costs.

Objective: The aim of this study was to explore the potential facilitators and barriers among health care providers (HCPs) regarding the use of mHealth interventions for COPD management.

Methods: This was a qualitative study. Semistructured individual interviews were conducted with HCPs, including nurses, pharmacists, and physicians who work directly with patients with COPD. A flexible prompts guide was used to facilitate discussions. Interview topics included the following: demographics, mHealth usage, perceptions toward challenges of mHealth adoption, factors facilitating mHealth adoption, and preferences regarding features of the mHealth intervention for COPD management. Interviews were conversational in nature, and items were not asked verbatim or in the order presented. The interviews were transcribed verbatim and compared against the digital recordings to ensure the accuracy of the content. After creating a codebook for analysis, 2 researchers independently coded the remaining interview data using pattern coding. They discussed commonalities and differences in coding until a consensus was reached.

Results: A total of 30 nurses, physicians, and pharmacists participated. The main facilitators to mHealth adoption are possible health benefits for patients, ease of use, educating patients and their HCPs, credibility, and reducing cost to the health care system. Alternatively, the barriers to adoption are technical issues, privacy and confidentiality issues, lack of awareness, potential limited uptake from the elderly, potential limited connection between patients and HCPs, and finances.

Conclusions: It is important to understand the perceptions of HCPs regarding the adoption of innovative mHealth interventions for COPD management. This study identifies some potential facilitators and barriers that may inform the successful development and implementation of mHealth interventions for COPD management.

(JMIR Mhealth Uhealth 2019;7(6):e13950) doi:10.2196/13950

KEYWORDS

mHealth; COPD; qualitative; smartphone; eHealth; technology

Introduction

Background

The surge in computing power and mobile connectivity has led to the emergence of mobile health (mHealth) that can transform the mode and quality of clinical research and health care [1]. mHealth is defined by the National Institutes of Health as the use of mobile and wireless devices to improve health outcomes, health care services, and health research. An mHealth intervention could also include the use of a medical device that is compatible with a smartphone. Evidence suggests that mHealth interventions may benefit patients with many chronic health conditions including chronic obstructive pulmonary disease (COPD) [2-5].

Although COPD is a preventable and treatable condition, it is estimated to be the third leading cause of death worldwide by 2020 [6]. According to the Canadian Institute for Health Information, COPD now accounts for the highest rate of hospital admission and readmission among major chronic illnesses in Canada [7]. The Conference Board of Canada has stated that the combined direct and indirect costs of COPD will increase from just under CAD $4 billion in 2010 to roughly $9.5 billion by 2030, an increase of 140% [8]. Dynamic modeling has shown that any intervention that can reduce the number of exacerbations in a population will have a substantial impact on morbidity and costs associated with COPD [8,9]. The authors previously published a systematic review and noted that the current literature on the role of smartphones in reducing COPD exacerbations is limited but does suggest that smartphone interventions may reduce COPD exacerbations [2].

Importance of Human-Centered Design

The International Organization for Standardization (ISO) 9241-210 standard defines human-centered design (HCD) as "an approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques" [10]. The ISO uses the term HCD instead of user-centered design to "address impacts on a number of stakeholders, not just those typically considered as users" [10]. However, in practice, these terms are often used synonymously.

There is increasing interest from academics and clinicians in harnessing smartphone apps as a means of delivering behavioral interventions for health; however, research on the development and evaluation of such apps is in the relatively early stages [11]. Many of the barriers to using mHealth can be avoided with better planning and collaboration [12]. Testing mHealth interventions with patients has revealed preferences and concerns unique to the tested population [13-15]. When developing an mHealth intervention, Hopkins et al [16] encourage including insights from key users to potentially improve the process and the outcome of the intervention.

Triantafyllidis et al used an iterative approach to refine a tablet computer–based home monitoring system for heart failure patients [17]. There was limited uptake of the system owing to usage difficulties and low levels of patient satisfaction. The authors recommended patient-centered approaches for sustainable delivery of remote health monitoring services [17]. Patient-centered care recognizes the complex, subjective, and changing nature of the patient's health status [18]; in addition, it links multiple episodes of care offered by diverse providers into continuous, integrated care trajectories unique to particular patients [19,20]. Developing a COPD mHealth intervention with insights from health care providers (HCPs) working with patients with COPD will potentially improve the process and outcome of the mHealth intervention.

Involvement of Health Care Providers

HCD is particularly suited to developing mHealth interventions, which generally involve multiple stakeholders. Bender suggests a collaborative care approach within teams comprising physicians, nurses, pharmacists, and patient advocates to lead to better care and higher patient adherence for complex and comorbid conditions [21]. Also, other researchers recommend the involvement of a multidisciplinary team in mHealth interventions to develop tailored messages [22], address patient medication needs [23], enhance physical activity in patients with COPD [24], and support the management of heart failure [17], diabetes [25], and cancer [26]. Chiang et al [27] stated that few studies have examined the obstacles faced by HCPs when carrying out telehealth interventions. Similar obstacles in mHealth need to be addressed.

Nursing, medicine, and pharmacy are some of the largest health professions in Canada. Nurses promote COPD management by supportive, preventive, therapeutic, palliative, and rehabilitative means to gain or maintain optimal function [28]. Physicians assess the condition of COPD and diagnose, treat, and prevent any disease, disorder, or dysfunction whereas pharmacists play a role in the promotion of health, prevention, and treatment of COPD through monitoring and management of medication therapy [28]. Furthermore, the role of pharmacists has shifted from drug dispensing responsibilities to the provision of direct patient care [29]. By obtaining the perspectives of nurses, physicians, and pharmacists, we hope to understand the facilitators and barriers affecting some of the largest health professions in Canada. Furthermore, it will enable us to understand the differences in requirements for an mHealth intervention.

Human-Centered Design in Chronic Obstructive Pulmonary Disease

Although mHealth is gaining popularity in recent years, patient and HCP perspectives toward using mHealth for COPD management are relatively unexplored [30]. One study provided insight into the perceptions of COPD patients and their HCPs toward using mHealth for COPD management. They stated that potential barriers to use mHealth include the following: patients avoiding confrontation with the disease, preference for personal contact with an HCP, difficulties with displaying feelings in an application leading to invalid patient measures and lack of trust in advising characteristics of an mHealth intervention, and lack of enthusiasm for mHealth by HCPs [30]. They also recommended including a larger sample of HCPs with more mHealth experience in future studies [30].

To improve the success of mHealth solutions in COPD management, we suggest including HCPs who work with patients with COPD in the development process. Lessons learned will bridge the knowledge gap of barriers and facilitators for mHealth uptake in COPD management. It will also be offered as a guide for research and technology developers working with COPD patients and their HCPs.

Methods

Purpose

This study was intended to explore and develop an understanding of potential facilitators and barriers that might influence HCPs using mHealth interventions for COPD management.

Study Design

We used a descriptive qualitative research design that was grounded in pragmatism [31,32]. Using a qualitative methodology allowed us to achieve an in-depth, contextualized picture of how a diverse sample of HCPs, in this case nurses, pharmacists, and physicians, think and feel about the possibilities and challenges of using mHealth. This has a pragmatic value as mHealth is an emerging option for delivering health care.

Recruitment and Study Setting

HCPs involved in the treatment of patients with COPD were eligible to participate. The primary investigator (PI) contacted the Newfoundland and Labrador Medical Association, the Association of Registered Nurses of Newfoundland and Labrador, and the Pharmacists' Association of Newfoundland and Labrador. These organizations were asked to forward a recruitment email to their mailing lists or post it in their websites. Interested HCPs contacted the PI via email or telephone, who then scheduled appointments to complete the consent forms and conduct the interviews. Our sample consisted of 30 HCPs: 10 nurses, 10 pharmacists, and 10 physicians. The study took place in St. John’s, Canada. We conducted some interviews at Memorial University and others at the participants’ offices or homes.

We used purposive typical case sampling to gather information that would reflect typical cases of mHealth use [32,33]. We also used a criterion-based selection [32] so that we could categorize participant characteristics such as age, familiarity with mHealth, health care profession, and years of experience. In addition, as the interviews progressed, some participants were recruited by snowball or chain sampling, where participants suggested other possible HCPs [32,34]. Snowball or chain sampling was used to ask a few information-rich participants for additional contacts to provide confirming or different perspectives, allowing for richer data [32].

Participants were recruited from April 2018 to August 2018. We first contacted nurses, and after interviewing 7 to 8 nurses, we reached saturation as we were not gathering new information. However, we continued interviewing until 10 nurses were interviewed. This was to strengthen the validity of inferences [35]. We used the same sampling strategy for the remaining professions, with similar saturation points and continuing to interview the 10 participants for each profession. Our final sample size was comparable with similar qualitative studies [30,36,37].

Ethical Considerations

Ethical approval for this study was obtained from the Newfoundland and Labrador Health Research Ethics Authority (HREB -2017-194). Before agreeing to participate, all subjects were informed about the nature of the research project, possible risks and benefits, and their rights as research subjects. All participants completed a written consent form. They were also given a copy of the consent form.

Data Collection

We conducted individual semistructured interviews to gain an understanding of the everyday life-worlds of HCPs in relation to using mHealth [38,39]. Using semistructured interviews allowed the interviewer to begin with a broad question to direct the focus of the interview and then to provide an opportunity for the HCPs to bring forth their thoughts and feelings about the phenomenon that they thought were important [38,39]. The interview prompts are available in Multimedia Appendix 1. If participants identified that they have not used mHealth, they were asked questions pertaining to why they had not used mHealth (barriers). However, we did not ask them about facilitators because they did not have the experience to answer these questions. To facilitate discussions, the interviews were conversational in nature and items were not asked verbatim or in the order presented. As the study progressed, emerging issues were explored with subsequent participants to refine the themes.

The prompts were informed by findings from the literature and input from the authors who have diverse backgrounds including mHealth, pharmacy, nursing, medicine, respirology, family medicine, education, and qualitative research.

The interviews were recorded to enable transparent and accurate transcription. Interview lengths ranged from 20 to 60 min. Topics included the following: demographics, mHealth usage, perceptions toward challenges of mHealth adoption, factors facilitating mHealth adoption, and preferences regarding features of the mHealth intervention for COPD management. Owing to the large amount of data, preferences regarding features of the mHealth intervention will be published in another article. Data consisted of more than 13 hours of interview time with approximately 300 pages of transcription.

Data Analysis

The interviews were transcribed verbatim and compared against the digital recordings to ensure the accuracy of the content. Identifying information (names) was removed to protect anonymity. We used NVivo (version 12; QSR International) to organize the data and examine the words, including frequency counts, as in classical content analysis [40]. All data were analyzed, but we only coded the data that were relevant for answering the research questions, as recommended by Saldana [41], Wolcott [42], and Yin [43]. An audit trail was created to keep track of all analytic decisions [44].

After using NVivo, we used first cycle coding with the nurses' data that were both structural and holistic [41], meaning that we used the interview prompts and the literature to guide some of the coding. One researcher analyzed the transcripts and developed a set of themes and subthemes and then obtained input from a second researcher. In the second cycle of coding, the 2 researchers independently coded the nurses' data using pattern coding to develop themes [41]. They then discussed commonalities and differences in their coding and theme development until a consensus was reached. The analysis of the nurses' data was mainly inductive and iterative throughout as we went back and forth among the data, the coding, and the themes [45].

After the nursing analysis was finished, we completed the same 2 cycles of analysis for the pharmacist and physician data. These 2 analyses included inductive and deductive analysis. However, the analysis was more deductive in nature as themes had already been developed from the nursing data. The iterative process continued as these analyses were conducted to find commonalities, differences, and new patterns in thinking in relation to the nurses' data. Once these 3 sets of analysis were complete, the 2 researchers discussed common and different trends among the 3 HCP groups to develop final themes that encompassed all the HCPs.

Results

Demographics

The sample included HCPs who worked with patients with COPD in various settings, including respirology clinics, cancer clinics, critical care, long-term care, and community health. Some HCPs founded a medical technology company or had a software programming background. About half of the HCPs had experience with an mHealth intervention to manage COPD. Participant demographics are outlined in Table 1.

The majority of HCPs thought that mHealth can play a role in COPD management; however, some HCPs had opposing views. One nurse who implements an mHealth intervention to manage COPD indicated that "...the majority of our patients are very sad to leave the programme." However, one physician expressed his concern:

> *There hasn't been a lot of evidence to prove that this makes a difference in terms of patient outcomes...I think those people are just happy to have another set of eyes watching them, right. I think it probably gives them reassurance.*

Finally, a pharmacist said,

> *There's obviously going to be some patients who don't want to do it who are technology averse in which case that's totally fine, they can use the traditional methods.*

We developed themes under 2 categories: facilitators and barriers that would influence the feasibility and use of mHealth. Table 2 summarizes the main facilitators and Table 3 summarizes the main barriers. We have also included details and examples to illustrate the HCPs' thoughts and beliefs.

Table 1. Participant demographics.

Demographics	Sample size	Age (years), mean (SD)	Years of experience, mean (SD)
Nurses	5	47.3 (6)	19.6 (9)
mHealth[a,b] nurses	5	40.6 (10)	15.8 (10)
Physicians	5	37 (9)	8.4 (8.7)
mHealth physicians	5	41.2 (12)	14.4 (11)
Pharmacists	7	35.7 (11)	11.4 (10)
mHealth pharmacists	3	27.5 (4)	3.6 (2)

[a]mHealth: mobile health.

[b]Experience in using a mobile health intervention.

Table 2. Themes with specific examples regarding the facilitators of mobile health (mHealth) adoption.

Theme	Specific examples for each theme
There are possible health benefits for patients	• Patients can become more readily educated about their disease; • In areas with limited access to health care, mHealth technologies can bridge the gap between patients and health care providers; • Patients can become more motivated, empowered, and accountable with managing their health care
The software needs to be easy to use	• The technology needs to be simple; • The language should be basic; • The software should be visually appealing
Health care providers and patients need to be educated on the use of mHealth	• Educational strategies are needed
The credibility of mHealth should be evident	• Evidence about the effectiveness of mHealth is important; • The credibility of the developer is important
mHealth should reduce the cost to the health care system	• It results in a decreased use of health care resources; • It is affordable owing to the reduced cost of medical devices, and it does not include a large physical infrastructure; • Partnering with private entities could facilitate uptake

Table 3. Themes with specific examples regarding the barriers of mobile health (mHealth) adoption.

Theme	Specific examples for each theme
There are technical issues with mHealth	• It may include equipment malfunction, password issues, and interoperability; • It requires internet access; • Many clinics are paper based
There may be privacy and confidentiality concerns	• People, other than the patients, might gain access to private information
Lack of awareness is a challenge	• Many HCPs[a] and patients are not aware of the current advancements in mHealth
There may be limited uptake from the elderly	• Some HCPs thought older age may be a barrier to technology adoption; • Some believed the upcoming generation will be more familiar with technology
mHealth may limit the personal connection between HCPs and patients	• Some thought personal connections are necessary; • Others thought the advantages of mHealth outweigh personal connections; • Others thought a hybrid approach might be optimal
There are possible financial barriers; There were a few challenges mentioned by a minority of HCPs	• This includes the high cost of the mHealth intervention, time consumption, and lack of billing codes for HCPs; • These included false sense of security, anxiety, lack of motivation, and loss to follow up.

[a]HCP: health care provider.

Facilitators

There are Possible Health Benefits for Patients

Pharmacists, nurses, and physicians agreed that mHealth has the potential to provide health benefits to patients. One nurse, who was experienced with mHealth remarked,

> *It could make life for them, you know, much easier, and improve their quality of life.*

Another nurse felt confident that patients would be "more educated about their diseases and about what things they should be looking for." A physician commented on his patients who were enrolled in a mHealth intervention program,

> *...I think those people are just happy to have another set of eyes watching them, right. I think it probably gives them reassurance.*

Some HCPs mentioned that mHealth could increase patient autonomy through simulating empowerment and motivation in patients. The following physician statement represents thoughts from several other HCPs, "it would give patients the power to then be a part of their management plan, which is better when patients are empowered, because they feel in control of their health." A few pharmacists also mentioned increased motivation as part of this same vein of thought and talked about "access to motivation or making the patient really feel like they were more kind of involved in their own healthcare." Some nurses indicated that mHealth interventions could provide a sense of accountability:

> *There's a sense of accountability I believe from the patients. The nurse is watching me this morning, I better do it because she'll be waiting or he'll be waiting, definitely.*

Access to health care in rural areas was also thought to be an important facilitator. Many HCPs highlighted the importance of mHealth in reducing travel time and improving access to rural areas (Newfoundland and Labrador, Canada). A nurse observed,

> *You look at all these small communities in and around the island, those people could certainly benefit from some kind of remote monitoring.*

This thought was reinforced by others, as in this physician statement,

> *I think that is probably the best benefit from Mobile Health in this province is that it can reach some of those rural communities where we can't go and see patients.*

As part of rural health care, it was consistently noted that mHealth would make it easier for HCPs to provide care. For example, a nurse pointed out that mHealth would help with management of time and perhaps allow for more patients to be monitored, as the following comment demonstrates,

> *...what they can achieve in a video appointment is sometimes quicker, and a bit more targeted and efficient, and they can fit them in within their other appointments.*

mHealth should reduce cancelled appointments and hospital visits as patients would not have to leave their homes, in urban as well as in rural areas. One physician expressed this concern about hospital visits, coupled with the advantage of mHealth:

> *...you can just send that from home. Not even have to go into a facility. And sometimes that's really onerous for people, especially people who are suffering from COPD, so they're going to have shortness of breath and exertion and find it even harder to get from the parking lot into the hospital, so the more you can do to make their lives easier, it's great.*

The Software Needs to Be Easy to Use

Usability was highlighted by the majority of HCPs as an important factor in increasing the uptake of mHealth. One nurse with experience in conducting mHealth interventions cautioned that patients may stop using the intervention owing to usability issues:

> *...they found it hard, I'd say largely related to the technology, not being able to handle it or finding it too much work. Too tiring, too much trouble, not for them, that kind of thing.*

Thus, most HCPs recommended the software to be easy to use. As another nurse pointed out:

> *...people are overwhelmed when they are diagnosed with something that is new and complicated, and affects something as important as your breathing. So, this has got to be something that is easy for them to access and, I think, easy for them to see benefits from.*

Some pharmacists recommended using simple language to enhance usability, as in "it needs to be kept useful, but also simple enough for them to be able to navigate and use." One nurse reinforced this notion and thought the language should be "set at a grade six reading level, so there's no issues with comprehension of what they're being asked or told." It was also thought that the software should be visually appealing, with color and perhaps daily progress or weekly tracking graphs. Font size was also raised as an issue. One nurse quipped, "people my age and above can't see. A lot of it is very tiny, so the need for reading glasses." This was apt as COPD generally develops in later stages of life.

In addition, one nurse with experience in mHealth interventions said HCPs may not use the intervention if it was difficult to use, "where the provider is getting all this information, doesn't feel that comfortable sorting through it, or using it to make clinical decisions, and then it just is going to no use it." So, users and providers need software that is easy to use as well as comprehensive. To streamline the physician workflow, one physician suggested that data collected by the mHealth intervention should be accessed via the electronic medical record: "I have an electronic medical record so it would be nice if it was actually in electronic format."

Health Care Providers and Patients Need to Be Educated on the Use of Mobile Health

It was recognized that mHealth is a different type of learning for many HCPs as it includes learning about technology instead of diseases. However, as one nurse rationalized, "we need to make sure we are staying up and current and on top of this." Many strategies were suggested for educating HCPs, such as integrating information about digital health and mHealth in school curricula; self-learning; Web-based learning; learning from coworkers, students, and sales representatives; attending educational sessions; and hiring coordinators for support.

The necessity to educate patients was also acknowledged. As one pharmacist suggested,

> *I guess most patients with COPD are older and would probably benefit from someone walking through the app with them and showing them how to use it.*

Hands-on learning, supplementary print materials, and a video tutorial were suggested as ways to teach patients how to use the

software. Others mentioned the convenience of having family support as an enabler.

In terms of who should teach patients, it was thought by some that HCPs should share the responsibility. As advocated by one pharmacist:

> *I guess anyone, if you're seeing a patient or person who is in need of that service could introduce it. I don't think one person should have to take all the responsibility, or one profession.*

However, this was not an agreed upon idea. Some thought there should be designated people to teach the necessary skills, but there were differing opinions about which group of HCPs should lead the patient education. It was also recommended by some that technical support staff be available as a resource for patients to call when they needed technical help.

The Credibility of Mobile Health Should Be Evident

HCPs thought that the credibility of mHealth needs to be made evident to HCPs and patients. This would help raise awareness to facilitate uptake by HCPs. A physician worded it like this,

> *...if I perceived that this is something that would help someone exercise a little bit more, control their weight, watch their diet, then I would recommend that.*

A nurse was even more specific in terms of evidence:

> *...it would be really important to have some solid, really good evidence to show that, in actual fact, we receive excellent outcomes in terms of quality of life indicators, activity levels, medication usage at a specific time point, be it within one or two years, to decide that this type of monitoring, and this type of connection with your provider is making a difference to your outcomes. I think that type of evidence is what's going to change my mind as a practitioner about whether it's worth using it or not.*

This sentiment was reiterated by a pharmacist who thought that "knowing if there's evidence to actually support its use" was essential.

In addition, the credibility of the developer was mentioned, as in this statement from a pharmacist, "it's also about the credibility of who's putting the app together." Added to this, recommendations from credible HCPs were also thought to be important. One physician commented:

> *I mean, the power of one's network. If I view something and I think that it's good, then me giving it a vote of confidence that would then get shared, and people would know that I am independently choosing to recommend something.*

Mobile Health Should Reduce the Cost to the Health Care System

It was thought that mHealth has the ability to provide the "clinical assessment and healthcare that was required in a more cost-effective manner", as recommended by one of the nurses. It should decrease emergency visits and hospital admissions, as explained by a nurse who thought it would "hopefully catch things in the earlier stage before these patients who were mostly elderly got in enough trouble that they would end up in the emergency department."

Advancements in mHealth can result in a decrease in expenses, as a third nurse explained,

> *I can send a patient a whole set of devices including a blood pressure cuff, O2 sat machine and a weigh scale for less than 300 dollars.*

Large physical infrastructure would not be required, and it was suggested that some of this could be outsourced to private entities that are already doing this type of work, thereby reducing expenses to taxpayers.

Barriers

There Are Technical Issues With Mobile Health

Many HCPs expressed that technical issues can be barriers for mHealth adoption. Specifically, equipment malfunction, password issues, and interoperability were mentioned frequently. For example, a nurse reported,

> *There's been issues with the technology not communicating because we have setups in four different ways.*

In addition, some technical specifications are required, such as the smartphone being Bluetooth compatible, along with cellular and Wi-Fi connections being available. One nurse elaborated,

> *...there are patients within little pockets of...that don't have cellular service or Internet connection, so unfortunately those patients will not be able to be referred to the program.*

Another limiting condition to sharing mHealth data via electronic medical records was mentioned by physicians, in that many clinics are still paper based or not up to date in technology use.

There May Be Privacy and Confidentiality Concerns

A few HCPs thought privacy and confidentiality could be a barrier to mHealth adoption. A pharmacist, echoing other HCPs, questioned,

> *How are patients confident that the information that's in that app is only going to stay with them and that other people are not going to see that data?*

The concern of family members viewing private information was raised,

> *Patients, you know, if they're competent they don't want their family members to see their information and that could be an issue.*

Also, the issue of stolen or lost phones that contained private information was raised. However, other HCPs thought these issues could be mitigated with security, as expressed by a pharmacist,

> *...if it is secure and the patient gets to decide who accesses it, I don't see it being an issue with confidentiality.*

And, some HCPs, as noted by a nurse, were ambivalent regarding privacy and confidentiality,

> *I wouldn't imagine that there are any more privacy concerns than there are with anything else within health care.*

Lack of Awareness Is a Challenge

Many HCPs indicated that lack of awareness is a major barrier to mHealth adoption. One physician with limited mHealth experience expressed this concern,

> *I think if that had been a part of my training more and I'd seen it more then it could definitely become part of my own training.*

Employers' lack of knowledge was also mentioned. For example, a nurse shared:

> *Our employer doesn't want to see us having them out, people will have the impression we are using it for personal use. That is one big factor. Our employer tells us, keep your phones hidden, don't have them out.*

There May Be Limited Uptake From the Elderly

HCPs had conflicting opinions regarding age and mHealth adoption. All physicians and some nurses and pharmacists agreed that the elderly may face issues in adopting these technologies, as indicated by this pharmacist,

> *A lot of the patients with COPD being older and maybe not as app-savvy as the group that you're aiming towards.*

This thought was reinforced by one physician's words, with a caveat of doubt,

> *I suppose I would assume that the elderly and the more frail would not be tech-savvy, though, I know smartphone use is increasing with the ageing population.*

This caveat was supported by some of the nurses with an mHealth experience, as expressed by one experienced nurse:

> *I had patients who are older than 90 who never owned a computer in their life and managed to do their sessions on their iPads and send it to me with no trouble. So, I think it depends on maybe education level and understanding, and maybe how things are explained to them.*

A couple of pharmacists experienced in mHealth even stated that some elderly people have embraced technology:

> *I've had a lot of kind of older generation patients that once we've kind of sat down they've said oh I've been tracking this or I have this app, and I was kind of shocked. So until you kind of try it out and recommend it to people you never know what they're open to using or what they're already using.*

It was also thought by some that the upcoming generation will be more familiar with technology, as a nurse surmised,

> *We have to be sensitive to the fact that technology is present in my world, it's present in yours, but it wasn't in my grandparents.*

Some physicians also thought that future generations will value and use mHealth more than the current generation,

> *I think the younger generation will, you know, take this in very easily and very much accept it, so I think going forward there's only going to be more of it, not less.*

It was also posited that some older HCPs may face issues when adopting mHealth, as put forward by one pharmacist:

> *I'm sure there'd be some potentially older pharmacists who are less familiar with smartphones and apps that might have more trouble, and may benefit from a tutorial type thing.*

This was reiterated by a physician:

> *I think that probably technology maybe gets pushed to the side. I think that a lot of the physicians too might be, not scared but reluctant to use technology and to learn a new skill, especially if they've been in practice for thirty years or something.*

Mobile Health May Limit the Personal Connection Between Health Care Providers and Patients

As with age, HCPs had conflicting opinions about mHealth and building personal connections between patients and HCPs. A nurse who worried that mHealth might limit the personal connections said:

> *I like to have a bit of actual contact and eye contact, and hear the tone of someone's voice, and a gentle touch sometimes can be so reassuring, you know. I think it's going to be lost with this type of technology.*

However, this same nurse added that even minimal contact could mitigate that barrier, as in "I think there needs to be some sort of human contact, even if it is just the face of the person who receives that information."

Some physicians also agreed that mHealth lacks this type of contact, as in "I don't think you're ever going to really replace that human element." However, although it was emphasized that interacting with patients face to face is better than online, some HCPs struggled with the advantages of human contact versus access. One nurse who was a champion of human contact, recognized that mHealth is:

> *...increasing access and to me, that would be a better benefit than the actual face to face, to be able to reach more people more often.*

Then there were nurses experienced in mHealth who thought it could enhance the personal connection, as in,

> *I think the bond is actually a bit more in this program than it was when I was a bedside nurse in some ways, because you're getting more personal with the patient about other aspects of their healthcare as well.*

One nurse reported that she had done surveys about patient satisfaction, provider satisfaction, and support staff satisfaction

and "the surveys, they do come back that it's similar if not better than a face-to-face."

The majority of physicians, and many pharmacists, thought mHealth has the potential to improve personal connections. This pharmacist's statement represents a commonly expressed example of how this could happen:

> *I think it would strengthen that relationship because you could ask them about their apps and go through it with them when they come other than just seeing if they're late using their prescriptions or late picking them up or anything.*

Some HCPs suggested a hybrid model so that mHealth could supplement the personal connection, as in this physician's comment,

> *Do I think that it could totally replace it, absolutely not...but can I see it being hand-in-and, absolutely.*

This sentiment was supported by another physician:

> *I would like to see them for the initial consultation, but I think for follow-up reports, you know, we could save them great distances from travelling.*

This was also supported by a nurse:

> *I think having regular face-to-face contacts intermittently is still a very important part of healthcare, and it's something that I think will never be completely removed.*

There Are Possible Financial Barriers

HCPs had conflicting opinions about financial implications. A few HCPs said some patients with COPD may not be able to afford mHealth and do not have access to smartphones. One physician expressed it this way,

> *Generally more patients with COPD are falling in the lower socio-economic grouping that wouldn't necessarily be able to afford this.*

A nurse with experience conducting mHealth programs endorsed this concern by adding that only about 10% of participants may remain in the mHealth intervention if the insurance company stopped paying for the service.

In addition to individual patient costs, there is the initial cost of establishing the infrastructure, including costs related to storing data in the cloud. In addition, costs related to the maintenance and replacement of outdated technology were discussed, as reinforced by one nurse,

> *There's a number of equipment across the province that are nine years old, and if they die then there's no replacement.*

In addition, some mHealth programs are limited to a certain period. Participants may get medical devices (eg, blood pressure monitors and pulse oximeters) that have to be returned for cleaning to be used by other participants. One experienced nurse complained that getting medical devices back from patients can be problematic, as in:

> *I have actually been at the plants, the facilities where we get them back and clean them. Andcockroachesin the boxes that were coming back and just swilled with feces and blood and so on. It is just... They have been horrendous.*

Most physicians thought lack of time was a major challenge. They mentioned time to learn about mHealth themselves, time to teach patients, and time to review the results of the mHealth intervention. One physician gave this example:

> *When you get a 12 page report on one patient and you're seeing 40 patients a day and you know time constraints with the amount of work that you do outside in terms of paperwork is already a burden.*

Pharmacists had contradictory views about time. A few pharmacists thought lack of time could be a barrier, as in:

> *Most pharmacists are quite busy as it is... I see the workload potentially going up because now if patients are using this they can't forget to write things down or lose what they documented. It's all there for them, so now they bring the information in.*

Alternatively, other pharmacists thought that mHealth could save time by collecting information required in advance "with the expanded pharmacists role we're building more time to spend with our patients and in that sense we will have that time to teach them and to monitor some of these new technologies that are coming up."

Lack of reimbursement and billing codes were also mentioned as barriers. One experienced pharmacist explained:

> *...there's not a whole lot of reimbursement for services like this and like that's the biggest barrier with most things within the pharmacy profession...doing like daily monitoring on patients like is time-consuming and we definitely want to do it but unfortunately like it does take time and resources and those resources aren't always available.*

In addition, it was emphasized that the lack of billing codes for mHealth is another financial barrier, as a physician insisted:

> *I mean we're all so busy that nobody wants to do anything for free because why would I do that for free if I get paid for it. So that's a barrier that has to be overcome is that how do you change some of the way physicians are paid. There's no incentivizing the optimized care as an example. If I do a poor quality of care for my COPD patient or if I do an excellent quality of care, it's the same payment. So there's a problem with the system in that sense and physicians in general would be resistant to sort of evaluate how well they're doing with their patients.*

There Were a Few Challenges Mentioned by a Minority of Health Care Providers

There were additional challenges that were mentioned by small numbers of HCPs. For example, one physician thought patients may gain "a false sense of security" about their health status, owing to technology. Another physician voiced concern that,

> *...some sub-groups of patients with anxiety might have impaired quality of life because then they become obsessed with that rather than actually just saying okay that's what they're saying, I'm okay.*

A pharmacist questioned validity:

> *...the validity of the data would be something that some people might question. I guess a lot of that would depend on how straightforward the devices are to use or how much training might be required to make sure that they are using it correctly.*

Motivation to continue using the intervention was also a concern. A pharmacist wondered,

> *I think getting patients to use it and use it often enough might be difficult, depending on the patient.*

A few pharmacists and physicians noted that many patients with COPD are not motivated to manage their disease. One pharmacist commented:

> *The biggest challenge I find with COPD patients, now that's the population that I deal with, is that they are smokers and continue to smoke, the majority of them. Their education level is probably a little bit on the lower side and that's related to the whole smoking, right, that kind of thing, the socio-economic status of the patient. So they're not necessarily invested in improving their health with a lot of effort, right. They'll take an inhaler, take a pill to help them get better, but really changing their lifestyle and their smoking is not high on their list.*

One nurse highlighted that about 30% of patients dropped out after using an mHealth intervention.

Discussion

Principal Findings

This qualitative study found that HCPs, in general, had a positive attitude toward mHealth adoption for COPD management, but several facilitators and barriers were identified. More barriers were identified than facilitators, indicating a need to address these barriers to optimize successful implementation of mHealth interventions.

To facilitate mHealth uptake, our thoughts, based on the data, are that both HCPs and patients need to understand the potential benefits of the mHealth intervention. The interventions must be easy to use for both patients and HCPs. This could reduce the time and resources required to teach patients and providers about the mHealth intervention. One physician stated that the use of mHealth interventions could provide a false sense of security, thereby keeping the user from seeking medical advice in a timely manner. This concern along with the lack of awareness concerned HCPs, an important finding is the need for HCPs to teach patients about mHealth interventions. Some HCPs thought there should be a designated person to teach patients. It is preferable that these professionals have a background in chronic disease management and technical support.

There were a few barriers identified by the HCPs. Most of these barriers have the potential to be resolved, as suggested by many of the HCPs. Technical issues continue to be a challenge for mHealth adoption, especially for rural areas and developing countries that have poor connection network.

Comparison With Previous Work

Although the numbers of HCPs using mHealth interventions are growing, studies focusing solely on the frontline staff perspective on mHealth are limited [36,46]. Some of the findings presented in this study confirm findings that have been reported previously in the context of mHealth for COPD management. As Damhus et al [36] noted, HCPs reported technical issues as a major challenge for mHealth adoption. Our findings are in agreement with Vorrink et al [47], who stress the importance of training patients and HCPs on the proper use of mHealth. In this study, as well as other studies, we have noted that mHealth will not replace face-to-face interactions [30,36,47]. In agreement with Damhus et al [36] and Korpershoek et al [30], we suggest that the expected benefits of using mHealth contribute to the success of mHealth uptake, although our study provides additional insight with regard to these perceptions.

Strengths and Limitations

There are several strengths of this study. First, this research is based on a diverse sample of participants. It includes various perspectives by presenting the views of nurses, pharmacist, and physicians, including a respirologist. This human-centered approach ensures that needs and challenges of different people involved in the management of COPD can be considered before developing an mHealth intervention. Second, some HCPs had experience in using an mHealth intervention to manage COPD which further increases the richness of the data. Third, all of the interviews were conducted in a similar manner to ensure consistency during the data collection and analysis. Finally, mHealth is particularly important in geographic locations with relatively large proportions of rural residents such as Newfoundland and Labrador. mHealth may enhance care provider access throughout sparsely populated rural areas. Newfoundland and Labrador has a substantial remote and rural population, therefore our results may be more applicable to rural areas.

There were also several limitations. First, not all the HCPs had experience with using mHealth. Thus, the perceptions of these participants were not based on actual interventions with patients. Second, we used only one data collection method, thus not triangulating data collection. Conducting focus groups with some of the participants following the individual interviews could have yielded richer information as participants would have been given the opportunity to compare their thoughts and confirm or expand upon each other's ideas. This would be a recommendation for a future study.

Implications for Practice

The findings of this study provide insights into the barriers and facilitators for using mHealth as a part of COPD management. This information may help a variety of stakeholders who are planning to use mHealth interventions for COPD management. Lessons learned include the importance of raising an awareness

among patients with COPD and HCPs regarding the potential of mHealth interventions in COPD management. Professional associations and universities could play a significant role in raising an awareness of, and even introducing, mHealth in undergraduate health professional curricula. It also may be beneficial to designate an HCP, with a background in chronic disease management and technical support, to teach patients about mHealth.

The findings emphasize the importance of developing a user-friendly mHealth intervention. This could reduce the time and resources required to teach patients and providers about the mHealth intervention. In addition, the lack of an internet connection limits access to mHealth interventions, so this should be taken into consideration when measuring access to health resources in rural communities.

In terms of credibility, health organizations such as the Food and Drug Administration, Health Canada, or the Canadian Association for Drugs and Technologies should take an active role in regulating mHealth interventions. These organizations can develop their own app stores, similar to the Veteran Affairs app store, to showcase credible mHealth interventions. In addition, when developing mHealth interventions, it is important to follow international guidelines for the exchange, integration, sharing, and retrieval of electronic health information [48]. This could help in addressing interoperability issues. Nevertheless, these regulations should be implemented in a manner that supports mHealth uptake.

Recommendations for Future Research

Future studies would benefit from conducting focus groups with some of the participants following the individual interviews. Focus groups could yield richer information as participants would be given the opportunity to compare their thoughts and confirm or expand upon each other's ideas. Furthermore, including the perspectives of allied HCPs, such as physiotherapists, social workers, and occupational therapists, would be beneficial to understand the perspectives of administrators (eg, information technology managers) who may be able to identify some of the challenges with using mHealth for COPD management. The authors have conducted a similar study with a focus on the perspectives of individuals with COPD. In addition, a future article will focus solely on the features of the ideal mHealth intervention for COPD management. After developing a user-centered mHealth intervention, the authors recommend using a mixed methods framework for usability testing [49].

Conclusions

It is important to understand the perceptions of HCPs regarding the adoption of innovative mHealth interventions for COPD management. This study identifies the facilitators and barriers that may aid in the successful development and implementation of mHealth interventions for COPD management. Lessons from this study may also be applied to other chronic diseases. Additional research is needed to investigate the conflicting opinions regarding mHealth adoption by the elderly, the personal communication between HCPs and patients, and the cost-effectiveness of mHealth interventions in COPD management.

Acknowledgments

The authors would like to thank Dr Gerard Farrell and Dr Hai Nguyen for their contribution. The findings presented here have been made possible by the overwhelming dedication of nurses, pharmacists, and physicians who openly provided their creative ideas and wishes. The authors would like to thank the Newfoundland and Labrador Medical Association, the Association of Registered Nurses of Newfoundland and Labrador, and the Pharmacists' Association of Newfoundland and Labrador for their efforts in recruiting HCPs. The authors would also like to acknowledge the support of Sequence bio. Research funding was provided by the Saudi Arabian Cultural Bureau and Mitacs.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Health care provider interview prompts.

[DOCX File, 15KB - mhealth_v7i6e13950_app1.docx]

References

1. Steinhubl S, Muse E, Topol E. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
2. 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 1;4(3):e105 [FREE Full text] [doi: 10.2196/mhealth.5921] [Medline: 27589898]
3. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954 [FREE Full text] [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]

4. Juen J, Cheng Q, Schatz B. A natural walking monitor for pulmonary patients using mobile phones. IEEE J Biomed Health Inform 2015 Jul;19(4):1399-1405. [doi: 10.1109/JBHI.2015.2427511] [Medline: 25935052]
5. Zhang J, Song Y, Bai C. MIOTIC study: a prospective, multicenter, randomized study to evaluate the long-term efficacy of mobile phone-based Internet of Things in the management of patients with stable COPD. Int J Chron Obstruct Pulmon Dis 2013;8:433-438 [FREE Full text] [doi: 10.2147/COPD.S50205] [Medline: 24082784]
6. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. Lancet 1997 May;349(9064):1498-1504. [doi: 10.1016/S0140-6736(96)07492-2]
7. Canadian Institute for Health Information. Health indicators URL: https://secure.cihi.ca/free_products/HealthIndicators2008_ENGweb.pdf [WebCite Cache ID 76gVKOWKJ]
8. Najafzadeh M, Marra CA, Lynd LD, Sadatsafavi M, FitzGerald JM, McManus B, et al. Future impact of various interventions on the burden of COPD in Canada: a dynamic population model. PLoS One 2012 Oct;7(10):e46746 [FREE Full text] [doi: 10.1371/journal.pone.0046746] [Medline: 23071626]
9. Theriault L, Hermus G, Goldfarb D, Stonebridge C, Bounajm F. The Conference Board of Canada. 2012. Cost risk analysis for chronic lung disease in canada URL: https://www.conferenceboard.ca/errors/500 [accessed 2019-06-04] [WebCite Cache ID 78tLvpJ8j]
10. International Organization for Standardization. 2010. Ergonomics of human-system interaction - part 210: Human-centred design for interactive systems (ISO 9241-210) URL: https://www.iso.org/standard/52075.html [accessed 2019-06-06] [WebCite Cache ID 78vLE24w2]
11. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
12. World Health Organization. 2017. A handbook on how to implement mBreatheFreely, mHealth for COPD and asthma URL: https://www.who.int/sorry/ [accessed 2019-06-06] [WebCite Cache ID 78vLTs12o]
13. Nelson L, Mayberry L, Wallston K, Kripalani S, Bergner E, Osborn C. Development and usability of REACH: a tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. JMIR Hum Factors 2016 Sep 8;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
14. Sarkar U, Gourley G, Lyles C, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
15. Gray SC, Khan A, Kuluski K, McKillop I, Sharpe S, Bierman AS, et al. Improving patient experience and primary care quality for patients with complex chronic disease using the electronic patient-reported outcomes tool: adopting qualitative methods into a user-centered design approach. JMIR Res Protoc 2016 Feb 18;5(1):e28 [FREE Full text] [doi: 10.2196/resprot.5204] [Medline: 26892952]
16. Hopkins IG, Dunn K, Bourgeois F, Rogers J, Chiang VW. From development to implementation—a smartphone and email-based discharge follow-up program for pediatric patients after hospital discharge. Healthcare 2016 Jun;4(2):109-115 [FREE Full text] [doi: 10.1016/j.hjdsi.2015.11.003]
17. Triantafyllidis A, Velardo C, Chantler T, Shah SA, Paton C, Khorshidi R, et al. A personalised mobile-based home monitoring system for heart failure: the SUPPORT-HF Study. Int J Med Inform 2015 Oct;84(10):743-753. [doi: 10.1016/j.ijmedinf.2015.05.003]
18. Upshur RE, Tracy S. Chronicity and complexity: is what's good for the diseases always good for the patients? Can Fam Physician 2008 Dec;54(12):1655-1658 [FREE Full text] [Medline: 19074692]
19. Boyd CM, Fortin M. Future of multimorbidity research: how should understanding of multimorbidity inform health system design? Public Health Rev 2010 Dec 10;32(2):451-474. [doi: 10.1007/BF03391611]
20. Miller D, Steele GC, Kuluski K, Cott C. Patient-centered care and patient-reported measures: let's look before we leap. Patient 2015;8(4):293-299 [FREE Full text] [doi: 10.1007/s40271-014-0095-7] [Medline: 25354873]
21. Bender BG. Nonadherence in chronic obstructive pulmonary disease patients: what do we know and what should we do next? Curr Opin Pulm Med 2014 Mar;20(2):132-137. [doi: 10.1097/MCP.0000000000000027] [Medline: 24452102]
22. Heffernan KJ, Chang S, Maclean ST, Callegari ET, Garland SM, Reavley NJ, et al. Guidelines and recommendations for developing interactive eHealth apps for complex messaging in health promotion. JMIR Mhealth Uhealth 2016;4(1):e14 [FREE Full text] [doi: 10.2196/mhealth.4423] [Medline: 26860623]
23. DiDonato KL, Liu Y, Lindsey CC, Hartwig DM, Stoner SC. Community pharmacy patient perceptions of a pharmacy-initiated mobile technology app to improve adherence. Int J Pharm Pract 2015 Jan 9;23(5):309-319. [doi: 10.1111/ijpp.12168]
24. Vorrink SN, Kort HS, Troosters T, Lammers JJ. A mobile phone app to stimulate daily physical activity in patients with chronic obstructive pulmonary disease: development, feasibility, and pilot studies. JMIR mHealth uHealth 2016 Jan 26;4(1):e11 [FREE Full text] [doi: 10.2196/mhealth.4741]
25. Jo S, Park H. Development and evaluation of a smartphone application for managing gestational diabetes mellitus. Healthc Inform Res 2016;22(1):11 [FREE Full text] [doi: 10.4258/hir.2016.22.1.11]

26. Smith SA, Claridy MD, Whitehead MS, Sheats JQ, Yoo W, Alema-Mensah EA, et al. Lifestyle modification experiences of african american breast cancer survivors: a needs assessment. JMIR Cancer 2015 Aug;1(2) [FREE Full text] [doi: 10.2196/cancer.4892] [Medline: 26380378]
27. Chiang K, Wang H, Chien I, Liou J, Hung C, Huang C, et al. Healthcare providers' perceptions of barriers in implementing of home telecare in Taiwan: a qualitative study. Int J Med Inform 2015 Apr;84(4):277-287 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.01.007]
28. Ministry of Health and Long-Term Care. 2018. Regulated health professions - health workforce planning branch - health care professionals URL: http://www.health.gov.on.ca/en/pro/programs/hhrsd/about/regulated_professions.aspx [accessed 2019-06-04] [WebCite Cache ID 78tMUz7wN]
29. Hepler CD, Strand LM, Tromp D, Sakolchai S. Critically examining pharmaceutical care. J Am Pharm Assoc (Wash) 2002;42(5 Suppl 1):S18-S19. [Medline: 12296538]
30. Korpershoek YJ, Vervoort SC, Trappenburg JC, Schuurmans MJ. Perceptions of patients with chronic obstructive pulmonary disease and their health care providers towards using mHealth for self-management of exacerbations: a qualitative study. BMC Health Serv Res 2018 Oct 4;18(1) [FREE Full text] [doi: 10.1186/s12913-018-3545-4]
31. Merriam S, Tisdell E. Qualitative Research. Fourth Edition. San Francisco, CA, US: Jossey-Bass; 2015.
32. Patton M. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. New York: Sage Publications; 2015.
33. Creswell J. 30 Essential Skills for the Qualitative Researcher. Los Angeles: Sage Publications; 2016.
34. Emerson RW. Convenience sampling, random sampling, and snowball sampling: how does sampling affect the validity of research? J Vis Impair Blind 2018 Dec 13;109(2):164-168 [FREE Full text] [doi: 10.1177/0145482X1510900215]
35. Maxwell J. Qualitative Research Design. Los Angeles: Sage Publications; 2013.
36. Damhus CS, Emme C, Hansen H. Barriers and enablers of COPD telerehabilitation - a frontline staff perspective. Int J Chron Obstruct Pulmon Dis 2018;13:2473-2482 [FREE Full text] [doi: 10.2147/COPD.S167501] [Medline: 30154650]
37. Macdonald GG, Townsend AF, Adam P, Li LC, Kerr S, McDonald M, et al. eHealth technologies, multimorbidity, and the office visit: qualitative interview study on the perspectives of physicians and nurses. J Med Internet Res 2018 Dec 26;20(1):e31 [FREE Full text] [doi: 10.2196/jmir.8983] [Medline: 29374004]
38. Brinkman S, Kvale S. InterViews: Learning the Craft of Qualitative Research Interviewing, Third Edition. Los Angeles: Sage Publications; 2015.
39. Rubin H, Rubin I. Qualitative Interviewing: The Art of Hearing Data, Third Edition. Los Angeles: Sage Publications; 2012.
40. Leech NL, Onwuegbuzie AJ. Qualitative data analysis: a compendium of techniques and a framework for selection for school psychology research and beyond. Sch Psychol Q 2008 Dec;23(4):587-604 [FREE Full text] [doi: 10.1037/1045-3830.23.4.587]
41. Saldaña J. The Coding Manual for Qualitative Researchers Third Edition. Los Angeles: Sage Publications; 2016.
42. Wolcott H. Writing Up Qualitative Research, Third Edition. Los Angeles: Sage Publications; 2009.
43. Yin R. Qualitative Research from Start to Finish, Second Edition. New York: Guilford Press; 2016.
44. Guest G, MacQueen K, Namey E. Applied Thematic Analysis. Los Angeles: Sage Publications; 2012.
45. Miles M, Huberman A, Saldaña J. Qualitative Data Analysis. Los Angeles: Sage Publications; 2014.
46. Brunton L, Bower P, Sanders C. The contradictions of telehealth user experience in chronic obstructive pulmonary disease (COPD): a qualitative meta-synthesis. PLoS One 2015;10(10):e0139561 [FREE Full text] [doi: 10.1371/journal.pone.0139561] [Medline: 26465333]
47. Vorrink S, Huisman C, Kort H, Troosters T, Lammers J. Perceptions of patients with chronic obstructive pulmonary disease and their physiotherapists regarding the use of an eHealth intervention. JMIR Hum Factors 2017 Sep 19;4(3):e20 [FREE Full text] [doi: 10.2196/humanfactors.7196] [Medline: 28928110]
48. HL7.org. 2018. Mobile health URL: http://www.hl7.org/Special/committees/mobile/index.cfm [WebCite Cache ID 76VIoguNi]
49. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed-methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656 [FREE Full text] [doi: 10.2196/11656] [Medline: 31025951]

Abbreviations

COPD: chronic obstructive pulmonary disease
HCD: human-centered design
HCP: health care provider
ISO: International Organization for Standardization
mHealth: mobile health
PI: primary investigator

Edited by G Eysenbach; submitted 08.03.19; peer-reviewed by C Jacob, Y Korpershoek, D Fitzsimmons; comments to author 03.04.19; revised version received 12.05.19; accepted 29.05.19; published 10.06.19.

Please cite as:
Alwashmi MF, Fitzpatrick B, Davis E, Gamble JM, Farrell J, Hawboldt J
Perceptions of Health Care Providers Regarding a Mobile Health Intervention to Manage Chronic Obstructive Pulmonary Disease: Qualitative Study
JMIR Mhealth Uhealth 2019;7(6):e13950
URL: http://mhealth.jmir.org/2019/6/e13950/
doi:10.2196/13950
PMID:31199330

Original Paper

Mobile Health Adoption in Mental Health: User Experience of a Mobile Health App for Patients With an Eating Disorder

Dimitra Anastasiadou[1,2], PhD; Frans Folkvord[2,3], PhD; Eduardo Serrano-Troncoso[4,5], PhD; Francisco Lupiañez-Villanueva[1,2], PhD

[1]Department of Information and Communication Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

[2]Open Evidence Research, Barcelona, Spain

[3]School of Humanities and Digital Sciences, Tilburg University, Tilburg, Netherlands

[4]Child and Adolescent Psychiatry and Psychology Department, Hospital Sant Joan de Déu of Barcelona, Barcelona, Spain

[5]Children and Adolescent Mental Health Research Group, Institut de Recerca Sant Joan de Déu, Barcelona, Spain

Corresponding Author:
Dimitra Anastasiadou, PhD
Department of Information and Communication Sciences
Universitat Oberta de Catalunya
Rambla del Poblenou, 156
Barcelona, 08018
Spain
Phone: 34 93 450 52 74
Email: danastasiadou@uoc.edu

Abstract

Background: Despite the worldwide growth in mobile health (mHealth) tools and the possible benefits for both patients and health care providers, the overall adoption levels of mHealth tools by health professionals remain relatively low.

Objective: This study aimed (1) to investigate attitudes of health care providers and mHealth experts toward mHealth tools in the health context in general, and this study aimed (2) to test the acceptability and feasibility of a specific mHealth tool for patients with an eating disorder (ED), called TCApp, among patients and ED specialists.

Methods: To this purpose, we conducted an explorative qualitative study with 4 in-depth group discussions with several groups of stakeholders: our first focus group was conducted with 11 experts on mHealth from the Catalan Association of Health Entities; the second focus group included 10 health care professionals from the Spanish College of Doctors of Barcelona; the third focus group involved 9 patients with an ED who had used the TCApp over a 12-week period, and the fourth and last focus group involved 8 ED specialists who had monitored such ED patients on the Web.

Results: The focus groups showed that health care providers and mHealth experts reported barriers for mHealth adoption more often than facilitators, indicating that mHealth techniques are difficult to obtain and use. Most barriers were attributed to external factors relating to the human or organizational environment (ie, lack of time because of workload, lack of direct interest on a legislative or political level) rather than being attributed to internal factors relating to individual obstacles. The results of the mHealth intervention study indicate that the TCApp was considered as easy to use and useful, although patients and the ED specialists monitoring them on the Web reported different adoption problems, such as the inability to personalize the app, a lack of motivational and interactive components, or difficulties in adhering to the study protocol.

Conclusions: In general, this paper indicates that both health professionals and patients foresee difficulties that need to be addressed before comprehensive adoption and usage of mHealth techniques can be effectively implemented. Such findings are in line with previous studies, suggesting that although they acknowledge their possible benefits and cost-effectiveness, health care providers are quite resistant and conservative about integrating mHealth technologies in their daily practice.

(JMIR Mhealth Uhealth 2019;7(6):e12920) doi:10.2196/12920

KEYWORDS

eating disorders; mental health; mHealth; mobile applications; focus groups

Introduction

Background

Over the last decade, the number of people who have obtained a mobile phone or some other portable electronic communication devices worldwide has increased exponentially [1]. As a consequence, innovative apps for such devices have been developed to address health issues, and this has evolved into a new field labeled as mobile health (mHealth) [2,3]. Central to the concept of mHealth is that by removing geographical, temporal, and other barriers that influence usage of interventions that are related to health behavior, information and resource services are more flexible and have the potential to reach anyone, anytime, and anywhere [4].

mHealth has many different applications, such as the facilitation of data collection, the encouragement of health care consumers to adopt healthy lifestyles, the self-management of chronic conditions, such as pain, psychological distress, fatigue, and sleep [5,6], and the improvement of the delivery of health care services by targeting health care providers or by easing communication between these providers and their patients [7,8]. Despite the reported benefits for the patient and the health care provider, as well as mHealth's positive impact on the economic and organizational delivery systems [9], the overall adoption of mHealth by health professionals remains relatively low [10].

Many factors influencing the acceptance, adoption, and usage of mHealth by health care professionals have been reported in the systematic review by Gagnon and colleagues [11]. Specifically, factors relating to mHealth characteristics, such as perceived usefulness and ease of use of the tool, have been identified as the main positive drivers for mHealth adoption by health professionals. In addition, individual-level factors are also considered decisive when assessing adoption levels, such as individuals' time availability (eg, overcoming disruption of workflow), a positive risk-benefit assessment, and the level of confidence in and agreement with mHealth implementation.

Meanwhile, important barriers relating to mHealth characteristics were also mentioned in the same review [11], such as security, privacy, and cost issues, as well as design and technical concerns, in addition to factors relating to the human environment (eg, patient-health professional interaction) and other factors relating to the organizational environment, for example, workload, time constraints, dysfunctional relations among colleagues, and lack of human resources to support Information and Communication Technologies (ICT). Next to that, lack of integration and interoperability and the need for high-risk investment in unsure markets still present major barriers for adoption in the international literature [12]. Last but not least, service availability (ie, lack of connectivity) has also been identified as both a driver and a barrier for adoption by health care professionals [12,13].

Owing to the advantages of the use of mHealth technologies, it is important to determine what types of drivers and barriers are considered relevant by health professionals to adopt mHealth technologies efficiently. In this study, we will focus on the drivers and barriers for adoption of mHealth techniques in general health care delivery processes, and, more specifically, in an intervention using an mHealth tool specifically designed for patients with eating disorders (EDs) and the health care providers treating them.

Mobile Health Adoption in Mental Health: An Experience With Patients With Eating Disorders

According to Holmes and colleagues [14], mHealth can transform the availability and efficacy of psychological treatments for mental health problems, given that individual therapy in the past few decades has not been able to meet the increasing demand for psychological services [15]. More specifically, the use of app-based treatments, both self-delivered and therapist-led, has the potential to improve traditional psychotherapy by facilitating monitoring of symptoms and outcomes, improving access to psychoeducation materials and skills training, and offering patients the opportunity to communicate with their therapist [16]. Other functionalities of apps include personalization, setting reminders, and the real-time collection and visual presentation of data, which could offer tremendous added value to traditional, *static* interventions. Correspondingly, *blended treatment*, referring to face-to-face treatments, which include a digital intervention or component, has been gaining in popularity over the last few years [17-19], and it has shown promising results in various mental health treatments [20,21].

Health professionals treating patients with EDs might be an important group that could benefit from mHealth interventions, either as an adjunct to standard treatment or as a method to providing existing evidence-based treatments on the Web [22]. According to National Institute for Health and Care Excellence guidelines [23], the gold-standard psychological therapy for adult patients with EDs is cognitive behavioral therapy (CBT) [24], which considers self-monitoring of dietary intake and associated thoughts and feelings as one of its main behavioral components [25]. Self-monitoring is a task that can easily be implemented through real-time self-records via mHealth techniques [26]. Literature has shown that both users and clinicians find mobile phone apps that support and facilitate ED symptom monitoring to be highly practicable as a component of ED treatment [27,28]. However, somewhat contradictory findings were reported by a recent study in which the advantages and disadvantages of the implementation of an mHealth tool in ED treatment were discussed [29,30]. Despite the enthusiasm surrounding mHealth technologies [31] and the fact that they may address limitations of face-to-face treatments (low motivation to practice, generalize and maintain therapeutic skills, and limited accessibility) [32], the necessity to develop guidelines for correct patient and clinician app usage is apparent [29,30].

Furthermore, factors that facilitate or hinder the effective adoption of these tools by patients and clinicians need to be identified through qualitative research. In a systematic review of the literature by Anastasiadou and colleagues [22] that focused on mHealth interventions for EDs, qualitative analyses showed that although most mHealth interventions were considered as acceptable, easy to use, and motivating by patients and therapists, other individual studies highlighted various

problems surrounding their use, such as high dropout rates, technical issues, and privacy and personalization concerns [27,28]. Work by Juarascio and colleagues [32] confirmed that the limited number of acceptability and feasibility studies in this emergent field prevents researchers from drawing firm conclusions to date. Finally, the complexity of EDs, involving low treatment adherence and the need for an intensive level of care, may not make mHealth interventions the most suitable treatment in all cases. To address such matters, the 2 objectives of this study were (1) to investigate attitudes of health care providers and mHealth experts regarding mHealth tools in the health care context in general and (2) test the acceptability and feasibility of a specific mHealth tool designed specifically as an adjunct to standard treatment for patients with an ED and their therapists. To explore these 2 objectives, we conducted an explorative qualitative study with in-depth group discussions with various groups of stakeholders.

Methods

Design

We conducted 4 separate in-depth focus groups for this study. The first was carried out with health care providers, and the second was carried out with mHealth experts. Both focus groups had the same structure, and both aimed at assessing important drivers and barriers, which these professionals perceived relating to mHealth adoption in different health care services. The third and fourth focus groups were based on data from a larger randomized controlled trial (RCT) assessing the effectiveness of an mHealth intervention (ClinicalTrials.gov Identifier: NCT03197519), and they sought to assess the user (patient and provider) experience of a specific mHealth tool, called TCApp, in supporting self-management for patients with an ED. Focus group discussions were performed in accordance with the generic inductive approach [33].

The TCApp

The TCApp is an mHealth tool specifically designed to bidirectionally connect patients with EDs and their therapists in the periods between medical consultations. The app has been designed and developed by the Spanish start-up company HealthApp. It offers 4 different language options (Spanish, Catalan, French, and English), and it is available through both Google Play and Apple Store markets, and it currently has more than 1000 patients as active users. Through the TCApp, patients can record their thoughts, actions, emotions, and whatever other information their therapists consider relevant for the therapy. The app can be customized for each patient according to therapy requirements. The main objective of the app is to replace paper self-records with Web-based records, as the former sometimes present a source of discomfort for patients [34]. In addition, the app also introduces gamification esthetics [35], including prizes, rewards, and reminders to improve patients' engagement. Finally, it includes a chat function to facilitate patient-therapist communication between their regular face-to-face sessions. To date, the app has been used by over 1000 patients, counting with the special guidance of over 80 registered clinicians who are currently using the app's clinician interface across 20 different centers in Spain.

Study Sample

Health Care Providers and Mobile Health Experts

A total of 11 experts on mHealth from the Catalan Association of Health Entities (Associació Catalana d'Entitats de Salut, ACES) and 10 health care professionals from the Spanish College of Doctors of Barcelona (Colegio Oficial de Médicos de Barcelona, COMB) with a special interest in mHealth were invited to participate in focus groups organized by our research group at the Open University of Catalonia (Universitat Oberta de Catalunya, UOC). The generic inductive approach, on which the design of the study was based, required a purposive sampling. To better identify barriers and facilitators for mHealth adoption from the point of view of 2 different stakeholders' groups, we considered it important to carry out the first focus group with experts on mHealth from ACES and the second with health care professionals from COMB. The approval of the Ethical Committee of the University leading the study (UOC) to conduct the focus groups was obtained on February 21, 2017. Inclusion criteria for both focus groups were professionals from several private medical centers in Barcelona (Spain) who were actively engaging in mHealth technique usage, including both health care providers as well as mHealth experts, such as technical staff (eg, data analysts, software developers), ICT management staff, marketing and communication managers, and human resources management staff.

Participants of the Mobile Health Intervention using the TCApp

The study's population comprised a larger RCT for the mHealth intervention study, involving 8 public and private ED units in Spain. A total of 108 patients were recruited from this larger study, the protocol of which has been published elsewhere [36]. It is worth mentioning that the Institutional Review Board approval was obtained for each participating institution and for the University leading the study (UOC). Participants were selected from this larger RCT, and they were invited to take part in our focus groups. Inclusion criteria for the patients included (1) having been part of the experimental group and used the TCApp for at least 8 out of the 12 weeks in which the app was available to participants and (2) not presenting any significant symptom worsening or showing relapse at the time in which the invitation was sent.

Inclusion criteria for the ED specialists' group included having monitored their patients on the Web and having performed the following actions at least once a week: follow the patient's daily self-records and generate personalized reports or graphs and communicate with him or her via the chat function. From a total of 21 patients who were invited to the focus group, 11 decided to participate, whereas 2 had to be excluded, as they did not meet all inclusion criteria. Next to that, all invited ED specialists from the 8 ED units took part in their respective group discussion. However, it is worth mentioning that only patients from 3 out of the 8 ED units involved in the larger study participated in the focus group study. The final focus group sample thus comprised 9 patients with an ED and 8 specialists who were monitoring such patients on the Web through the TCApp. Tables 1 and 2 show the clinical and demographic characteristics of the patients and ED specialists at baseline.

Table 1. Description of patients.

Patient characteristics (*N*=9)	Values
Age (years), mean (SD)	15 (0.50)
Body Mass Index (actual), mean (SD)	19.55 (1.35)
Illness Duration (months), mean (SD)	16.67 (8.75)
Sex, n (%)	
Male	0 (0)
Female	9 (100)
Highest level of education, n (%)	
Primary	7 (78)
Secondary	2 (22)
Tertiary	0 (0)
Diagnosis, n (%)	
AN[a]-restrictive	8 (89)
AN-purging	1 (11)
BN[b]	0 (0)
EDNOS	0 (0)
Current treatment type, n (%)	
Day Hospital	1 (11)
Outpatient	8 (89)

[a]AN: anorexia nervosa.

[b]BN: bulimia nervosa.

[c]EDNOS: eating disorder not otherwise specified.

Table 2. Description of eating disorder specialists.

ED specialist characteristics (N=8)	Values
Age (years), mean (SD)	34.63 (7.21)
Sex, n (%)	
Male	2 (25)
Female	6 (75)
Specialty, n (%)	
Psychiatry	2 (25)
Psychology	5 (63)
Nursing	1 (12)
Employment status, n (%)	
Public	4 (50)
Private	4 (50)
Duration actual employment status (years), mean (SD)	6.19 (4.29)

Focus Groups

Qualitative data were collected through four in-depth focus groups, as previously indicated. The first comprised health care providers from the COMB, the second focus group was conducted with mHealth experts from the ACES, a third focus group with patients who had used the mHealth tool, TCApp, and a fourth with ED specialists who had monitored such patients on the Web.

The interviews carried out in the first 2 focus groups included open-ended questions designed to address the following topics relating to drivers and barriers of mHealth adoption:

- Factors related to mHealth characteristics

- Individual and professional factors
- Human environment factors
- Organizational environment factors: internal and external

Open-ended questions for the focus group were designed following previous work of our team [37]. This broader study had the objective to measure ICT adoption among 25 General Practitioners, representing 20 different European countries. The framework used in this previous study integrated several behavioral models and hypotheses extracted from the international literature (ie, Technology Adoption Model and its revised version Technology Adoption Model 2, Universal Theory of Acceptance and Use of Technology, Theory of Reasoned Action, and Theory of Planned Behavior). However, when designing the questions for the first and second focus group, little was yet known about general mHealth adoption among health care professionals. Thus, we adapted questions developed in the previously mentioned study to suit mHealth characteristics, and we were also inspired by the adoption factors identified in the systematic review by Gagnon and colleagues [11].

For the third and the fourth focus group, the interviews with patients and ED specialists were completed 3 months after the study start date. When designing the questions, we were inspired by a multidimensional evaluation framework, called Model for ASessment of mHealth (MASH), previously described and implemented by a European project in which our team also participated, termed DoChange. The MASH model was constructed on the basis of the evaluation framework of the Model for ASsessment of Telemedicine methodology [38]. We followed 5 of the 8 domains of the MASH model, which were translated into the following research questions that guided our focus groups:

- Domain 3 (both patients and ED specialists). Effectiveness.

  To what extent did patients' ED symptoms and quality of life improve when treated with TCApp as a complement to face-to-face treatment?

- Domain 4 (both patients and ED specialists). User experiences.

  Which were professionals'/patients' needs in implementing the TCApp services in daily practice? Were users satisfied with the app/usability of the services offered by the TCApp? Which were the most important facilitators in implementing the TCApp in routine practice? Which were the most important barriers in implementing the TCApp in routine practice?

- Domain 5 (ED specialists). Economic aspects.

  Are you willing to pay for use of the TCApp from now on in your routine practice?

- Domain 6 (both patients and ED specialists). Organizational aspects.

  Participants had to think of a future ideal scenario in which decisions at their institution were made by them and describe the specific proposals they would make to their organization about a new service that implements, in an ideal way, the TCApp together with other new technologies.

- Domain 7 (both patients and ED specialists). Sociocultural aspects.

  Have any issues arisen regarding cultural accessibility (considering the different languages spoken by the users) and socioeconomic accessibility (different ages, users' digital health literacy)?

Study Procedure

Health Care Providers and Mobile Health Experts

First, for the focus groups with health care providers and mHealth experts, a formal invitation was sent to both institutions (COMB and ACES, respectively) by the principal investigator of the study, FL. Both focus groups, of 1 hour each, were conducted by FL at the headquarters of ACES and at the UOC, and they were audio recorded with permission. A second researcher, CF, was present during these events and transcribed the group discussions verbatim.

Mobile Health Intervention TCApp

The study procedure of the mHealth intervention was described in detail elsewhere [36]; therefore, we will only briefly describe the important aspects of the study protocol here. After an initial semistructured interview and a baseline evaluation using self-report questionnaires at T0, participants were randomly assigned to 1 of the 2 study conditions (experimental or control group). Therapists in the participating ED units were invited to introduce to the patients of the experimental group the option of using the TCApp as a complementary tool to their standard face-to-face CBT. Each therapist was given instructions on how to use the Web-based environment intended for therapists, on how to add patients as users, and each therapist was given instructions on the basic functionalities of the app specifically designed for patients. Therapists were asked to explain briefly to each patient in the experimental group how to use the TCApp, in addition to supplying them with a brochure with written instructions, as well as an encouragement to choose a nonidentifying username. Over a 12-week period, patients were expected to use the TCApp at least once a week, completing self-records or contacting their therapist via chat on a regular basis when this was considered necessary. The therapist responsible for the Web-based monitoring was asked to, at least once a week, connect to the Web-based platform and perform the following actions: follow the patient's daily self-records, generate personalized reports or graphs, and communicate with the patient via the chat function. Little, if any, additional instruction was given on how to use the app clinically, to allow for natural adoption by patients as well as clinicians. In turn, patients in the control group were told that access to the TCApp would be offered to them after a waiting period of 6 months. Thus, for a period of 12 weeks, each group of patients received the treatment that corresponded to the patients' experimental condition, in addition to the patients' regular treatment. At the end of the 12-week period, patients from the experimental group were told to stop using the TCApp (they were discharged from the service), and a posttreatment evaluation was carried out with both groups of patients (T1). Furthermore, 1 week later, an

email invitation was sent to patients from the experimental group and ED specialists who had monitored their patients on the Web, asking them to participate in the present focus groups. Both focus groups with patients and ED specialists were conducted in the respective participating hospitals by DA, and both they were audio-recorded with permission. The duration of these focus groups was an hour on average per group. During the focus group with patients and directly after the group discussion, self-report questionnaires were administered, whose intended purpose was to evaluate usability of and participants' satisfaction with the tool. Similarly, a self-report questionnaire designed by our research team was administered to ED specialists after the group interview. A researcher of our group, CF, transcribed the group discussions after listening to the audio recordings. Subsequently, DA read and validated these transcriptions.

Additional Measures Used in the Mobile Health Intervention TCApp

For Patients

Sociodemographic and Clinical Characteristics

Patient information, including age, sex, diagnosis, illness duration, body mass index (BMI), degree of education, and employment status, was retrieved from the medical history of each patient, which was provided to the research team by the ED specialists responsible for each participating center.

Users' Experience Using the TCApp

System Usability Scale (SUS) [39] is a 10-item questionnaire that measures usability of a range of systems and has been described elsewhere [36].

Client Satisfaction Questionnaire (CSQ-8) [40] is an 8-item questionnaire that is designed to measure client satisfaction with services and has been described elsewhere [36].

For Eating Disorder Specialists

Sociodemographic characteristics, including age, sex, employment status, and specialty, were collected through an introductory self-report questionnaire.

Data Analysis

Thematic content analysis was used to analyze the qualitative data from the 4 focus group discussions. [41]. First, members of the research team (DA and FF) independently read and reread the transcripts of the 4 group discussions to familiarize themselves with the data. They then sought to achieve consensus regarding a common coding scheme for the information collected during the group discussions. Common categories of meaning in the data relating to the study objectives were identified inductively [33]. In case of doubt during coding, researchers stepped back and reclassified the coded data to ensure that any contradictory information was not omitted. After having categorized drivers and barriers of mHealth adoption, we compared our findings with the classification proposed by Gagnon and colleagues [11], as well as with previous research conducted by our team [37]. The data were coded manually, and no software or tool was used in this procedural step. Finally, the most significant and representative examples for each theme were collected from the transcripts after discussion among researchers on the team. To better reflect patient and provider experiences, all information was presented in tables. Each key theme was classified depending on how frequently it had been mentioned by participants, using +++ for the most frequently mentioned themes, ++ for frequently mentioned themes, and + for themes that were only occasionally mentioned (and sometimes generated a debate because of contradictory opinions on the topic). In addition, advantages and disadvantages of the TCApp as perceived by patients and ED specialists were also classified in common themes and domains, together with the frequency with which each theme was reported by patients and specialists (again ranking from + to +++).

The analyses of all descriptive data regarding participants' sociodemographic and clinical characteristics, SUS, and CSQ-8 were carried out using SPSS 20.0 (SPSS Inc) [42]. Credibility and validity of patients' responses were ensured through cross-verification (ie, triangulation) of the outcomes of the focus groups relating to usability and satisfaction and the SUS and CSQ-8 questionnaires.

Results

Focus Groups With Health Care Providers and Mobile Health Experts

In total, 4 domains and 16 key themes were identified and were then classified as barriers to or drivers for mHealth adoption. Most of these themes, 11 (69%), were classified by professionals as barriers, and only 5 (31%) were perceived as facilitators for mHealth adoption. The complete list of domains and themes, together with professionals' most representative examples recounted during the focus groups, is shown in Tables 3 and 4. In the same table, we also present an analysis of the factors on the basis of the frequency with which these were reported by professionals.

Table 3. Barriers relating to mobile health adoption by health care providers and mobile health experts.

Domain and key themes	Examples
External factors: organizational environment	
Lack of time and workload (+++[a])	"I had no time to integrate any changes in my work schedule"; "I know the theory...that you have to do a first push at the beginning in order to integrate changes, for example change to e-consultations, and then the workload will become less. But then in practice, it is difficult to implement due to our workload."
Management: Lack of strategic plan to implement mobile health (+++)	"Leaders have to believe in innovation in order to promote the use of ICT[b] tools, otherwise projects testing their efficacy will not succeed."
Health care policies and sociopolitical context: Lack of budget and direct interest (+++)	"Technology costs...There is a considerable lack of budget to support studies assessing the impact of mHealth tools"; "At a political and legislative level, there is not enough interest in mHealth. Doctors complain that they cannot reimburse mHealth"; "Technology is costly and it is difficult to verify its return. Often, the ideas we propose seem good, but when the budget is specified, people become more resistant, you see that the investment has a certain cost. In addition, technology evolves fast and you may not have time to recover the initial investment...Thus, inversions for mHealth should not only be innovative, but also timely and always updated."
Insufficient training (++[c])	"It is important to have a properly trained team...It is not enough to have an IT-department specialised and dedicated to this. All the staff has to be properly trained"; "It is essential to offer continuous training to ensure that students of health sciences acquire digital competences."
Human resources: Lack of information technology support (++)	"There is a considerable lack of support by the technical staff of our institution when we integrate a new mHealth tool into our services...Due to this lack of support, users either use the new tools in the wrong way, or stop using them because they get frustrated when they do not know how to use them in a proper way."
Individual factors	
Age: Lack of familiarity and mobile health skills (+[d])	"Age-based digital divide is present in the health sector...Young health professionals have more digital minds...Instead, for many professionals who are older in age, the handling of the Internet and other ICTs may seem complex and they prefer to do things in the traditional manner (paper-and-pencil methods)"; "I do not think that age has a significant impact on adoption...In my institution, older professionals do not have more negative attitudes towards ICTs than younger ones, neither do they perceive their utility and usefulness differently."
Lack of agreement with mobile health: Resistance (++)	"Professionals are often resistant to change because of fear of the unknown and new."
Risk-benefit assessment (perception; +)	"The value of face-to-face contact with our patients is inherent to all of us...sometimes we are afraid of losing this when introducing a new technology."
External factors: Human environment	
Insufficient interaction: Patient-Health professional-information technology team (++)	"New technologies cannot be implanted unilaterally by the IT team. Health professionals are those who really know what patients need. There has to be an alignment of needs between IT team, patients and health professionals."
Mobile health characteristics	
Security and privacy issues (++)	"Sometimes bureaucracy is used as an excuse to stop the implementation of an innovation...Once patients sign their informed concern, there is freedom to conduct innovations"; "We have a very protectionist system...The new law on security / privacy of medical histories is very restrictive about which data from patients can be viewed and shared...this also obstructs the sharing of this information online using mHealth tools"; "It is important to approve Apps by an official authority taking into account both technological validation (must be useful) and functional validation (applicable to the context)."
Design and technical concerns (+)	"Technology can fail"..."Technical limitations of the mHealth tools"

[a]+++: Very frequently mentioned (≥4 times).

[b]ICT: Information and Communication Technology/ies.

[c]++: Frequently mentioned (2-3 times)

[d]+: Only occasionally mentioned, sometimes a debate was opened because of contradictory opinions (mentioned 1 time, or 2 contradictory opinions).

Table 4. Drivers relating to mobile health adoption by health care providers and mobile health experts.

Domain and key themes	Examples
Mobile health characteristics	
Quality standard (+++[a])	"It is very important to achieve a proper quality control and approval of a mHealth tool by a certified institution"
Individual factors	
Economic incentives for professionals (+[b])	"If I personally had to implement a mHealth tool as complementary to my face-to-face visits and charge each visit by 2 euros more, I would think about it"; "I would rather offer training incentives to professionals as economic incentives in the public sector are not always realistic."
External Factors: Human environment	
Support and promotion of mobile health by colleagues (++[c])	"All stakeholders should participate in the design of the mHealth tools (usability, acceptability, feasibility)."
External factors: organizational environment	
Training (+++)	"Offer continuous training...Create awareness and empowerment of patient and professional before mHealth implementation."
Management: Strategic plan to implement mobile health (+++)	"It's a top-down approach...Leaders have to believe in innovation and push for its implementation."

[a]+++: Very frequently mentioned (≥4 times).

[b]+: Only occasionally mentioned, sometimes a debate was opened because of contradictory opinions (mentioned 1 time, or 2 contradictory opinions).

[c]++: Frequently mentioned (2-3 times).

Regarding barriers, the most recurrent themes pertained to the *Organizational Environment* domain. First, the lack of time and workload were frequently mentioned by professionals as barriers (eg, *"I had no time to integrate any changes in my work schedule")*. Second, the lack of a strategic plan by leaders in health institutions to implement mHealth was also reported (eg, *"Leaders have to believe in innovation in order to promote the use of ICT tools, otherwise projects testing their efficacy will not succeed")*. Third, the lack of budget and direct interest at a legislative or political level supporting mHealth implementation was another common barrier (eg, *"Technology costs...There is a considerable lack of budget to support studies assessing the impact of mHealth tools"*). Finally, the lack of training of professionals (eg, *"It is essential to offer continuous training to ensure that students of health sciences acquire digital competences"*) and the lack of support by the information technology team in health institutions (eg, *"...Due to this lack of support, users either use the new tools in the wrong way, or stop using them because they get frustrated when they do not know how to use them in a proper way"*) were other main themes identified as barriers that belonged to the *Organizational Environment* domain, but they were less frequently mentioned by participants. Age and risk-benefit balance were identified as barriers in the *Individual* domain, and these generated diverse and sometimes contradictory opinions by professionals. For example, some of these professionals believed that a digital divide was a common problem found in health care institutions:

> *Young health professionals have more digital minds...Instead, for many professionals who are older in age, the handling of the Internet and other ICTs may seem complex and they prefer to do things in the traditional manner -paper-and-pencil methods*

However, others thought the opposite:

> *In my institution, older professionals do not have more negative attitudes towards ICTs than younger ones, neither do they perceive differently their utility and usefulness*

In relation to the risk-benefit balance, some professionals believed that the fear of losing face-to-face contact with patients was a reality difficult to overcome (eg, *"The value of face-to-face contact with our patients is inherent to all of us...sometimes we are afraid of losing this when introducing a new technology"*), although others did not agree. The individual resistance to implement mHealth services as a health professional was also seen as a barrier to the adoption of mHealth. In fact, many participants declared that they were resistant to change because of their fear of the unknown. The insufficient interaction among patient, health professional, and information technology team was another barrier to mHealth adoption for the domain of *External factors relating to the human environment*. For example, a health care provider reported the following:

> *New technologies cannot be implanted unilaterally by the IT team. Health professionals are those who really know what patients need. There has to be an alignment of needs between IT team, patients, and health professionals*

Finally, security and privacy concerns, as well as design and technical issues, all of which related to *mHealth characteristics*, were less frequently mentioned and sometimes generated contradictory opinions among respondents. For example, professionals were worried about the security and confidentiality of the data transferred through these technologies, and they declared that they had to cope with a protectionist (Spanish) health care system that is very restrictive with sharing patients' information on the Web using mHealth tools. Regarding drivers for mHealth adoption, the importance of professionals' training

first and the subsequent implementation of an mHealth-related strategic plan by leaders were the most recurrent themes pertaining to the *Organizational Environment* domain. For example, most professionals highlighted the importance of a continuous training offered to them by their institution so that they could become more knowledgeable and empowered before using an mHealth tool in their regular visits. Next, the importance of having leaders who believe in mHealth innovations was very frequently mentioned as a driver for adoption (eg, *"It's a top-down approach...Leaders have to believe in innovation and push for its implementation"*).

Moreover, the option to offer economic incentives to professionals (eg, *"If I personally had to implement a mHealth tool as complementary to my face-to-face visits and charge each visit 2 euros more, I would think about it"*) was an individual driver that generated contradictory opinions, as it could be counterproductive at times (eg, *"I would rather offer training incentives to professionals, as economic incentives in the public sector are not always realistic"*). Finally, the importance of submitting an app to quality control to get proper homologation before implementation in clinical practice was another theme relating to mHealth characteristics, which was frequently mentioned as a driver.

Mobile Health Intervention Study

First, as regards patients' scores on the 2 self-report questionnaires administered after the group interview, the mean of the total CSQ-8 scale for the 9 patients was 16 (SD 4.69; Range=11-22), whereas the mean score for the SUS scale was 81.56 (SD 22.52; Range=27.50-100). This indicated moderate usability of and high satisfaction of patients with the tool. Tables 5 and 6 summarize qualitative information from the posttreatment focus group for participants in the experimental group.

Table 5. Perceived advantages of the TCApp by patients.

Domain and Key themes	Patients who agree with example statements (N=9), n	Examples of patients´ statements
Mobile health characteristics		
Perceived ease of use	9	"At the beginning I needed some instructions and guidance but then it was very easy to use."
Perceived usefulness	8	"Paper food records are a source of discomfort because they can be lost or forgotten at home while online records are comfortable and useful"; "The option of taking photos of your meals and send them to your therapist was very useful"; "I would recommend the app to a friend with a similar problem."
Design: App	9	"I found the online platform very attractive"; "I liked the colours and the personalized avatar."
Satisfaction with content available: Motivational components	2	"I liked the option that we had to receive rewards a lot...I was looking forward to receiving prizes depending on my weekly performance and to comparing my ranking with others."
Content appropriate for users (relevance)	6	"The app facilitated a better understanding of the contextual variables that surrounded my eating behavior...I am now more aware of what happens before, during and after the problematic behavior I would like to change...more than I used to be with paper records"; "It is a good company during your treatment process, especially when you feel lonely or with the urge to carry out a problematic behavior...The option to share your thoughts, emotions or actions and be sure that your therapist is going to read them, relieves stress and guilt."

Table 6. Perceived disadvantages of the TCApp by patients.

Domain and Key themes	Patients who agree with example statements (N=9), n	Examples of patients´ statements
Mobile health characteristics		
Privacy and anonymity concerns	1	"Sometimes I did not know who was going to read my messages, something that stopped me from using the chat option."
Negative perception of usefulness	2	"It is not always useful to keep track of your problematic behavior and quantify it...I would have liked it more if the app had offered a free text option next to each question asking for the presence and frequency of several symptoms (vomiting, restriction, intensive exercise, laxative use, etc.) so that we could have the opportunity to write further explanations or observations of our behaviors."
Problems with the design: App	9	"There was a word limit when I was using the chat, I could not finish my messages so I was sending them incomplete to my online therapist or split into 2 or 3 different messages"; "There was a 24-hour limit in the app; this means that you could not register your activity after 24.00 PM...that appeared as activity of the following day"; "I would prefer if the app collected information retrospectively, meaning that the day after you register what you did the day before. By doing so, you can gain a better insight of your behavior"; "It would have been nicer to change the design of the chat (each message as a new e-mail in your inbox) and make it similar to the chat of Facebook or Whatsapp, in order to see the whole history of conversations with your therapist or be able to see when he/she reads your messages, etc."
Content inappropriate for users	3	"Patients receiving intensive treatment, i.e. day hospital, have enough of support and don't want more monitoring of their symptoms. The content of the app is not appropriate for their needs. Then, as regards outpatients, when they are asked to complete self-records daily, they feel as if they retrocede in their treatment process. The app is not useful for them either"; "I think the app can be more appropriate for patients at their early treatment stages, when they are expected to gain awareness of their ED-related behaviors."
Lack of satisfaction with content available: Lack of personalization	9	"It would have been better to personalize the app according to ED diagnosis, type of treatment that each patient is receiving and also treatment stage."
External factors: Human environment		
Patient and therapist limited and not-personalized interaction	3	"The professional who was following me online was not the same with the one with whom a had face-to-face sessions...Sometimes I perceived a lack of understanding of my problems by my online therapist"; "Sometimes it took him/her a lot to answer to my online messages and when I finally received an answer, this was no longer useful to me"; "I would like to receive more immediate and personalized answers to my messages."
External factors: Organizational environment		
Study design: Strict instructions	1	"There was a lot of pressure by researchers to use the app once a day, even if it was not always necessary."
Individual factors		
Negative benefit-risk balance	1	"It was easier for me to lie through the online records compared to the paper self-records...The app is more private, nobody (referring to her parents) has access to your records."

As regards patients' perceptions about *usability* of the TCApp, all participants (9/9) perceived that the app was easy to use and very practical. Participants were comfortable with the app, and none of them neither reported discomfort with the technology nor were they faced with technical problems. At the start of the intervention, some assistance was required by all participants to ensure that the app was working properly. Regarding *usefulness*, most participants (8/9) indicated that the app was very useful for the time between regular visits to health care professionals, and they asserted that they would recommend the tool to a friend faced with a similar problem. Most participants (8/9) agreed that keeping food records, which they used to track manually in their standard CBT, generally posed a source of discomfort, as they could be lost or forgotten at home, and most participants agreed that a change to Web-based food records was more comfortable and easy to fill in on a daily basis. A total of 2 participants voiced negative comments concerning the usefulness of the TCApp, which were related to the fact that quantification of ED symptoms through presence or absence option was not always useful for them. Instead, more *free text* options could have given patients the opportunity to describe the context around symptoms and their function more comprehensively. Several problems regarding the *design* of the app were reported by patients (9/9), including problems with

word limit, a 24-hour time limit (ie, entries for a day could only be entered on that particular day's 24-hour window and not the next day), or concerns regarding the design of the Web-based chat with the therapist, which used a rather awkward and unhelpful format (see patients' examples in Tables 5 and 6). In addition, 1 patient reported that the single option of collecting data in real time was not always convenient and helpful, given that patients were already struggling with negative feelings concerning their problematic behavior. Having to conduct another task at the same time made the experience even more difficult. However, positive comments were also raised by patients regarding the design of the app, with all participants endorsing it for its visual attractiveness and possibility to personalize settings. With regard to patients' *satisfaction with the TCApp's content*, the views differed. On the one hand, only 2 patients appreciated the rewards and prizes received and felt that this technique increased their motivation to keep engaged with the app. On the other hand, all patients viewed the lack of personalization of the tool according to their specific needs rather negatively (ie, type of diagnosis, type of treatment received, and individual readiness to change). Regarding the *appropriateness of the TCApp's content*, opinions among participants varied as well. Positive comments expressed by 6 patients included the ability of the app to make patients gain a better understanding of their problematic behaviors. In addition, the app was described as a good companion that helped patients to better manage their negative feelings. A patient specifically stated the following:

> *It is a good company during your treatment process, especially when you feel lonely or with the urge to carry out a problematic behavior. The option to share your thoughts, emotions or actions, and be sure that your therapist is going to read them, relieves stress and guilt*

On the other hand, negative comments (3 out of 9 patients) included the inability of the app to satisfy treatment needs of day-hospital patients and outpatients (see patients' examples in Tables 5 and 6). A total of 2 important issues were raised by 3 patients of 2 public institutions who seemed to encounter problems with fully complying with the study protocol. The first problem was related to *privacy and anonymity concerns:* Although patients were expecting to be monitored on the Web by the same therapist who was following them in their regular face-to-face visits, this was not always possible in the participating hospitals because of the work burden with which therapists were faced. As a consequence, messages were often read and answered by another therapist, who had no knowledge of the patient's clinical history. This was reason for some concern with certain patients regarding the privacy of their personal information and their expectation and desire to feel more secure when using the app. The second issue concerned the *limited and not-personalized patient-therapist interaction.* All 3 patients indicated that therapists at times did not answer their Web-based messages immediately and that when they did, answers were rather impersonal. As a result, these patients perceived a lack of guidance in their treatment when using the chat option. Finally, another negative comment expressed by 1 patient was related to organizational factors; specifically, it was related to the strict instructions given by the research team to follow the *study design*.

Tables 7 and 8 summarize findings from the posttreatment focus group with eating disorder specialists who monitored patients on the Web.

Table 7. Perceived advantages of the TCApp by eating disorder specialists.

Domain and Key themes	ED specialists who agree with example statements (N=8), n	Examples of ED specialists' statements
Mobile health characteristics		
Perceived ease of use	8	"The platform (for professionals) is easy to use, very practical, quick and intuitive"; "Patients found it (TCApp) easy to use, simple, and they learned fast how to use its different functionalities. They have never asked for more explanations than those given at the beginning."
Perceived usefulness	8	"I value the immediacy of the instrument and the ability to advance visits when things were not going well a lot...I believe that the app can be a good tool for therapists. They can have information prior to their face-to-face visits."; "The app facilitates our clinical practice a lot...The whole team feels more reassured with regard to each patient´s treatment"; "I value the possibility to give patients quick and valuable information (mostly resolve their doubts) during the time in between their regular sessions or advance visits, when there is something worrying in their records or messages...By doing so, we can also reduce the amount of visits of our patients to the emergency department"; "The online chat provides very valuable information as it facilitates contact with patients in-between sessions. However, it does require more time to explore and exploit the potential of these messages and how to use the information provided by each patient in face-to-face sessions."
Design: Platform	8	"The platform is attractive and does not overload visually"; "The possibility to receive photographs of the patients' meals is great!! I like how the food records described with words appear in the same screen next to the meal photographs."

Table 8. Perceived disadvantages of the TCApp by eating disorder specialists.

Domain and Key themes	ED specialists who agree with example statements (N=8), n	Examples of ED specialists´ statements
Mobile health characteristics		
Problems with the design: Platform	8	"A different new screen was opened for each of my patients...It was difficult for me to follow several patients at the same time."
Technical problems	1	"In some cases, the application stopped working; this meant a lot of extra work without getting good results. Surely it depended on a user's smartphone model (Chinese), or it occurred due to the fact that they had downloaded another application that was interfering with the use of the TCApp."
Lack of satisfaction with content available I: Inappropriate motivational components	2	"I would modify all content (she refers to the option offered to patients by the app to receive prizes and rewards according to their performance). Many patients with AN profile are stressed when receiving prizes; they want to be first in the list, corresponding to their high levels of perfectionism and competitiveness. It can be counterproductive for these patients"; "My patients decided not to share the awards received through social networks for confidentiality issues. They wanted to keep everything inside the application and not get out of this environment."
Lack of satisfaction with content available II: Lack of personalization	8	"The app should be personalised according to patients' clinical characteristics...for example, severe patients who presented a lack of consistency in their treatment in general, showed the same with the app"; "You cannot use the same instructions concerning the frequency of use of the app (once a day) for all patients. You should assess before whether this frequency is beneficial for them or not, according to the treatment stage, specific diagnosis, and other clinical characteristics of your patients";"We should try to accommodate the therapeutic objectives offered by the app to the patient's profile."
Lack of satisfaction with content available III: Monotonous activity	1	"Patients' online activity was monotonous...for example, emotions experienced during the day were static...a specific emotion felt at a specific moment during the day cannot represent the emotional status of the patient for the entire day...Some further specifications may be added, for example, next to each emotion, an objective for the following week or a further explanation of the context around emotions and their function could be provided, in order to make emotion records more motivating for them."
Lack of satisfaction with content available IV: App was not interactive	3	"The app was not dynamic and interactive enough and its functionalities depended on the initiative demonstrated by the patient. Therefore, at times, when patients were feeling worse, they had no initiative to interact actively with the app and thus, they did not use it at all"; "The use of the app depended exclusively on each patient. For the most resistant and difficult patients, we had to follow-up a lot and remind them during sessions that they had to use the app regularly."
External factors: Human environment		
Lack of appropriate collaboration among colleagues	1	"Lack of coordination among different health professionals participating in the study. Sometimes it was not clear who was doing what."
External factors: Organizational environment		
Study design: Strict instructions	1	"Some of my patients felt frustrated and under pressure with the fact that they had to use the app at least once a day."
Economic resources available	2	"In private institutions (ie, Dexeus), sometimes patients are willing to replace face-to-face sessions with their therapist with the chat option offered by the app There are economic reasons behind this...I think that in private hospitals, the app should be used differently; you cannot force a patient to have extra visits if you consider it important, as you would do in a public hospital"; "The implementation of the app in private practice is easier...You can charge each visit 2 euros more, for example, and offer patients complementary treatment with the app...In public hospitals this option does not exist at the moment, as there is a lack of budget and of direct interest in mHealth."
Lack of time and workload	6	"Due to workload or need to attend emergencies and other priorities, sometimes patient monitoring through the app was postponed"; "Unfortunately, sometimes the symptom monitoring was not regular and constant (by us), as initially agreed when we were given the study protocol."

All of the clinicians reported that the TCApp was helpful in their treatment, and all preferred this method to keeping paper records. In terms of *ease of use*, all clinicians agreed that the app was practical, quick, and intuitive; instructions were clear, and no further technical support was needed. In terms of *usefulness*, comments were positive and mostly focused on the

usefulness of the Web-based chat function. Clinicians valued the immediacy of the instrument and the possibility to anticipate visits when necessary or to use in standard therapy all the valuable information provided by the app regarding the patient's week. In the long run, it is mentioned that the use of the app can also be cost-effective, as it can reduce patient's visits to emergency departments. A therapist specifically stated the following:

> *The app facilitates our clinical practice a lot...The whole team feels more reassured with regard to each patient´s treatment.*

The *design of the Web-based platform* was attractive and not visually overloading, and all therapists positively valued the option to receive photographs of their patients' meals.

The most frequently reported disadvantages of the app concerned (1) the lack of personalization of the app, (2) some problems with the design of the platform, and (3) the lack of time to regularly monitor patients because of workload. First, regarding the *lack of personalization of the app*, all therapists agreed that the app should be personalized in the future according to patients' clinical characteristics and specific needs (ie, sex, age, living with caregiver, treatment type, ED profile, and private vs public institution). A therapist also proposed to validate a protocol of correct use of the app according to a patient's profile, specifying the frequency of use of the app and possible functionalities relevant for a specific group of patients. Second, the only problem with the *design of the platform* concerned the difficulty expressed by therapists to follow several patients at the same time, given that a new screen was opened for each patient. Third, most of the therapists (6/8) stated that they could not always comply with the study protocol (monitor each patient on the Web at least once a week) because of their *workload* or having to attend to emergencies and other priorities. A total of 3 out of 8 therapists stated that the *app was not interactive* enough and that the use of each functionality (food records, emotion/thought/action records, chat, etc) depended on patients' self-determination and willingness to use the tool. A therapist commented on this saying the following:

> *The use of the app depended exclusively on each patient. For the most resistant and difficult patients, we had to follow-up a lot and remind them during sessions that they had to use the app regularly.*

Contradictory opinions existed regarding whether the app would be more easily implemented in private institutions or public institutions. The fear of replacing face-to-face sessions by the chat option, which is of course less costly, was expressed by a therapist working in the private sector. Another therapist from a public hospital said that the implementation of the app should be easier in private practice, as there are more economic recourses available for the integration of a complementary service that includes the TCApp. The *inappropriateness of the motivational components* employed by the TCApp (rewards, prizes, and ranking of users according to their performance) for patients with EDs, especially those with a restrictive profile, was mentioned by 2 therapists. One of them asserted the following:

> *I would modify all their content...Many patients with AN profile are stressed when receiving prizes; they want to be first in the list, corresponding to their high levels of perfectionism and competitiveness. It can be counterproductive for these patients.*

Finally, other disadvantages related to mHealth characteristics and the organizational and human environment shown in Tables 7 and 8 were reported by therapists only once. Regarding future modifications and points for improvement, the following topics were discussed:

- Personalize the app according to patients' needs and specific characteristics.
- Offer retrospective record keeping of the problematic behavior, in conjunction with real-time records. For example, a therapist commented on this saying the following:

  > *Some of my patients would have preferred it if they had had the opportunity to register a problematic behavior (restriction, vomit, binge...) the day after its occurrence. Many of the patients from the control group that normally used paper records filled the record the day after the event or when they were feeling calmer or more distanced from the problematic behavior.*

- Adapt the TCApp to nutritionists' needs. A therapist said the following:

  > *I think the use of the app could be more appropriate for nutritionists and not for psychologists...an additional functionality assessing the caloric content of the meal records may be an option.*

- Add psychoeducational material and relaxation or mindfulness techniques in the form of modules that would be activated according to patients' profile and evolution. Improve the gamified environment to make the app more interactive and motivating, for example, by including vodcasts with motivational messages by recovered patients or messages by therapists, personalized reminders, objectives, or coping strategies close to each problematic behavior.

Discussion

Principal Findings

Emerging literature on the field has shown that app-based treatment suggests promising results for enhancing the quality of mental health provision and treatment outcomes, whereas, at the same time, it can improve engagement with respect to different mental disorders, including depression, anxiety, stress, substance use, and symptoms of EDs [43,44]. Despite reported advantages for patients and health care providers, a large group of health professionals still seems somewhat critical and reluctant in the adoption of such techniques [10]. In this study, we first focused on establishing the drivers and barriers for adoption of mHealth techniques in general health care delivery processes. Second, we examined the adoption of mHealth techniques in an intervention for patients with EDs and their therapists.

The results showed that in the focus groups with health care providers and mHealth experts, most of the recurrent themes were classified as barriers, less so as facilitators, for mHealth adoption. This indicates that most professionals considered mHealth techniques as difficult to obtain and use. In addition, most barriers were attributed to external factors relating to the human or organizational environment rather than internal factors relating to individual, personal obstacles. More specifically, most health professionals reported a lack of time because of workload, a lack of strategic plan by leaders, lack of budget, and direct interest at a legislative or political level to support the implementation of mHealth. Other external barriers, which were less frequently reported, concerned the lack of specific ICT training for health professionals and the lack of communication and support among patients, health professionals, and information technology teams. In addition, individual factors, including the age-related digital divide and the fear of losing face-to-face contact with patients, generated contradictory opinions among professionals. These findings are in line with previous studies [11,13], suggesting that health care providers are still somewhat resistant and conservative about integrating mHealth technologies in their daily practice. If they decide to adopt such tools, they report that they need to feel skilled, trained, and supported by the IT team and the leaders of their institution in an adequate fashion. Similarly, Lindgreen and colleagues [30] have shown that clinicians in such situations were primarily preoccupied with challenges that related to workload. For example, some of the clinicians in the study reported frustration because of the fact that they were not supposed to monitor patient app data outside office hours, and they expressed concerns regarding the deterioration of the patient-clinician relationship. In addition, an undesired consequence of adopting mHealth techniques may be reduced clinician work satisfaction—particularly when their technological self-efficacy levels are not considered and addressed through educational and training efforts where necessary.

The results of the mHealth intervention study indicate that the TCApp was considered easy to use and useful, although patients and ED specialists monitoring patients on the Web reported different problems in adoption. Patients valued as highly positive the Web-based food records and the role these played in helping them gain a better understanding of their symptoms. In contrast, the problems reported most frequently concerned the lack of personalization of the app according to their needs (eg, diagnosis, type, and stage of treatment), as well as the overwhelming quantification of symptoms through the app (eg, presence or absence options) instead of a more qualitative evaluation of them on the basis of cognitive-behavioral chain analyses through Web-based free-text options. Similarly, a previous qualitative study with focus groups by Juarascio and colleagues [28] for patients diagnosed with binge eating indicated that the app used (including self-help material, behavior monitoring, and provisions of real-time interventions) was deemed workable and acceptable by both patients and clinicians as a complementary tool to regular treatment, although concerns were expressed about the degree of personalization and customizability of the tool. Similarly, Darcy and colleagues [27] reported that a simple self-monitoring app based on the CBT principles for ED patients was feasible and acceptable to both patients and clinicians. Last but not least, the qualitative study by Basterfield and colleagues [45] explored how individuals with an ED used various forms of technology, and the study highlighted the need for personalization, convenience, and easy-to-follow design. Participants also expressed concerns regarding safety of ICT tools, including the presence of triggering Web-based material.

In turn, ED specialists in this study highly appreciated the Web-based environment that was offered through the TCApp, especially the Web-based chat option, the usefulness of the tool outside and inside the therapeutic context, its immediacy, and the possibility to prepare for visits in advance when they felt this to be necessary. This partly runs counter to results found in the short message service text messaging intervention by Mazzeo and colleagues [46], targeting adolescent patients diagnosed with binge eating. Although their intervention showed good feasibility as reported by therapists, the adolescents in the study expressed a lack of enthusiasm for the texting component of the intervention. These findings were rather surprising, given the positive results of other studies that have used text messaging as an element of Web-based treatment for bulimia nervosa (BN) [47] and anorexia nervosa (AN) [48,49].

Both patients and health professionals of the mHealth intervention study experienced problems in complying with the study protocol at times. One of the reasons for this was the fact that patients, mostly those who were already receiving a demanding treatment at their hospital (day hospital), felt overloaded with the Web-based tasks that they had to perform on a daily basis. Another problem was that ED specialists, mostly those working in public hospitals, because of their workload or having to attend to emergencies and other priorities, tended to not give immediate responses to their patients' Web-based messages. In fact, such responses were sometimes provided by other professionals who were available at that particular point in time, raising privacy and confidentiality issues with their patients. This somewhat contradicts the results found in the systematic review by Dowling and Rickwood [50] regarding counseling and therapy using Web-based chat, where it was shown that patients were satisfied and valued the anonymity and invisibility that could be positively gained through Web-based textual conversations.

Strengths and Limitations

One of the main strengths of this study is that we investigated how both health professionals and patients experience mHealth techniques, as well as what they consider important factors that need improving to use such techniques more frequently. Second, we focused this study on a heterogeneous sample of health care providers and mHealth experts who deal with health-related problems on a daily basis and who could benefit from mHealth techniques extensively. Selecting a group of professionals with such diverse backgrounds, we were able to obtain more generic views of mHealth adoption issues. In addition, by focusing on a specific mental health problem, ED, we were able to examine attitudes and ramifications of both patients and ED specialists regarding the implementation of a specific mHealth tool in their daily practice. As a result, we were able to gather a broad view

of mHealth adoption issues among public and private health institutions in Spain for this type of condition.

The first limitation of this study is that there seemed to be a notable homogeneity among our sample of patients in terms of clinical representation, age, and gender. In fact, only female outpatients with AN decided to participate in the focus group, all of whom were adolescents and presented good adherence to face-to-face treatment (this can be also noticed by taking a look at the almost normal BMI of the participants). The impact of the exclusively AN sample on our findings may be that this group of patients, characterized by rigidity, compulsivity, and focus on detail [51], may find the app more attractive and easy to use as a complementary tool to their regular psychological treatment than patients with other ED profiles. For example, BN patients, who are often characterized as showing mood instability and lower compliance to therapeutic tasks compared with patients with AN, could possibly demonstrate worse adherence to the additional tasks offered to them through the app. The inclusion of such patients could thus pose new challenges on how to improve the gamified environment, to get such patients more engaged in Web-based treatment. Therefore, our app does not seem to have the same transdiagnostic utility as found in a study examining the utility of another mobile phone app for EDs [27]. In addition, this study's results can neither be generalized to an adult population nor to patients receiving more intensive types of treatment, such as hospitalization or admission to day hospital. A second limitation concerned the degree to which patients were satisfied with the tool and the trial in general, which largely depended on the institution in which the trial took place and the time and effort invested by professionals in complying with the study protocol. Those patients who indicated being most satisfied all hailed from public institutions, and the same therapists who followed patients face-to-face also followed them on the Web. In turn, the least satisfied users were also patients in public institutions, but here, because of workload, the task to monitor patients on the Web was carried out by an external psychologist (a collaborator in the study) who did not see patients in face-to-face sessions, and thus the external psychologist had limited knowledge regarding the clinical history of each patient. In addition, the response rate for the focus groups was lower than desirable, despite the various efforts to achieve a larger and a more representative sample. Finally, regarding data analysis and interpretation of this study's findings, the way to decide whether a finding was deemed important for participants was determined purely by frequency, that is, the number of times a theme was mentioned. However, it is possible that these themes were mentioned more often not because of a high degree of importance but because of an existing social pressure to discuss a certain theme more often than other themes. Results based on the method of categorization thus applied should therefore be interpreted with some caution.

Future Recommendations

In terms of future recommendations, all organizational, technological, and individual levels mentioned appear important for a correct mHealth implementation in clinical practice. For the individual aspect, hospitals may need to provide incentives and continuous training to encourage professionals to integrate mHealth techniques when carrying out their regular tasks. If professionals feel more empowered, skilled, and supported by their institution in adopting new technologies, they should be more willing to make use of the opportunity to try new innovations.

At an organizational level, there is a need for a strategic plan that establishes a common framework for evaluating mobile mental health apps, which allows clinicians and patients (and importantly, not only information technology teams) to identify and choose among high quality and safe mobile phone apps in accordance with their needs. Although a great number of mental health apps are readily available, and there seems to be a major potential for such apps in psychiatric assessment and interventions [14,16,32,52,53], there is limited data on their efficacy and clinical utility, and little is currently known regarding their digital security [54]. As a result, clinicians and patients remain concerned about both efficacy and privacy issues. In particular, guidelines should be established for the correct use of the TCApp by both patients and professionals, and respective functionalities should be put in place in line with each patient's clinical profile and readiness to change, as well as professional's needs, which was also suggested by Lindgreen and colleagues [29,30].

As regards mHealth characteristics, ED specialists recommended that the TCapp should be better personalized according to a patient's clinical profile and that its gamified environment should be improved by integrating more useful motivational and interactive components. It was suggested that doing so would lead to improvements in engagement for the most difficult patients. A relapse prevention module should also be integrated into the TCApp for almost recovered patients who generally receive less regular visits to the hospital and who could benefit from some functionalities of the app (ie, online chat, symptoms monitoring). Other ideas for future improvement included the possibility to adapt the TCApp to nutritionists' needs, add psychoeducational material, add relaxation or mindfulness techniques, and add a group chat functionality with a therapist who coordinates the group conversation. Last but not least, taking into account that family therapy for children and young people is the gold-standard treatment for AN [23], a version of the app should be developed for families, and its efficacy should by tested in another RCT. In fact, in some ED units that participated in the trial, it was predominantly the parents of patients who were filed in the food records that were part of their children's treatment on a daily basis.

Conclusions

In sum, this study shows that health professionals and patients foresee some issues that need to be resolved to increase adoption and usage of mHealth techniques in the near future. Owing to the possible benefits and cost-saving opportunities of mHealth techniques in health care [16], possibilities to overcome the barriers perceived by health professionals and patients are nevertheless extremely relevant. Last but not least, the results obtained by this study indicate that blended treatments might offer a good solution for the treatment of patients with EDs.

Acknowledgments

The research leading to these results has received funding from RecerCaixa. The authors would like to thank Thjb Lubbers for his help with the English editing of the manuscript.

Conflicts of Interest

None declared.

References

1. Ofcom Communications. 2018 Aug 02. Market Report URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0022/117256/CMR-2018-narrative-report.pdf [accessed 2019-05-16] [WebCite Cache ID 78PI3nNrH]
2. Fiordelli M, Diviani N, Schulz P. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013 May 21;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
3. Putzer GJ, Park Y. The effects of innovation factors on smartphone adoption among nurses in community hospitals. Perspect Health Inf Manag 2010 Jan 01;7:1b [FREE Full text] [Medline: 20697467]
4. Zhang H, Cocosila M, Archer N. Factors of adoption of mobile information technology by homecare nurses: a technology acceptance model 2 approach. Comput Inform Nurs 2010 Jan;28(1):49-56. [doi: 10.1097/NCN.0b013e3181c0474a] [Medline: 19940621]
5. Hernandez Silva E, 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] [Medline: 30635865]
6. Marrie R, Leung S, Tyry T, Cutter G, Fox R, Salter A. Use of eHealth and mHealth technology by persons with multiple sclerosis. Mult Scler Relat Disord 2019 Jan;27:13-19. [doi: 10.1016/j.msard.2018.09.036] [Medline: 30296732]
7. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
8. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005;24(5):1103-1117. [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
9. Price M, Yuen E, Goetter E, Herbert J, Forman E, 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] [Medline: 23918764]
10. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Dec 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
11. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
12. Dehzad F, Hilhorst C, de Bie C, Claassen E. Adopting health apps, what's hindering doctors and patients? Health 2014;06(16):2204-2217. [doi: 10.4236/health.2014.616256]
13. Wu IL, Li JY, Fu CY. The adoption of mobile healthcare by hospital's professionals: an integrative perspective. Decision Support Systems 2011 Jun;51(3):587-596. [doi: 10.1016/j.dss.2011.03.003]
14. Holmes EA, Ghaderi A, Harmer CG, Ramchandani P, Cuijpers PG, Morrison AP, et al. The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. Lancet Psychiatry 2018 Dec;5(3):237-286. [doi: 10.1016/S2215-0366(17)30513-8] [Medline: 29482764]
15. Kazdin A, Blase S. Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspect Psychol Sci 2011 Jan;6(1):21-37. [doi: 10.1177/1745691610393527] [Medline: 26162113]
16. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485] [Medline: 69651175]
17. Wentzel J, van der Vaart R, Bohlmeijer E, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
18. Fairburn C, Patel V. The impact of digital technology on psychological treatments and their dissemination. Behav Res Ther 2017 Jan;88:19-25. [doi: 10.1016/J.BRAT.2016.08.012]
19. Aardoom JJ, Dingemans AE, Van Furth EF. E-health interventions for eating disorders: emerging findings, issues, and opportunities. Curr Psychiatry Rep 2016 Apr;18(4):42. [doi: 10.1007/s11920-016-0673-6] [Medline: 26946513]

20. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Dec 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
21. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer E, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
22. Anastasiadou D, Folkvord F, Lupiañez-Villanueva F. A systematic review of mHealth interventions for the support of eating disorders. Eur Eat Disord Rev 2018 Dec;26(5):394-416. [doi: 10.1002/erv.2609] [Medline: 29927004]
23. National Institute for Health and Care Excellence. Eating disorders: recognition and treatment URL: https://www.nice.org.uk/guidance/ng69 [accessed 2019-05-16] [WebCite Cache ID 78PdUdj18]
24. Fairburn C. Cognitive behavior therapyeating disorders. In: Cognitive Behavior Therapy And Eating Disorders. New York: The Guilford Press; 2008.
25. Bauer S, Moessner M. Harnessing the power of technology for the treatment and prevention of eating disorders. Int J Eat Disord 2013 Jul;46(5):508-515. [doi: 10.1002/eat.22109] [Medline: 23658102]
26. Dölemeyer R, Tietjen A, Kersting A, Wagner B. Internet-based interventions for eating disorders in adults: a systematic review. BMC Psychiatry 2013 Aug 06;13:207 [FREE Full text] [doi: 10.1186/1471-244X-13-207] [Medline: 23919625]
27. Darcy A, Adler S, Miner A, Lock J. How smartphone applications may be implemented in the treatment of eating disorders: case reports and case series data. Adv Eat Disord 2014 Aug 04;2(3):217-232. [doi: 10.1080/21662630.2014.938089]
28. Juarascio AS, Goldstein SP, Manasse SM, Forman EM, Butryn ML. Perceptions of the feasibility and acceptability of a smartphone application for the treatment of binge eating disorders: qualitative feedback from a user population and clinicians. Int J Med Inform 2015 Oct;84(10):808-816 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.06.004] [Medline: 26113461]
29. Lindgreen P, Lomborg K, Clausen L. Patient experiences using a self-monitoring app in eating disorder treatment: qualitative study. JMIR Mhealth Uhealth 2018 Jun 22;6(6):e10253 [FREE Full text] [doi: 10.2196/10253] [Medline: 29934285]
30. Lindgreen P, Clausen L, Lomborg K. Clinicians' perspective on an app for patient self-monitoring in eating disorder treatment. Int J Eat Disord 2018 Dec;51(4):314-321. [doi: 10.1002/eat.22833] [Medline: 29469979]
31. Fairburn CG, Rothwell ER. Apps and eating disorders: a systematic clinical appraisal. Int J Eat Disord 2015 Nov;48(7):1038-1046 [FREE Full text] [doi: 10.1002/eat.22398] [Medline: 25728705]
32. Juarascio AS, Manasse SM, Goldstein SP, Forman EM, Butryn ML. Review of smartphone applications for the treatment of eating disorders. Eur Eat Disord Rev 2015 Jan;23(1):1-11 [FREE Full text] [doi: 10.1002/erv.2327] [Medline: 25303148]
33. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
34. 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]
35. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Dec 04;6(10):e012447 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
36. Anastasiadou D, Lupiañez-Villanueva F, Faulí C, Arcal Cunillera J, Serrano-Troncoso E. Cost-effectiveness of the mobile application TCApp combined with face-to-face CBT treatment compared to face-to-face CBT treatment alone for patients with an eating disorder: study protocol of a multi-centre randomised controlled trial. BMC Psychiatry 2018 Dec 02;18(1):118 [FREE Full text] [doi: 10.1186/s12888-018-1664-4] [Medline: 29716580]
37. Codagnone C, Lupiañez-Villanueva F. European Commission. 2013. Benchmarking deployment of eHealth among General Practitioners URL: https://ec.europa.eu/digital-single-market/en/news/benchmarking-deployment-ehealth-among-general-practitioners-2013-smart-20110033 [accessed 2019-05-17] [WebCite Cache ID 78QwPnFiZ]
38. Kidholm K, Clemensen J, Caffery L, Smith A. The Model for Assessment of Telemedicine (MAST): a scoping review of empirical studies. J Telemed Telecare 2017 Oct;23(9):803-813. [doi: 10.1177/1357633X17721815] [Medline: 28758525]
39. Brooke J. SUS-A quick and dirty usability scale. In: Jordan P, Thomas B, Weerdmeester B, McClelland I, editors. Usability Evaluation In Industry. UK: Crc Press; 1996.
40. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979 Jan;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. IBM Corp. Armonk, NY; 2011. IBM SPSS Statistics for Windows, Version 20.0 URL: https://www.ibm.com/analytics/spss-trials?mhq=spss&mhsrc=ibmsearch_a
43. 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 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]

44. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
45. Basterfield A, Dimitropoulos G, Bills D, Cullen O, Freeman VE. "I would love to have online support but I don't trust it": positive and negative views of technology from the perspective of those with eating disorders in Canada. Health Soc Care Community 2018 Dec;26(4):604-612. [doi: 10.1111/hsc.12557] [Medline: 29498122]
46. Mazzeo SE, Lydecker J, Harney M, Palmberg AA, Kelly NR, Gow RW, et al. Development and preliminary effectiveness of an innovative treatment for binge eating in racially diverse adolescent girls. Eat Behav 2016 Dec;22:199-205 [FREE Full text] [doi: 10.1016/j.eatbeh.2016.06.014] [Medline: 27299699]
47. Shapiro JR, Bauer S, Andrews E, Pisetsky E, Bulik-Sullivan B, Hamer RM, et al. Mobile therapy: use of text-messaging in the treatment of bulimia nervosa. Int J Eat Disord 2010 Sep;43(6):513-519. [doi: 10.1002/eat.20744] [Medline: 19718672]
48. Fichter M, Quadflieg N, Lindner S. Internet-based relapse prevention for anorexia nervosa: nine- month follow-up. J Eat Disord 2013;1:23 [FREE Full text] [doi: 10.1186/2050-2974-1-23] [Medline: 24999404]
49. Fichter MM, Quadflieg N, Nisslmüller K, Lindner S, Osen B, Huber T, et al. Does internet-based prevention reduce the risk of relapse for anorexia nervosa? Behav Res Ther 2012 Mar;50(3):180-190. [doi: 10.1016/j.brat.2011.12.003] [Medline: 22317754]
50. Dowling M, Rickwood D. Online counseling and therapy for mental health problems: a systematic review of individual synchronous interventions using chat. In: Journal of Technology in Human Services. United Kingdom: Taylor and Francis; Jan 2013:1-21.
51. Treasure J, Sepulveda A, MacDonald P, Whitaker W, Lopez C, Zabala M, et al. The assessment of the family of people with eating disorders. Eur Eat Disord Rev 2008 Jul;16(4):247-255. [doi: 10.1002/erv.859] [Medline: 18240125]
52. Ben-Zeev D, McHugo G, Xie H, Dobbins K, Young M. Comparing retrospective reports to real-time/real-place mobile assessments in individuals with schizophrenia and a nonclinical comparison group. Schizophr Bull 2012 May;38(3):396-404 [FREE Full text] [doi: 10.1093/schbul/sbr171] [Medline: 22302902]
53. Tregarthen JP, Lock J, Darcy AM. Development of a smartphone application for eating disorder self-monitoring. Int J Eat Disord 2015 Nov;48(7):972-982. [doi: 10.1002/eat.22386] [Medline: 26213130]
54. Chan S, Torous J, Hinton L, Yellowlees P. Towards a Framework for Evaluating Mobile Mental Health Apps. Telemed J E Health 2015 Dec;21(12):1038-1041. [doi: 10.1089/tmj.2015.0002] [Medline: 26171663]

Abbreviations

ACES: Associació Catalana d'Entitats de Salut
AN: anorexia nervosa
BMI: body mass index
BN: bulimia nervosa
CBT: cognitive behavioral therapy
COMB: Colegio Oficial de Médicos de Barcelona
CSQ-8: Client Satisfaction Questionnaire
ED: eating disorder
ICT: information and communication technology
MASH: Model for ASessment of mHealth
mHealth: mobile health
RCT: randomized controlled trial
SUS: System Usability Scale
UOC: Universitat Oberta de Catalunya

Edited by G Eysenbach; submitted 26.11.18; peer-reviewed by P Lindgreen, C Jacob; comments to author 10.01.19; revised version received 19.03.19; accepted 23.04.19; published 31.05.19.

Please cite as:
Anastasiadou D, Folkvord F, Serrano-Troncoso E, Lupiañez-Villanueva F
Mobile Health Adoption in Mental Health: User Experience of a Mobile Health App for Patients With an Eating Disorder
JMIR Mhealth Uhealth 2019;7(6):e12920
URL: https://mhealth.jmir.org/2019/6/e12920/
doi:10.2196/12920
PMID:31199329

Original Paper

Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study

Na Wang[1,2,3], MScN; Zequn Deng[1,3], BS; Li Ming Wen[4,5], PhD; Yan Ding[2], PhD; Gengsheng He[1,3], PhD

[1]Department of Nutrition and Food Hygiene, School of Public Health, Fudan University, Shanghai, China

[2]Nursing Department, Obstetrics and Gynaecology Hospital of Fudan University, Fudan University, Shanghai, China

[3]Key Laboratory of Public Health Safety of Ministry of Education, Fudan University, Shanghai, China

[4]School of Public Health, University of Sydney, Sydney, Australia

[5]Health Promotion Unit, Sydney Local Health District, Sydney, Australia

Corresponding Author:
Gengsheng He, PhD
Department of Nutrition and Food Hygiene
School of Public Health
Fudan University
No. 130 Dongan Road
Shanghai, 200032
China
Phone: 86 21 54237229
Fax: 86 21 54237271
Email: gshe@shmu.edu.cn

Abstract

Background: Hospital-based health promotion resources to assist pregnant women in adopting a healthy lifestyle and optimizing gestational weight gain are important, but with limited effects. Increasingly, women are using mobile apps to access health information during the antenatal period.

Objective: The aims of the study were to investigate app-usage by Chinese women during pregnancy and to gain a better understanding of their views and attitudes toward apps containing health information.

Methods: A mixed methods study design was applied. Study participants were recruited from 2 maternity hospitals in Shanghai, China, between March and July 2018. A self-administered Web-based survey was conducted with 535 pregnant Chinese women on their sources of health information and reasons for using apps during pregnancy. A total of 4 semistructured focus groups were also conducted with the pregnant women (n=28).

Results: The use of pregnancy-related apps and the internet was common among the respondents. Almost half of the women had used pregnancy-related apps. Specifically, the use of apps for health information declined as pregnancy progressed from 70% (35/50) in the first trimester to 41.3% (143/346) in the third trimester. The main reason for using an app was to monitor fetal development (436/535, 81.5%), followed by learning about nutrition and recording diet in pregnancy (140/535, 26.2%). The women found that the apps were useful and convenient and can support lifestyle modifications during pregnancy. However, some apps also contained misinformation or incorrect information that could cause anxiety as reported by the participants. Many women expressed the need for developing an app containing evidence-based, well-informed, and tailored health information to support them during pregnancy.

Conclusions: The study suggests that apps were widely used by many Chinese women during pregnancy to monitor fetal development, to obtain diet and physical activity information, and to track their body changes. The women highly appreciated the evidence-based information, expert opinions, and tailored advice available on apps. Smartphone apps have the potential to deliver health information for pregnant women.

(JMIR Mhealth Uhealth 2019;7(6):e12631) doi:10.2196/12631

KEYWORDS

mobile applications; pregnancy; consumer health information; health promotion

Introduction

Background

The Developmental Origins of Health and Disease concept describes how, during early life (at conception and/or during fetal life, infancy, and early childhood), the environment induces changes in development that have a long-term impact on later health and disease risk [1]. Maternal lifestyle, such as diet and physical activity, and their health status (eg, obesity and gestational diabetes mellitus) can affect the potential risk of metabolic diseases, obesity, and diabetes of the next generation [2,3]. In general, women are especially receptive to advice about a healthy lifestyle during pregnancy, in which optimizing their lifestyle for the health benefits of their offspring is a powerful motivator. Despite this, physical inactivity, under nutrition or over nutrition, and suboptimal gestational weight gain (GWG) are commonly reported by studies with pregnant women [4-6].

Pregnancy is a life phase when women are adjusting to their new status as mothers. They may be dealing with uncertainty and anxiety, as well as learning how to care for themselves and their newborns [7]. Hospital-based resources to assist pregnant women in achieving better nutrition and physical activity levels and optimizing GWG are extremely important, but with limited effects [8,9]. Research also suggests that many people would seek health information via Web-based media following a medical visit when they felt dissatisfied or disagreed with the advice they received from their medical providers [10]. However, seeking more health information can also be positive; people may look for additional and complementary information to discuss with health care professionals.

Mobile health (mHealth), defined as the use of mobile phones and other wireless technology such as texting messaging, apps, and video messaging to support the achievement of health objectives [11], is a burgeoning field of public health. With the widespread application of digital health care, mHealth services are also thriving in the field of maternal and child health in China [12]. More women are turning to digital sources for health information and support during pregnancy and early motherhood [13]. In particular, thousands of maternal and child health-related apps have appeared in the Android app market and app store in China to cater to women's needs [12]. These apps cover a range of topics relating to maternal and fetal health, from general factual encyclopedias of pregnancy facts to more specific information on maternal diet for gestational diabetes [14]. Recently, a survey conducted in Hubei province, China, showed that women used apps mainly to monitor and record fetal development and their body changes as well [15]. However, gaps exist between functions of apps in the current market and women's expectations for functions of a pregnancy app. These gaps mainly reflect in the ability of an app to solve the common problems encountered and its guidance on diet and physical activity during pregnancy. Hearn et al conducted a study with perinatal women in Australia, which was focused on developing a website and an app for health promotion during the perinatal period [16]. Their study found that, in particular, women wanted information on issues, such as nutrition and diet, exercise, managing weight gain, sleeping problems, emotional fluctuations, allergies, breastfeeding, and gestational diabetes.

Smartphone apps may provide a novel way to provide health information, facilitate individual access to resources, and increase health care engagement. They could be used as a strategy to overcome some barriers associated with traditional face-to-face health care interactions [14]. Some commonly cited barriers to traditional health care include lack of referrals from health care professionals, program availability, inconveniences of meetings (eg, lack of time or transportation), patient embarrassment, and high financial costs [17]. Recently, several studies have investigated the use of apps as a tool to prevent excessive GWG and to manage or prevent gestational diabetes and some of them demonstrated the potential feasibility and efficacy of app-based interventions in promoting a healthy lifestyle in the antenatal period [18-20]. A systematic review [20] found that most studies on mHealth apps that aim to support lifestyle and medical care for high-income countries revealed the usability of these apps to reduce GWG, to increase intake of vegetables and fruit, to quit smoking but still needed further evidence to show the effective use of these apps during pregnancy. Ledford's study showed that, as a prenatal education tool, mobile apps effectively improved pregnant women's self-management for health [19].However, despite the increasing use of apps by pregnant women, to date, few published studies have focused on how and why women used apps for health promotion during pregnancy, particularly in the Chinese context.

Objectives

The aims of this study were to (1) investigate app usage of Chinese women during pregnancy, (2) gain a better understanding of their views and attitudes toward apps, (3) explore their concerns about data privacy and security, and (4) inform the development of a health promotion program through an app.

Methods

Study Design

A mixed methods study was used with both quantitative and qualitative studies, consisting of a cross-sectional Web-based survey and focus group (FG) interviews. In the survey, we focused on women's current status of app usage during pregnancy and their intention to use an app for health promotion information. To obtain a better understanding of pregnant women's attitudes and views of using apps, semistructured FGs were conducted. Both the survey and FGs were conducted in Shanghai, China, between March and July 2018.

Setting and Participants

This study was conducted in 2 maternity hospitals, the Obstetrics and Gynecology Hospital of Fudan University and Maternal and Child Health Hospital, Changning District, Shanghai, China. The former is a university-affiliated maternity hospital, a tertiary hospital with the total number of births around 15,000 annually. The latter is a district hospital ranked as a secondary hospital with the total number of births being 11,000 annually. Hospitals in China are categorized as primary, secondary, or tertiary institutions according to a 3-tier system that recognizes a

hospital's capacity to provide medical care and medical education and conduct medical research. For example, a tertiary hospital is responsible for providing specialist health services and performing a bigger role with regard to medical education and scientific research.

The purpose of recruiting pregnant women from the tertiary and secondary hospitals was to have a broader representative study sample. All the pregnant women who had registered in the clinics were required to attend the routine antenatal clinics and 5 face-to-face sessions of antenatal education throughout their pregnancy. At the end of the antenatal education sessions, women were approached by 2 research nurses for their eligibility and consent.

Women were eligible if they were (1) aged 18 years and above, (2) currently pregnant, (3) owning a smartphone, and (4) able to give written consent to participate.

Ethical Considerations

The study was approved by the Ethics Committee of the Obstetrics and Gynecology Hospital of Fudan University (Ethics Approval number 2017-74). Before conducting the survey or FGs, all potential participants were informed about the purposes and procedures of the study, and their written informed consent was obtained from those agreeing to participate. All participants were notified that they had the right to refuse to participate, and in addition, they could subsequently withdraw from the study at any time if they had participated.

Data Collection

A Cross-Sectional Survey

A self-administered Web-based survey using iPads was undertaken by pregnant women during their clinical visits using the survey website . A total of 582 women were approached by 2 research nurses for their eligibility and consent for the survey, and 535 respondents provided informed consent and completed the survey with a completion rate of 91.9% (535/582). The main reason for refusal was *not interested* or *uncompleted questionnaires*. To check for completeness, the mandatory items of the questionnaire would be highlighted. Participants were able to review and change their answers through a Back button before submission. No incentives were provided for participation in this survey.

The 19-item questionnaire (Multimedia Appendix 1) included maternal sociodemographic characteristics, medical history and current pregnancy information, health behaviors, and smartphone and app usage during pregnancy. The questionnaire was developed in consultation with experts in the field and pilot tested in 20 respondents before the survey. Physical activity levels during pregnancy were assessed based on the International Physical Activity Questionnaire [21]. In our study, we followed the American College of Obstetricians and Gynecologists (ACOG) guidelines [22] for physical activity for pregnant women to evaluate their daily moderate activity levels. The last 4 questions were about smartphone and pregnancy app usage. These questions had multiple response options, and the respondents could also add other options for the questions. These survey items were sourced from other surveys [15,23] for pregnancy app usage and were adapted for our survey.

Focus Groups

Following the Web-based survey, the women were invited to participate in the FG interviews if they indicated their interest in having further discussion on their views of using an app in the Web-based survey. Women who consented to the interview were purposively recruited to obtain a sample with a wide range of gestational ages. A total of 50 pregnant women were approached to participate in the FGs. Of them, 28 (28/50, 56%) agreed to participate, with their ages ranging between 24 and 37 years and gestational ages ranging between 13 and 40 weeks. During the FG session, the first author, as the moderator, conducted all the interviews with assistance from the second author (ZD) in a conference room located at the Obstetrics and Gynecology Hospital of Fudan University. The FGs were conducted in Mandarin and tape-recorded. A semistructured interview guide was used during the group discussion to ensure that all required information was sought [24], but the discussion remained open to any new issues or concerns related to app usage for supporting pregnancy. The interview guide with open-ended questions was developed based on the Web-based survey and a literature review of similar studies. The FGs began with broad issues and were then narrowed to more specific topics as the group discussion progressed. The main discussion questions covered topics including the following:

1. Their experience, attitudes, and views regarding using an app for supporting pregnancy;
2. Their sources of health information and reasons for selecting a particular app for pregnancy;
3. Their concerns about data privacy and security issues when using the apps;
4. What features they would like to see and use in an app for supporting pregnancy.

Each FG consisted of 5 to 8 women and took about 40 to 60 min. During the group discussion, field notes were also taken to facilitate the data analysis. Data saturation was reached after 4 FGs when no further unique or new themes were introduced by the FG participants.

Data Analysis

For the cross-sectional survey, data were analyzed using the SPSS software package (version 22.0, SPSS Inc). Body mass index (BMI) was calculated as weight (kg)/height (meters) squared and classified as underweight (BMI <18.5 kg/m^2), normal weight (BMI 18.5 to 23.9 kg/m^2), overweight (BMI 24.0 to 27.9 kg/m^2), or obese (BMI ≥28.0 kg/m^2) based on the BMI cut points for the Chinese [25]. Women were grouped into adequate GWG, inadequate GWG, and excessive GWG according to the Institute of Medicine (IOM) recommendations for week-specific weight gain during pregnancy [26]. Descriptive statistics were used to present the characteristics of the respondents, providing frequencies (numbers) and proportion (percentages) for categorical variables. Continuous variables were described with mean and SD or median and interquartile range (IQR) in the case of abnormal distribution. Women's sources of information and reasons for using apps during

pregnancy were analyzed and compared between different trimesters using the chi-square test.

A multivariable logistic regression was used to calculate adjusted odds ratios (ORs) and 95% CIs to identify factors associated with app usage of pregnant women in Shanghai. A few steps were taken to identify these factors. First, we used forward stepwise regression to remove those variables with *P* values >.20. Second, considering that education and house income were related to app usage as shown by previous studies [7,23], we also included these 2 variables in the models. In addition, maternal age as a potential confounder was also included in the final model. Finally, we included maternal age, education, household income, parity, prepregnancy BMI category, and trimester as independent variables in the multiple regression model.

Qualitative data from the 4 FGs were used to supplement the quantitative findings. All audio recordings were transcribed verbatim following each FG within 24 hours. Transcripts and field notes were reviewed and analyzed by 2 independent reviewers using a thematic analysis [27]. The goal of a thematic analysis is to identify themes, that is patterns in the data, that are important or interesting and use these themes to address the research or say something about an issue. We performed the analysis based on Braun and Clarke's 6-phase framework for doing a thematic analysis (Multimedia Appendix 2) [28]. After reading and rereading our transcript, we became familiar with our data. We worked through each transcript and coded every segment of the text that seemed to be relevant to or specifically address our research question. Specifically, line-by-line coding highlighted key comments depicting the attitudes and views regarding using apps and its related factors. Codes were then organized into broader themes that seemed to express something specific about our research question. Finally, we gathered together all the data that were relevant to each theme to review and refine the emerging themes. Any discrepancies in data analysis between the 2 authors involved were resolved by discussion until a consensus had been reached on the codes, subthemes, and themes.

Results

Quantitative Findings

Characteristics of The Survey Respondents

The demographic characteristics of the study population are presented in Table 1. The mean maternal age was 30.6 (SD 3.6 years) and the median (IQR) gestational age was 32 (17 to 33) weeks. Of the 535 respondents, 478 (478/535, 89.3%) were nulliparous. The prevalence of overweight and obesity was 19.4% (104/535), whereas underweight women were accounted for 15.9% (85/535). Compared with the IOM recommendations for GWG, only 43.7% of women had normal GWG. Only 22.1% (118/535) of women had a physical activity level meeting the guideline recommended by ACOG for physical activity for pregnant women. Anemia was also common among our population (23.7%, 127/535). The percentage of women smoking before pregnancy was 3.7% (20/535), but all quit during pregnancy. The respondents as a whole were well educated, with more than 90% (500/535, 93.5%) having completed a college degree.

Table 1. Characteristics of study participants in the survey, Shanghai, China, 2018 (N=535).

Characteristics	Values
Age (years), n (%)	
≤25	30 (5.6)
26-34	436 (81.5)
≥35	69 (12.9)
Education, n (%)	
≤Senior high school	35 (6.5)
College degree	398 (74.4)
Postgraduate degree or above	102 (19.1)
Marital status, n (%)	
Single/divorced/separated/widowed	6 (1.1)
Married	529 (98.9)
Household income[a,b], n (%)	
<¥10,000	115 (21.8)
¥10,000-¥30,000	302 (57.2)
>¥30,000	111 (21.0)
Health issues and physical activity, n (%)	
First degree family history of diabetes mellitus	73 (13.6)
Anemia before or during pregnancy	127 (23.7)
Smoking before pregnancy	20 (3.7)
Smoking during pregnancy	0 (0)
Physical activity (moderate level <30 min)	417 (77.9)
Parity (primipara)	478 (89.3)
Prepregnancy weight category, n (%)	
Underweight	85 (15.9)
Normal weight	346 (64.7)
Overweight	75 (14.0)
Obese	29 (5.4)
Gestational week (week), median (interquartile range)	32 (17-33)
Weight gain, n (%)	
Inadequate GWG[c]	172 (32.1)
Adequate GWG	234 (43.7)
Excessive GWG	129 (24.1)

[a]N=528.

[b]One Chinese Yuan (¥)=US $0.1437.

[c]GWG: gestational weight gain.

App Usage and Sources of Health Information During Pregnancy

Table 2 shows the respondents' sources of health information during pregnancy. More than 80% (438/535, 81.9%) of the respondents' smartphone operating system was iOS. Women reported that they received pregnancy-related health information mostly from Web-based media (73.3%, 392/535), followed by pregnancy apps (49.2%, 263/535), face-to-face consultations with health professionals (20.4%, 109/535), family and friends (18.1%, 97/535), and print materials (15.5%, 83/535), whereas their preferred or expected sources of information were Web-based media (68.4%, 366/535), followed by pregnancy apps (48.8%, 261/535) and face-to-face consultations with health professionals (34.4%, 184/535). We further compared the difference between women's expected sources of information and their current information sources. The results showed that the actual source of information from health professionals was

lower than that expected by the pregnant women (*P*<.001), whereas the source of information from family/friends was higher than that which was expected (*P*=.001). In addition, women's reported sources of information varied over the period of their pregnancy. Specifically, the use of an app for health information declined as pregnancy progressed from 70% (35/50) in the first trimester to 41.3% (143/346) in the third trimester. A logistic regression analysis also showed that, compared with those in the third trimester, the rates of app usage for health information were significantly higher in the first trimester (adjusted OR 3.057, 95% CI 1.575-5.932) and the second trimester (adjusted OR 2.206, 95% CI 1.453-3.349). In addition, nulliparous women were more likely to use an app for health information (adjusted OR 1.975, 95% CI 1.043-3.742; Table 3).

Table 2. Sources of health information by trimester of pregnancy of Chinese women in Shanghai, China, 2018.

Sources of health information	All women (N=535), n (%)	First trimester (n=50), n (%)	Second trimester (n=139), n (%)	Third trimester (n=346), n (%)	*P* value
Smartphone operating system					
iOS	438 (81.9)	39 (78)	111 (79.9)	288 (83.2)	.52
Android	97 (18.1)	11 (22)	28 (20.1)	58 (16.8)	.52
Expected sources of information for health promotion					
Pregnancy apps	261 (48.8)	35 (70)	75 (54.0)	151 (43.6)	.001[a]
Other Web-based media	366 (68.4)	37 (74)	96 (69.1)	233 (67.3)	.63
Television	36 (6.7)	6 (12)	17 (12.2)	13 (3.8)	.001[a]
Paper materials	84 (15.7)	14 (28)	33 (23.7)	37 (10.7)	<.001[a]
Face-to-face with health professionals	184 (34.4)	23(46）	57（41.0）	104（30.1）	.01[a]
Family/friends	59 (11.0)	4 (8)	4 (2.9)	51 (14.7)	.001[a]
Current sources of information for health promotion					
Pregnancy apps	263 (49.2)	35 (70)	85 (61.2)	143 (41.3)	<.001[a]
Other Web-based media	392 (73.3)	40 (80)	120 (86.3)	232 (67.1)	<.001[a]
Television	39 (7.3)	8 (16)	17 (12.2)	14 (4.0)	<.001[a]
Paper materials	83 (15.5)	15 (30)	35 (25.2)	33 (9.5)	<.001[a]
Face-to-face consultations with health professionals[b]	109 (20.4)	6 (12)	27 (19.4)	76 (22.0)	.25
Family/friends[b]	97 (18.1)	3 (6)	5 (3.6)	89 (25.7)	<.001[a]

[a]Represents a significant difference between the 3 groups.

[b]There were significant differences between current sources of information and expected sources of information from face-to-face consultations with health professionals (*P*<.001) and family/friends (*P*=.001) for health promotion.

Table 3. Results of a multivariable logistic regression analysis of factors associated with app usage of pregnant women in Shanghai, China (N=528; after exclusion of 7 cases for missing data on household income).

Factors[a]	Beta	SE	Wald X^2	Odds ratio	95% CI	*P* value
Age (years)	—[b]	—	0.286	—	—	.87
≤25	—	—	—	1.000 (reference)	—	—
26-34	.041	0.400	0.010	1.042	0.475-2.282	.92
≥35	−.113	0.474	0.057	0.893	0.352-2.263	.81
Education (1= Postgraduate degree or above, 0=≤College degree)	.036	0.238	0.024	1.037	0.651-1.653	.88
Household income[c]	—	—	3.787	—	—	.15
<¥10,000	—	—	—	1.000 (reference)	—	—
¥10,000-30,000	.222	0.233	0.910	1.249	0.791-1.970	.34
>¥30,000	.559	0.290	3.722	1.749	0.991-3.087	.05
Parity (1=multipara, 0=primipara)	.681	0.326	4.361	1.975	1.043-3.742	.04[d]
Prepregnancy weight category	—	—	6.994	—	—	.07
Underweight	.197	0.256	0.594	1.218	0.738-2.010	.44
Normal weight	—	—	—	1.000 (reference)	—	—
Overweight	−.396	0.264	2.248	0.673	0.401-1.129	.13
Obese	−.952	0.486	3.843	0.386	0.149-1.000	.05
Trimester	—	—	20.736	—	—	<.001[d]
First	1.117	0.338	10.91	3.057	1.575-5.932	.001[d]
Second	.791	0.213	13.802	2.206	1.453-3.349	<.001[d]
Third	—	—	—	1.000 (reference)	—	—

[a]Regression models included maternal age, education, household income, parity, pre-pregnancy body mass index category and trimester.

[b]Not applicable.

[c]One Chinese Yuan (¥)=US $0.1437.

[d]Represents the variable is significant in the logistic regression model.

As shown in Table 4, the most reported reason of app usage was for monitoring fetal development (436/535, 81.5%). Getting information related to nutrition and recording diet during pregnancy (26.2%, 140/535) was also a frequently reported reason for using an app. Other reasons for app usage included keeping track of information about their antenatal care appointments (128/535, 23.9%), tracking their own body (101/535, 18.9%), getting information related to physical activity and recording exercise during pregnancy (91/535, 17.0%), recording antenatal examination (87/535, 16.3%), having Web-based discussions with other pregnant women (83/535, 15.5%), storing photos of themselves during pregnancy (52/535, 9.7%), and uploading and storing fetal ultrasound images (27/535, 5.0%). These reasons for using apps during pregnancy also changed over the period of pregnancy.

Table 4. Reasons for using pregnancy apps by the trimester of pregnancy.

Reasons for using pregnancy apps	All women (N=535), n (%)	First trimester (n=50), n (%)	Second trimester (n=139), n (%)	Third trimester (n=346), n (%)	*P* value
Monitoring fetal development	436 (81.5)	36 (72)	89 (64.0)	311 (89.9)	<.001[a]
Tracking own body	101 (18.9)	16 (32)	37 (26.6)	48 (13.9)	<.001[a]
Learning information regarding nutrition during pregnancy and recording diet	140 (26.2)	16 (32)	48 (34.5)	76 (22.0)	.01[a]
Learning information regarding physical activity during pregnancy and recording exercise	91 (17.0)	14 (28)	25 (18.0)	52 (15.0)	.07
Understanding the content of antenatal care	128 (23.9)	15 (30)	40 (28.8)	73 (21.1)	.12
Storing photos of themselves	52 (9.7)	4 (8)	8 (5.8)	40 (11.6)	.14
Storing fetal ultrasound images	27 (5.0)	3 (6)	6 (4.3)	18 (5.2)	.88
Recording antenatal examination	87 (16.3)	7 (14)	18 (12.9)	62 (17.9)	.37
Web-based discussions with other pregnant women	83 (15.5)	7 (14)	21 (15.1)	55 (15.9)	.93

[a]Represents a significant difference between the 3 groups.

Table 5. Demographic characteristics of focus group participants in Shanghai, China.

Characteristics	Group 1 (n=5)	Group 2 (n=8)	Group 3 (n=7)	Group 4 (n=8)	Total (n=28)
Age (years), mean (SD)	29.8 (3.5)	29.3 (3.4)	30.1 (1.7)	29.3 (3.8)	29.6 (3.1)
Household income[a], n (%)					
<¥10,000	2 (40)	4 (50)	2 (29)	3 (38)	11 (39)
¥10,000-30,000	1 (20)	4 (50)	4 (57)	5 (63)	14 (50)
>¥30,000	2 (40)	0 (0)	1 (14)	0 (0)	3 (11)
Education, n (%)					
≤Senior high school	1 (20)	0 (0)	0 (0)	2 (25)	3 (11)
College degree	3 (60)	6 (75)	5 (71)	6 (75)	20 (71)
Postgraduate degree or above	1 (20)	2 (25)	2 (29)	0 (0)	5 (18)
Parity (primipara), n (%)	4 (80)	8 (100)	6 (86)	8 (100)	26 (93)
Prepregnancy body mass index (kg/m2), median (IQR[b])	18.6 (16.9-18.9)	21.8 (20.4-27.4)	22.6 (20.8-22.4)	20.6 (18.5-22.3)	21.0 (18.9-23.5)
Gestational age (week), median (IQR)	32 (13-36)	33 (15-40)	30 (18-37)	34 (17-39)	32 (15-38)

[a]One Chinese Yuan (¥)=US $0.1437.

[b]IQR: interquartile range.

Qualitative Findings

The characteristics of FG participants are shown in Table 5. In total, 4 overarching descriptive themes emerged from the data analysis (Multimedia Appendix 3) about app usage during pregnancy, including (1) accompany and support, (2) disadvantages, (3) data privacy and security issues, and (4) expectations for features of an app. Participants' quotations were used to illustrate the original text, and only codes were used to maintain anonymity.

Theme 1: Accompany and Support

Usefulness

In general, most of the participants (n=26) regarded that apps or other Web-based media play an indispensable role during the period of pregnancy, especially in their early stage of pregnancy. Participants valued apps with multifunctional features:

> *Most of the features are useful, such as what you can eat and what you can't eat during pregnancy, interpretation of baby ultrasound report, and how to deal with discomfort during pregnancy, and so on.* [Focus Group (FG) 1, Participant (P): C]

Some participants (n=3) used apps before pregnancy and continued to use apps after pregnancy. These behaviors reflected that apps were useful and helpful in their daily life. Apps provided them reference, companionship, and support:

> *I used "Crazy Creation" before pregnancy, it helped me conceive. Now I use "Pregnancy Butler Pro" to accompany pregnancy.* [FG2, P: C]

Participants (n=13) also valued the discussion forum platform provided by mobile apps, through which they could gain peer support, especially from experienced women or *someone who has just experienced this process*, who help a lot in their discussion forum:

> *I like online discussion and exchange feelings with other pregnant women and it can offer me some tips for my pregnancy and make me reassurance.* [FG3, P: A]

Convenience

The convenient nature of accessing health information through apps was favorably compared with the waiting time in clinics and the trip to a hospital for talking to health care professionals. Women sought apps and other Web-based media, partly because the antenatal care visit structure and the information delivered during antenatal care visits could not meet their needs. These needs were often reflected in questions such as "Do I need to take nutritional supplements?", "How do I deal with insomnia and constipation conditions during pregnancy?", and "Can I take medicines when I get a cold?". Women also expressed their understanding of doctors' heavy workload as doctors lacked time to deal with their queries. In China, Maternity Health Care is an obstetrician-led model, where there is no role for midwives to provide a health promotion service, which is different from many other countries. This became evident during the FGs as shown in their statements (n=10):

> *Usually, I can see doctors once a month for less than ten minutes conversation every time. Just imaging when I suffered from insomnia, the app and internet would provide me very detailed information regarding how to improve sleep immediately...You can see the advantages of it.* [FG4, P: B]

Another woman echoed this opinion:

> *Doctors like to say there is no other way than prescribing medicine. In that case, I don't want to bother him/her...I prefer to search the answers on my app.* [FG4, P: D]

Supporting Women's Lifestyle Modifications During Pregnancy

Pregnancy is widely viewed as a period when women are open to lifestyle changes. In our study, all the participants said they would change their behaviors that may harm their babies' health. To reach optimal pregnancy outcomes, an app can facilitate their lifestyle changes for a healthy pregnancy (n=20):

> *Yes, all the foods and exercise would be checked through app and internet. Sometimes, I can't help eating McDonald's, but it was really reduced a lot...I am also using the app to record my daily eating, physical activity and weight gain...the app can provide me a weight-gain curve, it is interesting.* [FG2, P: C]

Theme 2: Disadvantages in Apps

Negative Emotions

Although learning about the experiences of others may help individuals feel at ease, some women pointed out that they felt more worried and anxious after having read information on discussion forums through apps . Throughout the FG discussions (n=10), there was acknowledgment that many of the stories people posted on apps were often *horror stories*:

> *The grievances in some of the forums are too heavy, (discussions on) "Baby Tree" (one of apps) is full of negative emotions...Today a miscarriage of information suddenly jumped in front of me...no one want to see information like this.* [FG1, P: C]

Commercial Motives

As apps in the Chinese market can be freely downloaded by everyone, many developers would take advantage of this opportunity to make profits through advertisements:

> *There are advertisements in the "Baby Tree"...many contents are charged.* [FG1, P: A]

Even though women valued expert advice and expressed the desire for information and support offered by health care professionals, they did not want to pay for these Web-based services; however, they accepted the idea of patients paying for medical services in the hospital settings (n=8):

> *When I want to seek expert advice and some detailed explanation on this topic via apps...I am interested after scanning, but I was pointed in... you should pay for it.* [FG1, P: C]

Invalidated Information

Most pregnant women (n=18) were worried about the lack of a scientific and validated source of information available on apps:

> *I don't like anything about the messy apps in the app store, the information in the "Baby Tree" is quite contradictory sometimes, for example, the app that I used before said leeks and red beans could cause miscarriage, but it still recommended the recipe of red bean porridge to me...* [FG4, P: G]

Theme 3: Privacy and Data Security

Pregnant women (n=8) across all FGs said they were not worried about information security issues, because they were willing to sacrifice some privacy or personal information to gain their convenience and other uses. It was mentioned in 2 FGs (n=4) that when the hospital and developer cooperate to develop a pregnancy app, the hospital should be responsible for the privacy and security of the pregnant women's information:

> *You shouldn't find a place for information security right now, unless you don't use your phone or the internet.* [FG3, P: F]

Theme 4: Expectations for Features of an App

Professional and Evidence-Based Information

Participants (n=22) mentioned that an ideal app should provide them staged information and professional advice based on gestational age or different trimesters. When childbirth is approaching, they want to learn something about infant feeding practice in advance:

> *Apps should provide symptoms during each trimester that a (pregnant) women should have and give gestation-specific information...If only apps can tell me baby care such as breastfeeding and how to handle problems that may encountered.* [FG4, P: A and FG2, P: E]

> *I hope that the article can come from the doctor's hand, it is more convincing...Don't be confrontational (contradictory) between articles.* [FG4, P: G]

However, some women (n=4) had a different opinion toward information given by doctors. They wanted a gold standard or evidence-based guide that could be referenced and directed, rather than just doctors' opinions. The information acquired through apps helped them to take control of themselves:

> *Doctors have different opinions on the same report...one doctor told me to eat iron supplement, but the other said that I didn't need...it made me confused. I hope that an app provides the reference then I can read my report myself.* [FG3, P: B]

Generality and Individuality

Women wanted information from general factual encyclopedias of pregnancy facts to more specific domains. Particularly, when women suffered from complications during pregnancy, they desired an app to support disease management for conditions, such as gestational diabetes. This means that an app is expected to not only provide general information related to pregnancy but also give personalized support based on the individuals' conditions (n=6):

> *Can the app set up a diabetes zone? To tell the truth, I don't know how to eat after diagnosis of gestational diabetes until hospitalization...There are too many diabetes mums.* [FG3, P: G]

Integrating With Antenatal Care

Participants (n=10) favored integrating pregnancy apps with their antenatal care in the hospital and further receiving personalized advice from health professionals:

> *It is best to have an app that can be connected to the medical card and linked to the usual check-ups, then give suggestions based on personalized examinations and different gestational weeks. This is relatively better than anything else.* [FG4, P: E]

Discussion

Principal Findings

This study provides in-depth insights into the app usage of Chinese women during pregnancy. Our findings suggest that the use of apps and the internet during pregnancy was common among the respondents. Women found apps to be useful and convenient for accessing health information and can be used to promote lifestyle modifications during pregnancy. However, some apps also contained misinformation or incorrect information that could cause anxiety, confusion, and misguidance. The need for developing an app containing evidence-based, well-informed, and tailored health information was also expressed by many women in our study.

Pregnancy is a period that includes many changes which require physical, psychological, and social adaptation. The majority of the pregnant women in our study were first-time mothers, and they needed more information to cope with these changes during pregnancy. They sought information from various sources, such as their health providers, apps and the internet, books, and family and friends. Pregnant women are increasingly using mobile apps as a source of supplemental information [29]. In our study, nearly half of the women stated that they used at least one mobile app during pregnancy. Other studies reported that the rate of use of pregnancy-related mobile apps ranged from 41.3% to 73% [7,12,23,30], and this wide variability may relate to their local contexts, socioeconomic status, the time of the survey, gestational age, and parity. Researchers suggest that websites and apps may be particularly helpful for women from socioeconomically disadvantaged groups, who may lack access to traditional media sources [31].

Our survey also found that women's usage of apps declined from the first trimester to third trimester, which was also supported by the FGs indicating that they particularly needed pregnancy apps or other sources of information to guide them on what to do in their early pregnancy. Similar findings were also reported by Lupton and Pedersen [23]. In addition, women may turn to other sources for support, such as experienced women within their family or friend circle, as we found that the reported source of information from family and friends increased as pregnancy progressed.

Similar to findings from surveys conducted in Australia and Turkey [23,32], most women participating in this study used pregnancy apps to monitor fetal development, whereas only 26.2% and 17.0% of women used apps for managing their diets and exercise, respectively. A number of participants in the FGs expressed that they wanted to know what they can or cannot eat during pregnancy. The lower rate of using these functions may be due to a lack of evidence-based and credible information available on apps [33]. Womack pointed out that only 50% of their study sample cited pregnancy apps as a source for their recommendation and presented conflicting recommendations for alcohol and food intake during pregnancy. This lack of credibility alongside simultaneous conflicts of risk threatens women's ability to discern which recommendations should be followed [34]. Kaimal reported in their study of Web-based resources for obstetrics that only 3.6% of websites were created or sponsored by obstetrician-gynecologists [35]. In our FGs, information offered by health professionals was highly appreciated. When encountered with contradictive information between apps or between the app and their health professionals, they tended to believe what their doctors said. They expressed the need for information of future apps to be provided by health

professionals. However, doctors may also have different opinions toward the same medical report. As a highly educated group in our study, some women called for the *gold standard*, so they can judge right or wrong depending on the standard information. The information available on apps can empower them to actively engage in monitoring their own health.

In addition, participants in this study favored more interactive apps and recommended communication platforms for both pregnant women and medical staff. Client-to-client communication could meet the demand of users to seek peer support by communicating with women who had similar experiences or health issues. Users can share their experiences and knowledge with others and also gain emotional, social, or practical support from others [36]. However, it is hard to confirm the reliability of the information shared by experienced women. It is also possible that learning from the experiences of others on the Web may lead to greater anxiety or confusion about one's own condition [37], as reflected by our study. Lee showed that although a social networking function was important for pregnant women, interaction with health professionals remained limited [30]. This may be due to women's attitudes toward paying fees for consultations with health professionals.

Furthermore, most study participants in this study wanted pregnancy apps to have a Web-based linkage with hospital information systems so that more personalized and tailored information is provided by the apps in a convenient manner, without the need of women having to enter information such as their medical reports. However, most hospital information systems are closed networks; it will be really challenging to achieve free data transfer between apps and existing hospital electronic systems, given the concerns over data security and privacy issues. A recent survey, conducted in China, based on the app market showed that only a small number of apps had an internet connection with hospital information systems to support making appointments, making payments, obtaining hospital service guidance, or checking laboratory results [12], but without providing personalized information or suggestions. It is also noted that, in our FGs, no one expressed concerns about data privacy and security issues when using pregnancy apps. However, they pointed out that it should be the hospitals' responsibility to ensure data security. We believe that data privacy and security issues are always important in the digital world. Legal frameworks that govern the integrity of health data transfer and storage, in addition to identifying access control and medical liability, are critical to enabling mHealth in China [11].

Finally, it is worth mentioning that a considerable proportion of women in the survey suffered from anemia and abnormal prepregnancy BMI. In addition, more than half of the respondents gained abnormal GWG and over 70% did not reach the goal of physical activity levels recommended by the ACOG guideline [22], indicating poor maternal health status in this study population. This calls for the need for more effective interventions to support maternal and child health in the metropolitan area in China. To the best of our knowledge, this study was the first to investigate women's lifestyles, GWG, and their use of pregnancy apps simultaneously. Technological platforms, such as smartphone apps, have the unique capacity to integrate different functional modules to manage antenatal health. Recent studies conducted in some Western countries have demonstrated the effectiveness of an app-based lifestyle intervention on GWG and behavior changes [38]. However, in this complex and equivocal information environment, women are often provided with invalidated information by various prenatal apps [34]. This may limit the benefits that women would have got if apps containing quality content and evidence-based information were available.

Limitations

There were a few limitations in the study. First, the survey was completed by women from 2 hospitals in Shanghai, thus the findings may not be generalizable to pregnant women in other settings. In addition, prepregnancy BMI and GWG classification were based on self-reported body weight and height, in addition to self-reported anemia. We also acknowledge some potential limitations regarding the validity of the questions included in the survey. Furthermore, only women's views and attitudes were captured regarding their use of apps during pregnancy. Future studies should explore the views of health care professionals or policymakers regarding the feasibility or acceptability of an app for supporting women during pregnancy in the Chinese context.

Conclusions

The study provides insights into how pregnant women used apps for health information and their views and attitudes toward the apps. It suggests that apps were widely used by many pregnant women to monitor fetal development, to get diet and physical activity information, and to track their body changes. Women highly appreciated evidence-based information, expert opinions, and tailored advice available on the apps. With an increasing number of women using apps during pregnancy, development of quality apps may have the potential to improve prenatal outcomes by facilitating access to health information, reducing demands for face-to-face health services, and enabling the provision of personalized care. Efforts should be made by health professionals, app developers, and policymakers to ensure quality apps be developed for providing timely health promotion information. The apps can be integrated into maternal care to meet the needs of pregnant women, although it is extremely important to ensure privacy and security of individuals' data.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant No. 81773413) and the Nursing School of Fudan University (Grant No. FNF201717). The authors would like to thank all the women who participated in this study and pay a tribute to the 2 hospitals for supporting this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions used in the cross-sectional survey.

[PDF File (Adobe PDF File), 81KB - mhealth_v7i6e12631_app1.pdf]

Multimedia Appendix 2

Braun and Clarke's 6-phase framework for doing a thematic analysis.

[PDF File (Adobe PDF File), 48KB - mhealth_v7i6e12631_app2.pdf]

Multimedia Appendix 3

Codes, subthemes, and themes derived from the focus groups.

[PDF File (Adobe PDF File), 41KB - mhealth_v7i6e12631_app3.pdf]

References

1. Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? Physiol Rev 2014 Oct;94(4):1027-1076 [FREE Full text] [doi: 10.1152/physrev.00029.2013] [Medline: 25287859]
2. Yang IV, Zhang W, Davidson EJ, Fingerlin TE, Kechris K, Dabelea D. Epigenetic marks of in utero exposure to gestational diabetes and childhood adiposity outcomes: the EPOCH study. Diabet Med 2018;35(5):612-620. [doi: 10.1111/dme.13604] [Medline: 29461653]
3. Yajnik CS, Deshmukh US. Maternal nutrition, intrauterine programming and consequential risks in the offspring. Rev Endocr Metab Disord 2008 Sep;9(3):203-211. [doi: 10.1007/s11154-008-9087-z] [Medline: 18661241]
4. APrON and ENRICH Study Teams. Adherence to Canada's food guide recommendations during pregnancy: nutritional epidemiology and public health. Curr Dev Nutr 2017;1(7):e000356 [FREE Full text] [doi: 10.3945/cdn.116.000356] [Medline: 29955709]
5. Malek L, Umberger W, Makrides M, Zhou SJ. Adherence to the Australian dietary guidelines during pregnancy: evidence from a national study. Public Health Nutr 2016 May;19(7):1155-1163. [doi: 10.1017/S1368980015002232] [Medline: 26228526]
6. Liu FL, Zhang YM, Parés GV, Reidy KC, Zhao WZ, Zhao A, et al. Nutrient intakes of pregnant women and their associated factors in eight cities of China: a cross-sectional study. Chin Med J (Engl) 2015 Jul 05;128(13):1778-1786 [FREE Full text] [doi: 10.4103/0366-6999.159354] [Medline: 26112720]
7. Özkan Şat S, Yaman Sözbir S. Use of mobile applications and blogs by pregnant women in Turkey and the impact on adaptation to pregnancy. Midwifery 2018 Jul;62:273-277. [doi: 10.1016/j.midw.2018.04.001] [Medline: 29738988]
8. Lee A, Newton M, Radcliffe J, Belski R. Pregnancy nutrition knowledge and experiences of pregnant women and antenatal care clinicians: a mixed methods approach. Women Birth 2018 Aug;31(4):269-277. [doi: 10.1016/j.wombi.2017.10.010] [Medline: 29126796]
9. Whitaker KM, Wilcox S, Liu J, Blair SN, Pate RR. Provider advice and women's intentions to meet weight gain, physical activity, and nutrition guidelines during pregnancy. Matern Child Health J 2016;20(11):2309-2317 [FREE Full text] [doi: 10.1007/s10995-016-2054-5] [Medline: 27400917]
10. Raupach JC, Hiller JE. Information and support for women following the primary treatment of breast cancer. Health Expect 2002 Dec;5(4):289-301 [FREE Full text] [doi: 10.1046/j.1369-6513.2002.00191.x] [Medline: 12460218]
11. Ryu S. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Geneva: World Health Organization; 2012:231-233.
12. 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 08;6(6):e140 [FREE Full text] [doi: 10.2196/mhealth.9302] [Medline: 29884601]
13. Lupton D. The use and value of digital media for information about pregnancy and early motherhood: a focus group study. BMC Pregnancy Childbirth 2016 Jul 19;16(1):171 [FREE Full text] [doi: 10.1186/s12884-016-0971-3] [Medline: 27435182]
14. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
15. Cui H, Zheng Y, Shi J, Yang N, Wang Y, Li Q. Investigation on the status of maternal use of maternity APP [Chinese]. Journal of Chengde Medical College 2018;35(5):446-448. [doi: 10.15921/j.cnki.cyxb.2018.05.036]
16. Hearn L, Miller M, Lester L. Reaching perinatal women online: the healthy you, healthy baby website and app. J Obes 2014;2014:1-9 [FREE Full text] [doi: 10.1155/2014/573928] [Medline: 24872891]

17. Okorodudu DE, Bosworth HB, Corsino L. Innovative interventions to promote behavioral change in overweight or obese individuals: a review of the literature. Ann Med 2015 May;47(3):179-185 [FREE Full text] [doi: 10.3109/07853890.2014.931102] [Medline: 25011006]
18. Mackillop L, Hirst JE, Bartlett KJ, Birks JS, Clifton L, Farmer AJ, et al. Comparing the efficacy of a mobile phone-based blood glucose management system with standard clinic care in women with gestational diabetes: randomized controlled trial. JMIR Mhealth Uhealth 2018 Mar 20;6(3):e71 [FREE Full text] [doi: 10.2196/mhealth.9512] [Medline: 29559428]
19. Ledford CJ, Canzona MR, Cafferty LA, Hodge JA. Mobile application as a prenatal education and engagement tool: a randomized controlled pilot. Patient Educ Couns 2016 Apr;99(4):578-582. [doi: 10.1016/j.pec.2015.11.006] [Medline: 26610389]
20. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
21. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
22. ACOG Committee Obstetric Practice. ACOG committee opinion. Number 267, January 2002: exercise during pregnancy and the postpartum period. Obstet Gynecol 2002 Jan;99(1):171-173. [Medline: 11777528]
23. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. Women Birth 2016 Aug;29(4):368-375. [doi: 10.1016/j.wombi.2016.01.008] [Medline: 26874938]
24. McLafferty I. Focus group interviews as a data collecting strategy. J Adv Nurs 2004 Oct;48(2):187-194. [doi: 10.1111/j.1365-2648.2004.03186.x] [Medline: 15369499]
25. Zhou B, Coorperative Meta-Analysis Group Of China Obesity Task Force. [Predictive values of body mass index and waist circumference to risk factors of related diseases in Chinese adult population]. Zhonghua Liu Xing Bing Xue Za Zhi 2002 Feb;23(1):5-10. [Medline: 12015100]
26. Institute of Medicine, National Research Council. Weight Gain During Pregnancy: Reexamining the Guidelines. Washington (DC): National Academies Press (US); 2009.
27. Maguire M, Delahunt B. Doing a thematic analysis: a practical, step-by-step guide for learning and teaching scholars. AISHE-J 2017;9(3).
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
29. Wallwiener S, Müller M, Doster A, Laserer W, Reck C, Pauluschke-Fröhlich J, et al. Pregnancy eHealth and mHealth: user proportions and characteristics of pregnant women using web-based information sources-a cross-sectional study. Arch Gynecol Obstet 2016;294(5):937-944. [doi: 10.1007/s00404-016-4093-y] [Medline: 27084763]
30. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. Healthc Inform Res 2016 Apr;22(2):73-80 [FREE Full text] [doi: 10.4258/hir.2016.22.2.73] [Medline: 27200216]
31. O'Higgins A, Murphy OC, Egan A, Mullaney L, Sheehan S, Turner MJ. The use of digital media by women using the maternity services in a developed country. Ir Med J 2014;107(10):313-315. [Medline: 25556255]
32. Aksoy Derya Y, Timur Taşhan S, Duman M, Durgun Ozan Y. Turkish adaptation of the pregnancy-related anxiety questionnaire-revised 2: validity and reliability study in multiparous and primiparous pregnancy. Midwifery 2018 Jul;62:61-68. [doi: 10.1016/j.midw.2018.03.006] [Medline: 29655006]
33. Kratzke C, Cox C. Smartphone technology and apps: rapidly changing health promotion. Int Electron J Health Educ 2012;15:72-82 [FREE Full text]
34. Womack JJ, Anderson LN, Ledford CJ. Presence of complex and potentially conflicting information in prenatal mobile apps. Health Promot Pract 2018 Epub ahead of print(forthcoming). [doi: 10.1177/1524839918796216] [Medline: 30146903]
35. Kaimal AJ, Cheng YW, Bryant AS, Norton ME, Shaffer BL, Caughey AB. Google obstetrics: who is educating our patients? Am J Obstet Gynecol 2008 Jun;198(6):682.e1-682.e5. [doi: 10.1016/j.ajog.2008.03.030] [Medline: 18538152]
36. 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] [Medline: 27956659]
37. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
38. Kennelly MA, Ainscough K, Lindsay KL, O Sullivan E, Gibney ER, McCarthy M, et al. Pregnancy exercise and nutrition with smartphone application support: a randomized controlled trial. Obstet Gynecol 2018 May;131(5):818-826. [doi: 10.1097/AOG.0000000000002582] [Medline: 29630009]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
BMI: body mass index
FG: focus group
GWG: gestational weight gain
IOM: Institute of Medicine
IQR: interquartile range
mHealth: mobile health
OR: odds ratio

Edited by G Eysenbach; submitted 29.10.18; peer-reviewed by K Reuter, K Hameen-Anttila; comments to author 31.03.19; revised version received 17.05.19; accepted 20.05.19; published 18.06.19.

Please cite as:
Wang N, Deng Z, Wen LM, Ding Y, He G
Understanding the Use of Smartphone Apps for Health Information Among Pregnant Chinese Women: Mixed Methods Study
JMIR Mhealth Uhealth 2019;7(6):e12631
URL: http://mhealth.jmir.org/2019/6/e12631/
doi:10.2196/12631
PMID:31215516

Original Paper

Current Knowledge and Adoption of Mobile Health Apps Among Australian General Practitioners: Survey Study

Oyungerel Byambasuren[1], MD, MMedRes; Elaine Beller[1], MAppStat; Paul Glasziou[1], MD, PhD

Centre for Research in Evidence-Based Practice, Bond University, Robina, QLD, Australia

Corresponding Author:
Oyungerel Byambasuren, MD, MMedRes
Centre for Research in Evidence-Based Practice
Bond University
14 University Dr
Robina, QLD, 4226
Australia
Phone: 61 755951588
Email: obyambas@bond.edu.au

Abstract

Background: Mobile health (mHealth) apps can be prescribed as an effective self-management tool for patients. However, it is challenging for doctors to navigate 350,000 mHealth apps to find the right ones to recommend. Although medical professionals from many countries are using mHealth apps to varying degrees, current mHealth app use by Australian general practitioners (GPs) and the barriers and facilitators they encounter when integrating mHealth apps in their clinical practice have not been reported comprehensively.

Objective: The objectives of this study were to (1) evaluate current knowledge and use of mHealth apps by GPs in Australia, (2) determine the barriers and facilitators to their use of mHealth apps in consultations, and (3) explore potential solutions to the barriers.

Methods: We helped the Royal Australian College of General Practitioners (RACGP) to expand the mHealth section of their annual technology survey for 2017 based on the findings of our semistructured interviews with GPs to further explore barriers to using mHealth apps in clinical practice. The survey was distributed to the RACGP members nationwide between October 26 and December 3, 2017 using Qualtrics Web-based survey tool.

Results: A total of 1014 RACGP members responded (response rate 4.6% [1014/21,884], completion rate 61.2% [621/1014]). The median years practiced was 20.7 years. Two-thirds of the GPs used apps professionally in the forms of medical calculators and point-of-care references. A little over half of the GPs recommended apps for patients either daily (12.9%, 80/621), weekly (25.9%, 161/621), or monthly (13.4%, 83/621). Mindfulness and mental health apps were recommended most often (32.5%, 337/1036), followed by diet and nutrition (13.9%, 144/1036), exercise and fitness (12.7%, 132/1036), and women's health (10%, 104/1036) related apps. Knowledge and usage of evidence-based apps from the Handbook of Non-Drug Interventions were low. The prevailing barriers to app prescription were the lack of knowledge of effective apps (59.9%, 372/621) and the lack of trustworthy source to access them (15.5%, 96/621). GPs expressed their need for a list of safe and effective apps from a trustworthy source, such as the RACGP, to overcome these barriers. They reported a preference for online video training material or webinar to learn more about mHealth apps.

Conclusions: Most GPs are using apps professionally but recommending apps to patients sparingly. The main barriers to app prescription were the lack of knowledge of effective apps and the lack of trustworthy source to access them. A curated compilation of effective mHealth apps or an app library specifically aimed at GPs and health professionals would help solve both barriers.

(JMIR Mhealth Uhealth 2019;7(6):e13199) doi:10.2196/13199

KEYWORDS

mobile health; smartphone; mobile apps; mHealth; smartphone apps; general practice

Introduction

Over the past decade, smartphones have become an inseparable part of modern living, and mobile health (mHealth) apps have started to establish their place in health care [1]. If proven to help achieve measurable clinical improvements in patients' conditions, mHealth apps can be officially prescribed or recommended by general practitioners (GPs) [2]. However, with 350,000 apps available in the medical and health and fitness categories in the major app stores, it is challenging for GPs to find prescription quality mHealth apps from the app stores themselves to use in their clinical practice [3]. To overcome this issue, a number of initiatives such as the National Health Service (NHS) Apps Library in the United Kingdom [4] and Health Navigator app library in New Zealand [5] have been set up to help doctors. In Australia, official effort to support app prescription does not exist yet, but there are small initiatives such as the Victorian Health Promotion Foundation's Healthy Living Apps Guide aimed at the general public [6] and the Handbook of Non-Drug Interventions (HANDI) project by the Royal Australian College of General Practitioners (RACGP) [7], which is a repository of evidence-based nonpharmaceutical interventions for GPs.

Health care professionals' use of mHealth apps and mobile technologies have been explored in several recent reports from the United States [8], United Kingdom [9], France [10], and Turkey [11]. At least half of the surveyed GPs, specialists, dieticians, and pharmacists reported to recommend mHealth apps to patients, except for the French study of GPs, which reported half that rate. Barriers perceived by the health professionals regarding mHealth integration to their clinical practice include a variety of issues from infrastructure related problems such as Wi-Fi coverage and interoperability with the existing medical software to more specific data security, content reliability, and a universal lack of awareness of the effective apps to use. All of this echoes the findings of an earlier systematic review by Gagnon et al [12], which summarized the barriers and facilitators to mHealth adoption by health care professionals from around the globe.

Australian GPs' mHealth adoption has been explored only briefly as part of the annual technology survey by the RACGP since 2015. The purpose of this survey is to explore technological innovation and adoption in general practice, including mobile technology [13]. Following our qualitative study with GPs on the barriers and facilitators of mHealth app use in general practice, we collaborated with the RACGP to expand the mHealth section for 2017 to better understand the specific barriers to health app use in the wider Australian context. Thus, the objectives for this study were to (1) explore the knowledge and use of health apps of practicing GPs in Australia in more detail, (2) determine the barriers and facilitators to prescribing health apps in GP consultations in a wider cohort of GPs, and (3) explore the potential solutions to some of the barriers.

Methods

We used the Checklist for Reporting Results of Internet E-Surveys, recommended by the Journal of Medical Internet Research as a reporting guide for this study [14]. The data for this study were collected as part of the 2017 RACGP Technology Survey, which was conducted using a Web-based survey tool Qualtrics (Qualtrics, Provo, UT) between October 26 and December 3, 2017 in Australia [15]. We used convenience sampling and the survey link was emailed to all RACGP members, which included GP trainees, fellows, and vocationally registered GPs, as well as practice managers and clinic owners. Only the GPs currently practicing in Australia were able to advance and answer all questions. For GP registrars and GPs who were not practicing currently or not practicing in Australia, the survey ended after the relevant questions. Ethics approval was obtained from the RACGP National Research and Evaluation Ethics Committee.

The previous year's survey contained 6 questions regarding GPs' use of mobile devices and mHealth apps out of 46 questions [13]. On the basis of the findings of our semistructured interview study with 20 GPs who explored the barriers and facilitators to using mHealth apps in practice, we collaborated with the RACGP to expand the mHealth section questions for the 2017 survey making them more specific and informative. The questions were pilot-tested with the coauthors and academic GP colleagues and refined iteratively. This paper reports the analysis of 16 questions pertaining to demographic information, mobile device, and health apps usage out of the total 50 questions (Textbox 1).

Data were extracted from Qualtrics and descriptive statistics were conducted using Microsoft Excel (2016). Answers to the open-ended questions were coded according to their common themes (OB), checked by a second author (EB), and then summarized. Participation in the survey was voluntary and participants who completed the survey were invited to enter a draw to win one of 2 gift cards worth Aus $50.

Textbox 1. Survey questions analyzed as part of this study.

Screening and demographics questions

1. I am... (a general practitioner, a general practitioner registrar, other)

2. What is your role in the practice?

3. Where did you complete your training? (Australia, Overseas)

4. Do you currently practice in Australia? (Yes/No)

5. For how many years have you worked as a general practitioner? (1-5 years, 5-10 years, 10-20 years, 20-30 years, more than 30 years)

6. Please enter the postcode of your current practice location in which you spend the most time.

7. What is your age group? (Less than 35 years, 35-44 years, 45-54 years, 55-64 years, over 65 years)

Mobile health section

8. Do you use mobile devices in your day-to-day practice for patient-related work? (Yes/No)

9. I don't use mobile devices for patient-related work because... (choose all that apply)

- I am not confident in how to safely use mobile technology in my day-to-day practice
- I don't see how mobile technology can benefit my day-to-day practice
- My practice does not allow me to use my own mobile devices
- Other (please comment)

10. Which health apps do you use for yourself?

11. How often do you recommend the use of health apps to patients? (Daily, Weekly, Monthly, Rarely, Never)

12. Please share with us which health apps you have recommended.

13. Do you ever prescribe any of the following health apps to your patients (choose all that apply): (Quit Now: My QuitBuddy, Quit for You–Quit for Two, Sleepio, CBT-i Coach, SHUTi, Ankle, I do not prescribe the apps above)

14. Please rate the following barriers for health app integration into your daily clinical practice (where 1 is the most important barrier and 7 is the least important):

- lack of knowledge of effective apps
- lack of a trustworthy source to access effective apps
- lack of patient digital literacy
- lack of access to mobile devices
- lack of patient interest
- lack of practice incentives
- lack of understanding about benefits
- others (please specify)

15. What would help you to recommend health apps to patients more often?

16. How would you like to receive training on the use of effective health apps, including app evaluation? (eg, webinars, animations, podcasts)

Results

Of the 39,380 people on the RACGP mailing list, who were emailed the survey link, 21,884 were currently practicing GPs and 1014 of them responded to the survey (4.6%). The survey completion rate was 61.2% (621/1014). The median age was 51.4 years and the median years practiced was 20.7 years. Age and geographical distribution by state were representative of Australian GP workforce statistics [16]. About a quarter of the survey responders were trained overseas, which was half the rate of the national statistics (Table 1).

A half of the GPs reported not using mobile devices for patient related work (n=312; 50%). The main reasons for this included GPs not seeing how mobile technology could benefit their day-to-day practice (n=136; 39%), not being confident on how to use mobile technology safely in daily practice (n=68; 20%), and the practice they worked for did not allow the use of personal mobile devices in practice (n=51; 15%). Other reasons (n=91; 26%) included not wanting to use personal mobile devices in consultations, their desktop computer being sufficient and more convenient, and not needing to use mobile devices altogether. About two-thirds of the GPs used health apps themselves (n=440; 64%), mostly in the form of point-of-care references such as UpToDate, eTG, Medscape (n=298; 25%), and medical calculators (n=137; 11%).

Table 1. Participants' demographics.

Groups	This study, n (%)	National data (n=25,825), n (%)
Age (years; n=621)		
<35	46 (7.4)	2376 (9.2)
35-44	126 (20.3)	6361 (24.6)
45-54	196 (31.6)	7327 (28.4)
55-64	174 (28.0)	6637 (25.7)
65+	79 (12.7)	3124 (12.1)
Practice (years; n=621)		
<5	82 (13)	—[a]
5-10	81 (13)	—
10-20	136 (22)	—
20-30	166 (27)	—
>30	156 (25)	—
General physician training (n=621)		
Overseas	144 (23.2)	13207 (51.1)
Australia	477 (76.8)	12621 (48.9)
Geographic distribution (n=844)		
New South Wales	235 (27.8)	8468 (32.8)
Victoria	215 (25.5)	6506 (25.2)
Queensland	174 (20.6)	5525 (21.4)
Western Australia	78 (9.2)	2411 (9.3)
South Australia	70 (8.3)	1873 (7.3)
Tasmania	37 (4.4)	510 (2.0)
Northern Territory	21 (2.5)	212 (0.8)
Australian Capital Territory	14 (1.7)	320 (1.2)

[a]Not applicable.

A little over half of the GPs recommended apps for patients daily (n=80; 13%), weekly (n=161; 26%), or monthly (n=83; 13%). The other half rarely (n=210; 34%) or never (n=87; 14%) recommended apps. Figure 1 shows that the app recommendation frequency appears to decrease with the number of years practiced as a GP. GPs most commonly recommended mindfulness and mental health (n=337; 33%), diet and nutrition (n=144; 14%), exercise and fitness (n=132; 13%), and women's health (n=104; 10%) related apps to patients. Examples of the specific apps they use included Smiling Mind, Headspace, MyFitnessPal, and Easy Diet Diary.

The question about evidence-based apps from the HANDI project revealed that smoking cessation apps were reported as prescribed 119 times, insomnia apps 39 times, and the ankle exercise app 7 times. However, majority of the GPs did not recommend any of the 6 apps that are currently offered in HANDI (n=417; 72%).

GPs also rated the barriers to integrating health apps into their daily practice. The prevailing barriers were the lack of knowledge of effective apps (n=372; 60%) and the lack of trustworthy source to access them (n=96; 15%). Figure 2 shows the ranking of barriers rated by the combination of the first and second most important barriers. Most of the responders (n=555; 89%) also added their own barriers as other option. Among the additional barriers, consultation time constraint (n=24; 28%) and uncertain benefits of and interests in health apps (n=19; 21%) were the leading reasons as to why the use of apps in daily clinical practice would be a challenge, whereas cost (n=3; 3%) was not rated as a major barrier.

Figure 1. Frequency of app recommendation by years practiced.

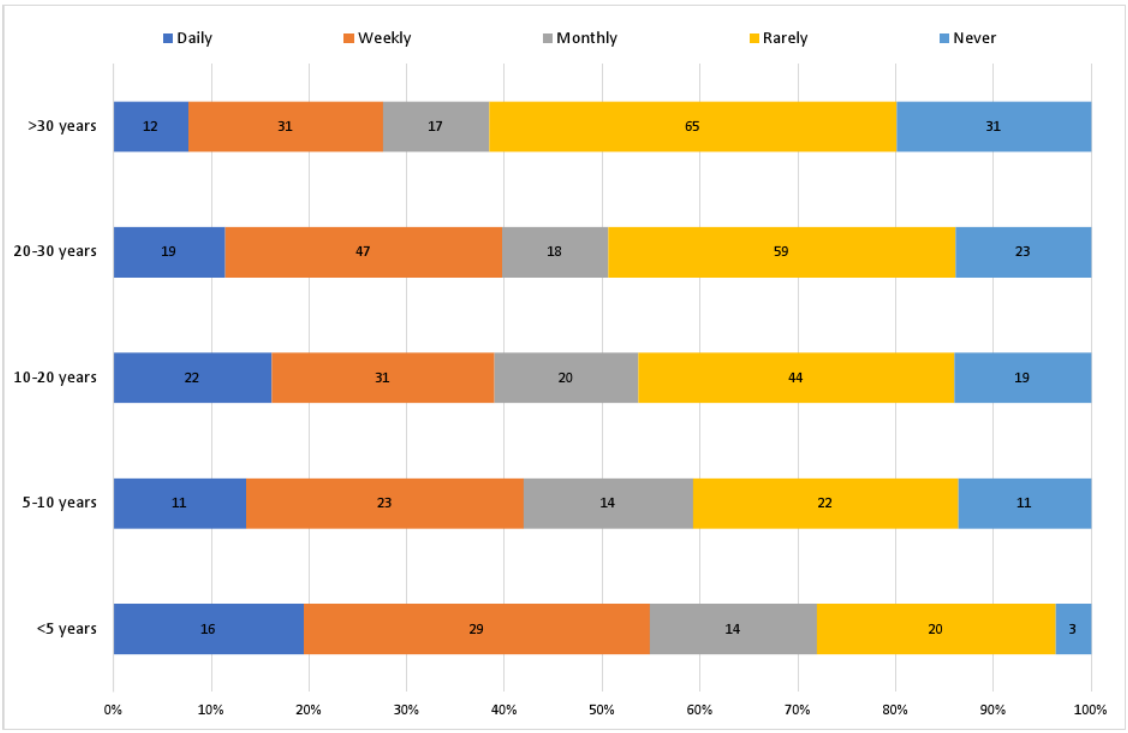
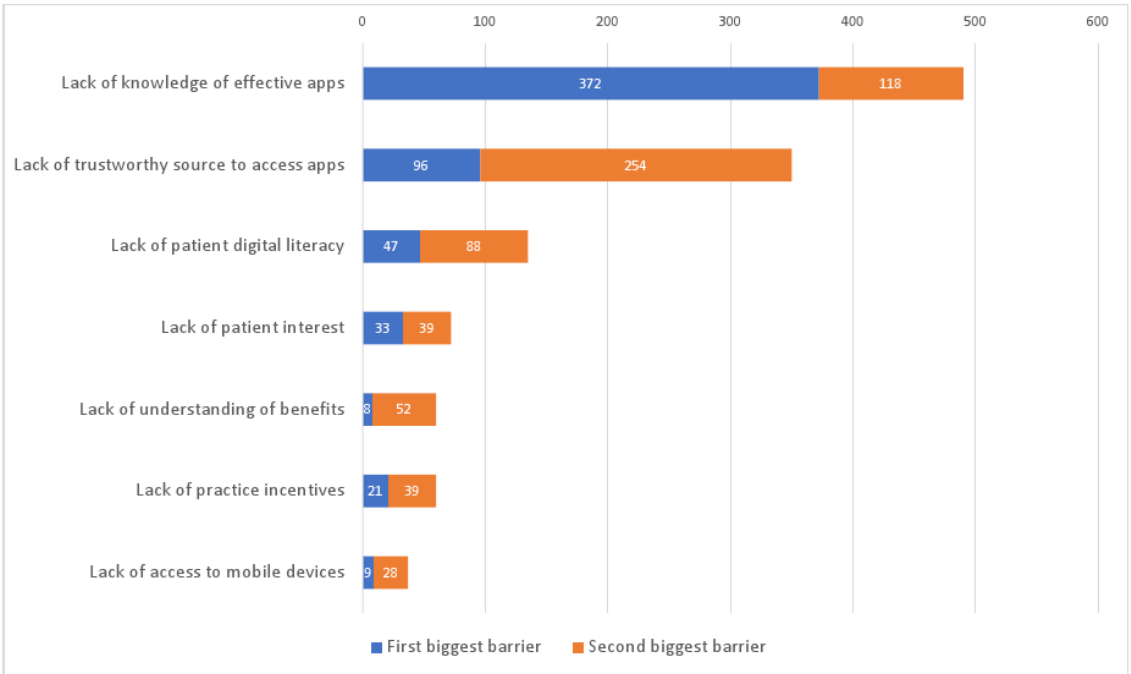


Figure 2. Barriers to app prescription. Numbers inside the bars show how many times the option was rated as the first and the second biggest barrier.



When asked what would help them to recommend health apps to patients more often, the top answers were more knowledge, awareness, and training on health apps (n=243; 30%), a list of safe and effective apps provided or endorsed by reliable authority such as the RACGP (n=224; 28%), and quality, benefits, and evidence of apps (n=92; 12%). Time, integration into electronic medical software, access to better internet and mobile services, patients' motivation, handouts, and practice incentives were all rated as of the least importance (less than 25 each, 1%-5%; Table 2).

The preferred ways to receive training on effective apps and app evaluation were permanent online video training material or webinars that the doctors could watch on their own time and pace (n=303; 33%). Other choices included podcasts (n=137; 15%), animation (n=101; 11%), face-to-face training (n=114; 12%), and other reading materials such as newsletters and articles (n=90; 10%).

Table 2. Facilitators of app prescription as responses to question "What would help you to recommend health apps to patients more often?" A total of 683 replies were coded into 800 answers. All other themes (n=9) had less than 25 answers each to support them.

Theme	Example comments
More awareness or knowledge or training (n=243)	"Educate us before we recommend to patients."; "Myself getting more familiar about it and for me to learn what it is all about."
List of approved or vetted apps (n=224)	"A clear directive, guideline about which are validated safe and useful"; "recommendation from respected advisors, for example, the Royal Australian College of GPs, National Prescribing Service"
Evidence or benefits or quality (n=92)	"Evidence for its use—case in the field."
Nothing (n=44)	"Nothing. I recommend to my patients that they get away from screens and go and do some exercise, appreciate nature, and breathe some fresh air and RACGP should do the same instead of apparently trying to encourage everyone, doctors, and patients to increase their screen time."
More time (n=39)	"More time during a consult to discuss benefits...however it starts to make us a 'Telstra[a] shop' and not a GP practice."
Practice incentive (n=29)	"Incentives—costs involved in recommending the apps—the ones I do recommend I research myself and spent time and money to do so."
Integration with practice software (n=26)	"If the health information software used could pick up on coded diagnoses in patient's clinical information system and recommend a trustworthy app for the relevant medical conditions, it would be most beneficial."

[a]Telstra is an Australian telecommunications company.

Discussion

This study provides insights into the current adoption of smartphone health apps by Australian GPs. We found that two-thirds of the GPs use apps professionally, and at least half are recommending apps to patients. Mindfulness and mental health apps were most commonly suggested. Majority of the GPs were not aware of, and thus not using, evidence-based health apps that are included in the free RACGP resource HANDI. The biggest barriers to app prescription were the lack of knowledge of effective apps and the lack of trustworthy source to access them. To overcome these barriers, GPs expressed their need for a list of safe and effective apps from a trustworthy source and more training on health apps in the form of permanent online video training material or webinar that they could watch in their own time and pace.

The 2017 RACGP Technology Survey results were similar to the preceding year's in terms of smartphone health app usage, most commonly used and recommended apps, and barriers to technology adoption, thus validating the trends [13]. Studies from several other countries reported that anywhere between 20% to 75% of health professionals use mHealth apps for their patients [8-11], which is comparable with the 50% use by Australian GPs. The main barriers we identified were also reflective of the barriers health professionals face around the world. The surveys indicate that health professionals recognize the potential benefit of smartphone health apps for self-management of health conditions and would like to use them in their work. However, they lack the knowledge, time, and skills to evaluate, find and recommend evidence-based apps, and therefore, they need help and guidance from the professional organizations and policy makers to overcome these barriers.

It is important that solutions to the barriers are unique to each country's health care structures and health care professionals' demands. For example, New Zealand's Health Navigator website hosts a health app library set up by a GP organization and health care providers who use the apps can curate and provide feedback [5]. United Kingdom's NHS not only provides Web-based apps library for doctors, but also introduced an app prescription platform for the GPs [17]. Similarly, for Australia, there is an opportunity for a professional organization such as the RACGP to lead the way to address the major barriers identified in this study. Although the inclusion of mHealth apps in the HANDI project is a starting place, more work needs to be done to raise awareness and profile of this initiative. Furthermore, integration of approved apps into the electronic medical systems to streamline the usability, as well as provide continuing professional development trainings for up-to-date information on mHealth apps, would enhance the use of evidence-based apps in clinical practice.

The strengths of this study include expanding on and improving previous year's mHealth questions with more specific questions regarding evidence-based app adoption and barriers in general practice based on our qualitative research on GPs to obtain more comprehensive data on a nationally representative sample. Our response rate of 4.6% was similar to that of other mHealth app surveys undertaken on health professionals [9,10]. The completion rate of the mHealth section was uniformly high, although skipping questions was allowed.

The limitations of our survey study are the selection bias inherent in survey studies and the low response rate. However, the median age, median years practiced, and geographical distribution data of the GPs in our study were comparable with those of the national GP workforce data [16], thus supporting the demographic representativeness of our study population. A reason for the low response rate could be survey fatigue because of the fact that the mHealth questions analyzed here were a part of a larger survey on technological innovation in general practice [18]. The challenge for conducting survey studies on medical professionals is a balancing act between conducting a dedicated

survey only focusing on single topics and the burdening of GPs with yet another survey. To increase the response rate of surveys that involve medical professionals, certain strategies could be undertaken such as offering more attractive incentives to participate and randomly sampling the cohort to send surveys and other study offers sparingly.

mHealth apps have a unique niche in the future of health care. However, the evidence of their effectiveness, safety, and usability issues are challenged by both the fast-evolving nature of the software and commercial aspects of the technology that can be easily exploited. Health care professionals need guidance on the quality of mHealth apps to assist in their adoption into clinical practice. In the absence of notable initiatives from government or private sectors to regulate app quality and safety, professional organizations must take the lead to address this challenge.

Acknowledgments

The authors wish to thank Gisele Rocha, RACGP project manager, for leading the development and facilitating modifications to the 2017 RACGP Technology Survey. OB is supported by the Australian Government Research Training Program scholarship.

Conflicts of Interest

None declared.

References

1. Aitken M, Clancy B, Nass D. Respiratory Effectiveness Group (REG). Parsippany, NJ, USA: IQVIA Institute for Human Data Science; 2017 Nov. The growing value of digital health: Evidence and Impact on Human Health and the Healthcare System URL: https://regresearchnetwork.org/wp-content/uploads/the-growing-value-of-digital-health.pdf [accessed 2017-12-01] [WebCite Cache ID 78gaiGMvS]
2. Byambasuren O, Sanders S, Bellers E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digit Med 2018 May 9;1(1):12. [doi: 10.1038/s41746-018-0021-9]
3. Research2Guidance. 2017 Nov. mHealth App Economics: Current statuts and future trends in mobile health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-05-27] [WebCite Cache ID 78gbXOeyA]
4. National Health Service. 2018. Apps library URL: https://www.nhs.uk/apps-library/ [accessed 2019-05-27] [WebCite Cache ID 78gblZr3h]
5. Health Navigator New Zealand. Health Navigator App library URL: https://www.healthnavigator.org.nz/apps-videos/ [accessed 2019-05-27] [WebCite Cache ID 78gbv03US]
6. The Victorian Health Promotion Foundation. 2018. Healthy living apps guide URL: https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps [accessed 2017-05-31]
7. The Royal Australian College of General Practitioners. Handbook of Non-Drug Interventions URL: https://www.racgp.org.au/your-practice/guidelines/handi/about/the-handi-project/ [accessed 2017-05-31]
8. American Medical Association. 2017. Integration of Mobile Health Applications and Devices into Practice URL: https://policysearch.ama-assn.org/policyfinder/detail/mobile%20health?uri=%2FAMADoc%2FHOD-480.943..xml [accessed 2019-05-28] [WebCite Cache ID 78gcpEHe9]
9. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J Hum Nutr Diet 2017 Dec;30(4):439-452. [doi: 10.1111/jhn.12446] [Medline: 28116773]
10. el Amrani L, Engberink AO, Ninot G, Hayot M, Carbonnel F. Connected health devices for health care in French general medicine practice: cross-sectional study. JMIR Mhealth Uhealth 2017 Dec 21;5(12):e193 [FREE Full text] [doi: 10.2196/mhealth.7427] [Medline: 29269336]
11. Sezgin E, Özkan-Yildirim S, Yildirim S. Investigation of physicians' awareness and use of mHealth apps: a mixed method study. Health Pol Technol 2017 Jul;6(3):251-267. [doi: 10.1016/j.hlpt.2017.07.007]
12. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
13. The Royal Australian College of General Practitioners. 2017. Views and attitudes towards technological innovation in general practice: Survey report 2016 URL: https://www.racgp.org.au/FSDEDEV/media/documents/Running%20a%20practice/Technology/Member%20engagement/RACGP-Technology-Survey-2016-Results.pdf [accessed 2019-05-28] [WebCite Cache ID 78geE0nlV]
14. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
15. The Royal Australian College of General Practitioners. 2018. Views and attitudes towards technological innovation in general practice: Survey report 2017 URL: https://www.racgp.org.au/FSDEDEV/media/documents/Running%20a%20practice/

Technology/Member%20engagement/RACGP-Technology-Survey-2017-Results.pdf [accessed 2019-05-28] [WebCite Cache ID 78geblS8v]
16. The Department of Health. 2017. GP Workforce Statistics 2001-02 to 2016-17 URL: https://www.health.gov.au/internet/main/publishing.nsf/Content/General+Practice+Statistics-1 [accessed 2019-05-27] [WebCite Cache ID 78getBWJB]
17. Building Better Healthcare. 2017 Dec 05. IQVIA Launches AppScript in the UK URL: https://www.buildingbetterhealthcare.co.uk/news/article_page/IQVIA_Launches_AppScript_in_the_UK/136934 [accessed 2019-05-27] [WebCite Cache ID 78gezGlRF]
18. Lavrakas P, editor. Encyclopedia of Survey Research Methods. Thousand Oaks, California: Sage Knowledge; 2008.

Abbreviations

GP: general practitioner
HANDI: Handbook of Non-Drug Interventions
mHealth: mobile health
RACGP: Royal Australian College of General Practitioners

Edited by G Eysenbach; submitted 19.12.18; peer-reviewed by E Sezgin, J Lander, A Paglialonga, L Lyzwinski; comments to author 14.02.19; revised version received 30.04.19; accepted 11.05.19; published 03.06.19.

Please cite as:
Byambasuren O, Beller E, Glasziou P
Current Knowledge and Adoption of Mobile Health Apps Among Australian General Practitioners: Survey Study
JMIR Mhealth Uhealth 2019;7(6):e13199
URL: https://mhealth.jmir.org/2019/6/e13199/
doi:10.2196/13199
PMID:31199343

Original Paper

Characteristics of Adopters of an Online Social Networking Physical Activity Mobile Phone App: Cluster Analysis

Ilea Sanders[1], MRES; Camille E Short[2], PhD; Svetlana Bogomolova[3], PhD; Tyman Stanford[4], PhD; Ronald Plotnikoff[5], PhD; Corneel Vandelanotte[6], PhD; Tim Olds[1], PhD; Sarah Edney[1], BA; Jillian Ryan[1], PhD; Rachel G Curtis[1], PhD; Carol Maher[1], PhD

[1]Alliance for Research in Exercise, Nutrition and Activity, School of Health Sciences, University of South Australia, Adelaide, Australia

[2]Freemasons Foundation Centre for Men's Health, Faculty of Health Sciences, The University of Adelaide, Adelaide, Australia

[3]Ehrenberg-Bass Institute, School of Marketing, University of South Australia, Adelaide, Australia

[4]School of Mathematical Sciences, The University of Adelaide, Adelaide, Australia

[5]Priority Research Centre for Physical Activity and Nutrition, University of Newcastle, Newcastle, Australia

[6]Physical Activity Research Group, Central Queensland University, Adelaide, Australia

Corresponding Author:
Carol Maher, PhD
Alliance for Research in Exercise, Nutrition and Activity
School of Health Sciences
University of South Australia
GPO Box 2471
Adelaide,
Australia
Phone: 61 8 8302 2315
Email: carol.maher@unisa.edu.au

Abstract

Background: To date, many online health behavior programs developed by researchers have not been translated at scale. To inform translational efforts, health researchers must work with marketing experts to design cost-effective marketing campaigns. It is important to understand the characteristics of end users of a given health promotion program and identify key market segments.

Objective: This study aimed to describe the characteristics of the adopters of Active Team, a gamified online social networking physical activity app, and identify potential market segments to inform future research translation efforts.

Methods: Participants (N=545) were Australian adults aged 18 to 65 years who responded to general advertisements to join a randomized controlled trial (RCT) evaluating the Active Team app. At baseline they provided demographic (age, sex, education, marital status, body mass index, location of residence, and country of birth), behavioral (sleep, assessed by the Pittsburgh Quality Sleep Index) and physical activity (assessed by the Active Australia Survey), psychographic information (health and well-being, assessed by the PERMA [Positive Emotion, Engagement, Relationships, Meaning, Achievement] Profile; depression, anxiety and stress, assessed by the Depression, Anxiety, and Stress Scale [DASS-21]; and quality of life, assessed by the 12-Item Short Form Health Survey [SF-12]). Descriptive analyses and a k-medoids cluster analysis were performed using the software R 3.3.0 (The R Foundation) to identify key characteristics of the sample.

Results: Cluster analyses revealed four clusters: (1) younger inactive women with poor well-being (218/545), characterized by a higher score on the DASS-21, low mental component summary score on the SF-12, and relatively young age; (2) older, active women (153/545), characterized by a lower score on DASS-21, a higher overall score on the SF-12, and relatively older age; (3) young, active but stressed men (58/545) with a higher score on DASS-21 and higher activity levels; and (4) older, low active and obese men (30/545), characterized by a high body mass index and lower activity levels.

Conclusions: Understanding the characteristics of population segments attracted to a health promotion program will guide the development of cost-effective research translation campaigns.

Trial Registration: Australian New Zealand Clinical Trial Registry ACTRN12617000113358; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=371463

International Registered Report Identifier (IRRID): RR2-10.1186/s12889-017-4882-7

(JMIR Mhealth Uhealth 2019;7(6):e12484) doi:10.2196/12484

KEYWORDS

k-medoid cluster analysis; social marketing; mobile phone app; physical activity; online social networking

Introduction

The risk of premature death caused by chronic diseases such as cardiovascular heart disease or diabetes has increased globally. Physical inactivity is a key risk factor for chronic disease, yet most adults in developed countries are insufficiently active to obtain health benefits [1,2]. Low cost, mass reach physical activity interventions are necessary to address the physical inactivity epidemic.

Technology-based interventions have become increasingly popular as the internet use has grown. It is estimated that internet use increased by 1052% between 2000 and 2018 with 54.4% of the world's population now having internet access [3]. In addition to the increase in internet usage, online social media has become increasingly popular. Currently, Facebook has 2.2 billion active monthly users [4]. This has created opportunities for health behavior change interventions to be delivered via the internet and online social networks with more personalized and engaging features, potentially reaching a much larger population than previously possible.

Systematic reviews and meta-analyses suggest that online-delivered interventions can significantly increase physical activity [5-7]; however, relatively few such interventions are translated after the efficacy trial, and what studies find in randomized controlled trial (RCT) conditions may be different than the effectiveness of the same intervention in real-life conditions [8]. For example, the systematic review of Wu et al [9] of diabetes self-management apps found that only 1 out of 12 apps included in their meta-analysis was publicly available in the Apple or Android app stores. Another recent review of online self-help interventions found that just 30% were publicly available [10]. A number of barriers exist that hinder the translation of physical activity interventions developed and evaluated by researchers to the real world. This includes lack of resources to maintain or support software beyond the RCT, lack of resources to modify software designed for an RCT evaluation for wide-scale release, and lack of expertise and financial resources to promote the existence of the physical activity program so that it may be adopted by a large number of new users. A promising approach in terms of reach and cost effectiveness is to collaborate with marketers to translate health interventions at scale [11].

Social marketing uses commercial marketing approaches to influence positive behavior changes for individuals and communities [12,13]. Market segmentation analysis, also known as cluster analysis, is a social marketing technique used to identify homogeneous market segments of people who have similar needs or characteristics. Segments could be based on demographic, geographic, psychographic, or epidemiological factors. Identified segments can then be targeted with products or services designed or marketed specifically to address the segments' needs, with the idea that such targeted products and services will be more appealing to specific population subgroups and therefore have higher adoption rates [14]. To date, few studies have used segmentation analysis in the context of adult physical activity. Friederichs et al [15] undertook a segmentation analysis based on physical activity motivational regulation of Dutch participants who had signed up to an eHealth (electronic health) physical activity intervention study. The researchers identified three groups (low motivation, controlled motivation, and autonomous motivation), and suggested that in the future, different intervention approaches might benefit individuals in different motivational clusters. Griffin et al [16] undertook a cluster analysis based on demographic and health behaviors (which included physical activity and other health behaviors such as smoking, alcohol consumption, and diet) of participants in a large Australian cohort study of older adults (the 45 and Up study). They identified six lifestyle behavior clusters, including one in which multiple risk behaviors clustered together, which they suggested highlighted the need for future targeted interventions to meet the needs of this group. Finally, Rundle-Thiele et al [13] undertook a market research study to identify social market segments based on physical activity behaviors, attitudes, and intentions in a sample of Australian adults. They identified four segments (young disinterested, successful enthusiasts, vulnerables, and happy retirees) and suggested that the study provided insights into key segments that could inform the development of future interventions. Thus, all three studies presented their findings as being potential targets for future intervention. However, to our knowledge, none has subsequently applied this information to directly inform the design or marketing of a real-world health promotion tool.

This study aimed to use a similar approach as previous segmentation studies to directly inform the development of a marketing strategy to promote an evidence-based online social networking physical activity program [17]. Active Team is a gamified, online social networking physical activity intervention that aims to increase physical activity by encouraging users to undertake a 100-day physical activity challenge with their friends [17]. The program incorporates social influence and gamification techniques and is used in conjunction with a step-counter. In collaboration with marketing experts, we are currently planning a translation study that will attempt to disseminate the program widely using an online social marketing campaign. This study aimed to inform the translational study by (1) examining the characteristics of adopters of the Active Team program and (2) identifying clusters among the adopters based on sociodemographic, psychographic, and behavioral characteristics. Findings will be useful to inform software improvements and develop segmented marketing strategies when Active Team is disseminated to the public.

Methods

Statement of Ethics

Ethics approval for this study was provided by the University of South Australia's Human Research Ethics Committee (protocol number 0000033967), and all participants provided informed consent. This study was undertaken in the context of an RCT that ran from late 2016 to late 2018 [17].

Participants and Data Collection

Participants were Australian adults aged 18 to 65 years recruited between October 2016 and February 2017 to participate in an RCT evaluating a social and gamified physical activity intervention entitled Active Team. Promotional efforts for the Active Team app and the research trial evaluating it included mainstream media news stories, flyers, and paid, nontargeted Facebook advertisements, with the vast majority of participants signing up via Facebook rather than the other promotional methods. The Facebook advertisements featured still images rather than videos, accompanied by text captions, and were designed to be gender neutral (ie, images of both men and women). The advertisements were set so that they would be shown to men and women aged 18 to 65 years located anywhere in Australia, with no other targeting. The ads were developed in consultation with an online marketing academic (Professor Karen Nelson-Field). From March 1, 2017, targeted recruitment efforts were commenced in an effort to increase enrollment of men into the study. Therefore, this study only includes the participants recruited through nontargeted advertisements, up to the end of February 2017.

Interested adults were directed to download the app from the Apple and Google Play App Stores before enrolling. The app contained a feature that allowed them to send online invitations to their friends who could then also register their interest. Potential participants were screened for initial eligibility (aged between 18 to 65 years, insufficiently active, use of Facebook at least weekly, fluent in English, and living within Australia) and asked to provide informed consent and complete a baseline questionnaire. Participants were required to use Facebook weekly since the Active Team software itself integrated with Facebook by connecting users to their Facebook friends within the Active Team app. More detail regarding the RCT is available in the protocol paper [17].

Following this initial phase, to be formally enrolled in the study, participants needed to successfully wear an accelerometer for one week and have at least two friends successfully complete the baseline assessments (including accelerometry). However, for inclusion in our analysis, participants only needed to pass initial screening, provide informed consent, and complete the baseline survey. By focusing on adopters, we aimed to gain insights into those groups the app is likely to appeal to the most.

The baseline questionnaire was administered via Qualtrics online survey software and captured demographic information including residential address, date of birth, sex, marital status, height and weight (to calculate body mass index [BMI]), country of birth, and highest education level.

Self-reported weekly minutes of moderate to vigorous physical activity was collected using the Active Australia Survey. This 8-item questionnaire captures walking and other physical activities undertaken in both leisure and household duty contexts including the intensity of the activity in the preceding week [18]. To calculate sufficient activity, the time spent walking or in moderate activity and twice the time spent in vigorous activity were summed. Sufficient activity was interpreted as a total of at least 150 minutes of activity per week. The survey has reasonable test-retest reliability (intraclass correlation coefficient .52) [19] relative to accelerometry (r=.49-.64) [20].

Self-reported quality of life was measured using the validated 12-Item Short Form Health Survey (SF-12). This scale assesses both physical and mental quality of life domains [21]. Summary scales were scored using US population norms creating two measures for physical and mental component summary scores. The survey's test-retest reliability (two weeks apart) was .89 for the physical component summary and .76 for the mental component summary [22]. Validity of the SF-12 physical and mental component summary scores was correlated with the physical component summary-36 (r=.951) and mental component summary-36 (r=.969) equating to an R^2 of .904 for the physical component summary and an .939 for the mental physical component summary [22].

Well-being was assessed using the Positive Emotion, Engagement, Relationships, Meaning, Achievement (PERMA) Profile measure. This 23-item scale measures well-being across well-being pillars (positive emotion, engagement, relationships, meaning, and accomplishment) using 11-point Likert-type items (0=not at all to 10=completely) [23]. The PERMA Profile measure for overall well-being was scored by calculating the mean for all 23 items [23]. Test-retest reliability (two weeks apart) was .80 for accomplishment, .86 for meaning, .83 for relationships, .78 for engagement, and .84 for positive emotions [23]. Validity correlation coefficients were .79 when compared with the Satisfaction with Life Scale, and .87 when compared with Flourishing Scale [23].

Depression, anxiety and stress were measured using the Depression and Anxiety Stress Scale (DASS-21). This scale comprises 21 questions using a 4-point Likert-type scale (0=not at all, 3=almost always) [24]. The DASS-21 was scored by summing the scores for depression, anxiety, and stress then multiplying the sums by two as the original version of the DASS has 42 items [24]. Test-retest reliability (three weeks apart) was .77 (95% CI .56-.88) for depression, .89 (95% CI .81-.94) for anxiety, and .85 (95% CI .51-.94) for stress [25]. Validity correlation coefficients were .79 for the depression scale when compared with Beck Depression Inventory, .85 for the anxiety scale when compared with Beck Anxiety Inventory, and .68 for the stress scale when compared with State-Trait Anxiety Inventory–Trait Version [26].

Sleep quality and quantity were collected using the Pittsburgh Sleep Quality Index (PSQI). Self-report sleep duration was recorded in minutes, and sleep quality was measured on a 4-point Likert-type scale (1=very bad, 2=bad, 3=good, and 4=very good) [27]. The results were then presented as categorical data in the analysis and presented as total count and

percentage of the sample. Test-retest reliability and validity were obtained: a score greater than 5 resulted in sensitivity of 89.6% and specificity of 86.5% (kappa=.75, *P*<.001) [27].

Participant goal-setting behavior, outcome expectations, self-efficacy, and intentions were assessed using the 21-item Social Cognitive Theory (SCT) scale [28-31]. A composite score for each variable was calculated by taking the mean of items for that variable. Test-retest reliability and validity were established for the decisional balance scales [28].

Selection of Cluster Inputs

The questionnaire items described above produced a total of 29 possible cluster inputs. However, the number of variables that can be used in a cluster analysis depends on the sample size. Formann [32] recommends that the minimal sample size should be 2^k, where k represents the number of variables. The number of cases with complete data was 459, therefore eight variables could be used in the cluster analysis. To determine which eight variables should be included, we first produced a correlation matrix to determine whether any of the potential cluster inputs were collinear. No collinearity was detected. Therefore, the cluster inputs were decided by discussion among the authors who represented a wide range of academic disciplines including health sciences, behavioral science, and marketing based on the following parameters: first, we prioritized variables that could be used to deliver targeted online advertising (eg, age, sex). Second, we prioritized variables that represented ancillary benefits of physical activity that could be used in marketing messaging such as stress or BMI (eg, for people with high levels of stress, the physical activity program could be marketed by highlighting the positive role of physical activity in managing stress) [33]. Third, we aimed to maximize variety in the types of included outcomes (eg, a range of sociodemographic, physical, behavioral, and psychological variables) to gain further insight and potential marketing strategies from the sample. As a result, the following eight cluster input variables were chosen: sex, age, physical activity level, education, BMI, overall stress, overall well-being, and physical activity self-efficacy.

Analysis

Descriptive statistics were inspected for all study variables. Continuous variables were examined for normality. The sample was described in terms of means and standard deviations. Categorical variables such as sex were described in terms of total count and percentage of the sample.

K-medoids cluster analyses were performed to identify segments within the sample. This cluster analysis approach was selected on the basis that it permits analysis of categorical data [34,35], can handle nonnormally distributed variables, and has been previously used in health-related market segmentation research [15]. In addition, it is more robust to noise and outliers than the k-means approach [34]. Analyses, conducted with R 3.3.0 (The R Foundation), used the partitioning around medoids (PAM) algorithm with the Gower metric. The analysis finds data points or medoids within the data whose average dissimilarity to all the objects in the cluster is minimized. It begins by finding an initial set of medoids, then iteratively replaces one medoid by one nonmedoid until it determines best fit [36]. Analyses were set to produce solutions for between two to eight clusters on the basis that this number of clusters allowed segmentation that would result in meaningful and interpretable clusters [34]. The optimal number of clusters was subsequently determined by identifying the cluster solution that resulted in the maximum average silhouette width. The influence of multivariate outliers on the cluster solution was examined by plotting the squared Mahalanobis distances of the principal components against the empirical chi-squared distribution and identifying data points beyond the 97.5 percentile. Sixteen data points were considered outliers for females and four for males. Outliers were removed, and the data were reclustered. Cohen kappa between cluster solutions with and without outliers showed high agreement (.86 females and .64 for males); therefore, the outliers were retained.

Cluster stability of the final models was examined by randomly generating 99 subsamples of 80% of the full sample and computing the average Rand index for the cluster solutions, using the R package clv [37]. The Rand index indicates the proportion of pairs of cluster allocations that agree between the full sample and the subsample, resulting in a value between 0 (complete disagreement) and 1 (complete agreement) [38].

Results

Participant Characteristics

The demographic details of the 484 participants are summarized in Table 1. In brief, the majority were female (392/484, 81.0%), living with a partner (348/484, 71.9%), had a university education (275/484, 56.8%), and lived in a major city (455/484, 94.0%). The average age was 41 years, and the majority were overweight or obese (≥25, 78%). Compared with available Australian population data, participants who signed up to use the online social networking physical activity program were more likely to be female (81% vs 51%), cohabit with a partner (72% vs 58%), have a university education (57% vs 22%), be born in Australia (72% vs 67%) [39], live in a major city (94% vs 69%) [40], and be overweight or obese (78% vs 63%) [41]. At the group level, physical and mental component summary scores were within the normal range for the SF-12 [21]. However, depression, anxiety and stress were relatively high compared with previously published Australian population data (depression mean 4.8 [SD 5.1] vs 2.6 [SD 3.9], anxiety mean 3.6 [SD 3.8] vs 1.7 [SD 2.8], and stress mean 7.5 [SD 4.9] vs 4.0 [SD 4.2]) [42].

Cluster Analysis

Two cluster analysis models were performed on a total of 459 participants who had complete data for the variables selected for inclusion in the analysis. An initial cluster analysis resulted in trivial clusters determined entirely by sex and educational level. Therefore, we removed education, stratified the population by sex, and performed separate cluster analyses for men and women using six cluster inputs due to the reduced sample by separating for sex. The silhouette widths were maximized when two clusters were derived for each sex (Table 2).

Table 1. Descriptive characteristics of the study sample (n=484).

Characteristics	Value
Age in years, mean (SD)	40.8 (12.1)
Sex, female, n (%)	392 (81.0)
Marital status, n (%)	
Partnered	348 (71.9)
No partner	121 (25.0)
Prefer not to disclose	15 (3.1)
Body mass index, mean (SD)	30.1 (6.7)
Education, n (%)	
High school or lower	73 (15.1)
Some post-high school (eg, trade or diploma)	136 (28.1)
University	275 (56.8)
Australian-born, n (%)	348 (71.7)
ASGS[a] remoteness, n (%)	
Major city	455 (94.0)
Inner regional	19 (3.9)
Outer regional	5 (1.0)
Remote	5 (1.0)
Very remote	0 (0)
Social cognitive theory, mean (SD)	
Goal setting (scale range 1-5)	3.0 (0.1)
Outcome expectations (scale range 1-5)	3.3 (0.4)
Intention (scale range 1-7)	4.1 (0.7)
Self-efficacy (scale range 1-5)	3.0 (0.7)
Positive Emotion, Engagement, Relationships, Meaning, Achievement, mean (SD) (scale range 1-10)	6.5 (1.5)
12-Item Short Form Health Survey, mean (SD) (scale range 0-100)	
Physical component summary	40.3 (6.1)
Mental component summary	48.0 (9.4)
PSQI[b] sleep quality, n (%)	
Very good	24 (5.0)
Good	223 (46.1)
Bad	198 (40.9)
Very bad	39 (8.1)
Depression, Anxiety, and Stress Scale, mean (SD)	
Depression	4.8 (5.1)
Anxiety	3.6 (3.8)
Stress	7.5 (4.9)

[a]ASGS: Australian Statistical Geography Standard [43].

[b]PSQI: Pittsburgh Sleep Quality Index.

Table 2. Female and male clusters sex, age, physical activity level, body mass index, stress, well-being, and self-efficacy per cluster.

Characteristic	Cluster 1 (n=218)	Cluster 2 (n=153)	Cluster 3 (n=58)	Cluster 4 (n=30)
Sex, male, n (%)	0 (0)	0 (0)	58 (100)	30 (100)
Age, mean (SD)	33.4 (7.8)	52.8 (5.9)	30.7 (7.7)	52.5 (6.5)
Weekly minutes PA[a], mean (SD)	226 (282)	204 (233)	297 (296)	159 (183)
Body mass index, mean (SD)	29.6 (7.5)	30.9 (6.4)	28.4 (5.5)	33.1 (4.3)
Stress, mean (SD)	8.4 (4.7)	6.0 (4.1)	8.9 (6.4)	6.2 (4.5)
Well-being, mean (SD)	6.4 (1.5)	6.6 (1.7)	6.3 (1.5)	6.7 (1.5)
Self-efficacy, mean (SD)	3.0 (0.7)	2.9 (0.7)	3.0 (0.8)	3.3 (0.9)

[a]PA: physical activity.

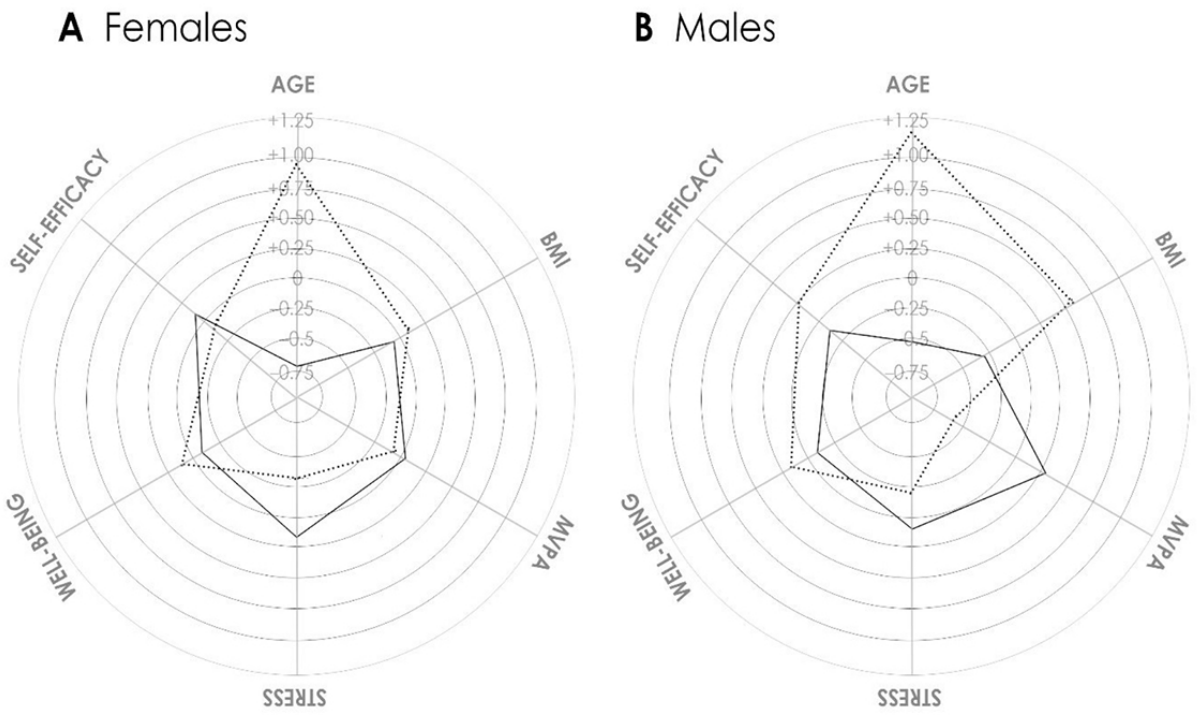
Among women, cluster 1 consisted of relatively younger inactive women with high stress, while cluster 2 was characterized by older women with lower stress and higher physical activity. Among men, cluster 3 was characterized by younger, stressed but physically active men, while cluster 4 consisted of older, inactive, and predominantly obese men. The median Rand index was 0.87 (IRQ 0.62-1.00) for the female cluster solution and 0.75 (IRQ 0.58-0.87) for the male cluster solution.

Clinical differences (eg, psychosociodemographic characteristics) of the clusters other than cluster inputs were then examined (Table 3). Women in cluster 1 were more likely to be single, have lower mental quality of life, and report higher depression and anxiety than women in cluster 2. Among the men, those in cluster 3 were more likely to be single compared with men in Cluster 4. The cluster groups are presented in Figure 1.

Table 3. Descriptive characteristics of the clusters.

Characteristic	Cluster 1 (n=218)	Cluster 2 (n=153)	Cluster 3 (n=58)	Cluster 4 (n=30)
Marital status, n (%)				
Partnered	146 (67.0)	121 (79.1)	32 (55.2)	28 (93.3)
No partner	61 (28.0)	28 (18.3)	25 (43.1)	2 (6.7)
Prefer not to disclose	11 (5.0)	4 (2.6)	1 (1.7)	0 (0)
Education, n (%)				
High school or lower	28 (12.8)	20 (13.1)	13 (22.4)	8 (26.7)
Some post-secondary	59 (27.1)	46 (30.1)	12 (20.7)	6 (20.0)
University	131 (60.1)	87 (56.9)	33 (56.9)	16 (53.3)
Australian-born, n (%)	181 (83.0)	127 (83.0)	39 (67.2)	26 (86.7)
Social Cognitive Theory, mean (SD)				
Goal setting	3.1 (1.0)	2.9 (1.0)	2.9 (1.0)	2.8 (0.9)
Outcome expectations	3.3 (0.4)	3.3 (0.4)	3.4 (0.4)	3.3 (0.4)
Intention	4.1 (0.7)	4.0(0.6)	3.9 (0.8)	4.2 (0.7)
12-Item Short Form Health Survey, mean (SD)				
Physical component	40.2 (5.8)	40.3 (6.4)	40.5 (6.1)	38.9 (7.0)
Mental component	46.6 (10.0)	50.3 (8.4)	46.8 (9.0)	48.7 (9.7)
Pittsburgh Sleep Quality Index, n (%)				
Very good	13 (6.0)	8 (5.2)	4 (6.9)	0 (0)
Good	105 (48.2)	70 (45.8)	25 (43.1)	14 (46.7)
Bad	87 (39.9)	64 (41.8)	23 (39.7)	12 (40.0)
Very bad	13 (6.0)	11 (7.2)	6 (10.3)	4 (13.3)
Depression, Anxiety, and Stress Scale, mean (SD)				
Overall depression	5.2 (5.0)	3.9 (4.8)	6.0 (5.6)	4.2 (5.5)
Overall anxiety	3.9 (3.8)	2.7 (3.1)	5.0 (5.2)	3.0 (3.5)

Figure 1. Radar graphs of the female and male cluster solution. BMI: body mass index, MVPA: moderate and vigorous physical activity.



Discussion

Principal Findings

This study aimed to determine the characteristics of adopters of an online social networking physical activity intervention and use segmentation analysis to identify homogenous market segments of users. Results revealed that the online social networking physical activity intervention attracted people who were well educated, urban, and female and had a higher BMI and higher depression, anxiety, and stress levels compared with population norms. The segmentation analysis identified four clusters within the adopters: younger inactive women with relatively poor mental well-being, older physically active women, younger active but stressed men, and older obese and inactive men.

The fact that the study attracted mostly urban residents, females, and people with a university education is consistent with previous health promotion research. For example, Duggan and Brenner [44] similarly reported that women and those with a higher level of education are more likely to engage with social media and online health interventions. In addition, a systematic review examining the effectiveness of online social networks to improve health behaviors conducted by Maher et al [6] reported high female participation rates (83.3%). It is also common for RCTs to have more difficulty engaging with rural participants [45]. People living in rural and remote areas are recognized to be underserved in terms of primary prevention and health care access. Online interventions appear to be a good avenue for addressing such access inequalities; however, our results suggest that access barriers remain. This could be because despite the avenue of delivery women are generally more engaged with their health and well-being [44].

In general, the key traits that determined cluster membership were age, sex, physical activity, BMI, and stress. This is reasonably consistent with previous studies of this nature with the variables age, physical activity, and BMI consistently appearing in cluster memberships with physical activity and BMI being age-related [13,15]. This study differed from previous research in that we included more psychographic measures. Cluster membership was segmented based on stress in both female and male clusters. In addition, among women, there were significant differences between the clusters for the SF-12 mental component and the depression and anxiety scales for the DASS-21. It is possible that power was an issue when attempting to identify statistical differences between the men's clusters relative to the women's.

Strengths and Limitations

Before considering the implications of these findings, study strengths and weaknesses should be acknowledged. To the best of the authors' knowledge, no technology-based physical activity programs to date have used segmentation analysis to inform translation efforts. Our approach was highly interdisciplinary, with the team including health, marketing, and statistical experts. We were able to use a wide range of outcome measures (sociodemographic, behavioral, psychographic, and physical), providing insights across many potential marketing targets. Furthermore, a statistical strength of this study was that we used the PAM algorithm with Gower metric, which improves the strength and validity of the clusters compared with conventional cluster analysis methods.

This study also has some limitations. First, due to the sample size, the number of cluster inputs had to be restricted to eight and then six. It is possible that choosing different cluster inputs may have resulted in different clusters. In addition, the sample

may limit generalization of our findings to other contexts. Specifically, despite using a nontargeted advertising approach, it is possible that the participant demographics were influenced by the recruitment strategies used and that a different advertising approach may have resulted in a different sample. Additionally, participants registered interest in an RCT of the online social networking intervention, and it is possible that this group is different from people who may download and use the app in a real-life situation. However, it is difficult to postulate in what ways they would differ—for example, it is possible that the RCT may attract relatively highly motivated users (since they are willing to take on the assessment burden associated with research participation). Conversely, it is possible that the RCT actually attracts less motivated people—for example, if they are motivated by a financial incentive offered by a study or are aware they lack intrinsic motivation and so are seeking the imposition of external discipline. Despite these limitations, this approach enabled us to examine a wider variety of participant characteristics than would be possible by examining users of a commercial app, providing valuable information for disseminating similar apps via online advertising in the future.

Implications

Technology-based health promotion using mobile phones is a growing field in research with great potential to impact the health and well-being of many. To date, however, relatively few evidence-based interventions developed and tested in research settings are ever attempted to be translated to the real world [7]. Successful translation will require complementary approaches that go beyond RCT designs to understand the true impact of public health interventions on the general population in everyday conditions.

The market segments identified in this study as being more likely to be attracted to this type of physical activity program used in a research setting were younger low-active women with poor mental well-being, older active women, young active but stressed men, and older low-active and obese men. All segments comprised people of relatively high educational status, with the majority being university educated. Such information may be used to inform future media and communication channels, given that media usage patterns differ by demographic segment. For example, an Instagram campaign may be useful to reach the young female cluster identified in the segment analysis, given that approximately 60% of Instagram users are aged between 18 and 35 years and the platform is equally popular for men and women [46].

The psychographic characteristics of the identified clusters provide guidance for the potential benefits people in these segments may be seeking by joining the program. These benefits (or solutions to their existing problems) can be used to inform marketing messages and translate those messages during promotional campaigns for physical activity programs. As an example, given that a major cluster comprised younger women with relative low well-being, marketing campaigns may focus on the positive effects physical activity has on mental health [33], or alternatively, suggest that additional program features targeted at improving well-being may be warranted.

The analysis also highlighted that this app largely did not reach certain demographics—for example, men and people of low educational status were underrepresented. This raises the possibility of two very different directions for future in-market program promotion efforts. The first option is to play to the program's current strengths and tailor marketing efforts for the ecological trial to the adopter market segments identified in this study. This approach could be expected to result in higher return on investment—that is, higher participation rate from the specific targeted groups (higher participation from the stressed younger women and almost no participation from other nontargeted segments). An alternative approach would be to go after a mass market including people who were not captured in this adopter study. While some of these segments might be much harder to reach (eg, males), keeping the approach and message more general might result in lower overall numbers of enrollments into the health program. The utility and cost effectiveness of each of these approaches can be tested in a well-planned ecological trial.

Conclusion

Technology-based health promotion programs offer great promise for delivering effective, appealing, and accessible interventions. However, for their full potential to be reached, evidence-based interventions must be scaled up, in particular to reach audience segments that health behavior research does not typically reach (eg, men, those of low socioeconomic status). This requires multidisciplinary efforts, teaming health researchers with marketers. This study serves as a novel example of how research translation might proceed following an RCT and demonstrates how a social marketing framework and market segmentation analysis approach may provide insights to guide future software improvements and targeted marketing efforts. Such judicious, evidence-based approaches are likely to be helpful, given that health researchers seek to deliver the greatest impact with limited program development and marketing budgets. Additional research aimed at identifying clusters of adopters of health apps in larger samples from more ecologically valid contexts would provide an important contribution to this emerging literature and further assist health researchers planning dissemination of health promotion programs.

Acknowledgments

This project is funded by the National Health and Medical Research Council, Australia, project grant 1125913. The funding body had no role in the design, collection, analysis, or interpretation of this study. The authors thank Portal Australia for assisting with the development of the Active Team app.

Authors' Contributions

CM, RP, CV, and TO conceived the original Active Team study, and JR and SE led data collection. IS, CM, CS, and SB conceived the cluster analysis. TS and RC oversaw the analysis. IS led drafting of the manuscript. All authors contributed to interpretation of study findings, assisted with drafting of the manuscript, and approved the final version.

Conflicts of Interest

None declared.

References

1. Dumith SC, Hallal PC, Reis RS, Kohl HW. Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. Prev Med 2011;53(1-2):24-28. [doi: 10.1016/j.ypmed.2011.02.017] [Medline: 21371494]
2. World Health Organization. Physical Activity URL: http://www.who.int/en/news-room/fact-sheets/detail/physical-activity [accessed 2018-10-10] [WebCite Cache ID 733XSgiUc]
3. Miniwatts Marketing Group. 2018. Internet World Stats URL: https://www.internetworldstats.com/stats.htm [accessed 2018-10-10] [WebCite Cache ID 733WvKWct]
4. Statista DMO. 2018. Number of monthly active Facebook users worldwide as of 4th quarter 2017 (in millions) URL: https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide [accessed 2019-04-26] [WebCite Cache ID 733X3YopR]
5. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9(1):52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
6. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
7. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, present, and future of eHealth and mHealth research to improve physical activity and dietary behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
8. Vandelanotte C, Duncan MJ, Kolt GS, Caperchione CM, Savage TN, Van Itallie A, et al. More real-world trials are needed to establish if web-based physical activity interventions are effective. Br J Sports Med 2018 Jul 03. [doi: 10.1136/bjsports-2018-099437] [Medline: 29970409]
9. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
10. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
11. Hastings G. Social Marketing: Why Should the Devil Have All the Best Tunes?. Oxford: Elsevier; 2007.
12. Fujihira H, Kubacki K, Ronto R, Pang B, Rundle-Thiele S. Social marketing physical activity interventions among adults 60 years and older. Social Marketing Quarterly 2015 Oct 05;21(4):214-229. [doi: 10.1177/1524500415606671]
13. Rundle-Thiele S, Kubacki K, Tkaczynski A, Parkinson J. Using two-step cluster analysis to identify homogeneous physical activity groups. Mrkting Intelligence Plan 2015 Jun;33(4):522-537. [doi: 10.1108/MIP-03-2014-0050]
14. McDonald M, Dunbar I. Market Segmentation: How to Do It and How to Profit from It. Chichester: John Wiley & Sons; 2012.
15. Friederichs SA, Bolman C, Oenema A, Lechner L. Profiling physical activity motivation based on self-determination theory: a cluster analysis approach. BMC Psychol 2015;3(1):1 [FREE Full text] [doi: 10.1186/s40359-015-0059-2] [Medline: 25678981]
16. Griffin B, Sherman KA, Jones M, Bayl-Smith P. The clustering of health behaviours in older Australians and its association with physical and psychological status, and sociodemographic indicators. Ann Behav Med 2014 Oct;48(2):205-214 [FREE Full text] [doi: 10.1007/s12160-014-9589-8] [Medline: 24500081]
17. Edney S, Plotnikoff R, Vandelanotte C, Olds T, De Bourdeaudhuij I, Ryan J, et al. "Active Team" a social and gamified app-based physical activity intervention: randomised controlled trial study protocol. BMC Public Health 2017 Nov 02;17(1):859 [FREE Full text] [doi: 10.1186/s12889-017-4882-7] [Medline: 29096614]
18. Australian Institute of Health and Welfare. 2003 Apr 24. The Active Australia Survey: a guide and manual for implementation, analysis and reporting URL: https://www.aihw.gov.au/getmedia/ff25c134-5df2-45ba-b4e1-6c214ed157e6/aas.pdf.aspx?inline=true [accessed 2019-04-26] [WebCite Cache ID 77ujvWc0O]
19. Brown W, Mummery K, Eakin E, Schofield G. 10,000 Steps Rockhampton: evaluation of a whole community approach to improving population levels of physical activity. J Physical Activity Health 2006 Jan;3(1):1-14. [doi: 10.1123/jpah.3.1.1]

20. Freene N, Waddington G, Chesworth W, Davey R, Cochrane T. Validating two self-report physical activity measures in middle-aged adults completing a group exercise or home-based physical activity program. J Sci Med Sport 2014 Nov;17(6):611-616. [doi: 10.1016/j.jsams.2013.11.002] [Medline: 24332192]
21. Avery J, Dal Grande E, Taylor A. Population Research and Outcome Studies Unit, Department of Human Services. 2004. Quality of life in South Australia as measured by the SF12 Health Status Questionnaire URL: https://tinyurl.com/y49rbtcd [accessed 2019-04-26] [WebCite Cache ID 77ukOzT8A]
22. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
23. Butler J, Kern ML. The PERMA-Profiler: a brief multidimensional measure of flourishing. Int J Wellbeing 2016 Oct 13;6(3):1-48. [doi: 10.5502/ijw.v6i3.526]
24. Lovibond S, Lovibond P. Manual for the Depression Anxiety Stress Scales. Sydney: Psychology Foundation; 1995.
25. Asghari A, Saed F, Dibajnia P. Psychometric properties of the Depression Anxiety Stress Scales-21 (DASS-21) in a non-clinical Iranian sample. Int J Psychol 2008;2(2):82-102 [FREE Full text]
26. Antony M, Bieling P, Cox B, Enns M, Swinson R. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychological Assessment 1998 Jun;10(2):176-181 [FREE Full text] [doi: 10.1016/j.jad.2008.01.023] [Medline: 18304648]
27. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [Medline: 2748771]
28. Plotnikoff RC, Blanchard C, Hotz SB, Rhodes R. Validation of the decisional balance scales in the exercise domain from the transtheoretical model: a longitudinal test. Measurement Phys Educ Exercise Sci 2001 Dec;5(4):191-206. [doi: 10.1207/S15327841MPEE0504_01]
29. Plotnikoff R, Lippke S, Courneya K, Birkett N, Sigal R. Physical activity and social cognitive theory: a test in a population sample of adults with type 1 or type 2 diabetes. Appl Psychol 2008;57(4):628-643. [doi: 10.1111/j.1464-0597.2008.00344.x]
30. Rhodes RE, Blanchard CM, Matheson DH, Coble J. Disentangling motivation, intention, and planning in the physical activity domain. Psychol Sport Exercise 2006 Jan;7(1):15-27. [doi: 10.1016/j.psychsport.2005.08.011]
31. Rovniak LS, Anderson ES, Winett RA, Stephens RS. Social cognitive determinants of physical activity in young adults: a prospective structural equation analysis. Ann Behav Med 2002;24(2):149-156. [Medline: 12054320]
32. Formann A. Die Latent-Class-Analyse: Einfuhrung in Theorie und Anwendung. Weinheim: Beltz; 1984.
33. Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. Curr Opin Psychiatry 2005 Mar;18(2):189-193. [Medline: 16639173]
34. Mooi E, Sarstedt M. Cluster analysis. In: Mooi E, Sarstedt M, editors. A Concise Guide to Market Research. Berlin: Springer-Verlag; 2011.
35. Park H, Jun C. A simple and fast algorithm for K-medoids clustering. Expert Syst Appl 2009;36(2):3336-3341. [Medline: 31018197]
36. Kaufman L, Rousseeuw P. Clustering by Means of Medoids. Sofia: North-Holland; 1987.
37. Nieweglowski L. R package version 0.3-2.1—clv: cluster validation techniques. 2013 Nov 11. URL: http://cran.r-project.org/web/packages/clv/index.html [accessed 2019-04-26] [WebCite Cache ID 77vYd7he8]
38. Rand WM. Objective criteria for the evaluation of clustering methods. J Am Stat Assoc 1971 Dec;66(336):846-850. [doi: 10.1080/01621459.1971.10482356]
39. Australian Bureau of Statistics. 2016. Census Quickstats URL: http://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/036 [accessed 2019-04-26] [WebCite Cache ID 733WXLRvA]
40. Australian Bureau of Statistics. 2008. Population Distribution URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Chapter3002008 [accessed 2018-10-10] [WebCite Cache ID 733VonT4x]
41. Australian Bureau of Statistics. Canberra: ABS; 2015. National Health Survey: 2014–15, Key Findings URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0.55.001 [WebCite Cache ID 734kopgJO]
42. Crawford J, Cayley C, Lovibond P, Wilson P, Hartley C. Percentile norms and accompanying interval estimates from an Australian general adult population sample for self‐report mood scales (BAI, BDI, CRSD, CES‐D, DASS, DASS‐21, STAI‐X, STAI‐Y, SRDS, and SRAS). Austr Psychol 2011;46(1):3-14. [doi: 10.1111/j.1742-9544.2010.00003.x]
43. Australian Bureau of Statistics. 2013. Population projections, Australia, 2012 (base) to 2101 URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Chapter3002008 [accessed 2018-10-10] [WebCite Cache ID 733WLGTnS]
44. Duggan M, Brenner J. The demographics of social media users, 2012. Washington: Pew Internet and American Life Project URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2013/PIP_SocialMediaUsers.pdf [accessed 2019-04-26] [WebCite Cache ID 77vZlVfTU]
45. Kenny A, Farmer J, Dickson-Swift V, Hyett N. Community participation for rural health: a review of challenges. Health Expect 2015 Dec;18(6):1906-1917 [FREE Full text] [doi: 10.1111/hex.12314] [Medline: 25470231]
46. Statista DMO. Distribution of Instagram users worldwide as of April 2018, by age and gender URL: https://www.statista.com/statistics/248769/age-distribution-of-worldwide-instagram-users/ [accessed 2018-10-10] [WebCite Cache ID 733XGjEyZ]

Abbreviations

BMI: body mass index
DASS-21: Depression, Anxiety, and Stress Scale
eHealth: electronic health
PAM: partitioning around medoids
PERMA: Positive Emotion, Engagement, Relationships, Meaning, Achievement
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
SCT: Social Cognitive Theory
SF-12: 12-Item Short Form Health Survey

Edited by G Eysenbach; submitted 17.10.18; peer-reviewed by S Li, R Tague, A Maeder; comments to author 12.02.19; revised version received 07.04.19; accepted 11.04.19; published 03.06.19.

Please cite as:
Sanders I, Short CE, Bogomolova S, Stanford T, Plotnikoff R, Vandelanotte C, Olds T, Edney S, Ryan J, Curtis RG, Maher C
Characteristics of Adopters of an Online Social Networking Physical Activity Mobile Phone App: Cluster Analysis
JMIR Mhealth Uhealth 2019;7(6):e12484
URL: https://mhealth.jmir.org/2019/6/e12484/
doi:10.2196/12484
PMID:31162130

Original Paper

A Feasibility Trial of Power Up: Smartphone App to Support Patient Activation and Shared Decision Making for Mental Health in Young People

Julian Edbrooke-Childs[1], PhD; Chloe Edridge[1], MPhil; Phoebe Averill[1], MSc; Louise Delane[1], PhD; Chris Hollis[2,3,4], PhD, FRCPsych; Michael P Craven[2,3,5], PhD; Kate Martin[6], PhD; Amy Feltham[6], BSc; Grace Jeremy[6], BSc; Jessica Deighton[1], PhD; Miranda Wolpert[1], DClinPsy

[1]Evidence Based Practice Unit, University College London and the Anna Freud National Centre for Children and Families, London, United Kingdom
[2]National Institute of Health Research MindTech MedTech Co-operative, Nottingham, United Kingdom
[3]National Institute of Health Research Nottingham Biomedical Research Centre, Nottingham, United Kingdom
[4]Division of Psychiatry and Applied Psychology, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom
[5]Bioengineering Research Group, Faculty of Engineering, University of Nottingham, Nottingham, United Kingdom
[6]Common Room Consulting Limited, London, United Kingdom

Corresponding Author:
Julian Edbrooke-Childs, PhD
Evidence Based Practice Unit
University College London and the Anna Freud National Centre for Children and Families
Clinical, Educational and Health Psychology, Gower Street
London, WC1E 6BT
United Kingdom
Phone: 44 (0) 20 7794 2313
Email: julian.edbrooke-childs@annafreud.org

Abstract

Background: Digital tools have the potential to support patient activation and shared decision making in the face of increasing levels of mental health problems in young people. There is a need for feasibility trials of digital interventions to determine the usage and acceptability of interventions. In addition, there is a need to determine the ability to recruit and retain research participants to plan rigorous effectiveness trials and, therefore, develop evidence-based recommendations for practice.

Objective: This study aimed to determine the feasibility of undertaking a cluster randomized controlled trial to test the effectiveness of a smartphone app, Power Up, co-designed with young people to support patient activation and shared decision making for mental health.

Methods: Overall, 270 young people were screened for participation and 52.5% (142/270) were recruited and completed baseline measures across 8 specialist child mental health services (n=62, mean age 14.66 (SD 1.99) year; 52% [32/62] female) and 2 mainstream secondary schools (n=80; mean age 16.88 [SD 0.68] years; 46% [37/80] female). Young people received Power Up in addition to management as usual or received management as usual only. Posttrial interviews were conducted with 11 young people from the intervention arms (specialist services n=6; schools n=5).

Results: Usage data showed that there were an estimated 50 (out of 64) users of Power Up in the intervention arms. Findings from the interviews indicated that young people found Power Up to be acceptable. Young people reported (1) their motivation for use of Power Up, (2) the impact of use, and (3) barriers to use. Out of the 142 recruited participants, 45.0% (64/142) completed follow-up measures, and the approaches to increase retention agreed by the steering group are discussed.

Conclusions: The findings of this study indicate that the app is acceptable, and it is feasible to examine the effectiveness of Power Up in a prospective cluster randomized controlled trial.

Trial Registration: ClinicalTrials.gov NCT02552797; https://clinicaltrials.gov/ct2/show/NCT02552797 (Archived by WebCite at http://www.webcitation.org/6td6MINP0)

(JMIR Mhealth Uhealth 2019;7(6):e11677) doi:10.2196/11677

KEYWORDS

telemedicine; patient participation; mental health; adolescent

Introduction

Background

A minority of the population (17%) will not experience mental health problems in their lifetime [1]. On the basis of the last prevalence study, 10% of children in the United Kingdom have a clinically diagnosable mental health problem, with one of the most prevalent being emotional problems including anxiety and depression, of whom 25% receive support from specialist mental health services [2]. Recent evidence suggests that levels of mental health difficulties in young people are increasing with, for example, 1 in 4 young women experiencing emotional problems [3].

Around 25% to 40% of the population have little knowledge, skills, and confidence to manage their own health and health care (referred to as patient activation) [4]. Empowering patients to be actively involved in the management of their health care and involving them in shared decision making are emphasized in the Health and Social Care Act 2012 [5]. Evidence suggests this may have a range of benefits to health and care [6]; for example, a systematic review found that patients were more likely to adhere to treatment when it was in line with their preferences [7,8]. However, empowering patients to actively manage their health care is not widely practiced [9] and clinicians report being unclear about how to facilitate it [10]. A systematic review of observer-rated studies showed that clinicians rarely facilitate patient involvement and adjust care to patient preferences even less often [11].

Young people want to be actively involved in making decisions about their health care and report feeling more in control of their care when they are included in decisions [12]. Parents also feel that their children should be involved in decisions about their care as it may increase their self-esteem and improve their overall welfare [12]. Furthermore, it is recognized under the United Nations Convention on the Rights of the Child that young people should be involved in all matters that affect them.

Interventions in child mental health settings which include empowering patients to be actively involved have been shown to improve quality of life and satisfaction [13,14], and child and parent experiences of shared decision making have been shown to be associated with higher levels of symptom improvement [15]. Evidence of promise has also emerged from an evaluation of tools supporting young people's mental health through preparing for discussions, mood tracking, and self-management [16]. Evidence from adult settings suggests that interventions targeting empowerment and patient activation may promote engagement with services and interventions [17]. Nonattendance of appointments in child mental health services is an estimated 15% to 28% [18-23]. Noncollaborative decision making is a key predictor of nonattendance [24]. Introducing resources that promote better accessibility to and integration of care, and that which correspondingly make clinicians' time more efficient, would be invaluable to guarantee that services such as specialist child mental health services can continue to provide good quality service to as many people as possible.

Despite these benefits, young people are presented with a lack of opportunities for reflection and involvement in the decisions that affect them and can often feel unskilled or unsupported in these situations. This is a barrier to their involvement. Correspondingly, there is a need for the development of appealing and acceptable patient activation and shared decision-making tools that support young people to ask questions independently and raise the issues they want to discuss.

Young people have advised that technology that is engaging, easy to access, informative, empowering, and provides support between sessions would be a particularly useful addition to therapy [25]. The use of technology in some areas of mental health care is recommended by the National Institute for Health and Care Excellence 2011 best practice guidance [26]. Indeed, young people report already using technology as an informal complement to treatment [25]. In Great Britain, 82% of adults use the internet daily and 70% of adults use the internet on smartphones [27]. Some 83% of young people aged 11 to 18 years own a smartphone and use mobile internet daily [28,29]. The growth in the smartphone and tablet market and the high levels of engagement within mobile app usage mean that there has been a rise in the adoption of mobile health care apps. The mobile health care market was estimated to be worth US $25.39 billion in 2017 [30].

Young people, carers, and clinicians report feeling positive about integrating the use of certain apps into interventions for young people in mental health settings [16]. However, the content of many youth mental health apps is not grounded in psychological theory or evidence-based practice [31]. There is a need for evidence from rigorous trials as to the effectiveness of digital interventions for mental health in young people [32], in addition to research investigating how best to integrate these apps into support provision [25]. To plan rigorous trials to examine effectiveness, feasibility trials are needed [33]. In particular, there is a need for feasibility trials to determine (1) the usage and acceptability of digital interventions for youth mental health and (2) the ability to recruit and retain research participants [34,35].

Aims and Objectives

The aim of this study was to determine the feasibility of examining the effectiveness of a smartphone and tablet app, Power Up, in a prospective cluster randomized controlled trial and to determine the usage and acceptability of Power Up. In addition, we examined the ability to recruit and retain research participants. The app was designed to increase a young person's patient activation related to their mental health by providing tools to support their voice in therapy, facilitate a more patient-centered approach, and increase shared decision making. The app was developed in partnership with young people and advocates to increase its acceptability to young people. Power Up enables young people to record their questions, plans,

decisions, and diary entries and supports young people to identify individuals in their support network with whom they would like to share these entries. By providing a digital space for young people to prepare what they want to bring to conversations about their mental health and well-being, Power Up was designed to empower young people to take an active role in decisions that impact their health and care. Both professionals and young people with lived experience were involved in the design of Power Up, ensuring that the views of all relevant groups were heard during app development.

The objective of the present feasibility trial was to collect the necessary parameters to plan a cluster control effectiveness trial of Power Up. In line with guidelines on conducting feasibility trials [36], the Trial Steering Committee developed Go and No Go criteria to determine whether or not the findings from this study indicated that it would be feasible to examine the effectiveness of Power Up in a full trial.

Methods

Overview

A protocol for the feasibility trial was published [37] and registered with trials registries. To determine the feasibility of examining the effectiveness of Power Up in a prospective cluster randomized controlled trial, the Trial Steering Committee agreed the following Go and No Go criteria for the study. In line with guidance on conducting feasibility trials [36], we did not set any criteria related to effectiveness as the aims of this study were to collect the parameters necessary to plan the full trial and ensure an analysis of effectiveness was adequately powered. The criteria for the feasibility study are presented in Textbox 1.

Textbox 1. Go and No Go criteria for the feasibility study.

Scoring key

- A total score of 0 represents all criteria were fully met, and the study is rated as green meaning the full trial can immediately proceed.
- A total score of 1 to 5 represents some criteria were partially met, and the study is rated as amber meaning the full trial can proceed once relevant criteria have been reviewed and a plan to increase adherence has been agreed by the Trial Steering Committee.
- If 1 criterion is not met (ie, rated as red), the trial is rated as red. The full trial can only proceed if (1) the total score is less than 8 meaning not more than 2 criteria can be rated red and (2) a plan to increase adherence to any red criteria has been agreed by the Trial Steering Committee.

Criteria

1. Ability to recruit and retain sites

- 70% to 100% of sites recruited and retained (green)
- 50% to 69% of sites recruited and retained (amber)
- 0% to 49% of sites recruited and retained (red)

2. Rates of sign up to study for young people approached

- 60% to 100% of young people signed up to participate (green)
- 40% to 59% of young people signed up to participate (amber)
- 0% to 39% of young people signed up to participate (red)

3. Rates of download and usage of Power Up

- 60% to 100% of young people download and use Power Up (green)
- 40% to 59% of young people download and use Power Up (amber)
- 0% to 39% of young people download and use Power Up (red)

4. Young people report Power Up as acceptable in interviews

- 70% to 100% young people report that Power Up is acceptable (green)
- 50% to 69% of young people report that Power Up is acceptable (amber)
- 0% to 49% of young people report that Power Up is acceptable (red)

5. Completion of study measures at baseline and follow-up

- 50% to 100% of young people and carers complete study measures at both baseline and follow-up (green)
- 20% to 49% of young people and carers complete study measures at both baseline and follow-up (amber)
- 0% to 19% of young people and carers complete study measures at both baseline and follow-up (red)

Changes to Protocol

We had initially planned to only conduct the feasibility trial in specialist Child and Adolescent Mental Health Services (CAMHS). We added a schools strand to the feasibility trial for 2 reasons. First, it became clear from feedback from service users, professionals, and researchers that Power Up was applicable to settings beyond specialist mental health services to empower young people to self-manage emotional well-being. Second, the rate of recruitment from specialist services was slower than anticipated, because of various reasons such as clinician workload, turnover, and either young people not attending appointments, not meeting study inclusion criteria, or both. Correspondingly, the target audience for Power Up expanded during the course of the study from just young people experiencing mental health problems to young people in schools to support self-management of emotional well-being.

Recruitment

In the feasibility trial, young people's experiences while using Power Up were compared with young people's experiences without using the app. Overall, 270 young people were screened for participation across 8 specialist services (n=79) and 2 mainstream secondary schools (n=191) (the demographics are reported in the Results section).

For the specialist services strand of the trial, a wait list control design was employed. Initially, 33 young people were recruited to the control phase of the trial where they received management as usual (average cluster size 4.13 (SD 3.48)). Subsequently, 30 different young people were recruited to the intervention phase of the trial where they were given Power Up to use alongside management as usual (average cluster size 3.75 (SD 3.20)). This study was given a favorable opinion by the Health Research Authority Research Ethics Committee (reference number: 192592). Clinicians identified young people who were aged 11 to 19 years and were in their initial assessment sessions for recruitment to the trial. Once consent had been given, young people completed a questionnaire containing a battery of measures. Study measures included (1) the Patient Activation Measure [4] to assess young people's empowerment and self-management of their mental health and well-being, (2) the CollaboRATE [38] and the Shared Decision Making Questionnaire 9 [39] to assess shared decision making, (3) the Youth Empowerment Scale—Mental Health [40] to assess young people's confidence to manage their mental health (ie, self-subscale) and the support they receive from services (ie, service subscale), (4) the Strengths and Difficulties Questionnaire [41] to assess young people's mental health, and (5) the Experience of Service Questionnaire [42] to assess young people's experiences within mental health services. In the intervention condition, young people were then provided instructions on how to download and use the app. After 3 months, all participants and clinicians were recontacted by the researchers and asked to complete the same questionnaires. Participants in the intervention phase were also asked if they would like to participate in a short semistructured interview about the acceptability of Power Up.

For the schools strand of the trial, a cluster randomized design was employed. Students in 12 clusters (classes) across 2 schools were randomized at the class level to either receive the app or not. This study was given a favorable opinion by University College London Research Ethics Committee (reference number: 6087/006). Randomization was achieved by using random number generation resulting in 6 intervention arm clusters, with 44 students allocated to receive Power Up (average cluster size 6.50 (SD 1.87)). The remaining 6 clusters were randomized to the control arm, with 50 students allocated to receive management as usual (average cluster size 7.50 (SD 4.04)). Researchers explained the nature of the research to the students before inviting them to take part. Participants gave their written consent and completed a questionnaire containing a similar battery of measures compared with that of the specialist services group, but excluding the shared decision-making measures: (1) the Patient Activation Measure [4] to assess young people's empowerment and self-management of their mental health and well-being, (2) the Strengths and Difficulties Questionnaire [41] to assess young people's mental health, (3) the Short Warwick-Edinburgh Mental Well-Being Scale [43] to assess young people's well-being, and (4) the child-friendly version of the EuroQol five dimension quality of life measure [44] to assess young people's quality of life and to inform the feasibility of conducting health economic analysis. Subsequently, those randomized to the intervention arm were given verbal instructions on how to use and download the Power Up app. After 6 weeks, all participants were contacted by researchers and asked to complete the same questionnaire measures. Participants in the intervention arm were also invited to participate in a short semistructured interview about the acceptability of Power Up.

Intervention

Power Up was developed based on the theory of patient activation [4] and a scoping review of tools to support young people to be actively involved in decisions about their care [45]. Power Up was designed to be a transdiagnostic and transtherapeutic intervention. To ensure it was accessible to young people, Power Up was co-designed with young people, carers, and clinicians through patient and public involvement workshops and interviews (see [46] for full details on the development of Power Up). A key topic of the codesign sessions was to ensure Power Up was simple and easy to use requiring minimal cognitive load so it could be used by a range of young people with different language skills, literacy levels, and experience of current distress. A mixture of text and icons is used and users can customize the iconography as straight-lined (aimed at older age ranges) or cartoon-style (aimed at younger age ranges). The main features of Power Up are described in Textbox 2 and Figure 1.

Textbox 2. Main features of Power Up.

- My People: At the center of Power Up, young people can add people in their support network. Users can flag information entered in other sections of the app to specific people in their support network, and all content flagged for sharing with a specific person from other areas of the app is displayed in My People. If an entry is not flagged to My People, it is stored chronologically; otherwise, a young person can prioritize which entries should appear first.
- My Diary: A space for users to express what is going on for them in their daily lives.
- My Plans: A section devoted to adding all plans and goals, including what to do in specific circumstances related to users' mental health, such as anxiety-provoking situations.
- My Questions: Young people can enter any questions they have or wish to discuss with carers, friends, teachers, or clinicians and keep a record of the answer after it has been discussed.
- My Decisions: A space for young people to work through decisions, weighing up the pros and cons associated with decisions using a visual weighted scale.
- There is the option for all entries in the above sections to be inputted in the form of photo, video, audio, or text.
- Entries such as photos or phone numbers can only be viewed within Power Up, not in the phone's main library or phonebook.
- Help and Support: A selection of resources that give the young person a series of links to websites and phone numbers. There are a set of prestored resources, however, the young person can also add his or her own. These can be called or visited directly from the app.
- Power Up is secure and password protected.

Figure 1. Power Up screenshots.

Statistical Analysis

To inform the planning of a prospective cluster randomized controlled trial, participant recruitment and retention was captured using the Consolidated Standards of Reporting Trials guidelines. Descriptive statistics were analyzed using SPSS (IBM Corp) [47]. Posttrial interviews were analyzed using thematic analysis [48] in NVivo (QSR International Pty Ltd.) [49].

Results

Recruitment and Retention

Recruitment and retention are reported in the Consolidated Standards of Reporting Trials diagram in Figure 2. During recruitment, 270 young people were assessed for eligibility. A number of young people were screened for participation but not recruited, because of reasons such as refusal and practical barriers such as not having enough time on the day of the research team's visit; for example, some young people and carers in specialist services were interested in the study but not able to stay to discuss the study as their parking was due to expire after their appointment. A further small proportion was excluded from the trial because of failing to complete all fields on their informed consent forms or failing to return study materials. In total, 142 participants were recruited and completed Time 1 measures (specialist services: n=62, mean age 14.66 (SD 1.99) years, 52% (32/62) female, 42% (26/62) white or white British; schools: n=80; mean 16.88 (SD 0.68), 46% (37/80) female, 26% (21/80) white or white British; see Table 1 for full demographic details). Of those who completed Time 1 measures, 64 (45%, 64/142) adhered to the study protocol and completed Time 2 follow-up assessments. All specialist services and school sites were retained in the study (N=10).

Figure 2. Consolidated Standards of Reporting Trials diagram: combined flow of participants through the study across both specialist services and schools strands of the trial.

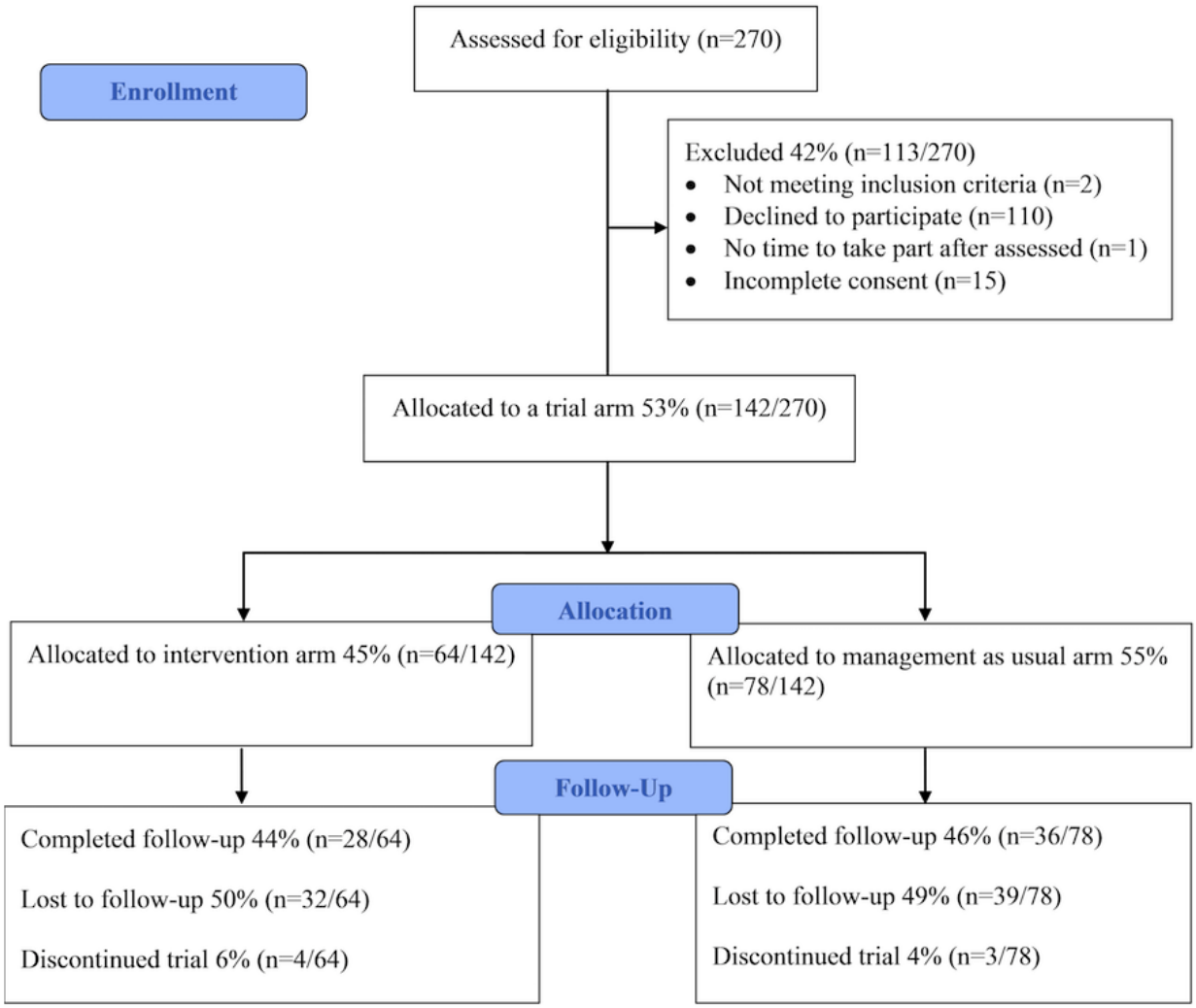


Table 1. Participant demographics.

Participant demographics	Child mental health services (n=62)	Schools (n=80)
Age (years)		
Mean (SD)	14.66 (1.99)	16.88 (0.68)
Range	11-18	16-18
Gender, n (%)		
Female	32 (52)	37 (46)
Male	26 (42)	40 (50)
Other	2 (3)	3 (4)
Prefer not to disclose	2 (3)	—[a]
Ethnicity, n (%)		
White	26 (42)	21 (26)
Black	7 (11)	14 (18)
Asian	15 (24)	27 (34)
Mixed	12 (19)	5 (6)
Other	1 (2)	12 (15)
Not reported	1 (2)	1 (1)
English as first language, n (%)		
Yes	52 (84)	52 (65)
No	7 (11)	26 (33)
Prefer not to disclose	—	1 (1)
Not reported	3 (5)	1 (1)
Registered as disabled, n (%)		
Yes	3 (5)	1 (1)
No	59 (95)	77 (96)
Prefer not to disclose	—	2 (3)

[a]No participants selected this response.

User Statistics

We could only determine a number of new users during the entire project time frame from January 2017 to February 2018. The number of active sessions and duration of sessions were available between November 2017 and February 2018 as the app developers upgraded the activity data capture system during the feasibility trial. App usage data are anonymous to comply with data protection and research ethics approvals. Overall, there were 70 new users between January 2017 and February 2018, of which we estimate 20 were nonintervention arm participants (ie, 5 members of the research team, 2 app developers, and 13 clinicians) resulting in an estimated 50 users of Power Up in the intervention arm, out of 64 participants allocated to this arm. There were 13 active users between November 2017 and February 2018 and these users used the app in 89 active usage sessions (corresponding to 6.8 sessions per active user) with an average of 8 min per session.

Posttrial Interviews

Posttrial interviews were conducted with 11 young people from the intervention arms of the feasibility trial (specialist services n=6; schools n=5; mean age 15.55 (1.86), range=11-17 years). Interviews were audio recorded and transcribed verbatim. Young people described their experiences of using Power Up and of its impact on self-management and well-being, including the context of app use and suggested amendments to the app. The analysis also provided understanding as to barriers faced by young people to downloading and using the app. Finally, interviews indicated the acceptability of Power Up, which will inform the prospective cluster control trial. Findings from the interviews are reported below relating to (1) motivation for use of Power Up, (2) impact of use, and (3) barriers to use (the themes, descriptions, and quotes from the posttrial interviews are fully reported in Textbox 3). We aimed to recruit 10 to 12 young people as we expected this would be sufficient for saturation, which was achieved. We report on these interviews in this study and discuss the views of clinicians, parents and carers on the development of Power Up elsewhere [46] and we have reported in detail on teachers' views and experiences of the implementation of another digital intervention [50].

Textbox 3. Themes, descriptions, and quotes from posttrial interviews.

1. Motivation for use

- Young people outlined different motivations for using Power Up, pertaining to the unique qualities of the app
  - "...I still did find myself using it, especially for the different aspects, like, how you can put key phone numbers on there, as well as, um, as I mentioned before, how you can weigh up decisions and so on, those, those are good aspects, which you wouldn't be able to do elsewhere…" [Young Person, aged 17 years]; "I used the Diary the most, like; I have quite a few, because I felt it easier to get my feelings down..." [Young Person, aged 14 years]; "I think that if you, like – the way that the Decisions work and the Plans work, it's really – it's good because it weighs it up for you." [Young Person, aged 14 years]; "Um, just if you have, like, any thoughts, so I got bad test results, so I, kind of, logged that, just so, like, I can, like, put it away and then reflect on it when I needed to." [Young Person, aged 17 years]; "During the holidays, I used the My Plans bit, so what I'm going to do. What I, like, aspirations that I want to do over the holiday." [Young Person, aged 16 years]

- The ability to make entries into the app via multiple modalities was evaluated positively by young people
  - "The decision making part was a key factor as well as diary entries, because I'm not a fan of writing, really, so typing it actually was a nice change for once." [Young Person, aged 17 years]; "I like the video function, definitely, because writing in a diary or something, you actually have to get a pen out and it takes time." [Young Person, aged 17 years]

- Young people expressed a need to be able to trust that the technology was a private platform, which could not be accessed by others, unless users chose to share it with them. Power Up represented an essential, secure space for users, allowing young people to feel comfortable to make entries, which were potentially personal and sensitive in nature
  - "Um, I liked that it was locked, because my parents have a tendency to go on my phone." [Young Person, aged 16 years]; "I guess the best word to describe it would be, I think, safe, because you can say honestly, you can say honestly anything you want, things you wouldn't share with other people, plans that you want to make for the better that you don't feel like telling everyone or your parents and it's a bit of, like, some alone time because some people work best when they're by themselves." [Young Person, aged 17 years]; "I like the fact that the PIN, because I feel that's really, really important because for young people, some aren't really open about how they feel and how…for someone that is worried the fact that they have a PIN on there is that no one can get in…If they did want to keep it secret, it's really discreet, that's what I like." [Young Person, aged 16 years]; "I think it's quite, it was, it was very discreet so when using it I didn't feel like, I didn't feel really weird...because there's a, there's a PIN, no one can really get into it, so I won't be worried about anyone seeing anything." [Young Person, aged 16 years]

- Beyond the inbuilt properties of the digital tool, the interviews elucidated several reasons for young people choosing to use Power Up. The app aided young people to remember important things to either share in a consultation with a health care professional, or to reflect on their thoughts over time
  - "I really loved the app…it's been so helpful with my appointments because I tend to have big gaps between my appointments… Of three to four weeks, and so I can hardly remember anything and so it was really easy for me to just pick up on little things and then, when I actually get to my appointment, I can just go in, like, chronological order and it was just amazing to be able to do that so I loved it." [Young Person, aged 16 years]; "Ah, I think the diary bit is good because when you're—when you're in that state of mind, you—my mind tends to run away with itself so it's good to look back on what I've thought about and what I've said." [Young person, aged 14 years]; "...because when I write the plans, um, I can write here and I can make sure I do it, because I mostly work in the plans. And in the diary I can look through back and see all the, all the stuff that happened before." [Young Person, aged 14 years]

- Young people also sought out Power Up when they could not speak to anyone else about their concerns and emotional experiences. Respondents reported using the app to express their feelings when they were alone
  - "Uh, well my first year score was quite bad, so when I came back home I, my parents were out, they weren't at home, so I just, like, used the app, like, [inaudible] I thought, like, a good way to get everything out, like immediately...And it just, it helped when I was frustrated and I didn't like, I felt there was no one to listen to me, but the phone will, so I just threw everything on there." [Young Person, aged 14 years]; "I think it normally be the later at night because as, you know, night can sometimes be a lonely time for most people…it's better than just bottling your thoughts up, you can actually, you'll have your phone, you can go on it and no one needs to know about it, you can just put anything down in the confidential app and it's just good to have that, especially later at night." [Young Person, aged 17 years]; "I don't have anyone that I can talk to, so writing it down, sort of, gives me a bit of relief, so I can just, like, release everything that I'm feeling and something that is written down, I can—it's there and I can forget about it until I actually have to talk about it. So that was really good." [Young Person, aged 16 years]

- Finally, the accessibility of Power Up in the moment was a key motivating factor for young people's engagement with the technology. Interviewees indicated that the portability of Power Up as an app on their mobile device was fundamental, meaning that they could access it immediately when required in a given situation
  - "I did also do a couple of ones for, like, plans how to deal with, like, situations… I'm very emetophobic, like as soon as I start feeling sick, like, yeah, that's it, so I did a couple of ones, like how I'm going to deal with it if it happens in a theatre, like how am I going to deal it with it happens in a restaurant, so I did a couple of those I think." [Young Person, aged 16 years]; "It's just good to have somewhere that's stored, like, if it was in a book, for example, you could, you wouldn't carry a book everywhere you go but your phone is with you roughly 24/7, so whenever you have anything it's all in this unit or place and it works very nicely." [Young Person, aged 17 years]; "Um, just to like put, so, like, if I had something that I was worried about, I'd, like, just talk myself through it, so then, I like, on the train home I can just go over what I'd do if that situation happened." [Young Person, aged 16 years]

2. Impact of use

- The analysis also illuminated the impact of using Power Up, from allowing young people to derive new insights from documenting their experiences, through to encouraging conversation with others in their support network. Young people explained how using the app had allowed them to see changes in their emotional state over time, marking progression in their psychological journey
  - "I could write it down and then I could look back and see how things have changed over, like, the week or the month and I could see, I went from a really bad phase to a really good phase and it was just, like, good to see progress through that, so that was great." [Young Person, aged 16 years]; "I've had a bit of a transition over these past few weeks, just mentally and everything regarding where I am at CAMHS. And so, so because also the fact that I have a replacement phone now, so after I get in contact with the new e-mail, I'll re-download it and I'll start actually afresh. Because then, almost like leaving my old self in the past, so hopefully I'll start with positive experiences and so on..." [Young Person, aged 17 years]; "I've gone through it on my own some nights when I just sit and go through and recall the things that have happened, and so, um, that's more of just me reviewing how much progress I've made and so it was really helpful in terms of that." [Young Person, aged 16 years]
- Participating young people identified that a further consequence of using the app was that it had helped them to gain greater understanding about themselves, as well as to clarify their thinking by weighing up the pros and cons of significant issues in a balanced, considered way
  - "...especially with Decisions and so on and the Diary entries because that, because they were there and it's in an easily accessible place, people do tend to go, go back on their thoughts a lot and because they have been written down it was very easily accessible to go back, change them and actually add them, so to, to, oh yes, get a deeper understanding of yourself." [Young Person, aged 17 years]; "Um, only when I was on My Diary and I was, like, looking through it and I had a few, like, you know, you can put little emojis, there was a few, like, sad ones, because I felt, like, I was using it more for when I was sad than, like, writing, oh I'm happy today. So I guess that was just more for myself, but then, like, I realised I was just, kind of, using it more for when I'm feeling emotional." [Young Person, aged 16 years]
- Young people highlighted that Power Up mediated communication with important people in their support network, facilitating conversation and helping them to share things with others, which they might not have otherwise
  - "I like that I can write everything, um, for myself and my mum can log into my account and check as well… (Young Person's parent: 'It's fine, I mean, there's things that he hasn't told me and then I've seen them, but he's writing them in there...') ...but I was happy that my mum can see all my diary entries and she can only see it." [Young Person, aged 11 years]; "I do have some, I do have someone in mind who I would mention it to, family wise, because they are going through a similar thing to me and they do, they are struck by boredom a lot, they have nowhere to really share their thoughts. And I'm the person they, kind of, go to, so I'm saying if I'm not always there, because I can't always be there for them, then there's this app. So if they want to, like, sit down and talk to me then that will always work and then we can actually share our thoughts on the apps what we have got and see how we are progressing, and get through it together." [Young Person, aged 17 years]; "I can allocate things to the, like, diary entries to the specific people and so afterwards, because I have, like, 54 diary entries, and so when I, if it got to a point where I want to speak to my therapist, I just click on, like, her name specifically, and all of the diary entries are allocated to her came up specifically." [Young Person, aged 16 years]

3. Barriers to use

- Despite young people describing largely positive experiences of Power Up, interviews indicated key barriers impacting engagement with the digital tool. Young people reported that a number of technological difficulties associated with the app occurred throughout the course of the feasibility trial
  - "So if you, so if you go to the Diary section and you stay in one journal without saving and you were just continuously writing thoughts as they flaring up, after five minutes, due to a safety feature implemented it would log you out and then that means that it wouldn't, um, there was no auto save or backup unfortunately, which means I would have to restart it." [Young Person, aged 17 years]; "It took a little bit longer to download than another app but I assume that wouldn't be, that wouldn't, kind of, affect it in its final form because you wouldn't need to download the test flight thing and remember code." [Young Person, aged 16 years]; "Sometimes it, well firstly it would take quite a long time to load the app. And it would, like it would, like, glitch with, like, colours of black and yellow and then it would go blue and then it comes up after that." [Young Person, aged 14 years]
- Barriers pertaining to context of app use were also raised. Young people reported disengaging from the app, because of experiencing difficulties in their personal circumstance
  - "Um, so I think I used it about, for like maybe the first month I just did, like, once every week. And afterwards, I think for me life got stressful and then I stopped using it..." [Young Person, aged 17 years]
- Additionally, it became clear that Power Up had not been fully embedded into mental health services and was seldom incorporated into clinical sessions between young people and their therapists, contrary to expectations
  - "I'm still being introduced into the whole process and I've only had two sessions, which would have been an entry session and a session regarding my medication. So possibly in future ones I'll ask for the integration of it, but I don't think there, there is any integration at the moment for some people who take CAMHS appointments and so on." [Young Person, aged 17 years]

Discussion

Principal Findings

The aim of this study was to determine the feasibility of examining the effectiveness of Power Up in a prospective cluster randomized controlled trial and to determine the usage and acceptability of Power Up and the ability to recruit and retain participants. A feasibility trial of Power Up was conducted in specialist services and school settings. The findings from this study indicate that it is feasible to examine the effectiveness of Power Up in a prospective trial: the overall score of the Go and No Go criteria was 2 and the Trial Steering Committee have agreed on the plan to increase adherence to the criteria that were partially met. For each criterion, we have reported achievement and described facilitators and barriers to achievement to inform planning of the future trial.

The total score of the Go and No Go criteria was 2, as there were 3 criteria fully met and 2 partially met. The study will be able to proceed as the Trial Steering Committee has agreed the plan to increase adherence to the 2 amber criteria. Recruitment and retention (criteria 2 and 5) will be increased by clarifying research sites' expectations from the outset, increasing ease of completing measures, and removing other barriers to participation (see Textbox 4). To further increase retention, barriers to downloading and using Power Up have also been removed by enabling participants to directly download the app from public app stores. As the app was still under development during the feasibility trial, Power Up was not fully available and participants were required to download Power Up using test software through the app developers; the new approach has removed this barrier. The developers have also upgraded the activity data capture system meaning we will be able to fully monitor adherence in the full trial.

The ultimate output of the project will be used by young people with emotional difficulties or other long-term conditions, aiming to empower self-management of problems. It is envisaged that Power Up will be developed for use with young people with other long-term conditions, and in other countries, however, future research is warranted to determine how Power Up should be modified for use with other conditions and in other contexts. Uptake of and engagement with novel digital technologies to support self-management of mental health difficulties and to promote decision making and well-being may be hindered if young people, their carers, teachers, and health professionals do not endorse them. The model of this study, cocreating Power Up with young people, carers, and professionals, will help to overcome this barrier by ensuring that Power Up meets the needs of these stakeholders and by providing young people with a sense of ownership in the knowledge that other young people have helped to create the resource.

Textbox 4. Adherence to the Go and No Go criteria.

1. Ability to recruit and retain sites
   - *70% to 100% of sites recruited and retained: 100% (N=10/10) of sites retained (green).*
     - Both specialist services and schools maintained high levels of engagement with the study. Clear communication about the requirements of the study from the outset was important to facilitate engagement, as was the research team attending frequent site visits, particularly in specialist services.
   - 50% to 69% of sites recruited and retained (amber).
   - 0% to 49% of sites recruited and retained (red).
2. Rates of sign up to study
   - 60% to 100% of young people signed up to participate (green).
   - *40% to 59% of young people signed up to participate: 53% (N=142/270) of participants recruited (amber).*
     - The reported recruitment rate from the school trial was conservative as our denominators were taken from overall class lists, not the number of students present on the day. In 1 school, this particularly deflated the recruitment rate as students from eligible classes were recruited from a morning assembly that was only attended by a small proportion of students.
   - 0% to 39% of young people signed up to participate (red).
3. Rates of download and usage of Power Up
   - *60% to 100% of young people download and use Power Up: 78% (N=50/64) new users between January 2017 and February 2018 (green).*
     - Barriers to downloading and using Power Up have been removed by enabling participants to directly download the app from public app stores (with a disclaimer that it is being used in a research trial). As the app was still under development during the feasibility trial, Power Up was not fully available and participants were required to go through a longer process to download Power Up using test software through the app developers; the new approach has removed this barrier. The developers have also upgraded the activity data capture system meaning we will be able to fully monitor adherence in the full trial.
   - 40% to 59% of young people download and use Power Up (amber).
   - 0% to 39% of young people download and use Power Up (red).
4. Young people report Power Up as acceptable in interviews
   - *70% to 100% young people report that Power Up is acceptable: 100% (N=11/11 young people) reported Power Up as acceptable in interviews (green).*
     - All young people interviewed reported the app as acceptable. Key themes from the interviews included young people's motivations for using the app, the impact of using the app, and barriers to use. Barriers identified related to downloading the app and technical bugs, which have been addressed by now enabling young people to download Power Up through app stores (with a disclaimer that it is being used in a research trial) and the commercial partner has resolved the technical bugs for the full deployment version of Power Up for the prospective trial. On the day of baseline assessments in 1 school, the app was not working during demonstration, likely impacting engagement and sign up (this issue has now been resolved). For participants with problems related to attention or anxiety, any technical bugs may have been of particular concern and likely to have caused frustration and disengagement.
   - 50% to 69% of young people report that Power Up is acceptable (amber).
   - 0% to 49% of young people report that Power Up is acceptable (red).
5. Completion of study measures at baseline and follow up
   - 50% to 100% of young people complete study measures at both baseline and follow up (green).
   - *20% to 49% of young people complete study measures at both baseline and follow up: 45% (N=64/142 young people) retained to follow up (amber).*
     - In addition to the above (see criteria 2 and 3) the research team has developed Web-based versions of the baseline and follow-up questionnaires to use during the trial reducing the burden on services and schools; electronic measures were also found to be preferable with a small number of young people who completed electronic follow-up measures in the feasibility trial. The research team has also developed videos to inform young people about the study and how to use the app. The research team is confident that the aforementioned approaches will increase retention, especially in schools as students will be able to complete follow-up measures in school or at home (eg, if they are absent on the day follow-up measures are administered). One of our main strategies to increase retention will be requiring services and schools to sign a Memorandum of Understanding (MoU) agreeing to the study, which will be useful to ensure clarity about what is expected of sites and researchers. This will also help to ensure that all senior leadership and individual staff members are aware of study before it starts. In signing the MoU, services and schools will be required to elect a key contact, increasing accountability.

In particular, research and well-being are high on schools' agenda, meaning the full trial is likely to remain a priority to schools. As reported in the interviews, clinician engagement with, and recommendation of, Power Up are also crucial to ensure young people's engagement (see Textbox 3). Ensuring young people, clinicians, and teachers fully understand the aims of the research, and that Power Up has been co-designed with young people, may also promote adherence to the study protocol.

- Clinicians have a pivotal role in the adoption of technologies in health care settings, as highlighted in a recent evidence framework proposed for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies (nonadoption, abandonment, scale-up, spread, and sustainability framework) [51]. Barriers such as competing priorities, busy schedules, and engagement with the intervention are particularly salient when considering clinicians' adoption of Power Up. Leadership and peer support for the importance of Power Up and the possible benefits for patients may help to keep Power Up a priority for clinicians. Training with a space for clinicians to work through how they would practically implement Power Up in *their* clinical practice with *their* patients may help facilitate easy integration in daily practice. Demonstrating the central role young people have played in the development of Power Up may help foster initial interest and sustained engagement over time with clinicians. Direct quotes from young people who have used it and found helpful are likely to be motivating also.

- 0% to 19% of young people and carers complete study measures at both baseline and follow-up (red).

Limitations

The scope of the feasibility trial was expanded to include young people from schools, in addition to young people from specialist services. Although this was based on feedback from young people, carers, and professionals that Power Up could be useful in settings beyond specialist services, it was also based on the slow recruitment rate from specialist services. We have developed plans to increase recruitment and retention for the full trial. This study was a feasibility trial, meaning definitive conclusions about the effectiveness of Power Up cannot be drawn. It was not possible to blind young people about their allocation in the feasibility trial, and it will not be possible to do so in the full trial. As with digital and psychotherapy research in general, a lack of allocation concealment and a reliance on self-report measures will be limitations that should be considered when interpreting any findings from future studies of Power Up. The Trial Steering Committee did consider using a restricted version of the app for the control condition, for example, only with the diary function. Although there is clinical equipoise, young people and clinicians indicated a strong preference not to include a *placebo* intervention, especially as young people may be unlikely to adhere to such a minimal intervention, minimizing the usefulness of comparisons. Nevertheless, in future research, independent randomization using a true randomization generator, intention-to-treat analysis, and monitoring and reporting of usage of Power Up will be methodological strengths. We will also be able to fully examine the relationship between a young person's characteristics (eg, age, gender, and presenting problems) and the effect of Power Up.

Conclusions

A feasibility trial of Power Up was conducted in specialist services and school settings. The findings from this study indicate that it is feasible to examine the effectiveness of Power Up in a prospective cluster randomized controlled trial: the overall score of the Go and No Go criteria was 2 and the Trial Steering Committee have agreed on the plan to increase adherence to the criteria that were partially met. This study addresses the call for feasibility trials of digital mental health interventions for young people [34,35]. Future research is needed to determine whether Power Up can be used by young people with emotional difficulties or other long-term conditions to empower them to self-manage difficulties.

Acknowledgments

The study was funded by the National Institute for Health Research (NIHR) Invention for Innovation Programme (i4i) project number: II-LA-0814-20005. CH and MC acknowledge the financial support of the NIHR MindTech MedTech Co-operative and NIHR Nottingham Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the National Health Service (NHS), the NIHR, or the Department of Health and Social Care. The research was supported by the NIHR Collaboration for Leadership in Applied Health Research and Care North Thames at Bart's Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. The research team acknowledges the support of the NIHR Clinical Research Network.

Conflicts of Interest

None declared.

References

1. Schaefer JD, Caspi A, Belsky DW, Harrington H, Houts R, Horwood LJ, et al. Enduring mental health: prevalence and prediction. J Abnorm Psychol 2017 Feb;126(2):212-224 [FREE Full text] [doi: 10.1037/abn0000232] [Medline: 27929304]
2. Green H, McGinnity A, Meltzer H, Ford T, Goodman R. Mental health of children and young people in Great Britain, 2004. Basingstoke: Palgrave Macmillan; 2005.

3. Patalay P, Fitzsimons E. UCL Institute of Education. London: Centre for Longitudinal Studies; 2017. Mental ill-health among children of the new century: trends across childhood with a focus on age URL: https://www.ncb.org.uk/sites/default/files/uploads/documents/Research_reports/UCL%20-%20NCB%20-%20Mental_Ill-Health%20FINAL.pdf [accessed 2019-01-31] [WebCite Cache ID 75qlTvbLv]
4. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4):1005-1026. [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
5. Department of Health. Department of Health. London; 2013. The Mandate: A mandate from the Government to the NHS Commissioning Board: April 2013 to March 2015 URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/256497/13-15_mandate.pdf [accessed 2019-01-31] [WebCite Cache ID 75qlh1AHo]
6. Hibbard J, Gilburt R. The King's Fund. London; 2014. Supporting people to manage their health: An introduction to patient activation URL: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/supporting-people-manage-health-patient-activation-may14.pdf [accessed 2019-01-31] [WebCite Cache ID 75qln337U]
7. Lindhiem O, Bennett CB, Trentacosta CJ, McLear C. Client preferences affect treatment satisfaction, completion, and clinical outcome: a meta-analysis. Clin Psychol Rev 2014 Aug;34(6):506-517 [FREE Full text] [doi: 10.1016/j.cpr.2014.06.002] [Medline: 25189522]
8. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]
9. Stevenson FA, Cox K, Britten N, Dundar Y. A systematic review of the research on communication between patients and health care professionals about medicines: the consequences for concordance. Health Expect 2004 Sep;7(3):235-245 [FREE Full text] [doi: 10.1111/j.1369-7625.2004.00281.x] [Medline: 15327462]
10. Abrines-Jaume N, Midgley N, Hopkins K, Hoffman J, Martin K, Law D, et al. A qualitative analysis of implementing shared decision making in Child and Adolescent Mental Health Services in the United Kingdom: stages and facilitators. Clin Child Psychol Psychiatry 2016 Jan;21(1):19-31. [doi: 10.1177/1359104514547596] [Medline: 25178946]
11. Couët N, Desroches S, Robitaille H, Vaillancourt H, Leblanc A, Turcotte S, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect 2015 Aug;18(4):542-561 [FREE Full text] [doi: 10.1111/hex.12054] [Medline: 23451939]
12. Coyne I. Consultation with children in hospital: children, parents' and nurses' perspectives. J Clin Nurs 2006 Jan;15(1):61-71. [doi: 10.1111/j.1365-2702.2005.01247.x] [Medline: 16390525]
13. Asarnow J, Jaycox LH, Tang L, Duan N, LaBorde AP, Zeledon LR, et al. Long-term benefits of short-term quality improvement interventions for depressed youths in primary care. Am J Psychiatry 2009 Sep;166(9):1002-1010. [doi: 10.1176/appi.ajp.2009.08121909] [Medline: 19651711]
14. Richardson L, McCauley E, Katon W. Collaborative care for adolescent depression: a pilot study. Gen Hosp Psychiatry 2009;31(1):36-45 [FREE Full text] [doi: 10.1016/j.genhosppsych.2008.09.019] [Medline: 19134509]
15. Edbrooke-Childs J, Jacob J, Argent R, Patalay P, Deighton J, Wolpert M. The relationship between child- and parent-reported shared decision making and child-, parent-, and clinician-reported treatment outcome in routinely collected child mental health services data. Clin Child Psychol Psychiatry 2016 Apr;21(2):324-338. [doi: 10.1177/1359104515591226] [Medline: 26104790]
16. Simons L, Valentine AZ, Falconer CJ, Groom M, Daley D, Craven MP, et al. Developing mHealth remote monitoring technology for attention deficit hyperactivity disorder: a qualitative study eliciting user priorities and needs. JMIR Mhealth Uhealth 2016 Mar 23;4(1):e31-e14 [FREE Full text] [doi: 10.2196/mhealth.5009] [Medline: 27009498]
17. Alegría M, Polo A, Gao S, Santana L, Rothstein D, Jimenez A, et al. Evaluation of a patient activation and empowerment intervention in mental health care. Med Care 2008 Mar;46(3):247-256 [FREE Full text] [doi: 10.1097/MLR.0b013e318158af52] [Medline: 18388839]
18. Cottrell D, Hill P, Walk D, Dearnaley J, Ierotheou A. Factors influencing non-attendance at child psychiatry out-patient appointments. Br J Psychiatry 1988 Feb;152(2):201-204. [doi: 10.1192/bjp.152.2.201] [Medline: 3167336]
19. Hoare P, Norton B, Chisholm D, Parry-Jones W. An audit of 7000 successive child and adolescent psychiatry referrals in scotland. Clin Child Psychol Psychiatry 1996 Jul 26;1(2):229-249. [doi: 10.1177/1359104596012005]
20. Kourany RF, Garber J, Tornusciolo G. Improving first appointment attendance rates in child psychiatry outpatient clinics. J Am Acad Child Adolesc Psychiatry 1990 Jul;29(4):657-660. [doi: 10.1097/00004583-199007000-00022] [Medline: 2387803]
21. Mathai J, Markantonakis A. Improving initial attendance to a child and family psychiatric clinic. Psychiatr Bull 2018 Jan 2;14(03):151-152. [doi: 10.1192/pb.14.3.151]
22. Stallard P, Sayers J. An opt-in appointment system and brief therapy: perspectives on a waiting list initiative. Clin Child Psychol Psychiatry 1998;3(2):-. [doi: 10.1177/1359104598032005]
23. Stern G, Cottrell D, Holmes J. Patterns of attendance of child psychiatry out-patients with special reference to Asian families. Br J Psychiatry 1990 Mar;156:384-387. [doi: 10.1192/bjp.156.3.384] [Medline: 2346839]

24. Mitchell AJ, Selmes J. Why don't patients attend their appointments? Maintaining engagement with psychiatric services. Adv Psychiatr Treat 2007;13(6):423-434. [doi: 10.1192/apt.bp.106.003202]
25. Montague AE, Varcin KJ, Simmons MB, Parker AG. Putting technology into youth mental health practice: young people's perspectives. SAGE Open 2015 Apr 15;5(2):215824401558101. [doi: 10.1177/2158244015581019]
26. National Institute for Health Care Excellence. National Institute for Health Care Excellence. London: NICE; 2013. Computerised cognitive behaviour therapy for depression and anxiety URL: https://www.nice.org.uk/guidance/ta97 [accessed 2019-01-31] [WebCite Cache ID 75qnnj3tT]
27. Office for National Statistics. 2016. Internet access – households and individuals, Great Britain: 2016 URL: https://tinyurl.com/zan568y [accessed 2019-01-30] [WebCite Cache ID 75odEoQq5]
28. Orange Exposure. Orange Exposure 2012/2013 research reveals stark constrast in mobile media habits between teenagers and adults. 2013. URL: http://exposure2012.orangeadvertisingnetwork.co.uk/pdf/exposure-2012-press-release-uk.pdf [WebCite Cache ID 75odMx3Wg]
29. Intersperience. Rising Internet use fules teen smartphone addiction. 2014. URL: https://blog.flexmr.net/ [accessed 2019-01-30] [WebCite Cache ID 75oddbWHD]
30. Statista. 2018. mHealth (mobile health) industry market size projection from 2012 to 2020 (in billion U.S. dollars) URL: https://www.statista.com/statistics/295771/mhealth-global-market-size/ [accessed 2019-01-30] [WebCite Cache ID 75odkGi4V]
31. Hides L. Are SMARTapps the future of youth mental health? Bulletin Aus Psychol Soc 2014;36:16-17 [FREE Full text]
32. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual research review: digital health interventions for children and young people with mental health problems-a systematic and meta-review. J Child Psychol Psychiatry 2017 Apr;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
33. National Institute for Health Research. National Institute for Health Research. 2017. Feasibility and pilot studies URL: https://tinyurl.com/y3ptvsdg [accessed 2019-01-30] [WebCite Cache ID 75oe5bPkk]
34. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 Dec 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
35. Grist R, Porter J, Stallard P. Acceptability, use, and safety of a mobile phone app (BlueIce) for young people who self-harm: qualitative study of service users' experience. JMIR Ment Health 2018 Feb 23;5(1):e16 [FREE Full text] [doi: 10.2196/mental.8779] [Medline: 29475823]
36. Avery K, Williamson PR, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P, Members of the Internal Pilot Trials Workshop supported by the Hubs for Trials Methodology Research. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. BMJ Open 2017 Dec 17;7(2):e013537. [doi: 10.1136/bmjopen-2016-013537] [Medline: 28213598]
37. Chapman L, Edbrooke-Childs J, Martin K, Webber H, Craven MP, Hollis C, et al. A mobile phone app to support young people in making shared decisions in therapy (Power Up): study protocol. JMIR Res Protoc 2017 Oct 30;6(10):e206 [FREE Full text] [doi: 10.2196/resprot.7694] [Medline: 29084708]
38. Elwyn G, Barr PJ, Grande SW, Thompson R, Walsh T, Ozanne EM. Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. Patient Educ Couns 2013 Oct;93(1):102-107. [doi: 10.1016/j.pec.2013.05.009] [Medline: 23768763]
39. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns 2010 Jul;80(1):94-99. [doi: 10.1016/j.pec.2009.09.034] [Medline: 19879711]
40. Walker JS, Thorne EK, Powers LE, Gaonkar R. Development of a scale to measure the empowerment of youth consumers of mental health services. J Emot Behav Disord 2009 Jun 05;18(1):51-59. [doi: 10.1177/1063426609337388]
41. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry 2001 Nov;40(11):1337-1345. [doi: 10.1097/00004583-200111000-00015] [Medline: 11699809]
42. Attride-Stirling J. Child Outcomes Research Consortium. London: Commission for Health Improvement; 2003. Development of methods to capture users' views of child and adolescent mental health services in clinical governance reviews (Project evaluation report) URL: https://www.corc.uk.net/media/1215/chi_projectevaluationreport.pdf [accessed 2019-01-31] [WebCite Cache ID 75qqzW7XD]
43. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007 Nov 27;5:63 [FREE Full text] [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
44. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res 2010 Aug;19(6):875-886 [FREE Full text] [doi: 10.1007/s11136-010-9648-y] [Medline: 20405245]
45. Cheng H, Hayes D, Edbrooke-Childs J, Martin K, Chapman L, Wolpert M. What approaches for promoting shared decision-making are used in child mental health? A scoping review. Clin Psychol Psychother 2017 Nov;24(6):O1495-O1511. [doi: 10.1002/cpp.2106] [Medline: 28752631]

46. Edridge C, Edbrooke-Childs J, Martin K, Delane L, Averill P, Feltham A, et al. Power Up: patient and public involvement in developing a shared decision-making app for mental health. J Health Design 2018 Mar 22;3(1):63-74. [doi: 10.21853/JHD.2018.40]
47. IBM. IBM Corporation. Armonk, NY: IBM Corp; 2012. SPSS Statistics 21.0 URL: https://www-01.ibm.com/support/docview.wss?uid=swg21608060 [accessed 2019-01-31] [WebCite Cache ID 75qrQi2sO]
48. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
49. QSR International Pty Ltd. QSR International. Warrington: QSR International Pty Ltd; 2012. NVivo qualitative data analysis URL: https://www.qsrinternational.com/nvivo/home [accessed 2019-01-31] [WebCite Cache ID 75qri7343]
50. Edridge C, Wolpert M, Deighton J, Edbrooke-Childs J. Implementation of ReZone: an mHealth intervention for young people to self-manage when feeling overwhelmed. JMIR Formative Res 2019 (forthcoming).
51. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Dec 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]

Abbreviations

CAMHS: Child and Adolescent Mental Health Service
MoU: Memorandum of Understanding
NHS: National Health Service
NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 25.07.18; peer-reviewed by J Oldenburg, S Rush; comments to author 06.09.18; revised version received 22.10.18; accepted 24.10.18; published 04.06.19.

Please cite as:
Edbrooke-Childs J, Edridge C, Averill P, Delane L, Hollis C, Craven MP, Martin K, Feltham A, Jeremy G, Deighton J, Wolpert M
A Feasibility Trial of Power Up: Smartphone App to Support Patient Activation and Shared Decision Making for Mental Health in Young People
JMIR Mhealth Uhealth 2019;7(6):e11677
URL: https://mhealth.jmir.org/2019/6/e11677/
doi:10.2196/11677
PMID:31165709

Original Paper

Using Stakeholder Values to Promote Implementation of an Evidence-Based Mobile Health Intervention for Addiction Treatment in Primary Care Settings

Andrew Quanbeck[1], PhD

Department of Family Medicine and Community Health, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Andrew Quanbeck, PhD
Department of Family Medicine and Community Health
University of Wisconsin-Madison
800 University Bay Drive
Madison, WI, 53705
United States
Phone: 1 608 609 7308
Email: arquanbe@wisc.edu

Abstract

Background: Most evidence-based practices (EBPs) do not find their way into clinical use, including evidence-based mobile health (mHealth) technologies. The literature offers implementers little practical guidance for successfully integrating mHealth into health care systems.

Objective: The goal of this research was to describe a novel decision-framing model that gives implementers a method of eliciting the considerations of different stakeholder groups when they decide whether to implement an EBP.

Methods: The decision-framing model can be generally applied to EBPs, but was applied in this case to an mHealth system (Seva) for patients with addiction. The model builds from key insights in behavioral economics and game theory. The model systematically identifies, using an inductive process, the perceived gains and losses of different stakeholder groups when they consider adopting a new intervention. The model was constructed retrospectively in a parent implementation research trial that introduced Seva to 268 patients in 3 US primary care clinics. Individual and group interviews were conducted to elicit stakeholder considerations from 6 clinic managers, 17 clinicians, and 6 patients who were involved in implementing Seva. Considerations were used to construct decision frames that trade off the perceived value of adopting Seva versus maintaining the status quo from each stakeholder group's perspective. The face validity of the decision-framing model was assessed by soliciting feedback from the stakeholders whose input was used to build it.

Results: Primary considerations related to implementing Seva were identified for each stakeholder group. Clinic managers perceived the greatest potential gain to be better care for patients and the greatest potential loss to be cost (ie, staff time, sustainability, and opportunity cost to implement Seva). All clinical staff considered time their foremost consideration—primarily in negative terms (eg, cognitive burden associated with learning a new system) but potentially in positive terms (eg, if Seva could automate functions done manually). Patients considered safety (anonymity, privacy, and coming from a trusted source) to be paramount. Though payers were not interviewed directly, clinic managers judged cost to be most important to payers—whether Seva could reduce total care costs or had reimbursement mechanisms available. This model will be tested prospectively in a forthcoming mHealth implementation trial for its ability to predict mHealth adoption. Overall, the results suggest that implementers proactively address the cost and burden of implementation and seek to promote long-term sustainability.

Conclusions: This paper presents a model implementers may use to elicit stakeholders' considerations when deciding to adopt a new technology, considerations that may then be used to adapt the intervention and tailor implementation, potentially increasing the likelihood of implementation success.

Trial Registration: ClinicalTrials.gov NCT01963234; https://clinicaltrials.gov/ct2/show/NCT01963234 (Archived by WebCite at http://www.webcitation.org/78qXQJvVI)

(JMIR Mhealth Uhealth 2019;7(6):e13301) doi:10.2196/13301

KEYWORDS

implementation models; implementation strategies; mHealth; behavioral economics; game theory; decision-framing; primary care; stakeholder engagement

Introduction

Context

The vast majority of practices shown to be effective by research remain unused in health care. It takes an estimated 17 years for an evidence-based practice (EBP) to be used in clinics, but only 14% of EBPs ever make it into use [1]. Mobile health (mHealth) technologies, in particular, hold great potential to transform health care. The evidence base for mHealth is limited but growing, with some technologies having been proven effective in randomized trials [2-5]. As of 2019, the degree to which mHealth technologies have been successfully implemented and integrated into the mainstream health care system in the United States remains limited.

The focus of this paper is a novel model for implementation that can generally be applied to EBPs. The model was developed through an exploratory analysis conducted in the context of an mHealth implementation research trial funded by the US National Institutes of Health (NIH) [6], making the results especially relevant to mHealth adoption. The trial involved 3 unaffiliated primary care clinics that enrolled 268 patients with substance use disorders to use a common mHealth system named "Seva," a Hindi word that means "selfless caring." As of 2019, the Seva implementation trial was among the most comprehensive mHealth implementation research trials reported in the US health care system, thus providing an instructive context for examining the emerging topic of mHealth implementation research. Little is known about the values and expectations stakeholders have regarding mHealth implementation [7]; this information needs to be brought forth and examined. Previous implementation research has focused on implementation frameworks and strategies [8-10], including frameworks specifically related to mHealth [11,12] and frameworks related to the definition and use of specific implementation strategies [10]. Substantial work has also been done to create tools to assess organizational readiness for change [13]. The decision-framing model contributes something new: a systematic approach that addresses the interactions between an mHealth intervention and the specific health system leaders, staff, and patients being asked to implement it.

Theoretical Foundations

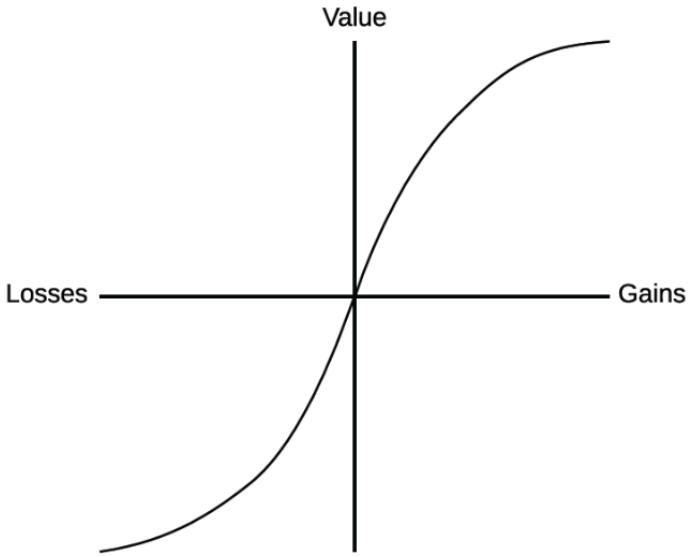
Fundamentally, implementing an EBP is a social process [14] involving human beings making decisions in the real world. Two areas of research provide important insights about the process of implementation but are rarely cited in implementation science: behavioral economics, which includes the concept of cognitive biases, and game theory. Both lines of research help explain how decision making works in the real world and why implementing an EBP is so challenging.

People frame their decisions on the basis of their own unique perspectives. Tversky and Kahneman's paper, "The Framing of Decisions and the Psychology of Choice," [15] defined a decision frame as "the decision maker's conception of the acts, outcomes, and contingencies associated with a particular choice. The frame that a decision maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision maker." This paper lays out a decision-framing model for the implementation of an EBP—an mHealth intervention—used in primary care.

Decision making involves not just a person's perception of the acts, outcomes, and contingencies related to a specific choice, but also a subjective evaluation that determines the perceived value associated with a given choice. The perceived value of a particular choice depends critically on each decision maker's unique perspective as a stakeholder in a health care system. In this context, *value* has a specialized meaning: It is the gains a person perceives in making a particular choice minus the losses the person perceives. Figure 1, which is adapted from Tversky and Kahnamen [15], illustrates the concept of perceived value in terms of the trade-off between perceived gains and losses.

This simple equation (wherein value equals perceived gains minus perceived losses) becomes complicated in light of Tversky and Kahneman's pioneering work showing that the everyday choices people make are typically not governed by rationality, as had been long assumed in classical economic theory [15]. Instead, the perception of a decision—the way a choice is framed—can influence the option a decision maker selects. For example, people are generally risk seeking when the consequences of a choice are framed in negative terms and risk averse when consequences are framed positively—people will make a risky choice to avoid losing money but lock in a choice that involves a monetary gain. Indeed, people's preferences can be reversed by the way a choice is put to them, and often be predicted by known cognitive biases (or heuristics). For example, agents of change (often, researchers) who encourage clinicians to adopt an EBP are likely to value the EBP more highly than prospective adopters do—sometimes because the agent has developed the practice himself or herself. This exemplifies a common decision-making bias called the *endowment effect*, that is, the tendency of individuals to ascribe inflated value to things for which they feel a sense of ownership. Decision makers are also highly *reference dependent;* inertia must be overcome to change from the status quo reference point that health care stakeholders begin with when considering a change in their routines. *Loss aversion* refers to decision makers' preference, in the face of uncertainty, for avoiding losses over acquiring gains. Research across several domains has shown that perceived gains must outweigh losses by a substantial margin for decision makers to favor changing from the status quo [16]. This implies that a change agent must convince a potential adopter that changing practice is going to be highly preferential when compared with maintaining the status quo.

Figure 1. Schematic representation of decision framing in terms of gains and losses (adapted from Tversky and Kahnamen [15]).



These decision-making biases are hard to address even among people well versed in them [17]. They combine to exert a powerful conservative force favoring the status quo with respect to clinical practice for both individuals and their work organizations.

Clearly, individual choices favor the status quo. Compounding the issue, implementation involves many different individuals making choices—all of them with different decision-making considerations. Game theory provides a framework for organizing implementation as a series of decisions made by members of different health care stakeholder groups [18]. According to game theory, all games comprise 4 key elements: players, actions, payoffs, and information. The central idea of this research is that success or failure in implementation is the result of decisions made by a diverse set of health care stakeholders, all of whom bring different perspectives and values to the decision of whether to adopt a given intervention. In health care, different stakeholders—payers (such as insurance companies), clinic managers, clinicians, and patients—are constantly confronted with choices about whether to adopt EBPs. Implementing an EBP may be conceptualized as a dynamic decision-making process that requires serial cooperation from all of these stakeholders to be successful.

In EBP adoption, the players can be defined as the 4 stakeholder groups named above (payers, clinic managers, staff, and patients); the players' actions are either to adopt or resist adopting the EBP; and each player's payoff corresponds to the perceived net benefits of adopting or not adopting *from that player's perspective*.

At each stage of implementation, different stakeholder groups make informal assessments of the value of adopting the EBP. Abstaining from decision making by failing to participate in the implementation process is common and tantamount to not adopting. If members of a stakeholder group do not perceive that they will benefit significantly by adopting, they may choose not to adopt and maintain the status quo instead. For example, if management promotes an EBP that staff members find onerous, staff will likely not adopt unless they are strongly compelled to adopt. The serial cooperation required for successful implementation will be broken at this level, and patients will not have access to the EBP because their access depends on the cooperation of clinical staff. Staff members are acting rationally in this example, because, by not adopting, they are selecting the option with greater value *from their perspective*.

The game theory conception of implementation helps further explain the dismal statistics on implementation success cited at the start of this paper [1]. Generally, the decision of any stakeholder group not to adopt an EBP is likely to prevent the adoption of the EBP in the organization. Admittedly, this conception of the implementation process does not fully account for the complexity of implementation, which is in fact less linear and straightforward than the conception suggests. Behavioral economics and game theory both suggest that implementation is a complex human enterprise because it is a social process. The process involves people with different perspectives making decisions about the same choice—managing the implementation process is both science and art, with *a priori* odds overwhelmingly tilted toward failure.

Aims

This paper provides a systematic model that implementers and researchers may use to gather input from health care stakeholders whose cooperation is essential to the successful implementation of EBPs. The model was applied specifically to an mHealth implementation study, and it therefore offers insights specific to mHealth, in addition to a method for designing and operating an effective implementation strategy. Someone who wants to introduce and implement an EBP into a system would benefit from understanding the considerations—the gains and losses—that potential adopters perceive as they think about adopting new practices. These considerations can then be used to modify the intervention or the implementation strategy (or potentially both) to better align with the considerations expressed by potential adopters and improve the likelihood of implementation success. Multimedia Appendix 1 describes the steps to implement the model.

Methods

The Focal Evidence-Based Practice

The mHealth intervention analyzed is called Seva, an evidence-based mHealth intervention designed to help prevent relapse in people recovering from substance use disorders [19,20]. Seva offers patients a discussion board used anonymously (with code names) by patients in the study; interactive modules that teach self-regulation, problem solving, and other skills; and health tracking tools and tools for coping with challenging situations, such as cravings and high-risk situations (eg, relaxation exercises, strategies from cognitive behavioral therapy, and links to local 12-step meetings). Seva gives clinicians a Web portal with a Clinician Report containing longitudinal information generated by patients' self-reported data about their substance use and well-being (eg, sleep, depression). Seva (under the name A-CHESS) was proven effective in a randomized trial of patients leaving residential treatment for alcohol use disorders [20]. It is currently being tested in other substance use treatment contexts.

Ethics Approval and Consent to Participate

The study protocol was designated minimal risk and approved by the University of Wisconsin's Health Sciences Institutional Review Board (protocol number: 2012-0937-CP019). The parent study is registered with ClinicalTrials.gov (NCT01963234).

Setting and Participants

The parent study [6] introduced Seva in 3 Federally Qualified Health Centers, which are primary care clinics in the United States that offer both primary and behavioral health care services to patients regardless of their ability to pay. At each of the 3 clinics, staff and patient participants were recruited for this exploratory analysis as a convenience sample from the staff and patients who consented to participate in the parent implementation study. Individual and group interviews were conducted with these stakeholders to elicit values related to the adoption of Seva and other EBPs. These interviews occurred during clinic visits from February 24 to 25, 2016; August 2, 2016; and September 21, 2016. These dates roughly corresponded to the transitional period between the stages of active implementation and maintenance in the parent study's implementation plan. In total, 6 clinic managers, 17 clinical staff, and 6 patients were involved in the individual and group interviews. Table 1 shows characteristics of the patients and staff who participated in the interviews.

Eliciting Stakeholder Considerations

The model used to frame decisions around EBP adoption is based on procedures for eliciting stakeholder considerations and defining a decision-analytic structure described by Edwards et al in their 2007 text, *Advances in Decision Analysis* [21]. The process is represented in broad terms by Figure 2. The type of decision analysis described here relies on a process of inductive reasoning in which the decision analyst constructs a model of the decision-making process by interviewing stakeholders. This model formulation is a first step; the result may then be tested prospectively in subsequent research and refined as necessary.

The series of stakeholder interviews took place one-on-one and in group interviews with clinic managers, clinic staff, and patients. The decision analyst (AQ) explained the premise of decision framing: that different groups of stakeholders have different considerations and contexts depending on their role in the health care system and that these considerations bear on their decisions about adopting EBPs. The objective of the interviews was to elicit the considerations that could be translated into values and serve as the foundation of a decision-framing model for mHealth implementation from the perspective of different stakeholder groups. The interviews were semistructured and exploratory. A series of planned questions were asked to promote the discussion of key issues around implementation from each stakeholder perspective, followed up with probing questions to understand the ideas that participants expressed—the potential gains (pros or advantages) and losses (cons or disadvantages) derived from implementing and using Seva.

Clinic Manager Interviews

Clinic managers at each implementation site were interviewed one-on-one by the decision analyst. Clinic managers continually make decisions about whether to undertake new projects, such as implementing Seva. Such decision making often occurs in the context of formal meetings intended to establish consensus around organizational goals (eg, monthly board meetings). An initial question posed during one-on-one interviews with managers from each site was, "What factors do you consider in deciding whether to introduce a new EBP like Seva to the staff and patients in your organization?" This initial inquiry was followed with specific questions about the factors the manager named, as well as questions that arose in the context of the discussion. Follow-up questions included the following: "At the organizational level, is there a process for deciding what new practices to implement? What barriers did you face in introducing Seva to your clinic? What would make it easier for you to implement Seva?"

Clinic Staff Interviews

Teams of staff members who participated in the implementation of Seva were interviewed in a group setting at each of the 3 implementation sites. When it comes to adopting a new EBP, clinic staff members can usually choose to adopt the new practice (such as Seva) or maintain the status quo.

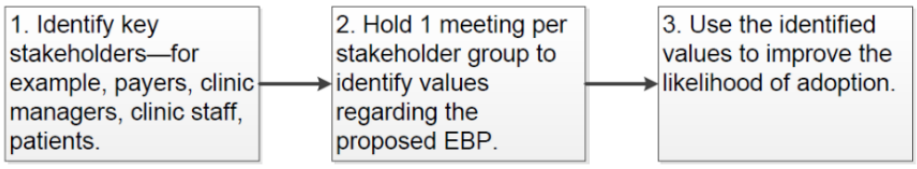
During group interviews with staff members, the decision analyst asked participants to reflect on the following question: "What do you think about when you are asked to do something new—a new procedure, a new technology, or some new evidence-based practice?" The decision analyst then gave staff members time to generate ideas individually. These ideas were then shared within the group in a round-robin fashion. After eliciting key considerations with respect to adopting new practices, the decision analyst used open-ended questions to expand on concepts presented by participants. Follow-up questions included the following: "In your different roles, how are you judged to be successful? Are metrics used (eg, number of patients seen, patient surveys, and other things)?"

Table 1. Baseline characteristics of participating clinics, clinic staff, and patients.

Characteristics	Site 1 (Madison, Wisconsin): Primary care and mental health	Site 2 (Missoula, Montana): Primary care, mental health, and addiction treatment	Site 3 (Bronx, New York): Primary care and mental health
Clinic staff and roles, n			
Participants	5	9	9
Manager	1	2	3
Physician	1	2	1
PhD psychologist	1	0	0
Therapist, counselor, or social worker	0	2	3
Care manager	0	2	1
Medical assistant	2	0	0
Clinic data manager	0	1	0
Other	0	0	1
Patients			
Participants, n	0	3	3
Age (years), range	—[a]	43-56	40-63
Gender (female), n	—	2	1
Highest education achieved, n			
Some high school	—	0	3
Some college	—	2	0
Associate's degree	—	1	0
Drug of choice, n			
Alcohol	—	2	0
Cocaine	—	0	1
Marijuana	—	0	2
Multiple drugs	—	0	1
Ethnicity, Hispanic/Latinx, n	—	0	1
Race, n			
White	—	3	1
African American/black	—	0	2

[a]Not applicable.

Figure 2. Decision-framing model. EBP: evidence-based practice.



Patient Interviews

Patients were interviewed in a group setting at both the Missoula and Bronx sites. (Owing to staggered implementation timing and turnover of a key staff member, patients at the first implementation clinic could not be reached for follow-up interviews). During these group interviews, patients were asked to reflect upon the considerations they had about adopting Seva and upon how Seva complemented other addiction treatment options, such as outpatient addiction treatment services offered by the clinic and traditional Alcoholics Anonymous/Narcotics Anonymous meetings. An initial question posed to patients was, "Assume you have a friend struggling with drug or alcohol problems who wants to know if you would recommend Seva to him or her. What would you say and why?" Each patient participant reflected on this question and shared responses,

prompting group discussion. Follow-up questions posed by the decision analyst included the following: "What problems (if any) did you have using Seva? To what extent was cost a barrier to your using Seva?"

Developing the Decision-Framing Model and Assessment of Face Validity

Input derived from this series of stakeholder interviews was used to establish the primary values (ie, trade-offs between perceived gains and losses) that governed stakeholders' decisions about the implementation of Seva. Considerations gathered through the interviews were systematically reviewed by the decision analyst and another researcher, and they were compiled into a decision-framing model that expressed the considerations as perceived gains and losses from the perspective of each stakeholder group.

The first-order approach to assessing the validity of any model focuses on face validity—that is, the degree to which the model concords with holistic judgments of validity by the stakeholders whose input was used to develop it [22]. After the initial decision-framing model was constructed, a draft version of the manuscript was emailed to 6 of the clinic managers and staff members from the 3 clinics whose feedback was central in development of the decision-framing model. (Owing to practical concerns, chiefly having to do with privacy, feedback from patients was not solicited.) The respondents were asked to independently rate the degree to which they thought the decision-framing model incorporated the most important values related to their implementation of Seva and whether it provided a reasonable representation of the implementation decision-making process. These 6 stakeholders, whose considerations informed the decision-framing model, were offered the opportunity to engage in follow-up phone and email correspondence to provide feedback on face validity.

Results

Results From the Parent Study

The subsequent results and discussion should be understood in the context of the parent implementation study [6], which showed that implementation and effectiveness outcomes were largely positive; management supported the use of Seva in all 3 clinics, 1 or more clinic champions emerged at each site to engage and support patients, and patients showed significant reductions in drinking and drug use. Adoption results were mixed; although patients adopted Seva with very high levels of use by normal mHealth standards, use of Seva did not penetrate primary care clinical processes beyond use by a handful of clinic champions at each site. Maintenance of the system was unsuccessful; each clinic's use of Seva ended when no long-term payer emerged to sustain the system after NIH grant funding ended.

Implementation Considerations Expressed by Stakeholder Groups

Table 2 summarizes the implementation considerations that emerged during individual and group interviews in the context of the choices that are available to each stakeholder group.

Table 2. Stakeholder implementation considerations.

Stakeholder group	Decision alternatives	Considerations: perceived gains and losses	Notes on implementation
Clinic managers	Support implementation of Seva versus allocate resources to competing projects	Gain: increased quality of patient care; Loss: additional clinical staff time required to implement and operate the intervention; Gain: advances organizational mission; Loss: uncertainty about sustainability potential of new intervention; Loss: opportunity cost of time for clinic champion to lead change efforts; Loss: lack of integration of new intervention into existing clinical workflows	Perceived gains were evident at outset. Clinics were compensated for staff time during grant period to offset costs. Management at all clinics supported introduction and use of Seva throughout the implementation period. Though management at 2 of 3 sites supported ongoing use of Seva, the challenges of transferring from grant funding to a long-term sustainable operational plan could not be successfully addressed, and system use ended at all 3 sites
Clinic staff	Adopt Seva or maintain status quo clinical practice for addiction	Loss: time required to learn and use a new system; Loss: disruption of current workflows, including integration with the electronic health record (EHR); Gain: improved quality of patient care; Loss: uncertainty about long-term sustainability; Gain: potential to automate clinical functions currently done manually	Seva was heavily used and valued by clinic champions, but penetration beyond clinic champions was limited. Failure to integrate Seva data into EHR made accessing Seva data infeasible for most clinicians
Patients	Use Seva (in addition to standard addiction treatment offered by the clinic) or continue with standard addiction treatment offered by the clinic or seek other treatment (eg, Alcoholics Anonymous)	Gain: access to a safe means of recovery support (anonymous and private, as well as coming from a trusted source); Gain: promotes access to resources and connections to similar others; Gain: promotes autonomy in recovery management (ie, voluntary use on patient's own time); Loss: cost to operate (including smartphone and data plan, covered by grant during intervention but transferred to patients after 12 months)	Patient out-of-pocket costs for Seva were paid with National Institutes of Health grant funding. Patient use during the study was high; use fell to zero when costs shifted to patients after grant funding ended. Logistical challenges made it difficult to transfer payment arrangements for data plans from the research team to individuals

Clinic Managers

In determining whether to adopt Seva, clinic managers considered the greatest potential gain (or *advantage* or *pro*) to be providing better care for patients and the greatest potential loss (or *disadvantage* or *con*) to be costs, which were expressed in such terms as staff time, sustainability, and opportunity cost required to implement the change. In 1 clinic, if a proposed EBP would be valuable to patients, it is assessed for its alignment with other values expressed in the clinic's mission—for example, the EBP should help build relationships with the community or foster the integration of medical, behavioral, and dental health services. This clinic follows a carefully designed process for deciding which innovations to adopt, using weekly management meetings of 5 clinic staff members. These meetings follow *Robert's Rules of Order*, a widely used parliamentary protocol for conducting meetings and reaching group decisions [23]. In reaching a decision, the group also actively solicits clinician feedback. Nevertheless, innovations that improve patient care and align with organizational mission face the constraint of cost, which is ultimately synonymous with sustainability.

According to 1 clinic leader who championed the Seva project, "If a program like Seva costs us money, on balance, it will be very hard for us to sustain. If it saves the organization money, it has a chance." As the director of behavioral health put it at another clinic, "We don't want to spend the resources to create something new that won't last." Many Federally Qualified Health Centers are accountable care organizations (or part of such organizations) that are responsible for the total cost of care for each patient. If an intervention like Seva helps patients maintain healthy and stable lives—and helps avoid the costly emergency room visits, detoxification stays, and hospitalizations often associated with addiction—the cost of implementing it may be worthwhile. If an innovation cannot be paid for by a grant or insurance reimbursement, it is, by definition, unsustainable, and an unsustainable innovation is not worth implementing. Managers also weigh costs in terms of staff time and integration with existing workflows, and small aspects of a proposed innovation can affect those costs. One management group uses a visual representation of clinic workflows to see how an intervention might fit because altering existing operations and workflow is expensive in labor and opportunity cost. As this manager put it, "How many hoops will staff have to jump through to do this?"

Clinic Staff

At all 3 clinics, virtually every clinic staff member interviewed cited time requirements as the foremost consideration in deciding to adopt a new practice. An intervention that costs staff extra time (expressed in terms of learning and using a new method or methods) is perceived negatively; a proposed intervention that might save staff time is perceived positively. An addiction psychiatrist whose patients used Seva shed light on what clinicians are trying to accomplish in the limited time they have with patients:

> *I've got 20 minutes with each patient every 6 weeks, if they show up. If you're my patient and you're not getting better, I want to know what is going on in your life. Are you not taking the medications I've prescribed? Are you using alcohol or other drugs? Are you having trouble with your family? Are you not sleeping? What's going on?*

This clinician values learning about a patient's problems as efficiently as possible—time is her most limited resource. An mHealth system like Seva has the potential to save clinicians time by continuously gathering and summarizing patient data so that they can quickly get an accurate picture of their patients' lives. Clinicians also said (like managers) that sustainability was important in their decision making, but to clinic staff, sustainability meant, in part, integration with the electronic health record (EHR). This EHR consideration arose in the group interviews because Seva data could not be integrated into the EHR, which caused inefficiency and frustration among clinicians. Clearly, the EHR is central to clinicians because it structures and monitors clinical work. This sets a high bar for EBP implementation because making changes to the EHR can be so difficult. In the health system that 1 clinic is part of, incorporating data into the EHR from external systems is reportedly so onerous that integration takes place only when changes are legally mandated (eg, a change to the EHR was enacted only after the state legislature passed a law requiring physicians to check the state's Prescription Drug Monitoring Program database before prescribing an opioid).

Patients

Because patients were asked about their decision making in the context of the Seva project, their responses reflected their thoughts about Seva and addiction treatment more specifically than the responses of managers and clinic staff, whose comments reflected their experience with the adoption of EBPs generally. Patients cited safety as their foremost consideration with regard to Seva. In the context of recovering from a substance use disorder, the concept of safety includes anonymity and privacy, as well as feeling confident that the innovation came from a trusted source (ie, on the basis of a recommendation from the patient's clinician). Patients chose to use Seva in part because they weighed it against specific alternatives—Alcoholics Anonymous or Narcotics Anonymous meetings or not receiving treatment—that felt less safe to them. Patients' comments also suggested that a successful innovation—all 6 patients interviewed regarded Seva as a valuable intervention—must promote connection to others and access to recovery resources. In both the rural setting of Missoula, Montana, and the urban setting of the Bronx, New York, patients reported feeling isolated. A patient in the Bronx said it was hard to get to Alcoholics Anonymous meetings, especially those held at night, because of concerns for her physical safety and the temptation to use drugs. In Missoula, getting to meetings was challenging because of difficulties with transportation. Seva addressed the isolation of these patients by enabling them to connect with peers and get help 24/7. A participant from the Bronx described how Seva helped him when he was on the brink of relapse. He decided to use Seva to reach out to one of the group's monitors (a member of the research team). "If I didn't have that phone, I don't know what would've happened," he reported. He also said the following:

> *Reaching out through Seva was the only thing that I could've done at that moment. I just needed someone to talk to; to listen to what was going on with me; to give me a push in the right direction.*

Finally, patients wanted to feel in control of their choices to use Seva rather than be coerced. "I like the fact that this is not something I'm forced to do," said one woman from the Bronx. She also said the following:

> *I can do it [use Seva] when I want to. This is my option. If I don't feel like listening, I won't listen!*

Face Validity

The 6 clinic managers and staff members who were consulted to provide face validity offered no corrections to the results presented.

Discussion

Principal Findings

This research conceptualizes the problem of implementing EBPs in a new way, borrowing key ideas from behavioral economics and game theory and integrating them with stakeholder feedback. This new conceptualization applies to the implementation of EBPs generally, but in this case, it was applied to one of the largest mHealth implementation trials to date, thereby specifically producing insights about mHealth implementation. The presumption of implementation researchers is that it is valuable to adopt EBPs *per se*. In truth, stakeholders have considerations, biases, and points of view that often limit the perceived usefulness of an EBP. It is the role of the implementer to frame decisions regarding intervention adoption for different stakeholders in the context of their considerations and values. This means presenting alternatives (ie, constructing subjective value functions) and adjusting the framing of the decision to maximize the probability of a positive choice for each stakeholder. The decision-framing model produced clear considerations that stakeholder groups used to evaluate the implementation of Seva, suggesting that decision-framing may be used to elicit multiple stakeholder perspectives. These considerations may in turn be used to (1) adapt the intervention to be implemented, or (2) tailor the implementation strategy used to deliver the intervention in ways that address stakeholders' most important considerations, or both. For example, the study revealed that finding a way to pay for Seva was essential to sustaining it after the study ended. If we had learned this before rather than after implementation, we could have tried to address this more robustly at the beginning of implementation. As an example of how stakeholder considerations can be incorporated into tailoring an implementation strategy, suppose clinical staff express concern about the additional time required to implement a change in practice. In response, an implementer could organize a meeting between management and clinicians to define the work required and then carve out dedicated time in a clinician's schedule—say, an hour every Friday morning—to commit to the work of implementation. This time expenditure would likely be viewed as a worthwhile investment by management if doing so leads to reduced hospitalizations for patients.

Payers are key stakeholders in health care systems, but payers were not directly interviewed for this study. In interviews with clinic managers and through interactions with a payer at 1 site, a single payer consideration was perceived as dominant: cost. An EBP may be perceived positively if it reduces the total cost of care—as Seva showed the potential to do through its effects on hospitalizations and emergency room visits [6]—or it may be perceived negatively if it increases the cost of care or has a cost that cannot be reimbursed.

The results of this inductively constructed model will now be tested prospectively, in a process using deductive reasoning, for its ability to predict the adoption of mHealth in a forthcoming implementation study funded by the US NIH (1R01DA04415901A1). In this test, the considerations reported here will be ranked (eg, clinic managers in the forthcoming trial will rank the 6 considerations reported by clinic managers in this paper), and then the decision alternatives will be rated (eg, clinic managers will rate how well the mHealth intervention addresses the considerations). See Figure 3 for an example section of the survey to be administered. The resulting rankings and ratings can then be weighted to determine how important they are to address, either by modifying the mHealth intervention or the strategy used to implement it. See the instructions for weighting considerations in Multimedia Appendix 1.

A nesting structure of values emerged through the interviews with stakeholders, both within each stakeholder group and between 1 stakeholder group and the next. For instance, management implied that for an intervention to be maximally appealing for adoption, it must first and foremost be valuable to patients, then palatable to staff, and then sustainable from a cost and reimbursement perspective. In a sense, management is implicitly incorporating the key values of patients, staff, and payers when deciding whether to approve implementation projects.

Figure 3. Illustration of prospective ranking and rating procedures.

Example Survey of Clinic Managers

Part I: Ranking decision-making considerations

In the context of your work, how important are the considerations below when you are deciding to adopt a new evidence-based practice? Please number each consideration, with 1 being the most important and 6 the least important.

_____ Likely sustainability of the new practice after the active implementation period ends
_____ Additional staff time required to implement and operate the new evidence-based practice
_____ Potential of the new practice to improve patient care
_____ Need to integrate the new practice into existing clinic workflows
_____ Potential of the new practice to advance our organizational mission
_____ Time required by a clinic champion to lead the implementation (vs work on a different project)

Part II: Rating decision alternatives

For each statement below, please circle the number that best reflects your response.

Compared with the status quo:

1. I think A-CHESS will improve the quality of patient care.

0	1	2	3	4	5	6	7	8	9	10
Strongly disagree					Neither agree nor disagree					Strongly agree

2. I think A-CHESS will require additional staff time to implement and operate.

0	1	2	3	4	5	6	7	8	9	10
A small amount					Neither a small nor large amount of time					A lot of time

3. I think A-CHESS will advance our organizational mission.

0	1	2	3	4	5	6	7	8	9	10
Strongly disagree					Neither agree nor disagree					Strongly agree

4. I think A-CHESS will be sustainable after the active implementation period ends.

0	1	2	3	4	5	6	7	8	9	10
Strongly disagree					Neither agree nor disagree					Strongly agree

5. I think a clinic champion will need to lead the implementation of A-CHESS, and this will mean that person cannot work on something else.

0	1	2	3	4	5	6	7	8	9	10
Strongly disagree					Neither agree nor disagree					Strongly agree

6. I think A-CHESS will need to be integrated into existing clinical workflows.

0	1	2	3	4	5	6	7	8	9	10
Strongly disagree					Neither agree nor disagree					Strongly agree

Comparison With Previous Work

Although many potentially useful instruments and frameworks are available from the implementation research literature to aid in the implementation process [13], the decision-framing model (1) provides a systematic approach for assessing the perceived value of an intervention from multiple stakeholder perspectives, (2) is concise and pragmatic, and thus suitable for widespread application (in contrast to more comprehensive, research-oriented models and questionnaires), and (3) offers an intervention-specific model that accounts for the complex interactions among organizational leaders, staff, the intervention itself, and patients. (Models of implementation [8] suggest that all these factors are relevant to implementation success.) Multimedia Appendix 1 provides practical guidance on how to use the decision-framing model in implementation and implementation research.

The need for systematic tailoring of implementation strategies has been identified as essential in the implementation research literature [24,25], but determining exactly *how* to conduct effective tailoring is still an understudied area. Decision framing can provide an organizing structure to gather information at the start of implementation—for instance, during the Exploration and Adoption/preparation stages of Aarons' framework [9]—as well as during active implementation. The process may yield useful information for tailoring implementation strategies on the basis of stakeholder values. External change agents, such as organizational coaches or facilitators, often lead implementation efforts. Research has shown that the effectiveness of facilitation can vary [26]. External change agents may benefit by using a systematic model, such as decision framing, because it provides a set of operating principles from which to orchestrate an implementation process that helps ensure that stakeholders have been involved and their values have been heard. Doing so could minimize the chance that implementation fails because of inconsistency in approaches to stakeholder engagement, a prospect that is virtually inevitable when left to the variability of human change agents.

Limitations

For logistical reasons, the decision-framing model was constructed retrospectively, at roughly the end of the implementation period at each site. Stakeholders' responses may have been different if the process had been undertaken before Seva was adopted. Prospective application of decision framing will take place in a forthcoming randomized trial, an NIH-funded implementation trial that was funded in 2018 (1R01DA04415901A1).

The face validity of decision framing was established in the context of a single study involving 1 type of health care setting (Federally Qualified Health Centers) and 1 EBP (an mHealth intervention for substance use disorders). Further research will be needed to validate the model and examine its usefulness with other interventions in other settings. The data reported also represent small samples, especially with only 6 patients interviewed, warranting caution about the generalizability of the findings.

Decision framing is a simple model that seeks to capture essential decision-making processes related to implementation research. More quantitatively robust decision-analytic techniques certainly exist (eg, multiattribute utility theory), but trade-offs are inevitable between pragmatism and research sophistication in selecting a model. Decision framing was selected in part because it is simple and intuitive enough for wider uptake.

Finally, decision modeling of any type invariably simplifies the complexity of any actual implementation process. Implementation does not always unfold in an orderly fashion, and assigning accurate weights to considerations can be difficult. For example, unreimbursed cost sealed the fate of Seva, despite patients' positive perceived value and the efforts of leadership in 1 clinic to find funding, and it may be that unreimbursed cost commonly plays such a role in implementation. In addition, the implementation of some practices may not require cooperation from all stakeholder groups—for example, patients may choose to use certain EBPs (such as mHealth apps) without any support or involvement from clinic management or staff. Demand for innovations can bubble up from patients and staff; indeed, such origins may bode more favorably for successful implementation than the top-down approach to implementation that is common in the health care system.

Conclusions

Though the decision-framing model is new to implementation research, the rationale for it is both simple and pragmatic: implementing an EBP is a fundamentally social process [8], and the inescapable biases associated with human decision making apply in implementation research just as they do in every other aspect of life. Decision-framing techniques have been exhaustively studied, validated, and applied in many fields, including psychology, business, and management. Innovation often lies in scanning many disciplines, making logical connections, and matching the most appropriate solutions available to the problem at hand. Newly applied to implementation research, decision-framing offers a potential tool for implementers to use in speeding the adoption mHealth interventions and other EBPs.

Acknowledgments

The author wishes to thank Chantelle Thomas, Mary Jane Nealon, Virna Little, Thomas McCarry, and Victoria Ward for their research collaboration. The author also wishes to thank David Gustafson, Randall Brown, John Mullahy, Barbara Bowers, Oguzhan Alagoz, and Ramon Aldag for their mentoring, Mark McGovern and Joann Kirchner for their valuable feedback on the manuscript, and Roberta Johnson and Nick Schumacher for their ongoing support.

Conflicts of Interest

AQ has a shareholder interest in CHESS Health, a public benefit corporation that disseminates Web-based health care intervention for patients and family members struggling with addiction. This relationship is extensively managed by the author and the University of Wisconsin–Madison's Conflict of Interest Committee.

Multimedia Appendix 1

Decision-framing to incorporate stakeholder perspectives in implementation.

[PDF File (Adobe PDF File), 524KB - mhealth_v7i6e13301_app1.pdf]

References

1. Balas E, Boren S. Managing clinical knowledge for health care improvement. In: Bemmel J, McCray AT, editors. Yearbook of Medical Informatics 2000: Patient-Centered Systems. Stuttgart, Germany: Schattauer Verlagsgesellschaft mbH; 2000:65-70.
2. Johnston L, Zemanek J, Reeve MJ, Grills N. The evidence for using mHealth technologies for diabetes management in low- and middle-income countries. J Hosp Manag Health Policy 2018 Jul;2:35. [doi: 10.21037/jhmhp.2018.07.01]
3. Peiris D, Praveen D, Johnson C, Mogulluru K. Use of mHealth systems and tools for non-communicable diseases in low- and middle-income countries: a systematic review. J Cardiovasc Transl Res 2014 Nov;7(8):677-691. [doi: 10.1007/s12265-014-9581-5] [Medline: 25209729]
4. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
5. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
6. Quanbeck A, Gustafson DH, Marsch LA, Chih M, Kornfield R, McTavish F, et al. Implementing a mobile health system to integrate the treatment of addiction into primary care: a hybrid implementation effectiveness study. J Med Internet Res 2018 Jan 30;20(1):e37 [FREE Full text] [doi: 10.2196/jmir.8928] [Medline: 29382624]
7. Wallis L, Blessing P, Dalwai M, Shin SD. Integrating mHealth at point of care in low- and middle-income settings: the system perspective. Glob Health Action 2017 Jun;10(sup3):1327686 [FREE Full text] [doi: 10.1080/16549716.2017.1327686] [Medline: 28838302]
8. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
9. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011 Jan;38(1):4-23 [FREE Full text] [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]
10. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
11. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
12. Wallis L, Hasselberg M, Barkman C, Bogoch I, Broomhead S, Dumont G, et al. A roadmap for the implementation of mHealth innovations for image-based diagnostic support in clinical and public-health settings: a focus on front-line health workers and health-system organizations. Glob Health Action 2017 Jun;10(sup3):1340254 [FREE Full text] [doi: 10.1080/16549716.2017.1340254] [Medline: 28838310]
13. Weiner BJ, Amick H, Lee SY. Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields. Med Care Res Rev 2008 Aug;65(4):379-436. [doi: 10.1177/1077558708317802] [Medline: 18511812]
14. Rogers E. Diffusion of Innovations. New York: Free Press; 2010.
15. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [Medline: 7455683]
16. Kahneman D, Lovallo D, Sibony O. Before you make that big decision. Harv Bus Rev 2011 Jun;89(6):50-60, 137. [Medline: 21714386]
17. Kahneman D. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux; 2013.
18. Myerson R. Game Theory: Analysis of Conflict. Cambridge, MA: Harvard University Press; 1991.
19. McTavish FM, Chih M, Shah D, Gustafson DH. How patients recovering from alcoholism use a smartphone intervention. J Dual Diagn 2012;8(4):294-304 [FREE Full text] [doi: 10.1080/15504263.2012.723312] [Medline: 23316127]
20. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
21. Edwards W, Miles R, von Winterfeldt D. Advances in Decision Analysis: From Foundations to Applications. New York: Cambridge University Press; 2019.
22. von Winterfeldt D, Edwards W. Decision Analysis and Behavioral Research. New York: Cambridge University Press; 2007.
23. Robert 3rd HM, Evans WJ, Honemann DH, Balch TJ, Seabold DE, Gerber S. Robert's rules of order newly revised. 11th edition. Boston, MA: Da Capo Press; 2011.

24. Powell BJ, Beidas RS, Lewis CC, Aarons GA, McMillen JC, Proctor EK, et al. Methods to improve the selection and tailoring of implementation strategies. J Behav Health Serv Res 2017 Apr;44(2):177-194 [FREE Full text] [doi: 10.1007/s11414-015-9475-6] [Medline: 26289563]
25. Lewis CC, Scott K, Marriott BR. A methodology for generating a tailored implementation blueprint: an exemplar from a youth residential setting. Implement Sci 2018 Dec 16;13(1):68 [FREE Full text] [doi: 10.1186/s13012-018-0761-6] [Medline: 29769096]
26. Ritchie MJ, Kirchner J, Parker L, Curran G, Fortney J, Pitcock J, et al. Evaluation of an implementation facilitation strategy for settings that experience significant implementation barriers. Implementation Sci 2015 Aug 14;10(S1):294 [FREE Full text] [doi: 10.1186/1748-5908-10-S1-A46]

Abbreviations

EBP: evidence-based practice
EHR: electronic health record
mHealth: mobile health
NIH: National Institutes of Health

Edited by G Eysenbach; submitted 24.01.19; peer-reviewed by P Dulin, E Szigethy, A Ramsey, J Lindsay, MY Chih, J Magee, T Giardina, S Gunther; comments to author 02.03.19; revised version received 26.04.19; accepted 11.05.19; published 07.06.19.

Please cite as:
Quanbeck A
Using Stakeholder Values to Promote Implementation of an Evidence-Based Mobile Health Intervention for Addiction Treatment in Primary Care Settings
JMIR Mhealth Uhealth 2019;7(6):e13301
URL: http://mhealth.jmir.org/2019/6/e13301/
doi:10.2196/13301
PMID:31237841

Original Paper

Weight Loss Following Use of a Smartphone Food Photo Feature: Retrospective Cohort Study

Daniela Ben Neriah[1,2], MS; Allan Geliebter[3,4], PhD

[1]Institute of Human Nutrition, Columbia University, New York, NY, United States

[2]Touro College of Osteopathic Medicine, New York, NY, United States

[3]Department of Psychology, Touro College and University System, New York, NY, United States

[4]Department of Psychiatry, Mount Sinai St. Luke's, Icahn School of Medicine at Mount Sinai, New York, NY, United States

Corresponding Author:
Daniela Ben Neriah, MS
Institute of Human Nutrition
Columbia University
630 W 168th St #1512
New York, NY, 10032
United States
Phone: 1 2125234000
Fax: 1 212 595 9441
Email: d.benneriah@caa.columbia.edu

Abstract

Background: Tracking of dietary intake is key to enhancing weight loss. Mobile apps may be useful for tracking food intake and can provide feedback about calories and nutritional value. Recent technological developments have enabled image recognition to identify foods and track food intake.

Objective: We aimed to determine the effectiveness of using photography as a feature of a smartphone weight loss app to track food intake in adults who were overweight or obese.

Methods: We analyzed data from individuals (age, 18-65 years; body mass index≥25 kg/m^2; ≥4 days of logged food intake; and ≥2 weigh-ins) who used a mobile-based weight loss app. In a retrospective study, we compared those who used the photo feature (n=9871) and those who did not use the feature (n=113,916). Linear regression analyses were used to assess use of the photo feature in relation to percent weight loss.

Results: Weight loss was greater in the group using the photo feature (Δ=0.14%; 95% CI 0.06-0.22; *P*<.001). The photo feature group used the weight loss app for a longer duration (+3.5 days; 95% CI 2.61-4.37; *P*<.001) and logged their food intake on more days (+6.1 days; 95% CI 5.40-6.77; *P*<.001) than the nonusers. Mediation analysis showed that the weight loss effect was absent when controlling for either duration or number of logged days in the program.

Conclusions: This study was the first to examine the effect of a food photo feature to track food intake on weight loss in a free-living setting. Use of photo recognition was associated with greater weight loss, which was mediated by the duration of app use and number of logged days in the program.

(JMIR Mhealth Uhealth 2019;7(6):e11917) doi:10.2196/11917

KEYWORDS

food intake; digital photography; app tracking; dietary assessment; free-living

Introduction

More than two-thirds of adults are considered to be overweight or obese in the United States [1]. Individuals with obesity are at an increased risk of developing cardiovascular disease, type 2 diabetes, and hypertension and are at a greater risk of mortality [2,3]. Weight loss can reduce the severity of these comorbidities or help prevent them [4,5]. Lifestyle modifications including calorie restriction can be effective in reducing body weight in the short term, although such approaches are less successful in the long term: Only about 20% of overweight individuals were successful at long-term weight loss, defined as losing at least 10% of the initial body weight and maintaining the weight for at least 1 year [6]. The current obesity epidemic has generated

a large market for weight loss programs. Commercial weight loss programs offer various aids, including mobile apps that allow tracking of food intake.

The majority of the US adult population (77%) reported owning a smartphone device in 2018 [7]. Tracking of food intake by using a mobile app can provide instant feedback about the calories and nutrients of the meal. Dietary self-monitoring is a key component in weight loss programs [8], and frequency of self-monitoring is strongly correlated with weight loss [9,10]. Some studies have reported a modest, although significantly greater, weight loss associated with the use of mobile apps compared to more traditional methods such as pen and paper [11-15], whereas others have found no difference [16-18]. Meta-analyses and systematic reviews of the recent literature also report mixed results, partly due to variation in the inclusion criteria [19,20]. In one meta-analysis, Flores-Mateo et al [19] reported significantly greater weight loss among people using mobile apps as compared to those using other methods, whereas Semper et al [20] reported no significant difference in their meta-analysis.

Self-monitoring requires time and effort, and many find tracking of dietary intake tedious, which contributes to attrition [21]. Obesity researchers and weight loss companies have attempted to improve the ease of tracking by offering new features. One such feature employs photography of food items to monitor dietary intake both in the clinical setting and everyday life [22-25]. Since the technology has only been developed recently, the literature on this subject is limited. However, a few studies have examined the effect of using photography to record food intake. A method developed by Martin et al [26], called Remote Food Photography Method (RFPM), uses semiautomatic computer analysis performed by researchers to obtain nutritional value from food images submitted by users. However, due to technological limitations, this process needs to be overseen by trained professionals [27,28]. Recent studies using RFPM conducted in pregnant women and preschoolers with obesity found that this method was not accurate in estimating energy intake compared to doubly labeled water [29,30]. Doumit et al conducted a cross-over study in college students, mostly of normal weight, to examine the effects of recording food intake from memory or with the aid of cell phone photos on energy intake and food choices [31]. They reported a nonsignificant trend for lower energy intake in the group using cell phone photos, suggesting increased awareness of food choice and portion size. The study did not assess weight change and was of a short duration, with 3 days for each assessment period. A retrospective cohort study examined factors related to the use of food photography with an app that promotes healthy eating [32]. Active users (with 10 photos or more) had, on an average, used the app for longer and had higher healthy ratings per photo than nonactive users (1 photo) and semiactive users (2-9 photos). However, this study did not monitor weight change.

Recently, a new tool, Snap It™, became available that allows participants to take photos of their food, and through image recognition technology attempts to match the food item to a large food photo database [33]. We investigated the effect of using a food image recognition feature as part of a mobile app for tracking food intake on weight loss. We hypothesized that the photo feature users would lose more weight than nonusers.

Methods

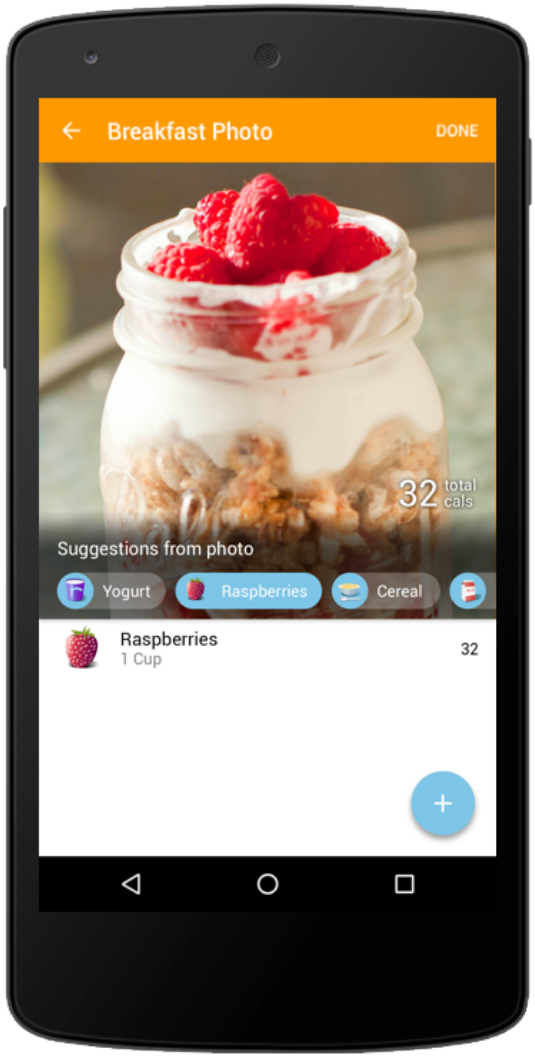
Mobile App

Lose It! is a free weight loss mobile app launched in 2008 by FitNow (Boston, MA), which allows users to record their daily food intake. Users enter their self-reported weight, height, gender, and age when they sign up. They choose their goal weight and the rate at which they would like to lose weight (0-2 lbs per week). The app allows users to record food intake and exercise, calculates calories consumed, and estimates calories expended. All users have a calculated calorie budget for the day and can see whether their intake was within their budget. Users can also choose to pay US $39.99 per year for the premium features, which offer meal and exercise planning, macronutrient tracking, and recipes. In November 2016, a new free feature was added to the app that utilizes a food image tool called Snap It, which recognizes food items and requests the user to confirm their meal items from a list of potential matching foods and to add estimated portion sizes. The feature became available to all users of the app, but only a fraction have used it. We investigated weight loss outcomes among those who used the app versus those who did not. The data were de-identified before they were provided to us by the company FitNow, which had no role in the development of the protocol, the interpretation of the data, or the preparation of the manuscript.

Photo Feature

The Snap It photo feature is shown in Figure 1 and is available to all app users at no additional cost. The user takes a photo of food items and then obtains a list of food items that the software recognizes as potential matches. The user subsequently chooses the correct match and specifies the quantity consumed from the list. Thereafter, the calories and macronutrients are calculated and displayed.

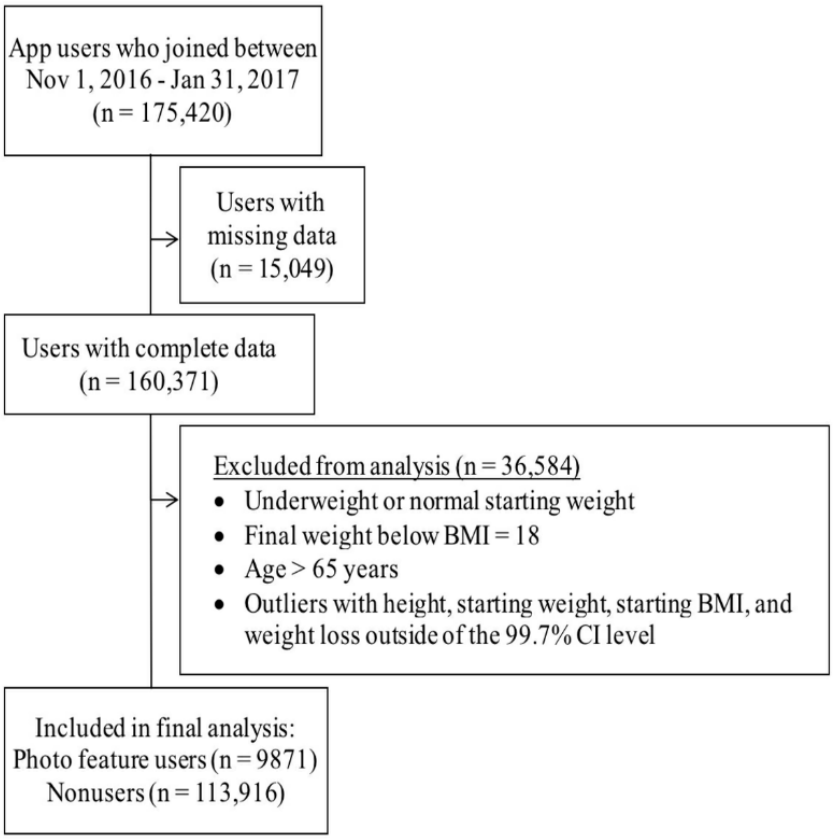
Figure 1. App photo feature. Users can take a snapshot of their food items before eating and select correct matches among a list of suggested items [33]. The calories and nutrients for the items are then displayed.



Study Sample

We received the dataset from FitNow, Inc, which included 175,402 users who joined between November 1, 2016, and January 31, 2017. Data collection for this sample ended in April 2017 and included individuals who had logged food intake in the app for a minimum of 4 days; had at least two reported weigh-ins; had a date of birth between January 11, 1936, and April 21, 1999; and reported >3 feet in height and a body weight between 70 and 750 pounds. The study was approved by the Touro College and University System Institutional Review Board (HSIRB# 1746E). We analyzed individuals who were overweight or obese, defined by a starting body mass index (BMI)≥25 kg/m^2 [34], with a final BMI≥18 kg/m^2 and age of 18-65 years. To exclude users with unrealistic data, we removed outliers outside the 99.7% CIs for height, starting weight, starting BMI, and %weight change. Only users who used an iPhone or an Android phone were included, thus excluding Web users, as the photo feature is not available on the LoseIt! website. Usage duration was calculated based on the number of days between the first and last reported weigh-ins, and the number of "logged days" was based on the number of days that intake was entered in the app. The groups were defined as photo feature users if they used the photo feature to log food intake for at least 1 or more days, and as nonusers if they did not use the photo feature at all. Only 7.8% (13,663/175,402) had used the photo feature, likely because the feature was new. We then excluded users who did not meet the inclusion-exclusion criteria or had missing or outlier data. We analyzed 123,787 users overall (Figure 2), grouped by those who used the photo feature (n=9871) and those who did not use the feature (n=113,916). Among the photo feature users, we examined weight change in relation to the number of days the photo feature was used to log food intake.

Figure 2. Flow chart of Lose It! app users in the study. BMI: body mass index.



Statistical Analysis

We applied regression analysis using RStudio (version 1.0.14, RStudio Team, Boston, MA) to compare the %weight loss for the photo users and nonusers as well as duration in the program and number of logged days. We included the covariates, starting BMI, gender, age, premium status, and type of user operating system (Android or iPhone). We then reanalyzed the %weight loss data and controlled for duration and number of logged days. Chi-squared tests and *t* tests were used for baseline characteristics and comparisons between groups. Results were considered significant if the *P* value was ≤.05. The SD or 95% CI was used to represent the variance around the mean.

Results

Study Sample

Table 1 shows the composition of the study sample and the baseline characteristics of the participants, including age, starting weight, starting BMI, type of operating system (Android or iPhone), premium subscription status, and goal weight. Nonphoto feature users were younger, had a lower BMI and a lower body weight, used iPhones more often (vs Android), and paid for the premium version of the LoseIt! app less often than photo feature users.

Table 1. Baseline characteristics and comparisons between groups.

Characteristic	Nonusers (n=113,916)	Users (n=9871)	*P* value
Age (years), mean (SD)	36.1 (12.1)	36.5 (11.9)	<.001
Starting weight (kg), mean (SD)	93.6 (20.5)	94.0 (20.6)	.051
Starting body mass index (kg/m^2), mean (SD)	32.7 (6.2)	33.1 (6.3)	<.001
Female, n (%)	83,150 (72.99)	7153 (72.46)	.26
Android, n (%)	28,041 (24.62)	3438 (34.83)	<.001
iPhone, n (%)	85,875 (75.38)	6433 (65.17)	<.001
Premium, n (%)	13,283 (11.66)	1834 (18.58)	<.001
Goal weight loss (kg), mean (SD)	8.7 (6.2)	8.9 (6.3)	.001

Weight Loss

Reported body weight decreased over time across the two groups, with greater weight loss in the photo feature group. The photo feature users lost 0.14% more (mean 3.42%; 95% CI 3.27-3.58) than the nonusers (mean 3.28%; 95% CI 3.24-3.32; *P*<.001), which translated to a 0.15 kg difference (photo feature users: mean weight loss=3.29 kg, 95% CI 3.14-3.45; nonusers: mean weight loss=3.14 kg, 95% CI 3.10-3.19). The difference in the %weight loss remained significant after adjusting for starting BMI, age, gender, user operating system, and premium status (*P*=.002), which were all significantly associated with the %weight loss. When adjusted for duration, photo feature use was not significantly associated with %weight loss (*P*=.40), and when adjusted for the number of logged days, the weight loss effect was reversed, where photo feature use was associated with an increase in body weight (+0.18 kg; *P*<.001). Duration and number of logged days were both significantly associated with %weight loss (*r*=0.33, *P*<.001 and *r*=0.46, *P*<.001, respectively). Collinearity between duration and number of logged days was not significant (variation inflation factor=1.37).

Within group analysis of photo feature users only showed that the number of days the photo feature was used was significantly associated with the %weight loss (*P*<.001). The difference was associated with 0.04% (95% CI 0.02-0.06) weight loss for every additional day of using the photo feature.

Premium Version

Of the 123,787 users analyzed, 15,117 (12.2%) were premium users and 1834 (12.1%) of the premium users used the photo feature. Premium users had a higher starting weight (96.4 kg; 95% CI 95.6-97.2) than nonpremium users (93.2 kg; 95% CI 93.0-93.4; *P*<.001). Premium users also lost more weight (mean 3.77 kg; 95% CI 3.60-3.93) than nonpremium users (mean 3.23 kg; 95% CI 3.19-3.27; *P*<.001). The difference remained significant even after adjusting for starting BMI, age, gender, and user operating system.

Duration

The photo feature group used the app for 3.5 days more than nonusers (photo feature users: mean 59.0 days, 95% CI 57.3-60.7; nonusers: mean 55.5 days, 95% CI 55.1-56.0; *P*<.001). The difference remained significant (3.2 days; *P*<.001) after adjusting for starting BMI, age, gender, user operating system, and premium status (*P*<.001).

Number of Logged Days

Photo users logged 6.1 more days than nonusers (users: mean 43.1 days, 95% CI 41.7-44.5; nonusers: mean 37.0 days, 95% CI 36.6-37.4). The difference remained significant (5.4 days, *P*<.001) after adjusting for starting BMI, age, gender, user operating system, and premium status.

Mediation Analysis

As reported above, %weight loss was significantly associated with photo feature use (*P*<.001). Duration was also significantly associated with photo feature use (*P*<.001) and %weight loss (*P*<.001). When controlling for duration in the program, the effect of photo feature use on weight loss became nonsignificant (*P*=.40), indicating that duration was a mediator (Figure 3). The number of logged days was also significantly associated with photo feature use (*r*=0.05; *P*<.001) and %weight loss (*P*<.001). The effect on %weight loss was reversed (ie, photo feature use was associated with an increase in body weight) when the number of logged days was controlled for (+0.18 kg; *P*<.001), indicating mediation (Figure 4). Thus, both duration and number of logged days were significant mediators of the photo feature effect on %weight loss.

Figure 3. Duration as a mediator. The photo feature use was significantly associated (P<.001) with weight loss before adding the potential mediator, duration, as a covariate. The photo feature use was also significantly associated with duration (P<.001), and duration was significantly associated with weight loss (P<.001). When adjusting for duration, the photo feature use was no longer significantly associated with weight loss (P=0.40), indicating that duration was a mediator.

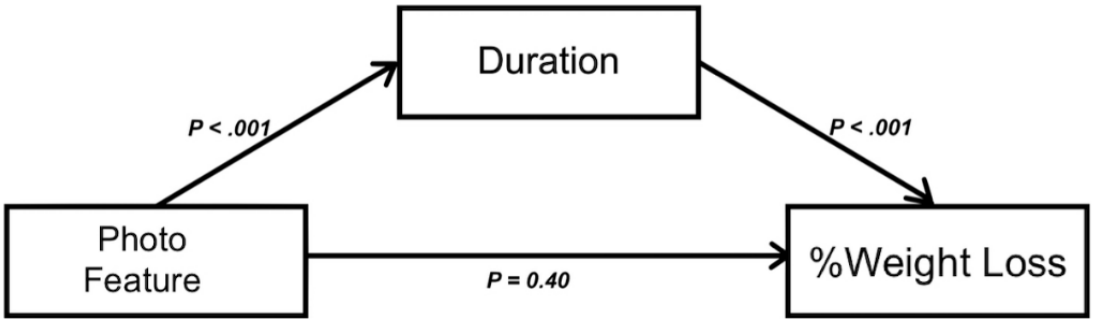
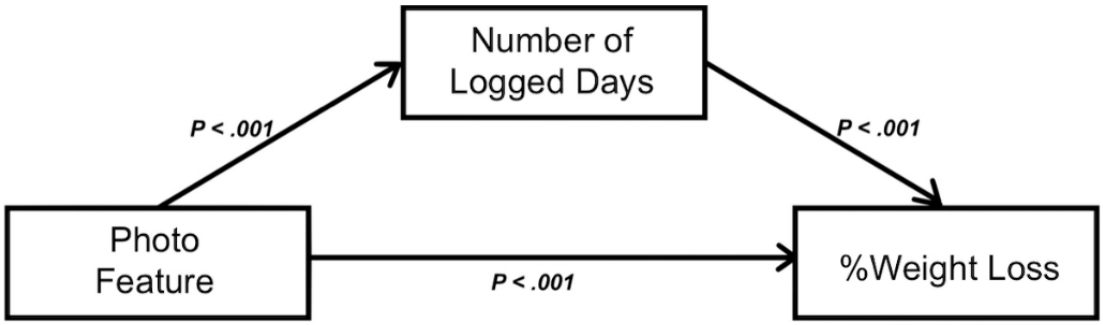


Figure 4. Number of logged days as a mediator. The photo feature use was significantly associated with weight loss (P<.001) before adding the potential mediator, number of logged days, as a covariate. The photo feature was also significantly associated with the number of logged days (P<.001), and the number of logged days was significantly associated with weight loss (P<.001). Direction of significance was reversed when adjusted for the number of logged days (P<.001), indicating that the number of logged days was a mediator.



Discussion

Primary Outcomes

In this study, we analyzed data from 123,787 users of a weight loss app, which included an optional photo feature. Photo feature users lost significantly more weight than nonusers after adjusting for starting BMI, age, gender, user operating system, and premium status, which confirmed our hypothesis. Use of the photo feature might be less time consuming, more motivating, and more interactive than typing in food items.

Although statistically significant, the difference in weight loss between the groups was not clinically significant, indicating only a very small benefit of capturing food images by a phone camera. In addition, the photo feature users had more logged days and longer duration in the program, which may be due to enhanced motivation derived from taking photos, such as greater awareness of choice of food and portion size. Photographing food might also improve memory of the food eaten, which has been associated with reduced food intake [35]. On the other hand, individuals who were more motivated initially may have opted to use the photo feature.

The weight loss effect was not significant when adjusted for duration in the program and was reversed when adjusted for number of logged days. Mediation analysis showed that the weight loss differences were mediated by duration and number of logged days. Since entering the number of logged days as a covariate reversed the direction of significance (ie, photo feature was associated with weight gain when adjusted for number of logged days), the number of logged days appears to more strongly mediate the effect than duration in use of the photo feature and weight loss. As previously shown, increased self-monitoring, such as that by recording food intake, correlates with weight loss [36]. Thus, as the photo users remained in the program for a longer duration and logged more days, they spent more time recording their food intake.

An additional possible benefit to recording food intake by taking photos as compared to typing it into an app is that the photo is taken before the food is eaten, whereas typing in a food item is customarily done at end of the meal or end of the day, based on memory [37]. Thus, taking a photo prompts the user to acknowledge the nutritional content of the food item before consuming it, which may have a greater influence on their choice of food and portion size.

The frequency of photo feature use, reflected by the number of days of use of the photo feature, was significantly associated with weight loss, indicating that the more the photo feature was used, the greater the amount of weight lost. The photo feature users' adherence, reflected in the number of logged days, was also higher than that of nonusers, suggesting that either the use of the photo feature motivated the users to continue with the program or that the users were more motivated to start with and used this feature as an additional tool. The latter theory is also supported by the increase in the proportion of photo feature users that subscribed to the premium features. As reported above, increased frequency of app use is correlated with increased weight loss and better maintenance [38]. Use of the photo feature was associated with increased weight loss, as hypothesized, possibly due to greater frequency of app use leading to a more successful outcome. This is the first study to examine the effect of a food image recognition app feature to track food intake on weight loss in a large cohort in a naturalistic, free-living setting.

Limitations

There were several limitations, including limited demographic information, which did not include income and race. This study was also based on self-reported values of body weight and height. Additionally, the groups were self-selected and the numbers were uneven. In future studies, we would add a survey to gather demographic information and feedback on the ease of use and motivation to use the app with the photo feature. We observed a significant correlation between %weight loss and photo feature use, but this observational study does not allow determination of cause and effect. The next logical step would be a randomized controlled trial to help determine causal direction.

Conclusions

We showed that use of a photo feature as part of a weight loss app was associated with greater weight loss, an effect mediated by increased duration and more logged days.

Acknowledgments

We thank Lose It!, a component of FitNow, Inc. In particular, we thank Sarah Molhan for providing us with the dataset. We also thank Katie Chase Martin for assistance with submission of the protocol to the institutional review board. Funding was provided by the Obesity Society through the Weight Watchers Karen Miller-Kovach Research Grant to AG.

Conflicts of Interest

None declared.

References

1. Fryar C, Ogden C. National Institute of Diabetes and Digestive and Kidney Disease. Overweight & Obesity Statistics URL: https://www.niddk.nih.gov/health-information/health-statistics/overweight-obesity [accessed 2018-08-08] [WebCite Cache ID 71WngdVvO]
2. 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]
3. Grieve E, Fenwick E, Yang H, Lean M. The disproportionate economic burden associated with severe and complicated obesity: a systematic review. Obes Rev 2013 Nov;14(11):883-894. [doi: 10.1111/obr.12059] [Medline: 23859626]
4. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
5. Wadden TA, Volger S, Sarwer DB, Vetter ML, Tsai AG, Berkowitz RI, et al. A two-year randomized trial of obesity treatment in primary care practice. N Engl J Med 2011 Nov 24;365(21):1969-1979 [FREE Full text] [doi: 10.1056/NEJMoa1109220] [Medline: 22082239]
6. Wing RR, Phelan S. Long-term weight loss maintenance. Am J Clin Nutr 2005 Jul;82(1 Suppl):222S-225S [FREE Full text] [Medline: 16002825]
7. Pew Research Center: Internet & Technology. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-12-26] [WebCite Cache ID 74wNkYQCN]
8. 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] [Medline: 21185970]
9. Wharton CM, Johnston CS, Cunningham BK, Sterner D. Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. J Nutr Educ Behav 2014 Oct;46(5):440-444. [doi: 10.1016/j.jneb.2014.04.291] [Medline: 25220777]
10. Kruger J, Blanck HM, Gillespie C. Dietary and physical activity behaviors among adults successful at weight loss maintenance. Int J Behav Nutr Phys Act 2006 Jul 19;3:17 [FREE Full text] [doi: 10.1186/1479-5868-3-17] [Medline: 16854220]
11. Linde JA, Jeffery RW, French SA, Pronk NP, Boyle RG. Self-weighing in weight gain prevention and weight loss trials. Ann Behav Med 2005 Dec;30(3):210-216. [doi: 10.1207/s15324796abm3003_5] [Medline: 16336072]
12. Chin SO, Keum C, Woo J, Park J, Choi HJ, Woo J, et al. Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. Sci Rep 2016 Nov 07;6:34563 [FREE Full text] [doi: 10.1038/srep34563] [Medline: 27819345]
13. Goldstein DJ. The Management Of Eating Disorders And Obesity (nutrition And Health). Totowa, NJ: Humana Press; 2005.
14. Allen JK, Stephens J, Dennison Himmelfarb CR, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. J Obes 2013;2013:151597 [FREE Full text] [doi: 10.1155/2013/151597] [Medline: 24392223]
15. 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;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
16. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. Obesity (Silver Spring) 2015 May;23(5):935-942. [doi: 10.1002/oby.21063] [Medline: 25919921]
17. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12. [doi: 10.7326/M13-3005] [Medline: 25402403]
18. Turner-McGrievy G, Tate D. Tweets, Apps, and Pods: Results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]

19. Flores Mateo G, 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;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
20. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Sep;17(9):895-906. [doi: 10.1111/obr.12428] [Medline: 27192162]
21. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
22. Winzer E, Luger M, Schindler K. Using digital photography in a clinical setting: a valid, accurate, and applicable method to assess food intake. Eur J Clin Nutr 2018 Jun;72(6):879-887. [doi: 10.1038/s41430-018-0126-x] [Medline: 29563639]
23. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. "Snap-n-Eat": Food Recognition and Nutrition Estimation on a Smartphone. J Diabetes Sci Technol 2015 May;9(3):525-533 [FREE Full text] [doi: 10.1177/1932296815582222] [Medline: 25901024]
24. Kong K, Zhang L, Huang L, Tao Y. Validity and practicability of smartphone-based photographic food records for estimating energy and nutrient intake. Asia Pac J Clin Nutr 2017 May;26(3):396-401 [FREE Full text] [doi: 10.6133/apjcn.042016.05] [Medline: 28429902]
25. Hill C, Weir BW, Fuentes LW, Garcia-Alvarez A, Anouti DP, Cheskin LJ. Relationship Between Weekly Patterns of Caloric Intake and Reported Weight Loss Outcomes: Retrospective Cohort Study. JMIR Mhealth Uhealth 2018 Apr 16;6(4):e83 [FREE Full text] [doi: 10.2196/mhealth.8320] [Medline: 29661750]
26. Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. Br J Nutr 2009 Feb;101(3):446-456 [FREE Full text] [doi: 10.1017/S0007114508027438] [Medline: 18616837]
27. Martin CK, Nicklas T, Gunturk B, Correa JB, Allen HR, Champagne C. Measuring food intake with digital photography. J Hum Nutr Diet 2014 Jan;27 Suppl 1:72-81 [FREE Full text] [doi: 10.1111/jhn.12014] [Medline: 23848588]
28. Martin CK, Correa JB, Han H, Allen HR, Rood JC, Champagne CM, et al. Validity of the Remote Food Photography Method (RFPM) for estimating energy and nutrient intake in near real-time. Obesity (Silver Spring) 2012 Apr;20(4):891-899 [FREE Full text] [doi: 10.1038/oby.2011.344] [Medline: 22134199]
29. Most J, Vallo PM, Altazan AD, Gilmore LA, Sutton EF, Cain LE, et al. Food Photography Is Not an Accurate Measure of Energy Intake in Obese, Pregnant Women. J Nutr 2018 Apr 01;148(4):658-663. [doi: 10.1093/jn/nxy009] [Medline: 29659958]
30. Nicklas T, Saab R, Islam NG, Wong W, Butte N, Schulin R, et al. Validity of the Remote Food Photography Method Against Doubly Labeled Water Among Minority Preschoolers. Obesity (Silver Spring) 2017 Dec;25(9):1633-1638 [FREE Full text] [doi: 10.1002/oby.21931] [Medline: 28758370]
31. Doumit R, Long J, Kazandjian C, Gharibeh N, Karam L, Song H, et al. Effects of Recording Food Intake Using Cell Phone Camera Pictures on Energy Intake and Food Choice. Worldviews Evid Based Nurs 2016 Jun;13(3):216-223. [doi: 10.1111/wvn.12123] [Medline: 26646555]
32. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. J Med Internet Res 2014;16(4):e109 [FREE Full text] [doi: 10.2196/jmir.3084] [Medline: 24735567]
33. Lose It!, FitNow, Inc. URL: https://www.loseit.com/snapit/ [accessed 2018-08-08] [WebCite Cache ID 71WtIEUUA]
34. World Health Organization. Obesity: preventing and managing the global epidemic: report of a WHO consultation URL: http://www.who.int/nutrition/publications/obesity/WHO_TRS_894/en/ [accessed 2018-08-08] [WebCite Cache ID 71WsKvNfN]
35. Brunstrom JM, Burn JF, Sell NR, Collingwood JM, Rogers PJ, Wilkinson LL, et al. Episodic memory and appetite regulation in humans. PLoS One 2012;7(12):e50707 [FREE Full text] [doi: 10.1371/journal.pone.0050707] [Medline: 23227200]
36. Painter SL, Ahmed R, Hill JO, Kushner RF, Lindquist R, Brunning S, et al. What Matters in Weight Loss? An In-Depth Analysis of Self-Monitoring. J Med Internet Res 2017 May 12;19(5):e160 [FREE Full text] [doi: 10.2196/jmir.7457] [Medline: 28500022]
37. Burke LE, Sereika SM, Music E, Warziski M, Styn MA, Stone A. Using instrumented paper diaries to document self-monitoring patterns in weight loss. Contemp Clin Trials 2008 Mar;29(2):182-193 [FREE Full text] [doi: 10.1016/j.cct.2007.07.004] [Medline: 17702667]
38. Burke LE, Warziski M, Starrett T, Choo J, Music E, Sereika S, et al. Self-monitoring dietary intake: current and future practices. J Ren Nutr 2005 Jul;15(3):281-290. [Medline: 16007557]

Abbreviations

BMI: body mass index
RFPM: Remote Food Photography Method

Edited by G Eysenbach; submitted 10.08.18; peer-reviewed by PH Lin, S Painter; comments to author 01.11.18; revised version received 25.12.18; accepted 21.01.19; published 29.05.19.

Please cite as:
Ben Neriah D, Geliebter A
Weight Loss Following Use of a Smartphone Food Photo Feature: Retrospective Cohort Study
JMIR Mhealth Uhealth 2019;7(6):e11917
URL: https://mhealth.jmir.org/2019/6/e11917/
doi:10.2196/11917
PMID:31199300

Original Paper

Development of a Mobile Health Intervention to Promote Papanicolaou Tests and Human Papillomavirus Vaccination in an Underserved Immigrant Population: A Culturally Targeted and Individually Tailored Text Messaging Approach

Hee Yun Lee[1*], PhD; Mi Hwa Lee[2*], PhD; Monica Sharratt[3*], MSW; Sohye Lee[4*], PhD; Anne Blaes[5*], MD

[1]School of Social Work, The University of Alabama, Tuscaloosa, AL, United States

[2]School of Social Work, College of Heath and Human Performance, East Carolina University, Greenville, NC, United States

[3]School of Social Work, College of Education and Human Development, University of Minnesota, Twin Cities, MN, United States

[4]Loewenberg College of Nursing, University of Memphis, Memphis, TN, United States

[5]Division of Hematology, Oncology and Transplantation, Department of Medicine, University of Minnesota, Twin Cities, MN, United States

[*]all authors contributed equally

Corresponding Author:
Hee Yun Lee, PhD
School of Social Work
The University of Alabama
1022 Little Hall Box 870314
Tuscaloosa, AL, 35401
United States
Phone: 1 205 348 6553
Email: hlee94@ua.edu

Abstract

Background: Disparities in cervical cancer incidence and mortality signify the need for intervention efforts targeting Korean American immigrant women.

Objective: The purpose of this study was to demonstrate how a culturally targeted and tailored mobile text messaging intervention, mobile screening (mScreening), was developed to promote the uptake of Papanicolaou tests and human papillomavirus vaccine among young Korean American immigrant women.

Methods: Guided by the Fogg behavior model, the mScreening intervention was developed through a series of focus groups. Braun and Clarke's thematic analysis was used to identify core themes.

Results: Overall, 4 themes were identified: (1) tailored message content (ie, basic knowledge about cervical cancer), (2) an interactive and visual message format (ie, age-appropriate and friendly messages using emoticons), (3) brief message delivery formats to promote participant engagement, and (4) use of an incentive to motivate participation (ie, gift cards).

Conclusions: This study demonstrated the processes of gathering culturally relevant information to develop a mobile phone text messaging intervention and incorporating the target population's perspectives into the development of the intervention. The findings of the study could help guide future intervention development targeting different types of cancer screening in other underserved racial or ethnic groups.

(JMIR Mhealth Uhealth 2019;7(6):e13256) doi:10.2196/13256

KEYWORDS

uterine cervical cancer; papanicolaou test; papillomavirus infections; papillomavirus vaccines; text messaging; Asian American; immigrants

Introduction

Background

Although cervical cancer has become a largely preventable disease with recent reductions in cervical cancer mortality in the United States [1], it still remains a major concern in the field of women's health. Approximately 12,000 women are newly diagnosed with cervical cancer and 4000 mortalities occur annually [2]. Although there is lower overall cervical cancer incidence and mortality across the aggregate Asian American group when compared with other ethnic or racial groups [3], certain Asian subgroups report some of the highest rates of cervical cancer across all ethnic or racial groups. For example, Korean American immigrant women were found to have an incidence rate of 11.9 per 100,000, much higher in contrast with non-Latino whites (7.1), Japanese (6.2), and Chinese (5.8) samples [4].

The Papanicolaou (Pap) test is a screening procedure for cervical cancer that helps to facilitate earlier detection and treatment. The US Preventive Services Task Force recommends a Pap test for all women between the ages of 21 and 65 years every 3 years [5]. Although 72.8% of non-Latino white women aged 18 years and older had reported completing a Pap test within the past 3 years in 2010, only 68.0% of Asian women had done the same [6]. One potential reason behind the disproportionate burden of cervical cancer faced by Korean American immigrant women might be the low uptake of both the Pap tests and the human papillomavirus (HPV) vaccines. In accordance with that disparity, Korean American immigrant women have consistently demonstrated a much lower likelihood of being up to date with recommended Pap tests than non-Latina white women [7,8]. A previous study in California reported that 35% of Korean American immigrant women aged 18 to 65 years have not received a Pap test within the past 3 years [9].

The HPV vaccine prevents infection from the strains of HPV that cause cervical cancer. The HPV vaccine was introduced in 2006 for women and in 2008 for men and is recommended for all adolescents starting at the age of 9 years, as well as young women and men up to the age of 26 years [10]. Despite the availability of the HPV vaccine, an effective mechanism for preventing cervical cancer, the rate of vaccination in Asian American immigrant women is significantly lower (38.6%) than in non-Latina white women (60.7%) [11]. For example, only 24% of Korean American immigrant mothers reported that their vaccine-eligible daughters had initiated the HPV vaccine series of 3 doses [12]. Therefore, an effective intervention is needed to promote HPV vaccination to prevent cervical cancer for Korean American immigrant populations.

The disparities in utilization of the Pap tests and HPV vaccine highlight the need for intervention efforts to specifically target Korean American immigrant women. Previous efforts to reduce cervical cancer were focused on promoting Pap tests in Asian American women, such as community-based navigation [13], lay health workers [14], and lay health worker outreach combined with media intervention [15]. However, these studies reported mixed outcomes in improving Pap testing. Additional interventions to promote Pap tests and HPV vaccines specifically targeting underserved women such as Korean American immigrant women are needed [16-18].

Incorporating mobile health (mHealth) technology into these interventions might improve outcomes. Text messaging has previously been shown to be a useful and efficient means of educating patients on sensitive health-related issues that require confidentiality and security, such as HIV prevention [19], mental health [20], sexually transmitted infection management [21], and smoking cessation [22-24]. As such, it is possible that the incorporation of mHealth technology could help improve the efficacy of efforts to promote Pap tests and HPV vaccination.

Objectives

Although mobile phone–based interventions appear to hold promise as effective vehicles promoting behavioral change, little knowledge has been established regarding how to successfully develop such intervention programs for cervical cancer prevention. This study aimed to illustrate how a culturally targeted and individually tailored mobile text messaging intervention, mobile screening (*mScreening*), was developed. The *mScreening* is a 7-day educational program to promote the uptake of the Pap tests and HPV vaccines among young Korean American immigrant women aged between 21 and 29 years. Our *mScreening* studies showed the effectiveness and feasibility of the intervention, including improved knowledge of cervical cancer and Pap test rates [25] and raising knowledge, attitude, and uptake rates of HPV vaccine [26]. This study retrospectively shares the design methodology based on the success we had with the solution.

We targeted young Korean American immigrant women aged between 21 and 29 years because of age guidelines of receipt of Pap tests (21 to 65 years) and HPV vaccines (18 to 26 years). Given that young Korean American immigrant women have a high accessibility to mobile phones [27], we targeted young Korean American immigrant women as an intervention group. We also targeted both the Pap test and the HPV vaccine as primary outcomes given that they are highly important prevention strategies for cervical cancer, with the Pap test helping to facilitate earlier detection and treatment of cervical cancer and the HPV vaccine helping to reduce the likelihood of cervical cancer.

Methods

Conceptual Framework

The procedure for developing *mScreening* was guided by the Fogg behavior model (FBM) [28]. The FBM consists of 3 components: identifying barriers, developing motivators, and providing triggers to act. This theory basically posits that to make positive health behavior happen, researchers should identify barriers to a specific health behavior, convert the barriers to motivators, and finally, provide triggers to make the action occur.

Figure 1. Conceptual framework.

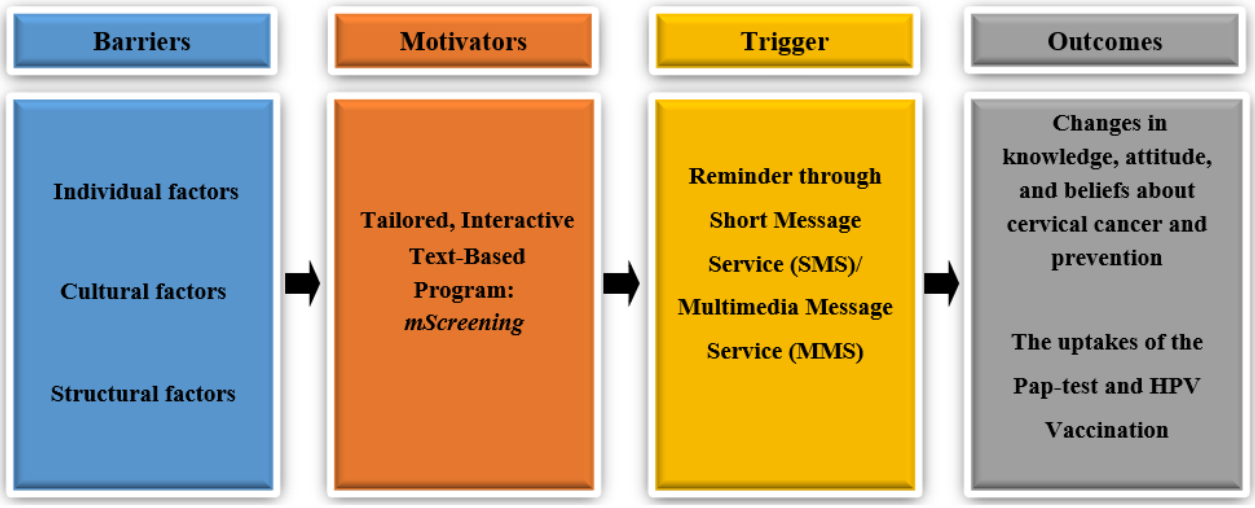


Figure 1 shows the specific approach that we proposed using FBM. To identify the barriers and develop motivators and triggers for Pap test uptake and HPV vaccine receipt in young Korean American immigrant women, a series of focus groups was conducted. Through the focus groups, 3 barriers were identified and previously published, including limited knowledge about cervical cancer and preventive behaviors, cultural barriers, and low accessibility to health care services [29]. On the basis of these identified barriers, motivators and triggers were also developed.

In addition to the focus group meetings, the research team had regular meetings with community advisory board members and mobile technology developers to get feedback and refine the *mScreening.* Through periodic meetings, strategies and topics—such as what messages and pictures would motivate young Korean American immigrant women, what real life stories exist, how the stories should be presented, and in what format the content should be delivered–were discussed and finalized. The community advisory board consisted of Korean American immigrant women in their 20s, a Korean American physician, a Korean American nurse, and religious and community leaders. Furthermore, 2 mobile phone technology experts also reviewed the developed texting program and helped the research team to revise and finalize it.

Research Design and Sampling

As mentioned earlier, we conducted a series of focus groups with young Korean American immigrant women to identify their barriers and to develop motivators and triggers for the use of Pap tests and HPV vaccine receipt. Overall, 5 focus groups were conducted between April 2011 and March 2012 with a total of 22 young Korean American immigrant women residing in the Twin Cities area of Minnesota, and the details of the procedure were published previously [25]. A total of 5 women participated in the first focus group, 7 in the second group, 4 in the third group, 3 in the fourth group, and 3 in the fifth group. In addition, 2 women participated twice; thus, we only count 20 women for the total number. The participants were recruited from postings on Korean American student associations' websites, flyers at Korean churches, and the research team members' personal networks. They were able to attend the focus groups regardless of their Pap test or HPV vaccine history. Each participant received US $20 for their time commitment.

Data Collection

All of the focus groups were conducted in Korean and lasted approximately 1.5 to 2 hours. The first author (HL), along with 2 research staff, facilitated each focus group. The first author started the focus groups with welcoming messages, used ice breakers to build a rapport (eg, talking about the weather), and introduced the purpose of the meeting. The first author mainly asked questions to the participants and the other 2 research staff supported the first authors by asking follow-up questions and taking notes. The participants were informed about confidentiality, and they had a chance to introduce themselves to the group. Verbal and written consent was obtained from all participants before starting the groups. The topics of the focus groups included (1) barriers and facilitators of cervical cancer screening and HPV vaccination, (2) effective ways to increase Korean American immigrant women's awareness of cervical cancer screening and HPV vaccination, (3) patterns of mobile phone use and ways to use mobile phone text messaging as an intervention medium to increase Pap testing and HPV vaccination, (4) ways to effectively deliver messages and age-appropriate designs for the intervention, and (5) potential implementation methods that would maximize participation and engagement. All sessions were digitally recorded to ensure that no information was missed during the sessions. Topic 1 (barriers to cervical cancer screening and HPV vaccination) was published [29]. This study reports on topics 2 to 5. The University of Minnesota institutional review board approved this study.

Data Analysis

Core themes from the focus groups were identified using Braun and Clarke's thematic qualitative method of analysis [30]. This method involves 6 phases: (1) becoming familiar with the data

through transcription of verbal data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. For the first phase, all recordings were transcribed by a bilingual Korean research assistant and then reviewed by the research team. For the second phase, to make sure that no particular framework was imposed, the transcripts were coded by 3 different Korean bilingual researchers. Each researcher identified and highlighted every codable unit of text in the transcripts, then compared their analyses and agreed upon a set of codes, subcategories, and categories. Subsequently, based on the generated codes, the research team identified and reviewed themes present in the transcripts. Themes were compared across each transcript to ensure that they were representative and inclusive of all transcripts. Through this process, clear definitions and names for each theme were generated. Finally, the most representative quotes were selected to present in this paper and translated into English. For the translation process, a bilingual research assistant first translated the quotes into English and then another bilingual research team member back-translated them into Korean to ensure that the meaning was not lost in translation. All translations were then finalized by the first author.

Results

Study Participants

Table 1 presents the sociodemographic information of the focus group participants. The average age of the participants was 26 years, and 95% of the participants reported that they were unmarried. Most of the participants indicated that their families lived in Korea, and all participants were born and raised in Korea. They were undergraduate or graduate students in Minnesota at the time of this study. Participants reported majoring in various disciplines, including economics, biology, global studies, psychology, music, English literature, and social work. The average length of time they had resided in the United States was 3.9 years. All participants had health insurance at the time of the study.

Themes From Focus Groups

Overall, 4 themes were identified from the focus group data about developing motivated messages to make health behavior change: (1) culturally targeted and tailored messages, (2) an interactive and visually appealing message format, (3) brief message delivery formats to promote participant engagement, and (4) the use of an incentive to motivate participation. In the rest of this section, each theme will be described and how it informed the development of *mScreening* will be explained. To protect the confidentiality of each participant, we only described each participant's number in the study when presenting quotes.

Theme 1: Culturally Targeted and Individually Tailored Messages

Participants agreed on the importance of disseminating specific information on cervical cancer and ways to detect or prevent it to young Korean American immigrant women. The participants suggested that the following 3 content areas should be included and tailored when disseminating knowledge: (1) tailored statistics about cervical cancer incidence and mortality to raise awareness of the severity of cervical cancer, (2) culturally targeted message development to deliver knowledge of what the Pap tests and HPV vaccine are, and (3) tailored content to improve knowledge related to health care access on how to get the Pap tests and HPV vaccine.

Table 1. Sociodemographic characteristics of the sample (n=20).

Categories	Frequency, n (%)
Age (years; mean=26 years)	
20-24	10 (40)
25-29	12 (25)
Marital status	
Unmarried	19 (95)
Married	1 (5)
Education[a]	
Undergraduate	13 (65)
Graduate	7 (35)
Years in the US (mean=3.9 years)	
Less than 5 years	13 (65)
More than 5 years	7 (35)
Health insurance	
Yes	20 (100)
No	0 (0)

[a]Majors: biology, economics, English literature, global studies, music, psychology, and social work.

On the basis of the participants' suggestions, a 7-day text message program was developed. First, *mScreening* introduced information about cervical cancer. Then, the Pap test was explained, followed by information on clinics and health professionals, as well as the cost of the Pap test. Next, HPV and the HPV vaccine were introduced and described, followed by information on the cost of the HPV vaccine and possible cultural barriers. The last day consisted of summarizing what they have learned through the 6 days using quizzes and games. The *mScreening* messages for each content area were developed using the message-framing techniques that affect health decision making based on the FBM model [28]. The FBM model particularly emphasizes that to make positive health behavior occur, barriers to specific health behavior (eg, Pap test and HPV vaccine) should be converted to motivators and triggers to action.

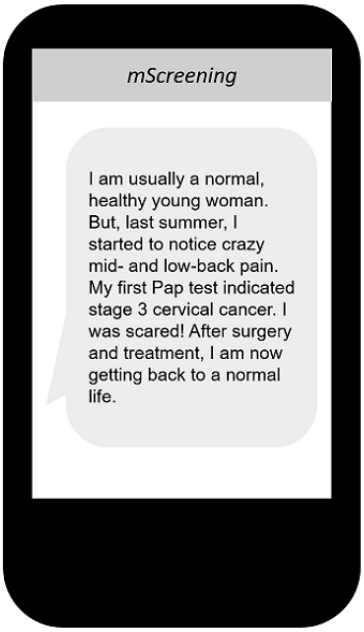
Subtheme 1.1: Tailored Statistics About Cervical Cancer Incidence and Mortality to Raise Awareness of the Severity of Cervical Cancer

One of the main content areas identified was the importance of educating the target population to improve their limited knowledge of and indifference to the topic of cervical cancer prevention. Participants expressed their desire to learn the definition of cervical cancer, risk factors for cervical cancer, and the incidence and mortality rates of cervical cancer among young women by race and ethnicity. In addition, they suggested that some of the information should be presented in a way that is *shocking* or *surprising*:

> *A shocking method! Young people are not easily shocked. Provide shocking information...something that can alert the young people to think that it is possible for them to have cervical cancer.* [Participant #15]

The research team sought to integrate participants' suggestions about creating shocking or surprising messages on the topic of basic knowledge about cervical cancer to capture the target population's attention, thereby increasing their awareness. For example, the messages included images that compared a healthy cervix with an unhealthy cervix, statistics on mortality, incidence, and screening rates for cervical cancer in Korean American immigrant women compared with other racial or ethnic groups, and testimony from a young Korean American immigrant woman who was diagnosed with cervical cancer (Figure 2).

Figure 2. Mobile screening message example: used a testimony from a young Korean American immigrant woman.



Subtheme 1.2: Culturally Targeted Message Development to Deliver Knowledge of the Pap Tests and Human Papillomavirus Vaccine

The participants expressed expectations that the text messages would include information on the guidelines, process, cost, and pain level of Pap testing. They also advised that the messages should be targeted to revise cultural misconceptions about cervical cancer screening and prevention behaviors. Almost all the participants reported that they had not heard of cervical cancer or screening before the study or, even if they had, that they had paid little attention to the need for prevention. Interestingly, some participants stated that they wanted information related to the HPV vaccine and its side effects to be provided by a doctor because of a lack of trust they held toward other sources of vaccination information, which was the result of the excessive promotion of vaccinations via media in Korea.

As such, *mScreening* messages regarding the Pap tests and the HPV vaccine were developed focusing on enhancing motivation for carrying out cervical cancer prevention–related behaviors (Figure 3). For instance, messages were created to educate participants about the benefits of receiving the Pap test and HPV vaccination using a gain message frame [31]. This is one way of converting the barrier (lack of knowledge on Pap test and HPV and HPV vaccine) to motivator (empowering through education) to make positive health behavior change according to the FBM model [28]. These messages sought to promote Pap

testing and HPV vaccine uptake by using a female doctor as the spokesperson. All participants suggested to use a female doctor to educate the content given that talking about women's reproductive organs with a male doctor is embarrassing in Korean culture. Thus, it was decided that a female doctor would be a preferable spokesperson compared with a male doctor.

In addition, a majority of participants also expressed wanting text messages to be targeted to overcome cultural misinformation (eg, hymen breaking during a Pap test) and apprehension about cervical cancer screening (eg, feeling uncomfortable in a gynecology clinic as an unmarried female). The participants suggested several methods toward achieving this goal, such as including powerful emotional messages or life experience examples from people who have successfully overcome these obstacles:

> *Showing a real example...real person...like a student who got the shot can share what she thought about it before, what it really was like, and how she feels afterwards. That kind of example...[or] if a person who went to have [the Pap] test explains how she felt.* [Participant #4]

To reflect the participants' suggestions, testimonial text messages were created and incorporated into the *mScreening* by working with young Korean American immigrant women who had received the Pap test. Given that obstetric and gynecological health issues are culturally sensitive in the Korean American community, especially so for unmarried young women [32], the testimonial messages were designed to reframe the Pap test more positively for young Korean American women (Figure 4). In addition to being influenced by most of the participants' suggestions, this approach is supported by previous research that suggests that sharing the personal experiences of individuals in the same ethnic group can be influential in changing participants' cultural beliefs [33].

Figure 3. Mobile screening message example: focused on enhancing motivation.

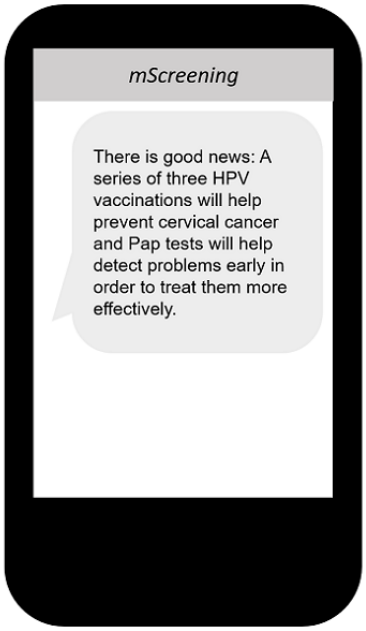
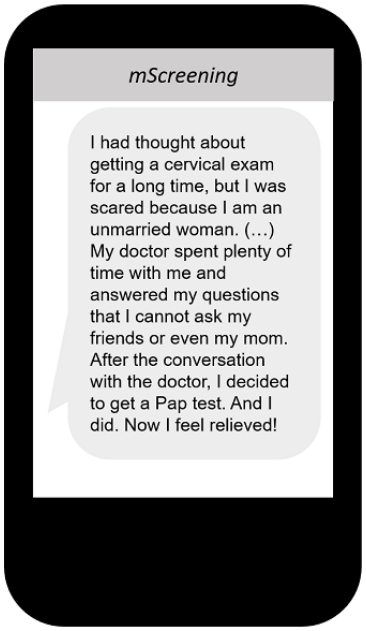


Figure 4. Mobile screening message example: reframed the Pap test more positively for young Korean American women.



Subtheme 1.3: Tailored Content to Improve Knowledge Related to Health Care Access

Participants suggested including practical information about the costs of receiving the Pap test and HPV vaccination, clinics they could visit to receive these procedures, and how to make appointments in *mScreening.* Participants shared their views that practical information was needed because of their lack of knowledge about the US health care system:

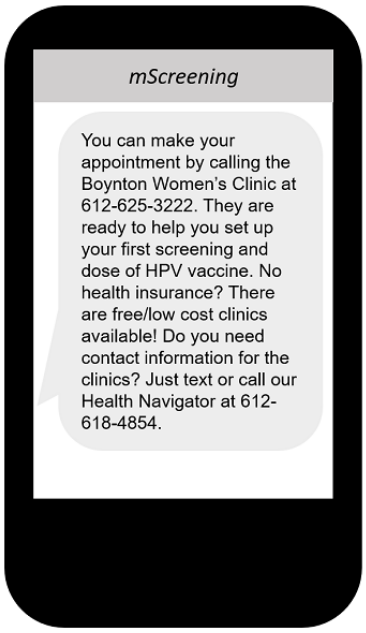
> *I was surprised that the cost for the Pap test is free. [Korean] People may think medical costs are very high in the US system. They may go and get the screening if they know it is for anyone.* [Participant #10]

Participants also mentioned that it would be helpful to have messages from a doctor about the procedures related to Pap testing and HPV vaccination. Participants felt that this would increase their familiarity with the process and reduce their uneasiness. Furthermore, participants wanted to know about the availability of female doctors who performed Pap tests at the clinics because of their reluctance to engage in conversations about their periods or personal sexual activities with male doctors:

> *If I know that these [female] doctors are the ones that I meet with for an appointment, I would feel more comfortable making an appointment [to have a Pap test].* [Participant #16]

As a result, the *mScreening* messages included information on the price of the HPV vaccine, phone numbers of women's clinics that employed female doctors and nurses, and a list of available free or low-cost clinics in Minnesota. One example of a targeted message is shown in Figure 5.

Figure 5. Mobile screening message example: included information on the price of the human papillomavirus vaccine and phone numbers of women's clinics that employed female doctors and nurses.



Theme 2: Message Format—Interactive and Visually Appealing Messages

All participants agreed that mobile phones are a significant part of their daily lives. The participants' endorsement of the high accessibility of mobile phones suggests that developing a texting program for the target population could be successful. Participants agreed with this stance when queried and suggested 3 types of messages that could be used together to maximize engagement: visual messages, interactive messages, and dynamic characters or emoticons. On the basis of this feedback, *mScreening* disseminated information using 3 types of messages as described below.

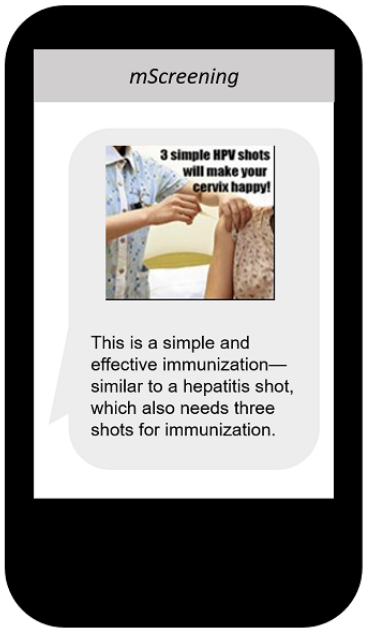
Subtheme 2.1: Visual Messages

Participants indicated a preference for visual messages that integrated video clips or images, such as a picture of a cervix and charts of mortality rates, over plain text:

> *Rather than texting, [it would be better to receive messages] through videos, images, posters, or cartoons because web cartoons are popular these days...* [Participant #19]

As a result of this input as well as previous studies that have demonstrated the power of visual messages to persuade people [34], *mScreening* included a variety of graphs, pictures, and drawings (Figure 6).

Figure 6. Mobile screening message example: demonstrated the power of visual messages to persuade people.



Subtheme 2.2: Young Age–Friendly Messages With Graphics

A majority of participants also highly recommended the inclusion of age-friendly messages that used characters or emoticons, stating that they would be more attractive and dynamic to the young Korean American population than simple texts, helping the participants feel more comfortable. As participants recommended, age-friendly messages with graphics (eg, emoticons, logos, and cartoons) were created to target young Korean American women. Given the fact that cervical cancer screening and prevention are not a common topic of discussion in the Korean American community, using age-appropriate text messages with cartoons was suggested as a way to culturally target the intervention and decrease discomfort around having conversations regarding screening and prevention:

> *If it is through a smartphone, it would be good to use cartoons or videos. This will keep me more focused and engaged [to learn cervical cancer screening and prevention methods.* [Participant #15]

To help achieve this suggestion, we first created the study's logo, which symbolizes a happy and healthy cervix. The focus group participants provided multiple ideas of how the logo should look and what color and shape would represent the cervix in a culturally appropriate way. The study's logo is shown in Figure 7.

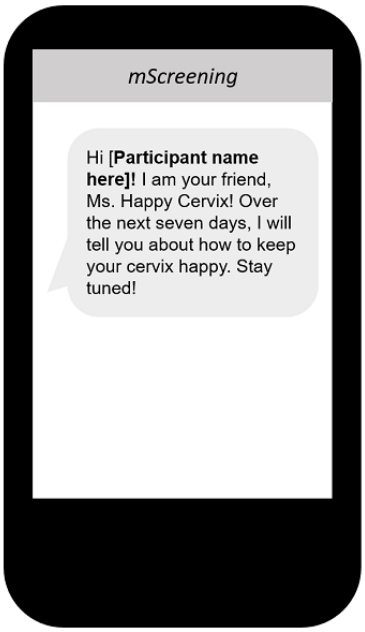
With input from focus group participants and community advisory board members, we revised the logo and developed multiple characters depicting various facial expressions and used them in mScreening messages to help the target population feel more positively engaged with the intervention and at ease with the information presented (Figure 8).

In addition, the first text of each day was a tailored, welcoming message that addressed the participant by her first name or nickname (Figure 9).

Figure 7. Mobile screening message example: created a study project logo.

Figure 8. Mobile screening message example: developed multiple characters to help target population feel more positively engaged with the intervention.

Figure 9. Mobile screening message example: tailored welcoming message that addressed the participant by her first name or nickname.



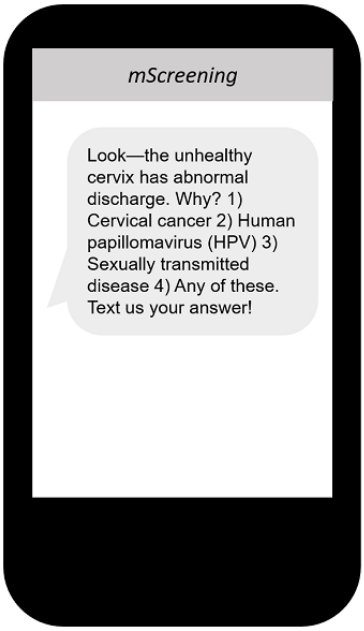
Subtheme 2.3: Interactive Messages

Participants suggested that interactive messages involving question and answer (Q&A) or games should be included in the intervention to minimize barriers of a 1-way text messaging delivery system. Participants expressed wanting the ability to interact with others involved in the project or the content provider so that they would able to communicate their questions or any confusion they had. Participants also pointed out some benefits of Q&A format messages:

> *It would be good to mix information and questions. How about we receive related information when we answer a question? In this way, messages can be tailored for each individual and this would increase study engagement.* [Participant #18]

To ensure the inclusion of interactive messages with the intent to promote participants' attentiveness and motivation during the intervention, 10 interactive Q&A messages were developed (Figure 10).

Figure 10. Mobile screening message: developed interactive question and answer messages.



If a participant texted answer (1), for example, then an appropriate response message was sent to the participant so that the tailored message was relevant for the participant. A gynecologic doctor who served on the community advisory board and worked at the University Health Center reviewed all the texts for medical accuracy and confirmed that all text message content used was accurate. We also had a bilingual health navigator, who is a registered nurse, available for questions and comments as part of *mScreening.* We provided the phone number of the health navigator each day at the end of the intervention and asked if each participant had a question or concern to discuss with the health navigator.

Theme 3: Message Delivery Formats

Questions centering on how to deliver suggested content were also discussed in the focus groups. Participants were asked about their preferred length, frequency, interval between messages, and duration of messages, as well as potentially appealing program names. Participants suggested that short and concise messages with online resources should be included, given that short message service text messages are generally limited to 160 characters maximum for a regular mobile phone. As 1 participant indicated:

> *A message through a regular mobile phone is limited to 160 characters. It could be hard to provide information in detail. What about sending a short message with a Facebook or website link where we can explore information further as needed?* [Participant #2]

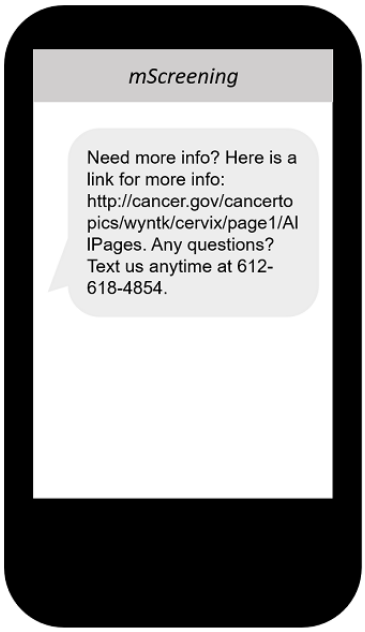
To address the limit on the length of text messages, succinct messages were created and incorporated into *mScreening.* In addition, on the last day of the intervention, participants were provided with 1 message that contained an online link for further information (Figure 11).

In terms of the number of messages received and the duration of the intervention, approximately 3 messages sent daily from *mScreening* over the course of a week to a month was considered acceptable. The participants also expressed that they wanted to be asked about their preferred time of day for receiving messages before the program started. After discussion with community advisory board members and mobile phone technology experts, the research team decided to develop a 1-week program with a more intensive frequency of messages and assess participants' preferred time for message delivery through the pretest questionnaire. More specifically, based on focus group participants' input, *mScreening* was designed to deliver each participant 15 to 20 messages every day, depending on each day's topic, over the course of 7 days. One concern was that messages to their mobile phones would message their spam instead. Participants suggested that creating an appealing name for the intervention would help the target population recognize the messages as coming from a welcome source:

> *It might be better to have a program name such as "big sister." I prefer to have the feeling that a sister (who has known me for a while) takes care of me and gives me that kind of information. A sister sounds friendlier than a mother or doctor.* [Participant #5]

The research team used *mScreenin* g as the intervention name, which indicates a mobile phone–based cervical cancer screening and prevention program. The intervention was designed to ensure that a welcoming message incorporating the name would be sent as the first message each day to prevent the participants from confusing the intervention with any potential spam messages.

Figure 11. Mobile screening message example: provided one message that contained an online link for further information.



Theme 4: Use of an Incentive for Study Engagement

Another theme present in the focus groups was that of possible remuneration. Participants recommended providing an incentive to maintain high levels of engagement in the intervention. In addition, the participants endorsed the belief that an incentive would help encourage advertisement among participants, leading to the recruitment of more potential participants. In terms of the type of incentive, the participants discussed several possibilities, such as gift cards for coffee or small gifts such as cell phone attachments (which are a trend among young Korean American women). Participants also suggested sending a celebratory message to intervention participants who received a Pap test or a dose of the HPV vaccine after their participation in *mScreening.* The research team created a cell phone attachment with a mini mirror that features the study's logo. This attachment was provided to participants upon completion of the intervention, whereas a celebratory message was sent after participants received a Pap test or dose of the HPV vaccine.

Discussion

Principal Findings

This study sought to demonstrate how a culturally and individually tailored mobile text messaging intervention program, *mScreening,* was developed to promote the uptake of Pap testing and HPV vaccination in young Korean American immigrant women. The message contents, format, and delivery format of *mScreening* were suggested by participants through a series of focus groups. And then, community advisory board members and mobile technology experts reviewed the initial program and refined for intervention use. We identified 4 themes and we used them to inform the development of *mScreening.*

Comparison With Previous Work

The first theme centers on developing and delivering culturally targeted and tailored messages. Through the use of focus groups, various types of culturally tailored motivating messages were suggested and created to help increase basic knowledge and awareness about cervical cancer, Pap testing, and HPV vaccination. Participants in focus groups suggested that the presentation of information might be sometimes shocking or emotionally affecting in ways consistent with a previous study that promoted sexual health for young adults in Australia [35]; in this study, fear and the use of grave statistics were recommended by young participants as especially effective strategies. A recent study that used guilt-fear appeal for HPV vaccination reported that fear is a mediator between perceived susceptibility and HPV intentions [35].

At the same time, participants suggested that the *mScreening* messages should highlight the benefits of Pap testing and HPV vaccine uptake by framing the messages in terms of potential gain [31]. The recommendation for this gain-framed messaging departs from previous studies that have employed messages framed in terms of potential losses. For example, 1 study suggested that loss-framed cancer screening messages (eg, the costs of not undergoing screening) are more effective and powerful at promoting screening than gain-framed messages (eg, benefits of undergoing screening) would be [31]. The study similarly found that loss-framed messages were more persuasive at promoting HPV vaccination [36]. Other studies have noted that cultural contexts should be considered when messages are framed [37,38]. These studies suggest that message framing alone without consideration of individuals' cultural characteristics will not be effective. On the basis of these perspectives, we decided to use both gain and loss frames in developing targeted and tailored messages in our study.

As one of the subthemes in Theme 1, culturally tailored content delivery to improve literacy related to health care access was suggested. In addition to culturally tailored information regarding the Pap test and HPV vaccine, *mScreening* messages should deliver practical information on participants' health insurance coverage and information on physicians and clinics to help enhance health care access. The participants' suggestions further implied that increasing familiarity with the US health

care system might facilitate young Korean American immigrant women's participation in cervical cancer screening and prevention behavior. This finding reflects the findings from a previous study with Vietnamese American women that demonstrated that access to health care for cancer screening is improved with health navigator service [14]. On the basis of this suggestion, we added a bilingual Korean health navigator service into the *mScreening* intervention so that any barriers to health care service access can be minimized, which may ultimately enhance health care accessibility for cervical cancer screening and HPV vaccine promotion.

The second and third themes focused on an interactive and visually appealing message development. When participants were asked about message formatting, they suggested various methods to help increase study participants' attention and engagement. Reflecting participant suggestions, the research team developed various types of messages that included images and emoticons, as well as interactive Q&A messages. Previous studies provide supporting evidence that message formats could potentially enhance people's understanding of information and create a relaxed atmosphere that would also help maintain engagement [39]. For example, as emoticons can be used to express nonverbal communication, such as humor, solidarity, support, positive feelings, and appreciation [39], messages that include emoticons may help create a comfortable atmosphere in which young Korean American immigrant women may feel more willing to explore cervical cancer and talk more easily to peers about their health. In addition, the Q&A messages address the shortcoming of a 1-way text messaging system.

Regarding message delivery format, participants agreed that short and concise messages were preferable, although this could potentially impact the accuracy of the messages. To deliver accurate information despite the short messages, the research team created messages that would lead participants to a reliable online resource for further information. The last theme is about retention strategy for the 3-month intervention period. Participants suggested using an incentive system to maintain high levels of participation, engagement, and retention. Previous studies also reported effectiveness of monetary incentives (eg, gift cards or cash bills) in recruiting potential participants, participant engagement, and reduction of attrition [40,41]. Similarly, a review conducted by Tishler and Bartholomae [42] indicated that financial rewards could play a critical role in the motivation levels for volunteers deciding to participate in research studies.

Limitations

Although this study has yielded various pieces of critical information, there are several limitations. First, given that the study participants were all young Korean American immigrant women residing in the upper Midwest, the findings cannot be generalized to Korean American immigrant women in other locations at large. Furthermore, although we made an effort to recruit Korean American immigrant women residing in the community, all participants were current college or graduate students who had resided in the United States for brief periods of time, further limiting generalizability of the findings. In recent years, however, the population of Asian American immigrants has raised, including college students and college graduates [43]. As most of participants were fairly new to the United States, English proficiency functioned as a barrier to understand the actual meaning of the text messages. Some of the participants suggested to provide the texts in both Korean and English by providing a language choice. Although they are in college and graduate schools, when it comes to health information, understanding via mother tongue makes them accurately understand each health-related message. Future researchers should further explore the impact of language and importance of tailoring interventions according to participants' health literacy and acculturation factors. Finally, it is important to mention that the age range of focus group participants was 21 to 29 years—with an average age of 26 years—which is the upper age limit for obtaining the HPV vaccine in girls and women according to age guidelines (9 to 26 for girls and women). In general, the HPV vaccine is promoted to the younger adolescent girls aged 11 to 17 years, even as young as 9 years; thus, the messaging from this study may not be generalizable to the common populations targeted with HPV vaccine messaging.

Conclusions

Despite the limitations, this study provides critical information in developing culturally targeted and individually tailored motivational messages to bring about positive health behavior change. This study sought to provide guidelines and suggestions to researchers interested in developing text message–based interventions for underserved minority populations. The results of the study highlighted creating and delivering culturally tailored and targeted messages, the use of intervention tools such as visual images and graphic pictures as an effective way of delivering content on sensitive topics, and provision of navigation services for study participants who are immigrants. Furthermore, the participants expressed preferences for interactive messages over 1-way communication that primarily seeks to increase participants' engagement in the study. In addition, the desired format of messages that seek to improve knowledge of cervical cancer are short and precise messages that provide links to reliable online resources. To the best of the authors' knowledge, this study is the first that provides information on the process of gathering culturally relevant information for the purpose of designing a mobile phone text messaging intervention to promote Pap tests and HPV vaccines. The study also provided information on the process of developing the intervention based on the target population's perspectives. Given that mobile phones are not only becoming the most common means of communication among young people but also that mHealth technology has demonstrated its effectiveness as a useful tool for behavioral changes in disease prevention and self-management in chronic diseases [44-50], the findings of this study could potentially guide future development of cancer screening interventions to promote Pap testing and HPV vaccination for other underserved minority groups and other types of cancer screening, including breast, colorectal, and lung cancer.

Acknowledgments

This research was funded by the National Cancer Institute R21 (5R21CA155531-02). The research team appreciates the funding to pilot this intervention development study.

Authors' Contributions

All authors listed have contributed sufficiently to the project to be included as authors.

Conflicts of Interest

None declared.

References

1. National Institutes of Health. Fact Sheet: Cervical Cancer URL: https://report.nih.gov/nihfactsheets/Pdfs/CervicalCancer(NCI).pdf [accessed 2019-05-19] [WebCite Cache ID 74cygacbl]
2. Centers for Disease Control and Prevention. Vital sign: Cervical Cancer is Preventable URL: https://www.cdc.gov/vitalsigns/cervical-cancer/ [accessed 2019-05-19] [WebCite Cache ID 74cyne1Fs]
3. National Cancer Institute. Table 5.1 Cancer of the Cervix Uteri (Invasive) URL: https://seer.cancer.gov/archive/csr/1975_2012/results_merged/sect_05_cervix_uteri.pdf [accessed 2019-05-19] [WebCite Cache ID 78VcJ01II]
4. Wang SS, Carreon JD, Gomez SL, Devesa SS. Cervical cancer incidence among 6 asian ethnic groups in the United States, 1996 through 2004. Cancer 2010 Feb 15;116(4):949-956 [FREE Full text] [doi: 10.1002/cncr.24843] [Medline: 20029972]
5. US Preventive Service Task Force. Final Update Summary: Cervical Cancer: Screening URL: https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/cervical-cancer-screening [accessed 2019-05-19] [WebCite Cache ID 74czstvk9]
6. Centers for Disease Control Prevention. Table 84 (page 1 of 5) Use of Pap smears among women aged 18 and over, by selected characteristics: United States, selected years 1987-2010 URL: https://www.cdc.gov/nchs/data/hus/2013/084.pdf [accessed 2019-05-19] [WebCite Cache ID 74d00ekqS]
7. Lee EE, Fogg L, Menon U. Knowledge and beliefs related to cervical cancer and screening among Korean American women. West J Nurs Res 2008 Dec;30(8):960-974 [FREE Full text] [doi: 10.1177/0193945908319250] [Medline: 18539883]
8. Yoo GJ, Le MN, Vong S, Lagman R, Lam AG. Cervical cancer screening: attitudes and behaviors of young Asian American women. J Cancer Educ 2011 Dec;26(4):740-746 [FREE Full text] [doi: 10.1007/s13187-011-0230-2] [Medline: 21553330]
9. Lin MK, Moskowitz JM, Kazinets G, Ivey SL, Kim Y, McDonnell DD. Adherence to Pap test guidelines: variation among Asians in California. Ethn Dis 2009;19(4):425-432. [Medline: 20073144]
10. Centers for Disease Control and Prevention. HPV Vaccines: Vaccinating Your Preteen or Teen URL: https://www.cdc.gov/hpv/parents/vaccine.html [accessed 2019-05-19] [WebCite Cache ID 74d0Mw6i9]
11. Bastani R, Glenn BA, Tsui J, Chang LC, Marchand EJ, Taylor VM, et al. Understanding suboptimal human papillomavirus vaccine uptake among ethnic minority girls. Cancer Epidemiol Biomarkers Prev 2011 Jul;20(7):1463-1472 [FREE Full text] [doi: 10.1158/1055-9965.EPI-11-0267] [Medline: 21602307]
12. Lee HY, Kwon M, Vang S, DeWolfe J, Kim NK, Lee DK, et al. Disparities in human papillomavirus vaccine literacy and vaccine completion among Asian American Pacific Islander undergraduates: implications for cancer health equity. J Am Coll Health 2015;63(5):316-323. [doi: 10.1080/07448481.2015.1031237] [Medline: 25836058]
13. Wang X, Fang C, Tan Y, Liu A, Ma GX. Evidence-based intervention to reduce access barriers to cervical cancer screening among underserved Chinese American women. J Womens Health (Larchmt) 2010 Mar;19(3):463-469 [FREE Full text] [doi: 10.1089/jwh.2009.1422] [Medline: 20156089]
14. Taylor VM, Jackson JC, Yasui Y, Nguyen TT, Woodall E, Acorda E, et al. Evaluation of a cervical cancer control intervention using lay health workers for Vietnamese American women. Am J Public Health 2010 Oct;100(10):1924-1929 [FREE Full text] [doi: 10.2105/AJPH.2009.190348] [Medline: 20724673]
15. Mock J, McPhee SJ, Nguyen T, Wong C, Doan H, Lai KQ, et al. Effective lay health worker outreach and media-based education for promoting cervical cancer screening among Vietnamese American women. Am J Public Health 2007 Sep;97(9):1693-1700. [doi: 10.2105/AJPH.2006.086470] [Medline: 17329652]
16. Fang CY, Ma GX, Handorf EA, Feng Z, Tan Y, Rhee J, et al. Addressing multilevel barriers to cervical cancer screening in Korean American women: a randomized trial of a community-based intervention. Cancer 2017 May 15;123(6):1018-1026 [FREE Full text] [doi: 10.1002/cncr.30391] [Medline: 27869293]
17. Moskowitz JM, Kazinets G, Wong JM, Tager IB. "Health is strength": a community health education program to improve breast and cervical cancer screening among Korean American Women in Alameda County, California. Cancer Detect Prev 2007;31(2):173-183. [doi: 10.1016/j.cdp.2007.02.002] [Medline: 17418978]

18. Wismer BA, Moskowitz JM, Min K, Chen AM, Ahn Y, Cho S, et al. Interim assessment of a community intervention to improve breast and cervical cancer screening among Korean American women. J Public Health Manag Pract 2001 Mar;7(2):61-70. [doi: 10.1097/00124784-200107020-00009] [Medline: 12174401]
19. van Velthoven MH, Brusamento S, Majeed A, Car J. Scope and effectiveness of mobile phone messaging for HIV/AIDS care: a systematic review. Psychol Health Med 2013;18(2):182-202. [doi: 10.1080/13548506.2012.701310] [Medline: 22788357]
20. Clarke G, Yarborough BJ. Evaluating the promise of health IT to enhance/expand the reach of mental health services. Gen Hosp Psychiatry 2013;35(4):339-344 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.013] [Medline: 23701698]
21. Lunny C, Taylor D, Memetovic J, Wärje O, Lester R, Wong T, et al. Short message service (SMS) interventions for the prevention and treatment of sexually transmitted infections: a systematic review protocol. Syst Rev 2014 Jan 16;3:7 [FREE Full text] [doi: 10.1186/2046-4053-3-7] [Medline: 24433348]
22. Johnston V, Westphal DW, Glover M, Thomas DP, Segan C, Walker N. Reducing smoking among indigenous populations: new evidence from a review of trials. Nicotine Tob Res 2013 Aug;15(8):1329-1338. [doi: 10.1093/ntr/ntt022] [Medline: 23519776]
23. Toll BA, Rojewski AM, Duncan LR, Latimer-Cheung AE, Fucito LM, Boyer JL, et al. "Quitting smoking will benefit your health": the evolution of clinician messaging to encourage tobacco cessation. Clin Cancer Res 2014 Jan 15;20(2):301-309 [FREE Full text] [doi: 10.1158/1078-0432.CCR-13-2261] [Medline: 24436474]
24. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
25. Lee HY, Koopmeiners JS, Rhee TG, Raveis VH, Ahluwalia JS. Mobile phone text messaging intervention for cervical cancer screening: changes in knowledge and behavior pre-post intervention. J Med Internet Res 2014 Aug 27;16(8):e196 [FREE Full text] [doi: 10.2196/jmir.3576] [Medline: 25164545]
26. Lee HY, Koopmeiners JS, McHugh J, Raveis VH, Ahluwalia JS. mHealth pilot study: text messaging intervention to promote HPV vaccination. Am J Health Behav 2016 Jan;40(1):67-76 [FREE Full text] [doi: 10.5993/AJHB.40.1.8] [Medline: 26685815]
27. Lee SK, Katz JE. Bounded solidarity confirmed? How Korean immigrants' mobile communication configures their social networks. J Comput-Mediat Comm 2015 Nov 19;20(6):615-631. [doi: 10.1111/jcc4.12142]
28. Fogg B. A behavior model for persuasive design. 2009 Apr Presented at: 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, California, USA. [doi: 10.1145/1541948.1541999]
29. Lee H, Lee M. Barriers to cervical cancer screening and prevention in young Korean immigrant women: implications for intervention development. J Transcult Nurs 2017 Dec;28(4):353-362. [doi: 10.1177/1043659616649670] [Medline: 27193752]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Okuhara T, Ishikawa H, Okada H, Kiuchi T. Identification of gain- and loss-framed cancer screening messages that appeared in municipal newsletters in Japan. BMC Res Notes 2014 Dec 11;7:896 [FREE Full text] [doi: 10.1186/1756-0500-7-896] [Medline: 25494623]
32. Juon H, Seung-Lee C, Klassen A. Predictors of regular Pap smears among Korean-American women. Preventive Medicine 2003 Dec;37(6):585-592. [doi: 10.1016/j.ypmed.2003.09.006] [Medline: 14636792]
33. Champion VL, Springston JK, Zollinger TW, Saywell RM, Monahan PO, Zhao Q, et al. Comparison of three interventions to increase mammography screening in low income African American women. Cancer Detect Prev 2006;30(6):535-544. [doi: 10.1016/j.cdp.2006.10.003] [Medline: 17110056]
34. Fogg B. Persuasive technologie301398. Commun ACM 1999;42(5):26-29 [FREE Full text] [doi: 10.1145/301353.301396]
35. Gold J, Lim MS, Hellard ME, Hocking JS, Keogh L. What's in a message? Delivering sexual health promotion to young people in Australia via text messaging. BMC Public Health 2010 Dec 29;10:792 [FREE Full text] [doi: 10.1186/1471-2458-10-792] [Medline: 21190584]
36. Nan X. Communicating to young adults about HPV vaccination: consideration of message framing, motivation, and gender. Health Commun 2012;27(1):10-18. [doi: 10.1080/10410236.2011.567447] [Medline: 22276999]
37. Han K, Jo S. Does culture matter? A cross-national investigation of women's responses to cancer prevention campaigns. Health Care Women Int 2012;33(1):75-94. [doi: 10.1080/07399332.2011.630117] [Medline: 22150267]
38. Uskul AK, Sherman DK, Fitzgibbon J. The cultural congruency effect: culture, regulatory focus, and the effectiveness of gain- vs. loss-framed health messages. J Exp Soc Psychol 2009 May;45(3):535-541. [doi: 10.1016/j.jesp.2008.12.005]
39. Wolf A. Emotional expression online: gender differences in emoticon use. CyberPsychol Behav 2000;3(5):827-833 [FREE Full text]
40. McCoy SI, Shiu K, Martz TE, Smith CD, Mattox L, Gluth DR, et al. Improving the efficiency of HIV testing with peer recruitment, financial incentives, and the involvement of persons living with HIV infection. J Acquir Immune Defic Syndr 2013 Jun 01;63(2):e56-e63. [doi: 10.1097/QAI.0b013e31828a7629] [Medline: 23403860]

41. Patel MX, Doku V, Tennakoon L. Challenges in recruitment of research participants. Adv Psychiatr Treat 2018 Jan 2;9(3):229-238. [doi: 10.1192/apt.9.3.229]
42. Tishler C, Bartholomae S. The recruitment of normal healthy volunteers: a review of the literature on the use of financial incentives. J Clin Pharmacol 2002 Apr;42(4):365-375. [Medline: 11936560]
43. Pew Research Center. 2012. The Rise of Asian Americans URL: http://www.pewsocialtrends.org/2012/06/19/the-rise-of-asian-americans/2/ [accessed 2019-05-20] [WebCite Cache ID 74d5I9Z1q]
44. Abebe NA, Capozza KL, Des Jardins TR, Kulick DA, Rein AL, Schachter AA, et al. Considerations for community-based mHealth initiatives: insights from three Beacon Communities. J Med Internet Res 2013 Oct 15;15(10):e221 [FREE Full text] [doi: 10.2196/jmir.2803] [Medline: 24128406]
45. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database Syst Rev 2013 Dec 5(12):CD007458 [FREE Full text] [doi: 10.1002/14651858.CD007458.pub3] [Medline: 24310741]
46. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database Syst Rev 2012 Mar 14(3):CD009756 [FREE Full text] [doi: 10.1002/14651858.CD009756] [Medline: 22419345]
47. Kharbanda EO, Stockwell MS, Fox HW, 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] [Medline: 19833982]
48. Nglazi MD, Bekker L, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. BMC Infect Dis 2013 Dec 2;13(1):566 [FREE Full text] [doi: 10.1186/1471-2334-13-566] [Medline: 24295439]
49. Shaw R, Bosworth H. Short message service (SMS) text messaging as an intervention medium for weight loss: a literature review. Health Informatics J 2012 Dec;18(4):235-250 [FREE Full text] [doi: 10.1177/1460458212442422] [Medline: 23257055]
50. 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] [Medline: 22635061]

Abbreviations

FBM: Fogg behavior model
HPV: human papillomavirus
mHealth: mobile health
mScreening: mobile screening
Pap: Papanicolaou
Q&A: question and answer

Edited by G Eysenbach; submitted 28.12.18; peer-reviewed by C Jacob, C Zhang; comments to author 07.02.19; revised version received 19.04.19; accepted 26.04.19; published 06.06.19.

Please cite as:
Lee HY, Lee MH, Sharratt M, Lee S, Blaes A
Development of a Mobile Health Intervention to Promote Papanicolaou Tests and Human Papillomavirus Vaccination in an Underserved Immigrant Population: A Culturally Targeted and Individually Tailored Text Messaging Approach
JMIR Mhealth Uhealth 2019;7(6):e13256
URL: https://mhealth.jmir.org/2019/6/e13256/
doi:10.2196/13256
PMID:31199340

Original Paper

Feedback on Physical Activity Through a Wearable Device Connected to a Mobile Phone App in Patients With Metabolic Syndrome: Pilot Study

Up Huh[1,2*], MD; Young Jin Tak[2,3*], MD; Seunghwan Song[1,2], MD; Sung Woon Chung[1,2], MD; Sang Min Sung[2,4], MD; Chung Won Lee[1,2], MD; Miju Bae[1,2], MD; Hyo Young Ahn[1,2], MD

[1]Department of Thoracic and Cardiovascular Surgery, Pusan National University Hospital, Busan, Republic of Korea

[2]Biomedical Research Institute, Pusan National University Hospital, Busan, Republic of Korea

[3]Department of Family Medicine, Pusan National University Hospital, Busan, Republic of Korea

[4]Department of Neurology, Pusan National University Hospital, Busan, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Seunghwan Song, MD
Department of Thoracic and Cardiovascular Surgery
Pusan National University Hospital
179 Gudeok-Ro, Seo-gu
Busan, 49241
Republic of Korea
Phone: 82 51 240 7267
Fax: 82 51 243 9389
Email: song77.sh@gmail.com

Abstract

Background: Little is known of the effect of wearable devices on metabolic impairments in clinical settings. We hypothesized that a wearable device that can monitor and provide feedback on physical activity may help resolve metabolic syndrome.

Objective: This study aimed to examine the objective effects of the use of these devices on metabolic syndrome resolution.

Methods: Patients diagnosed with metabolic syndrome were recruited. Participants were prescribed regular walking using a wearable device (Coffee WALKIE +Dv.3, GC Healthcare CI, Korea) on their wrist for 12 weeks. Participants received self-feedback on the amount of their exercise through an app on their mobile phone. The information on physical activities of the participants was uploaded automatically to a website. Thus, a trained nurse could provide individuals with feedback regarding the physical activity via telephone consultation on alternate weeks. Blood pressure (BP), body composition, fasting plasma glucose, and lipid profiles were recorded. The primary outcome was metabolic syndrome resolution. The secondary outcome was an improvement in the components of metabolic impairment.

Results: Of the 53 participants recruited, 20 participants with a median age of 46 (range 36-50) years completed the trial. There was no significant difference in the amount of calorie expenditure at weeks 4, 8, and 12. After 12 weeks, metabolic syndrome was resolved in 9 of 20 participants (45%), and the mean number of metabolic impairment components per person decreased from 3.4 to 2.9. Particularly, the mean systolic and diastolic BP decreased from mean 136.6 (SD 18.5) mm Hg to mean 127.4 (SD 19.5) mm Hg and from mean 84.0 (SD 8.1) mm Hg to mean 77.4 (SD 14.4) mm Hg (both P=.02), respectively.

Conclusions: This study found that a 12-week intervention via feedback, based on a wearable physical activity monitor, helped metabolic syndrome patients to be more engaged in regular walking and it improved impaired metabolic components, especially in BP. However, some practical challenges regarding patients' adherence and sustained engagement were observed.

(JMIR Mhealth Uhealth 2019;7(6):e13381) doi:10.2196/13381

KEYWORDS

electronic activity monitor; wearable devices; metabolic syndrome; physical activity

Introduction

Metabolic syndrome is a constellation of cardiovascular disease risk factors, such as abdominal obesity, hyperglycemia, hypertension, and hyperlipidemia [1]. With the increasing prevalence of obesity, metabolic syndrome has been reported as one of the most common health conditions among obese people worldwide [2]. Insufficient physical activity is a major cause of obesity and metabolic syndrome, and even small increases in physical activity can have a significant beneficial impact on metabolic syndrome and the prevention of cardiovascular disease [3]. Exercise is known to enhance insulin sensitivity by increasing the AMP-activated protein kinase activity, promoting translocation of glucose transporter type 4 to the cell membrane, thereby boosting glucose uptake. The decreases in intramuscular saturated fatty acids and stimulated beta cell activity also contribute to attenuating insulin resistance [4]. Despite evidence of improvement in metabolic impairment with regular exercise [5], the number of individuals involved in physical activity remains low, and sedentary lifestyle is prevalent, with less than 25% of people estimated to be engaging in regular physical activity [6].

As activity-tracking devices have become smaller, cheaper, and more readily wearable, it is predicted that they will be used extensively for various purposes [7]. Advances in wearable devices and gathering of personal data provide patients with chronic diseases the potential to engage in self-management. However, the data collected by the wearable devices are rarely integrated into the programming of regimens for impaired metabolic conditions. Moreover, evidence supporting the sustained use of data derived from wearable devices or their positive effects on health outcomes is lacking as most studies have mainly focused on establishing the feasibility of the devices and the association between measured physical activity and short-term benefits [8-10]. Recent research indicates that feedback on activity monitoring can successfully increase physical activity levels and lead to beneficial outcomes in the management of target diseases [11]. However, they focused on patients with diabetes mellitus [12,13], heart failure [14], or chronic pulmonary disease [15]. Little has been demonstrated regarding whether use of wearable devices may be a pragmatic option for metabolic syndrome in a clinical setting.

Mobile phones allow users to track their path by connecting to the internet with apps. Internet-based interventions seem to motivate people to increase physical activity with a relatively low cost, time, and effort [16]. Mostly used throughout the day, mobile phones are considered to be a good tool for tracking physical activity in real time. Electronic activity monitors play a potential role as a delivery medium by replicating most aspects of pedometer-based interventions. These monitors measure physical activity or behavior indicators, such as heart rate, and are connected with a mobile device through an app or personal computer to provide extensive feedback. The feedback can be more individualized than that offered in clinical assessments and can include social comparisons, multiple charts, and markers of progress toward individual goals. Considering the ubiquity of mobile phones, few clinical trials have been conducted to assess the impact of the application of mobile phone-wearable step trackers on improvement in metabolic impairment. A number of previous studies have been designed to explore the effectiveness of mobile phone apps in weight loss [17,18] or in increasing physical activity [19,20]. However, they failed to include people with metabolic syndrome and to incorporate a wearable device into their interventions, and none of them were done in a clinical setting.

Given that many people try to use wearable devices compatible with mobile phones to boost physical activity with little evidence-based practice, it is clinically pressing to demonstrate the potential of feedback via this technology for the management of outpatients with metabolic syndrome. There is also a need to assess whether there is any obstacle when applied to real clinical situations. Therefore, we conducted a pilot study aimed to explore the potential and barriers of the application of these devices for metabolic syndrome management in a clinical setting.

Methods

Study Participants

The study included patients diagnosed with metabolic syndrome based on the results of a comprehensive health examination at a health promotion center at Pusan National University Hospital (Busan, South Korea) between March and December 2016. The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board at Pusan National University Hospital (review reference number: E-2015030). All participants consented to the study protocols and the subsequent publication of their respective findings.

Inclusion criteria were as follows: (1) diagnosed with metabolic syndrome by a doctor, (2) aged between 20 and 64 years, (3) possession of a mobile phone and a daily mobile phone user for the past 3 months, and 4) no plans to change medication within 3 months after the start of the study.

Definition of Metabolic Syndrome

Metabolic syndrome was diagnosed by a family doctor according to the 2009 Joint Interim Statement issued by a number of international organizations and expert groups [21]. The definition requires the presence of three or more of the following five components: (1) central obesity (ie, waist circumference of ≥85 cm for women and ≥90 cm for men according to the Korean Society for the Study of Obesity [22]), (2) hyperglycemia (ie, fasting plasma glucose [FPG] ≥100 mg/dL or the use of antidiabetes medication), (3) high blood pressure (BP) (ie, systolic BP ≥130 mm Hg or diastolic BP ≥85 mm Hg or the use of antihypertensive medication), (4) hypertriglyceridemia (ie, fasting plasma triglycerides ≥150 mg/dL or the use of lipid-lowering medication), and 5) low high-density lipoprotein cholesterol (HDL-C) (ie, fasting plasma HDL-C<50 mg/dL for women, <40 mg/dL for men). Patients with any of the following conditions were excluded: (1) any cancer or uncontrolled metabolic or cardiovascular disease, (2) prescriptions for medication that may affect metabolism, and (3) disability or difficulty with regular walking. All participants were recommended for regular exercise of at least five times

per week by a doctor as part of the treatment for metabolic syndrome, according to the current guideline [21].

Data Collection

All study participants underwent an 8 hour-fasting blood test as part of a comprehensive health examination. Data on medical history and anthropometric measurements of participants were collated. Participants also completed a questionnaire about demographics, medical history (diagnosis of or medication for hypertension, diabetes mellitus, or dyslipidemia), and health-related habits (smoking, drinking, and alcohol). To assess physical activity at baseline, participants were asked about the frequency and duration of vigorous and light/moderate physical activity. Individuals who were classified as inactive reported no sessions of light/moderate or vigorous activity of at least 10 minutes' duration. Those classified as having some activity reported at least one session of light/moderate or vigorous physical activity of at least 10 minutes' duration but did not meet the definition of regular exercise. Those classified as having regular activity reported three or more sessions per week of vigorous activity lasting at least 20 minutes or five or more sessions per week of light/moderate activity lasting at least 30 minutes in duration [6,21]. Participants were divided into nonsmokers, former smokers, or current smokers, and into nondrinkers (0-98 g/week) or drinkers defined as drinking of an average of seven cups for men and five or more cups for women, more than two times per week [23].

Participants' BP was tested three times in the sitting position after a 15-minute rest using a BP-203 RVII (Colin Corp, Aichi, Japan), and the average measurement was recorded. Body weight and height were measured using a digital scale and stadiometer (BSM370, Biospace Co Ltd, Seoul), with patients wearing a light gown without shoes. Body mass index (BMI) was calculated as weight (kg) divided by height squared (m^2). Waist circumference was assessed by trained examiners (following a normal expiration) at the midpoint between the lower costal margin and the iliac crest, to the nearest millimeter. Body composition, including percentage body fat and muscle mass, were calculated via bioelectric impedance analysis (Inbody 720, Biospace Co Ltd, Seoul). Blood samples were collected from an antecubital vein after an 8-hour fast. Samples were then analyzed at a laboratory in our hospital. Lipid profiles were tested using an autoanalyzer (Hitachi 747, Hitachi Corp, Japan) and an enzymatic colorimetric method. FPG levels were evaluated using a glucose oxidase method and a Synchron LX 20 (Beckman Coulter, Fullerton, CA). All participants were assessed again at 12 weeks after the start of the intervention.

Intervention Using the Wearable Device

All participants were advised to walk aiming at the consumption of a minimum of 150 kcal per day, which was set based on the standard recommendations for metabolic syndrome encouraging a daily minimum of 30 minutes of moderate-intensity (such as brisk walking) physical activity [21]. To estimate the required amount of physical activity for our metabolic syndrome participants, we referenced the study that determined the effects of exercise amount and intensity of metabolic syndrome [24]. The intervention they used for the exercise training group with a low amount of moderate-intensity physical activity was equivalent to 1221 (SD 222) kcal/week with no significant difference between genders. Considering their participants' mean BMI of 29.9 (SD 3.2) kg/m^2, and the estimated mean BMI of 23.1 (SD 0.1) kg/m^2 in Koreans with metabolic syndrome (referenced to the study on 6561 metabolic patients among a nationally representative sample of Koreans) [22], it made sense to set the calorie goal of 150 kcal per day loss for our participants to consume by exercising. Rather than informing the participants about the number of steps to walk, we had them put the minimum calorie goal of 150 kcal loss on their mobile phone app, which was connected to the wearable device that automatically calculated the required step counts to be taken depending on the weight and height of the participants. Then when the user walked, the app measured the estimated calorie consumed, taking into account the walking speed.

Participants were given a wearable device (Coffee WALKIE +Dv.3, GC Healthcare CI, Korea) fitted on the wrist or waist. This wearable device has been certified by the Korea Testing Laboratory (Certification Number: MSIP-CRM-NSJ-Coffee). The device used was validated on the wrist with 86.7% accuracy for the calories burned and 90.5% for the step counts. Considering previous data showing that the waist attachment site detects step counts better than that on the wrist [25,26], we expected similar accuracy regarding the waist placement of the device in our study. Moreover, increased wear times can occur with the wrist placement site [27] and the US National Health and Nutrition Examination Survey also used a step counter worn on the wrist (2008-2014) or the waist (2003-2006) [25]. Therefore, we permitted the participants to wear the device either on the waist or wrist, according to their preference. However, all participants ultimately chose their nondominant wrist as the attachment site, saying that it would be more comfortable for all-day wear, handling, and checking of the device. Detailed instructions on wearable device usage were provided during face-to-face/one-on-one contact by the study coordinator with each participant. Demonstration and instruction sheets including on how to turn on/off and how to operate the wearable device, download the app, and connect the mobile phone and the wearable device to Bluetooth were performed. Subsequently, the compatible app was installed on the mobile phones of the participants, and this was connected to the wearable device. To ensure that the participants become used to the device, they were asked to individually try using it several times under supervision. This practice lasted for approximately 30 minutes. Each participant also received a log-in identity and password that granted them access to the mobile phone app and website where they could track their physical activity over the 12-week study duration. The number of steps taken and calories consumed were displayed on the device's screen. In addition, participants could check if they had achieved the amount of daily walking. Their workout records were also registered on an online program. An administrative webpage granted the researchers access to check and track the step counts of the participants on a daily basis.

Participants were instructed to wear the device for as many waking hours as possible except while swimming and bathing. All days of exercise were included in the analysis except when the number of step counts exceeded 25,000 in one day, which

is considered an extremely high count, or when participants withdrew consent [28]. If the data from the wearable device were not transmitted to the Web server for three consecutive days, an automatic text message was sent to the participant to encourage the use of the device and to see if they had any issue with the device (in which case, they were asked to contact the researcher to address the issue). We then checked the Web server again to see whether step count data were being detected. If no response was received from the user for 6 days in a row, the researcher called to ask if they had any issue with the device or whether it was being used appropriately. Participants who failed to respond to the researcher and those who rejected participation were considered to be dropouts. Thus, we were able to include step records of at least one day of every week. Then we estimated the daily step counts by dividing the total step counts per week by seven. To be included in the analysis, at least one day with any step counts was required. Laboratory-based studies typically require 4 days or more of valid data to be included as a study sample. However, quite a few studies required at least one valid day to be included in the analysis, a method that is consistent with the original examinations of the National Health and Nutrition Examination Survey physical activity data.

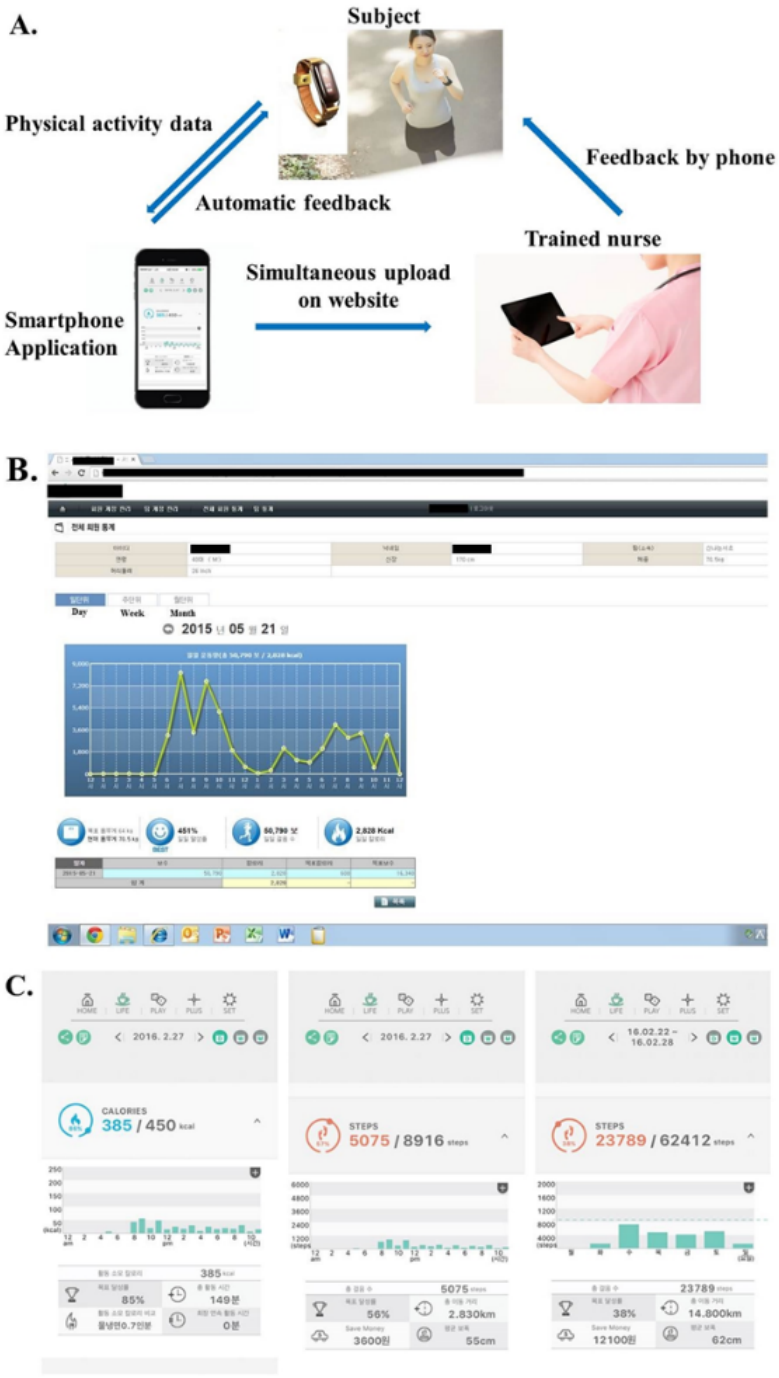
Over the 12 weeks of the trial, participants were provided with feedback by the nurse via phone on alternate weeks. A trained nurse phoned each participant for personal counseling on their physical activity every other week at an agreed-on time. This included counseling regarding exercise practices and encouragement to continue with the exercise regimen based on current guidelines and recommendations on physical activities. The telephone feedback followed a standardized script (Multimedia Appendix 1) but was flexible depending on the individual, including goal achievement. The consultant answered questions or discussed problems regarding the use of the wearable device, the app (Figure 1), and the exercise, and also provided encouragement to continue with the exercise. This intervention was designed to help improve problem solving, goal setting, and self-monitoring. Participants also had access to additional resources online.

Statistical Analysis

The sample size for the study was calculated based on the results of a randomized weight-loss intervention using self-monitoring mobile devices [29] that showed an effect size of 0.53. The sample size was 24 for two-sided tests of significance at alpha=.05 and power 1−beta=80%. To account for attrition, we attempted to recruit approximately 30 participants in total.

Data are presented as frequency or proportion and mean and standard deviation (SD) for normally distributed values, or medians and interquartile ranges (IQRs) for nonnormally distributed values. Changes in the components of metabolic syndrome for each participant were examined from baseline to follow-up visit (week 12). The Shapiro-Wilk test was employed to test the normality assumption. To compare baseline characteristics between the study completed and uncompleted groups, we employed the independent *t* test for age and the chi-square test for categorical variables. Comparisons of metabolic components between baseline and at week 12 were conducted using the paired *t* test (waist circumference, body fat mass, and body fat rate). Because some metabolic parameters were nonnormally distributed, the Wilcoxon signed-rank test was used to detect differences between baseline and at week 12 (FPG, systolic/diastolic BP, triglycerides, HDL-C, low-density lipoprotein cholesterol, and total cholesterol). A repeated-measures analysis of variance (ANOVA) was conducted to verify the difference in step count and calorie expenditure changes by walking, at weeks 4, 8, and 12. All analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA), and a *P* value of <.05 was considered statistically significant.

Figure 1. Overview of the intervention using the wearable device and a mobile phone app. (A) Flowchart of participants' physical activity data collection. (B) Administrative webpage allowing researchers to track the participants' physical activity. (C) App screen on the participants' mobile phones presenting the goal setting of physical activity, progress toward the daily goal, total daily steps taken, and estimation of calories burned by physical activity performed.



Results

Participant Characteristics

Of the 53 participants recruited, 20 participants with a median age of 46 years completed the exercise monitoring schedule over 12 weeks (Figure 2). The baseline characteristics of the participants who completed the intervention are presented in Table 1. There was a statistically significant difference in participants' age between those who did and did not complete the intervention (mean 51.97, SD 8.49 years vs mean 44.20, SD 9.55 years; *P*=.003) (Table 1). The majority of participants ceased monitoring their exercise within the first few weeks of the study (Figure 3) primarily due to miscommunication between their mobile device and the wearable device. Only five of the participants who completed the trial (n=20) responded that they had regularly exercised at baseline, whereas 55% (11/20) of them had been physically inactive.

Figure 2. Flowchart of patient enrollment and the reasons for dropout. Right-about: withdrawal of consent within 24 hours; lost to follow-up: participant neither picked up the counseling call nor provided any reasons.

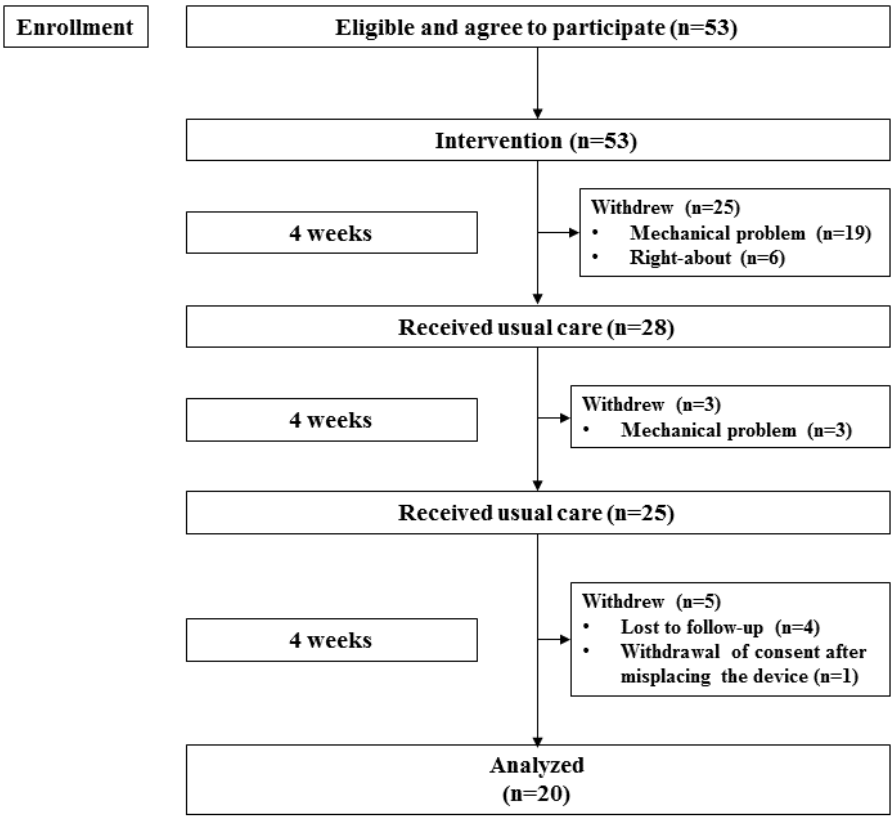


Table 1. Clinical characteristics of the study participants.

Variables	Total	Incomplete group	Complete group	*P* value[a]
Participants, n (%)	53 (100)	33 (63.3)	20 (37.7)	—
Age (years), mean (SD)	49.04 (9.60)	51.97 (8.49)	44.20 (9.55)	.003
Male, n (%)	39 (73.6)	22 (66.7)	17 (85)	.20
Current smoker, n (%)	11 (20.8)	5 (15.2)	6 (30)	.17
Former smoker, n (%)	10 (18.9)	6 (18.2)	4 (20)	.30
Alcohol drinker,[b] n (%)	18 (34.0)	12 (36.4)	6 (30)	.77
Regular exercise,[c] n (%)	19 (35.8)	14 (42.4)	5 (25)	.20
Physically inactive,[d] n (%)	29 (54.7)	18 (54.5)	11 (55)	.81
Medication, n (%)				
Antihypertensive	25 (47.2)	16 (48.5)	9 (45)	.81
Antidiabetic	15 (28.3)	11 (33.3)	4 (20)	.30
Antihyperlipidemia	18 (34.0)	13 (39.4)	5 (25)	.28
Total number of abnormal metabolic components, n (%)				.19
3	31 (58.5)	18 (54.5)	13 (65)	
4	13 (24.5)	7 (21.2)	6 (30)	
5	9 (17.0)	8 (24.2)	1 (5)	
Have abnormal metabolic component, n (%)[e]				
Central obesity	44 (83.0)	26 (78.8)	18 (90)	.26
Hyperglycemia	29 (54.7)	16 (48.5)	13 (65)	.59
High blood pressure	48 (90.6)	31 (93.9)	17 (85)	.12
Hypertriglyceridemia	44 (83.7)	30 (90.9)	14 (70)	.002
Low high-density lipoprotein cholesterol (HDL-C)	29 (54.7)	16 (48.5)	13 (65)	.24

[a]*P* value was obtained by chi-square test or independent *t* test.

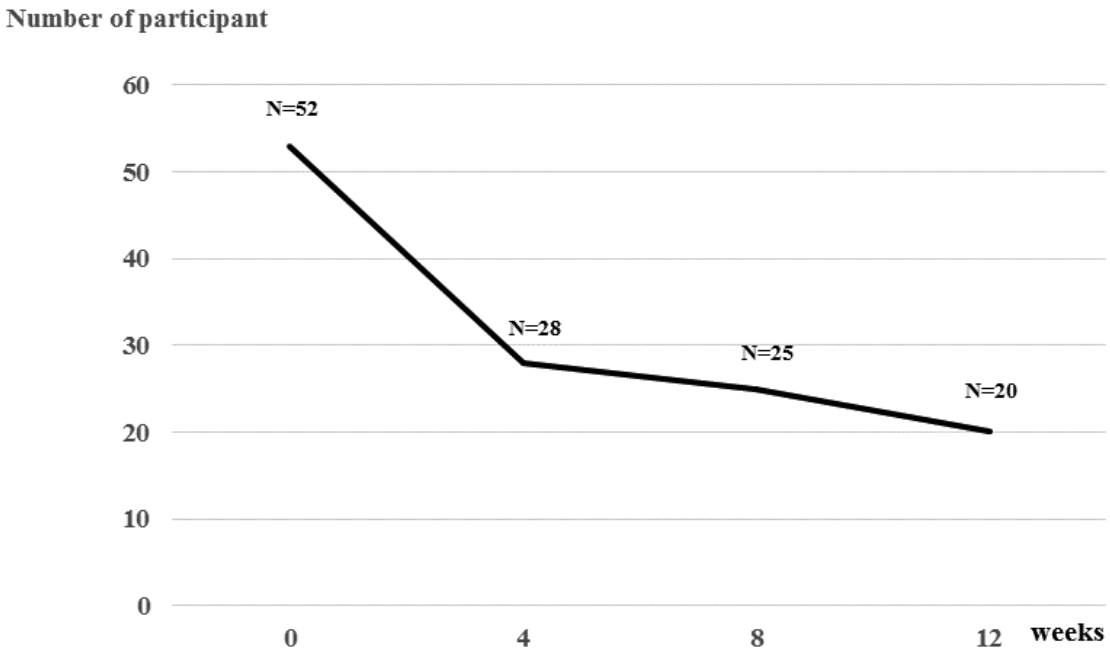
[b]Alcohol drinking was defined as drinking of an average of seven cups for men and five or more cups for women, more than two times per week.

[c]Regular exercise was defined as having regular activity three or more sessions per week of vigorous activity lasting at least 20 minutes, or five or more sessions per week of light/moderate activity lasting at least 30 minutes in duration.

[d]Physically inactive was classified as participation in no sessions of light/moderate or vigorous activity of at least 10 minutes' duration.

[e]Central obesity: waist circumference ≥90 cm for men, ≥85 cm for women. Hyperglycemia: fasting plasma glucose ≥100 mg/dL or the use of antidiabetes medication. High BP: systolic BP ≥130 or diastolic BP ≥85 mm Hg or the use of antihypertensive medication. Hypertriglyceridemia: fasting plasma triglyceride ≥150 mg/dL or the use of lipid-lowering medication. Low HDL-C: fasting plasma HDL-C<50 mg/dL for women, <40 mg/dL for men.

Figure 3. Timeline of participant dropout.



Mean Number of Steps and Amount of Calorie Expenditure

The mean number of steps per week on weeks 4, 8, and 12 are described in Table 2. For the first 4 weeks, the participants walked 7616 steps daily and then increased the step counts to 8244, which was equivalent to a 17.5% increase from week 4, for the next 4 months. Between week 8 and week 12, the number of steps remained at the same level as that of week 4. Although there was no statistically significant upward tendency in step counts over time, the results indicated that the participants walked more for the first 8 weeks after enrollment in the trial and maintained the amount of walking at the same level for the last 4 weeks. Thus, it was assumed that the participants consumed a mean of approximately 420 calories daily by walking. Additionally, there was no significant difference in week 4 (P=.90), week 8 (P=.52), and week 12 (P=.63) in step counts between the participants who had regularly exercised before the study and those who had not (Table 2).

Resolution of the Metabolic Syndrome Status

At 12 weeks following the intervention, metabolic syndrome was resolved in 9 of 20 (45%) participants. We also noted a significant reduction in metabolic syndrome prevalence between baseline and follow-up (Figure 4). The number of metabolic abnormalities decreased in 11 of 20 (55%) participants. In 7 of 20 (35%) participants, there was no change in the number of metabolic abnormalities following the intervention. Metabolic syndrome was aggravated in 1 (5%) participant.

Table 2. The daily mean number of steps walked and estimated calories burned (N=20) for groups 1 and 2[a].

Week and measure	Total (N=20), mean (SD)	Group 1 (n=5), mean (SD)	Group 2 (n=15), mean (SD)	*P* value[b]
Week 4				
Daily steps	7615.8 (3669)	7811.3 (2622)	7550.7 (4036)	.90
Daily calories burned	410.4 (171)	468.0 (117)	391.2 (184)	.40
Week 8				
Daily steps	8244.4 (3746)	9202.0 (4062)	7925.2 (3727)	.52
Daily calories burned	445.9 (208)	548.7 (258)	411.7 (187)	.30
Week 12				
Daily steps	7510.4 (3525)	8194.0 (3440)	7282.6 (3640)	.63
Daily calories burned	413.7 (211)	448.6 (160)	398.2 (172)	.57

[a]Group 1: the participants who regularly exercised at baseline; group 2: the participants who did not regularly exercise at baseline.

[b]Derived from a *t* test to compare step counts and calories burned between the participants who regularly exercised at baseline and those who had not.

Mean Change of Each Metabolic Component

The total number of metabolic impairment components in the 20 participants decreased from 68 to 58, indicating that the mean number of metabolic impairment components per person decreased from 3.4 to 2.9 (Figure 4B). Figure 5 shows the changes in the data of each metabolic component and body composition. After the intervention, systolic BP and diastolic BP significantly decreased by 6.71% and 7.98%, respectively, compared to the baseline assessment (both *P*=.02). The levels of FPG (9.24%) and triglyceride (7.49%) also decreased considerably compared to the baseline; however, these changes were not statistically significant. Furthermore, there were no significant differences in body weight, waist circumference, and body fat between baseline and week 12.

Figure 4. Change in the metabolic abnormality status in study participants from baseline to follow-up at 12 weeks. (A) Resolution of metabolic syndrome. (B) Changes in components of metabolic syndrome.

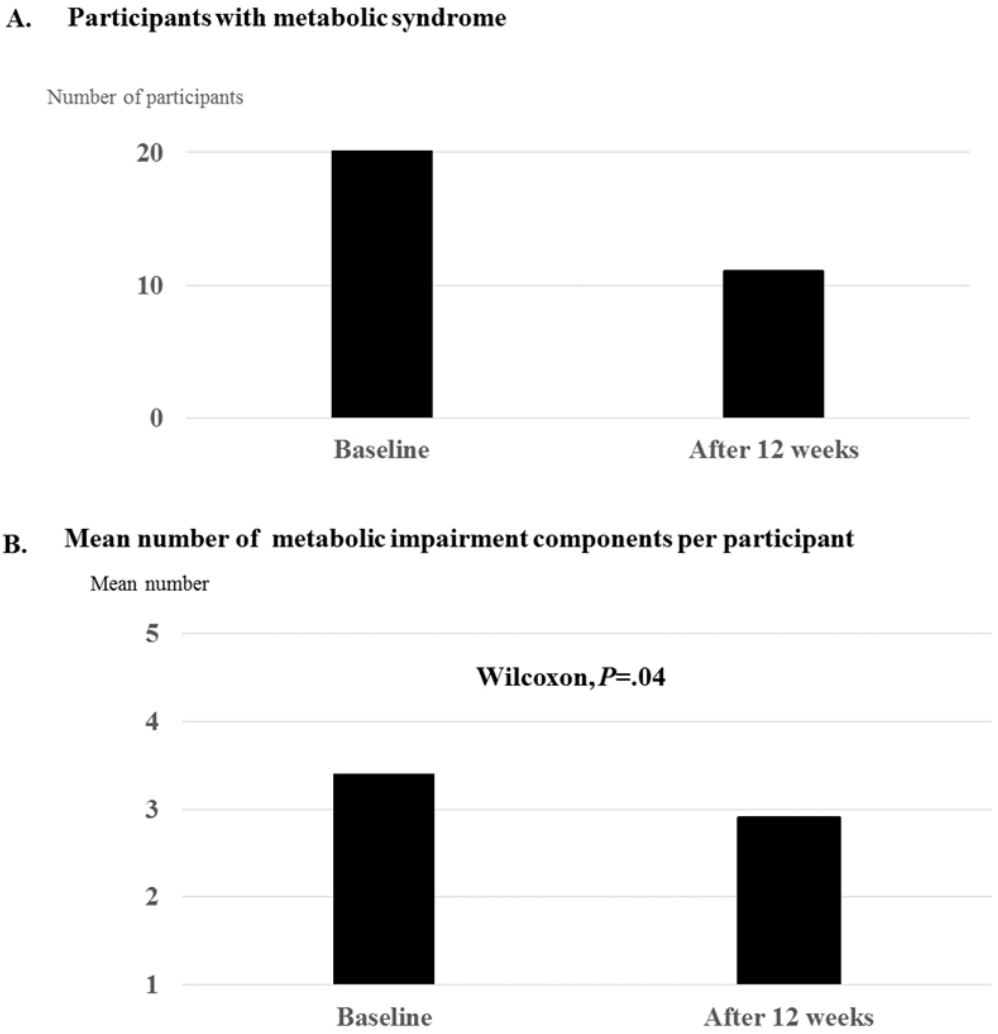
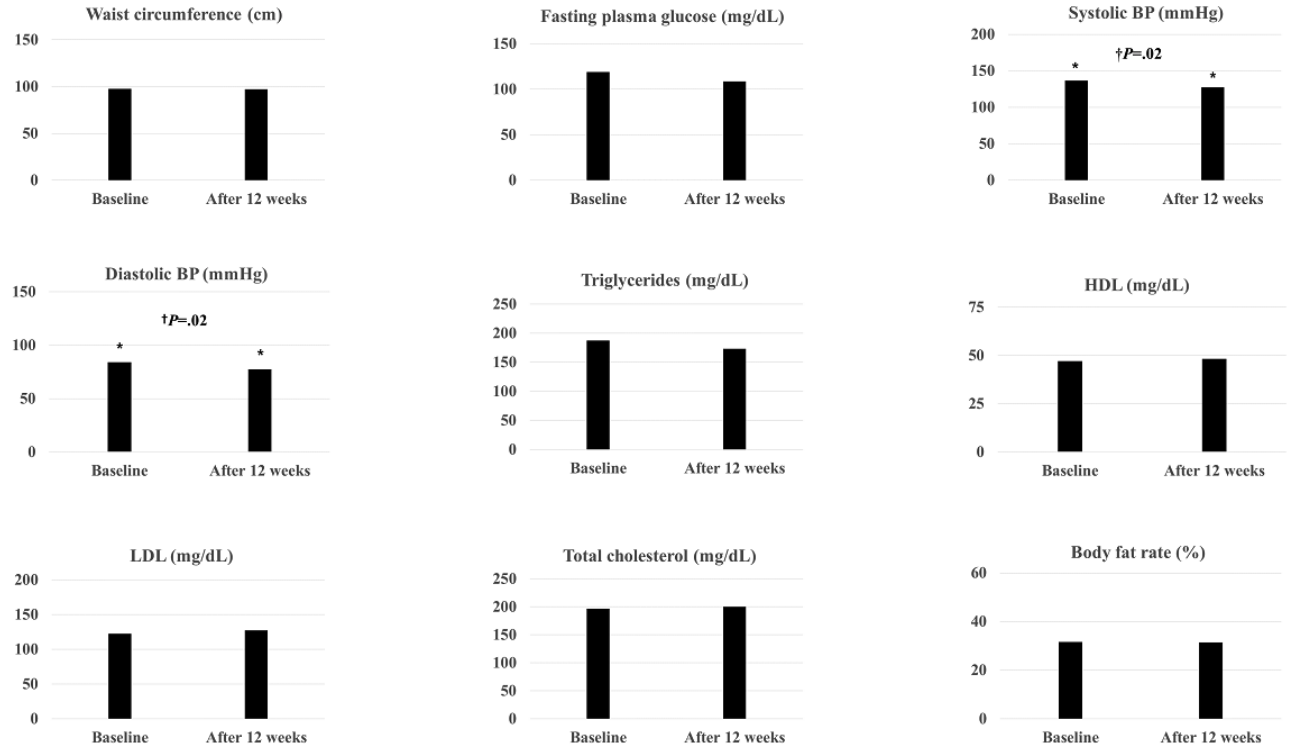


Figure 5. Changes in metabolic components and body compositions after the 12-week intervention. †: Wilcoxon signed-rank test; BP: blood pressure; HDL: high-density lipoprotein; LDL: low-density lipoprotein.



Discussion

Principal Results and Comparison With Prior Work

To the best of our knowledge, we conducted the first pilot study to explore the potential and barriers of adoption of connected care technologies using wearable activity trackers with a mobile phone app for metabolic syndrome patients' management in a clinical setting. Our results showed that this intervention promoted physical activity for 12 weeks, resulting in significant improvement in metabolic impairments—especially in BP—leading to a 45% metabolic syndrome resolution rate. However, we saw a high dropout rate, revealing substantial obstacles in applying this technology in real life.

The main finding of this study was a significant decrease in systolic BP by 9.2 mm Hg (6.71%) and diastolic BP by 6.65 mm Hg (7.98%) 12 weeks after the first use of the device. The degree of reduction in systolic BP observed in our study is equivalent to that obtained through antihypertensive medications with a mean of 8.8 mm Hg (95% confidence interval [CI] −9.58 to −8.02). It was also superior to that obtained through structured exercise interventions with a mean of 4.84 mm Hg (95% CI −5.55 to −4.13) according to a recent meta-analysis of 391 randomized controlled trials comparing the effect of medications and exercise regimens on systolic BP [30]. Additionally, systolic and diastolic BP were the most improved components among the metabolic indexes in our study, which is consistent with prior study, showing that systolic and diastolic BP were most inversely related with steps per day tracked by a pedometer [31]. It is well established that regular walking is effective in BP control by improving vasodilatory function, lowering vasoconstrictor tone, and mainly decreasing endothelin-1 endogenous bioavailability. However, walking is required to be almost daily for a significant improvement in BP [32]. However, many people reportedly find it difficult to maintain daily walking without any motivation or feedback [16]. In this context, the observed step counts and estimated calorie expenditure at weeks 4, 8, and 12 were quite impressive given that more than half the participants had not engaged in any type of exercises at baseline.

For the mean number of steps taken throughout the trial, there was obviously a consistency in the walking behavior among the participants who completed the 12-week intervention. For the first 4 weeks after wearing the device, the participants walked on average 7616 steps daily. This value is equivalent to the "moderately active" level according to the steps-per-day categories and classification system of Tudor-Locke and Bassett [33]. Furthermore, the participants took 17.5% more steps to reach 8244 in the next 4 months. Between weeks 8 and week 12, the number of steps remained at the same level as that shown in the first 4 weeks. Thus, these results indicate that the participants fulfilled the amount of physical activity during the 12 weeks of trial that was recommended in the current guidelines for metabolic syndrome management. Given that only five participants in our study conducted regular exercise at baseline, it was assumed that at least 55% of the participants started being more physically active than before enrollment in the trial.

Although the number of steps that they had walked before enrollment in the study was not assessed, our data showed no significant difference in step counts between those who had regularly exercised before the study and those who had not. It is clinically meaningful for people who were once physically inactive to maintain a moderately active level for 12 weeks, which can be considered a driving contributor to the 45% resolution in metabolic syndrome in our study.

On the other hand, this study failed to see a significant reduction in waist circumference, which is the main indicator of metabolic syndrome. This is in line with the finding of previous studies that saw no improvement in waist circumference or body weight despite having promoted physical activity through feedback using wearable devices [34,35]. The result is possibly due to the increase in physical activity without dietary intervention, which is not sufficient to bring about weight loss.

Traditional metabolic syndrome therapies are focused on reducing insulin resistance through dietary control, physical exercise, and lifestyle modification. However, the majority of the general population have no or difficult access to personally supervised exercise programs or fitness centers with personal health trainers [11]. Moreover, it is challenging for patients with metabolic syndrome to correct their sedentary lifestyle and sustain health-related changes in their everyday lifestyle [21]. Although doctors can prescribe medications, they cannot constantly oversee the routine of their patients. Therefore, it is important to implement strategies to encourage or support physical activity on a daily basis among patients with metabolic syndrome. Mobile phones have evolved as an integral part of people's lives, and innovative apps support consumers in various ways. With an upsurge in mobile phone distribution, Korea is one of the countries with the highest number of mobile phone users worldwide [36]. The development of mobile phones with advanced computer technology to support Web browsing, third-party apps, wireless connectivity, and sensors (eg, pedometers, accelerometers, and global positioning system trackers) has opened up a new era for personal health self-management [16]. Along with a wide distribution of mobile phones and innovative mobile apps, several low-cost or free health care self-management apps have been developed for continuous and personalized tracking of daily activities. Historically, validated medical devices implanted with sensors have been adopted in clinical studies and targeted research conducted in medical settings. However, technological advances in the measurement of activity (eg, steps), biochemistry (eg, pH), and physiology (eg, blood oxygen saturation) have now fostered the development of patient care and research outside of the hospital setting [37].

Currently, most popular consumer-accessible wearable devices estimate movements via accelerometers that apply algorithms to calculate activity levels (generally in the form of steps taken) and calories expended. These wearable devices differ from traditional pedometers by adding a variety of techniques related to health behavior changes, such as social support or comparisons, goal setting, and rewards [38]. There are various types of pedometers, ranging from simple and inexpensive ones that quantify steps to technologically advanced accelerometers that can measure the amount and intensity of physical activity and total consumed calories in daily life [33-35]. These activity monitors can be worn without major inconvenience, are easy to use, and are compatible with most daily activities [31]. Previous reviews have already reported that physical activity counseling is associated with a significant increase in self-reported daily physical activity levels [39]. However, self-reported measurements of physical activity cannot provide valuable data because they are inaccurate, and there may be a social desirability bias [31].

Studies have found that interventions consisting of counseling combined with activity monitoring have a positive effect on daily physical activity levels in participants with chronic disease [7,9-11]. However, information regarding the effects of counseling with activity monitoring on objective measures of physical activity and health-related outcomes (eg, BMI, BP, and lipid profiles) remains limited to date. Although the use of wearable devices, such as wristbands, smart watches, and biomonitors, has considerably increased worldwide, comparatively few studies have investigated their effect on the resolution of metabolic syndrome. The most important limitation of wearable devices is that for the majority of users, most devices fail to drive long-term sustained engagement. A recent Pew Internet and American Life survey showed that over 59% of customers in the United States who own an activity tracker no longer use it and one-third of customers who own a device no longer use it after 6 months [40]. Designing a strategy to ensure sustained engagement is the key to the long-term success of wearable devices in health care. Most products and services offering a range of uses fail to cause a meaningful change in users' health-related behaviors and habits; for example, activity tracker users rapidly abandon devices that do not help them to make positive changes and, ultimately, they fail to achieve self-care in health [41,42]

Our study observed a significant dropout rate that was higher than expected. One of the possible reasons was that our participants were relatively older (median age 46 years) compared to previous studies [16,36]. The participants who failed to complete the trial were older than those who completed it. This is consistent with previous research that investigated the potential sociodemographic characteristics of individuals using wearable activity trackers or mobile phone apps for physical activity surveillance [38]. We included participants who were daily mobile phone users of more than 3 months to minimize the dropout rate caused by unfamiliarity with handling of the app or of the Bluetooth on mobile phones required for connecting the wearable device. Relevant research showed that the older participants were more likely to be unwilling to try to address the issues with mobile phones when that happens [41]. This can be an explanation for why most of those who dropped out in our study did so during the first 4 weeks after they started using the device. Despite efforts of researchers to solve the technical issues when they happened and encourage the users to try again, the majority of participants who experienced the problems withdrew their consent due to dissatisfaction with the device. Particularly, older participants had trouble learning how to operate the device or getting used to it. Additional education should have been provided for participants aged 50 years and older to resolve any technical problems. This phenomenon illustrates the importance of making smart medical devices user-friendly in clinical settings, and suggests the main point to be addressed when health providers consider applying this digital health self-tracking system to their patients. Devices that incorporate a function allowing users to change their habits will encourage sustained behavioral changes and result in long-term self-management of health care. Further study is needed to

identify factors that influence users' intention to continue using connected care technologies and the reasons for usage discontinuance.

Limitations

This analysis has several limitations that warrant consideration when interpreting the results. First, the sample size was small, and the intervention period was relatively short. The participants were individuals undergoing health examination in Korea; therefore, it is not possible to generalize our findings to other ethnicities or geographic regions. Secondly, a significant proportion of the participants were lost to follow-up. Moreover, this study included only a few middle-aged women due to the low number of women volunteers. During the study, we were unaware if any participant received another commercialized program for dietary control. Finally, our study had no control group in which participants were provided with educational programs to increase their physical activities instead of using the wearable device. As a result, the positive findings seen in this study might have come from participants' behavior changes such as healthy dietary choice or motivation by the nurse's telephone feedback rather than the wearable device itself.

Despite these limitations, our study is considerably valuable owing to several strengths. Firstly, to the best of the authors' knowledge, this is the first study to explore the potential and possible barriers to the adaptation of a wearable device connected to a mobile phone app for the management of metabolic syndrome in a clinical setting. Another primary strength of this study is the comprehensive health examination-based screening performed that resulted in an accurate diagnosis of metabolic syndrome by a family medicine doctor, unlike previous studies in which the diagnosis of metabolic syndrome was based on self-reports. Therefore, recall bias and the risk of disease misclassification were minimized. Furthermore, the study collated detailed information on participants' medical history, demographics, and laboratory results, which allowed the exclusion of ineligible participants and adjustments for important potential confounders during the analysis.

Conclusion

In conclusion, this study found that a 12-week intervention via feedback based on a wearable physical activity monitor helped metabolic syndrome patients to engage in more regular walking and improved impaired metabolic components especially in terms of BP. However, some practical challenges need to be addressed with respect to patients' adherence and sustained engagement.

Acknowledgments

We would like to thank the Research & Development Project of Pusan National University Hospital and KT for providing the wearable device (CMITKT-06).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Content of the telephone feedback by a trained nurse.

[DOCX File, 15KB - mhealth_v7i6e13381_app1.docx]

References

1. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet 2005;365(9468):1415-1428. [doi: 10.1016/S0140-6736(05)66378-7] [Medline: 15836891]
2. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation 2005 Nov 15;112(20):3066-3072. [doi: 10.1161/CIRCULATIONAHA.105.539528] [Medline: 16275870]
3. Sattelmair J, Pertman J, Ding EL, Kohl HW, Haskell W, Lee I. Dose response between physical activity and risk of coronary heart disease: a meta-analysis. Circulation 2011 Aug 16;124(7):789-795 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.010710] [Medline: 21810663]
4. Bird SR, Hawley JA. Update on the effects of physical activity on insulin sensitivity in humans. BMJ Open Sport Exerc Med 2016;2(1):e000143 [FREE Full text] [doi: 10.1136/bmjsem-2016-000143] [Medline: 28879026]
5. Woodcock J, Franco OH, Orsini N, Roberts I. Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. Int J Epidemiol 2011 Feb;40(1):121-138. [doi: 10.1093/ije/dyq104] [Medline: 20630992]
6. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
7. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med 2015 Apr 09;13:77 [FREE Full text] [doi: 10.1186/s12916-015-0319-2] [Medline: 25889598]

8. Takemoto M, Lewars B, Hurst S, Crist K, Nebeker C, Madanat H, et al. Participants' perceptions on the use of wearable devices to reduce sitting time: qualitative analysis. JMIR Mhealth Uhealth 2018 Mar 31;6(3):e73 [FREE Full text] [doi: 10.2196/mhealth.7857] [Medline: 29599105]
9. Bavan L, Surmacz K, Beard D, Mellon S, Rees J. Adherence monitoring of rehabilitation exercise with inertial sensors: a clinical validation study. Gait Posture 2019 Dec;70:211-217. [doi: 10.1016/j.gaitpost.2019.03.008] [Medline: 30903993]
10. Fernández Peruchena CM, Prado-Velasco M. Smart sensors and virtual physiology human approach as a basis of personalized therapies in diabetes mellitus. Open Biomed Eng J 2010;4:236-249 [FREE Full text] [doi: 10.2174/1874120701004010236] [Medline: 21625646]
11. Vaes A, Cheung A, Atakhorrami M, Groenen M, Amft O, Franssen F, et al. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: A systematic review and meta-analysis. Ann Med 2013 Sep;45(5-6):397-412. [doi: 10.3109/07853890.2013.810891] [Medline: 23952917]
12. Umpierre D, Ribeiro PA, Kramer CK, Leitão CB, Zucatti AT, Azevedo MJ, et al. Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. JAMA 2011 May 04;305(17):1790-1799. [doi: 10.1001/jama.2011.576] [Medline: 21540423]
13. Quinn CC, Butler EC, Swasey KK, Shardell MD, Terrin MD, Barr EA, et al. Mobile diabetes intervention study of patient engagement and impact on blood glucose: mixed methods analysis. JMIR Mhealth Uhealth 2018 Feb 02;6(2):e31 [FREE Full text] [doi: 10.2196/mhealth.9265] [Medline: 29396389]
14. Alharbi M, Straiton N, Gallagher R. Harnessing the potential of wearable activity trackers for heart failure self-care. Curr Heart Fail Rep 2017 Dec;14(1):23-29. [doi: 10.1007/s11897-017-0318-z] [Medline: 28181075]
15. Robinson SA, Shimada SL, Quigley KS, Moy ML. A web-based physical activity intervention benefits persons with low self-efficacy in COPD: results from a randomized controlled trial. J Behav Med 2019 Apr 12:e1. [doi: 10.1007/s10865-019-00042-3] [Medline: 30980223]
16. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
17. 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]
18. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 01;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
19. 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 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
20. van Dantzig S, Geleijnse G, van Halteren A. Toward a persuasive mobile application to reduce sedentary behavior. Pers Ubiquit Comput 2012 Jul 12;17(6):1237-1246. [doi: 10.1007/s00779-012-0588-0]
21. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, International Diabetes Federation Task Force on EpidemiologyPrevention, Hational Heart, Lung,Blood Institute, American Heart Association, World Heart Federation, International Atherosclerosis Society, International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009 Oct 20;120(16):1640-1645 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192644] [Medline: 19805654]
22. Lee SY, Park HS, Kim DJ, Han JH, Kim SM, Cho GJ, et al. Appropriate waist circumference cutoff points for central obesity in Korean adults. Diabetes Res Clin Pract 2007 Jan;75(1):72-80. [doi: 10.1016/j.diabres.2006.04.013] [Medline: 16735075]
23. Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. Prev Med 2004 May;38(5):613-619. [doi: 10.1016/j.ypmed.2003.11.027] [Medline: 15066364]
24. Johnson JL, Slentz CA, Houmard JA, Samsa GP, Duscha BD, Aiken LB, et al. Exercise training amount and intensity effects on metabolic syndrome (from Studies of a Targeted Risk Reduction Intervention through Defined Exercise). Am J Cardiol 2007 Dec 15;100(12):1759-1766 [FREE Full text] [doi: 10.1016/j.amjcard.2007.07.027] [Medline: 18082522]
25. Bassett DJ, Toth L, LaMunion S, Crouter S. Step counting: a review of measurement considerations and health-related applications. Sports Med 2017 Jul;47(7):1303-1315 [FREE Full text] [doi: 10.1007/s40279-016-0663-1] [Medline: 28005190]
26. Tudor-Locke C, Barreira T, Schuna J. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc 2015;47(4):839-842. [Medline: 25121517]
27. Troiano R, McClain J, Brychta R, Chen K. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]

28. Tudor-Locke C, Bassett D, Shipe M, McClain J. Pedometry methods for assessing free-living adults. J Phys Act Health 2011 Mar;8(3):445-453. [Medline: 21487145]
29. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011 Dec;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
30. Naci H, Salcher-Konrad M, Dias S, Blum MR, Sahoo SA, Nunan D, et al. How does exercise treatment compare with antihypertensive medications? A network meta-analysis of 391 randomised controlled trials assessing exercise and medication effects on systolic blood pressure. Br J Sports Med 2018 Dec 18:bjsports-2018-099921. [doi: 10.1136/bjsports-2018-099921]
31. Jahan N, Shenoy S. Relation of pedometer steps count & self reported physical activity with health indices in middle aged adults. Diabetes Metab Syndr 2017 Dec;11 Suppl 2:S1017-S1023. [doi: 10.1016/j.dsx.2017.07.033] [Medline: 28780228]
32. Sabbahi A, Arena R, Elokda A, Phillips S. Exercise and hypertension: uncovering the mechanisms of vascular control. Prog Cardiovasc Dis 2016;59(3):226-234. [doi: 10.1016/j.pcad.2016.09.006] [Medline: 27697533]
33. Tudor-Locke C, Bassett D. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med 2004;34(1):1-8. [doi: 10.2165/00007256-200434010-00001] [Medline: 14715035]
34. Pal S, Cheng C, Egger G, Binns C, Donovan R. Using pedometers to increase physical activity in overweight and obese women: a pilot study. BMC Public Health 2009 Aug 25;9:309 [FREE Full text] [doi: 10.1186/1471-2458-9-309] [Medline: 19703317]
35. Yang Y, Wang C, Wang J, Lin C, Yang Y, Wang J, et al. The Effects of an Activity Promotion System on active living in overweight subjects with metabolic abnormalities. Obes Res Clin Pract 2017;11(6):718-727. [doi: 10.1016/j.orcp.2017.06.002] [Medline: 28729003]
36. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. J Exerc Rehabil 2017 Feb;13(1):3-11 [FREE Full text] [doi: 10.12965/jer.1732928.464] [Medline: 28349027]
37. Sarasohn-Kahn J. California Healthcare Foundation. 2013 Feb 27. Making sense of sensors: how new technologies can change patient care URL: https://www.chcf.org/publication/making-sense-of-sensors-how-new-technologies-can-change-patient-care/ [accessed 2019-06-01] [WebCite Cache ID 78nm7j5BU]
38. 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;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
39. Orrow G, Kinmonth A, Sanderson S, Sutton S. Effectiveness of physical activity promotion based in primary care: systematic review and meta-analysis of randomised controlled trials. BMJ 2012 Mar 26;344:e1389 [FREE Full text] [doi: 10.1136/bmj.e1389] [Medline: 22451477]
40. Fox S, Duggan M. Pew Research Center. 2013 Jan 28. Tracking for health URL: https://www.pewinternet.org/2013/01/28/tracking-for-health/ [accessed 2019-06-01] [WebCite Cache ID 78nmIT6a3]
41. Paré G, Leaver C, Bourget C. Diffusion of the digital health self-tracking movement in Canada: results of a national survey. J Med Internet Res 2018 Dec 02;20(5):e177 [FREE Full text] [doi: 10.2196/jmir.9388] [Medline: 29720359]
42. Strain T, Wijndaele K, Brage S. Physical activity surveillance through smartphone apps and wearable trackers: examining the UK potential for nationally representative sampling. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11898 [FREE Full text] [doi: 10.2196/11898] [Medline: 30694198]

Abbreviations

ANOVA: analysis of variance
BMI: body mass index
BP: blood pressure
FPG: fasting plasma glucose
HDL-C: high-density lipoprotein cholesterol
IQR: interquartile range
TG: triglyceride

Edited by G Eysenbach; submitted 12.01.19; peer-reviewed by K Ng, MS Aslam, M Orme; comments to author 03.04.19; revised version received 06.05.19; accepted 25.05.19; published 18.06.19.

Please cite as:
Huh U, Tak YJ, Song S, Chung SW, Sung SM, Lee CW, Bae M, Ahn HY
Feedback on Physical Activity Through a Wearable Device Connected to a Mobile Phone App in Patients With Metabolic Syndrome: Pilot Study
JMIR Mhealth Uhealth 2019;7(6):e13381
URL: http://mhealth.jmir.org/2019/6/e13381/
doi:10.2196/13381
PMID:31215513

Original Paper

Using the Extended Parallel Process Model to Examine the Nature and Impact of Breast Cancer Prevention Information on Mobile-Based Social Media: Content Analysis

Liang Chen[1], PhD; Xiaodong Yang[2], PhD; Lunrui Fu[1], BA; Xiaoming Liu[1], BE; Congyi Yuan[1], BS

[1]Key Laboratory for Big Data Analysis and Simulation of Public Opinion, School of Communication and Design, Sun Yat-sen University, Guangzhou, China

[2]School of Journalism and Communication, Shandong University, Jinan, China

Corresponding Author:
Xiaodong Yang, PhD
School of Journalism and Communication
Shandong University
No.57 Shanda South Road
Jinan, 250100
China
Phone: 86 531 88361159
Email: XYANG012@e.ntu.edu.sg

Abstract

Background: With the rise of mobile technology, an increasing number of people use mobile-based social media to access health information. Many scholars have explored the nature of health information on social media; however, the impact of such information on people was understudied.

Objective: This study aimed to examine the nature and impact of health information on mobile-based social media. Specifically, we investigated how the levels of threat and efficacy of breast cancer prevention information affect individuals' engagement with the information, such as readings and likes.

Methods: Breast cancer prevention articles posted on a Chinese mobile-based social media platform (ie, WeChat Subscription Account [WeChat SA]) from January 1 to December 31, 2017, were extracted using the Python Web Crawler. We used content analysis and analysis of covariance to analyze our data.

Results: The results revealed that the vast majority of titles and main bodies of the articles involved one of the extended parallel process model components: threat or efficacy.

Conclusions: Breast cancer prevention information on WeChat SA was well designed. Both threat and efficacy significantly affected the number of readings, whereas only efficacy had a significant effect on the number of likes. Moreover, breast cancer prevention information that contained both high levels of threat and efficacy gained the largest number of readings and likes.

(JMIR Mhealth Uhealth 2019;7(6):e13987) doi:10.2196/13987

KEYWORDS

breast cancer; prevention information; mobile social media; EPPM

Introduction

Background

With the rise of mobile technology, an increasing number of people use mobile-based social media to access health information, for example, seeking social support from others and inquiring of medical professionals. In China, WeChat has become the most widely and frequently used mobile-based social media platform and is a crucial part of Chinese people's daily lives [1]. In 2012, WeChat provided a platform called *Subscription Account* (SA). Similar to Facebook pages, people automatically receive information that they are interested in by following a certain SA. With the high incidence and mortality rate of breast cancer in China [2], many breast cancer–related SAs were created on WeChat to provide a variety of information regarding breast cancer prevention.

A number of studies have explored the nature of health information on social media. For example, Jiang and Beaudoin

examined the content of antismoking information on social media based on 3 concepts: subjective norms, perceived risk, and self-efficacy [3]. They found that most content is related to perceived risk, followed by the other 2 factors. Zhang and He examined 1352 messages posted in a Facebook diabetes group and found that health information mainly comprises personal experience, opinions, and advice [4]. Shi and Chen examined the content of all 7215 messages posted in an HIV/AIDS Weibo group and found that the health information in People living with HIV/AIDS Weibo group mainly included emotional, informational, and instrumental support [5]. Although many scholars have explored the nature of health information on social media, the impact of such information on people was understudied.

Therefore, this study intended to examine both the nature and impact of breast cancer–related information on WeChat SA based on the extended parallel process model (EPPM) [6,7]. Some scholars have suggested that media content may not directly affect people's attitudes or behaviors but may instead stimulate them to pay more attention to, and engage more in, related information first [8]. Thus, instead of exploring the direct effect of social media on individuals' breast cancer preventive behavior, this study sought to investigate how the level of threat and efficacy of breast cancer prevention information affects individuals' engagement with the information.

The Extended Parallel Process Model

The fundamental aim of health communication is to deliver knowledge to people to assist them in response to potential diseases or health problems. In modern society, the media is acknowledged as the most prevalent provider of health information, and it is effective in raising public awareness of diseases and promoting precautionary action [9]. Considering the potential impact of the media on public reactions to health issues, this study endeavors to investigate how health messages are constructed in the media and how health information influences people's engagement with information. Thus, this study adopts the EPPM as a theoretical basis.

The EPPM is a predominant message design theory for investigating individuals' reactions to fear appeals [6]. According to the EPPM, to persuade specific groups of people, threat messages should be first provided to attract individuals' attention and stimulate their perception of threat. More specifically, perceived threat includes 2 elements: perceived severity and susceptibility [10]. Perceived severity refers to the perception of seriousness or magnitude of a threat, whereas perceived susceptibility refers to the likelihood of suffering from the threat based on one's belief. After individuals evaluate the extent to which they are likely to experience the threat and the seriousness of the threat, they should receive feasible recommendations that could increase their perceived efficacy. Perceived efficacy also comprises 2 dimensions: perceived self-efficacy and response efficacy. To be specific, perceived self-efficacy refers to an individual's beliefs regarding his or her ability to perform recommended behaviors, whereas perceived response efficacy refers to personal beliefs regarding the effectiveness of the recommended behaviors to avoid hazards. Ideally, when the audience has a higher level of perceived efficacy than that of threat, they will enter the danger control process in which they perform protective behaviors to avert hazards.

A number of empirical studies have documented that fear appeal messages are effective in motivating individuals to perform certain health behaviors. For example, Chen and Yang (2018) found that messages containing a high level of threat and efficacy increased women's intentions to adopt recommended practices, such as breast self-examination [11]. Kotowski et al found that well-designed fear appeal messages can motivate people to use hearing protection aids [12]. Some studies have indicated that people who are exposed to these messages only once may not change their behaviors directly. However, the effectiveness of fear appeal messages might be realized by multiple interactions with the information instead of one exposure to the information [13]. Several studies have pointed out that fear appeal messages may not directly lead people to protective behavior but may motivate them to seek and share relevant information and engage in discussion about a particular issue first [8,14,15]. Thus, we applied the EPPM to examine how breast cancer–related information can affect people's engagement with information.

Engagement With Social Media Content

Social media content usually provides more information than does traditional media. The message itself, the social media content, usually includes some social cues that refer to system- and user-generated cues, such as the number of readings and the number of likes on WeChat SA [16]. Some scholars have indicated that social cues, including the number of readings and likes, have become common indicators of social media campaigns' reach and awareness [17,18]. A higher number of readings and likes indicate a high level of media effects. Thus, social cues have been used as proxies to represent the effectiveness of mobile-based social media campaigns [16]. On the basis of the above, this study intended to examine how the threat and efficacy components in breast cancer prevention messages affect the number of readings and likes. We proposed the following research questions (RQs):

- *RQ1*: How do the titles of articles on breast cancer–related WeChat SAs present (1) threat and (2) efficacy?
- *RQ2*: How do the main bodies of articles on breast cancer–related WeChat SAs present (1) threat and (2) efficacy?
- *RQ3*: How does the threat of article titles affect the number of readings?
- *RQ4*: How does the efficacy of article titles affect the number of readings?
- *RQ5*: How does the threat of the main bodies of the article affect the number of likes?
- *RQ6*: How does the efficacy of the main bodies of the article affect the number of likes?

Methods

Data Extraction

All articles regarding breast cancer prevention posted in breast cancer–related WeChat SAs from January 1 to December 31,

2017, were extracted using the Python Web Crawler in September 2018. The unit of analysis for this study was a single article. Commercial advertisements and articles not related to breast cancer prevention were excluded from the sample. Hence, the final sample size was 537 articles. The content, length of the title and main body, number of readings and likes, and video were retrieved.

Content Analysis

Manual content analysis was used in the current research. Specifically, 3 trained coders, who are native Chinese speakers, were recruited to conduct the coding procedure. First, the 537 article titles were coded as 0 (absent) and 1 (present) in the 2 variables: threat and efficacy. Krippendorff alpha tests revealed an acceptable level of intercoder reliability for the 2 variables: .87 for threat and .91 for efficacy. The main bodies of the article were coded as 0 (low), 1 (medium), and 2 (high) for the 2 variables: threat and efficacy. The values of Krippendorff alpha were .81 for threat and .83 for efficacy, which is also acceptable.

Results

The Nature of Breast Cancer Prevention Information on WeChat Subscription Accounts

First, the results showed that 82.1% (441/537) article titles involved one of the EPPM components: threat or efficacy. To be specific, 63.9% (343/537) article titles contained efficacy and 26.4 % (142/537) article titles contained threat. In terms of the main bodies of the articles, 75.2% (404/537) articles provided medium and high levels of threat, whereas only 49.7% (267/537) articles contained medium and high levels of threat.

The Impact of Breast Cancer Prevention Information on WeChat Subscription Accounts

In addition, a series of analysis of covariance (ANCOVAs) was conducted to address *RQ3* to *RQ6*. First, an ANCOVA was performed to examine how the threat and efficacy of article titles affect the number of readings, with title length serving as a covariate. The results revealed a significant main effect for threat, where titles with the threat component had a greater number of readings (mean 8448.73, SD 8286.11) and those without the threat component had a lower number of readings (mean 5957.64, SD 4852.52), $F_{1,532}$=62.99, P<.001, partial η^2=0.11 (see Tables 1 and 2). The results also showed a significant main effect for efficacy, which indicated that titles with the efficacy component tended to have a greater number of readings (mean 7477.39, SD 6346.04) than did those without the efficacy component (mean 5094.04, SD 5155.45), $F_{1,532}$=85.87, P<.001, partial η^2=0.14 (see Tables 1 and 2). Furthermore, the results revealed a significant interaction effect between threat and efficacy, $F_{1,532}$=51.51, P<.001, partial η^2=0.09. Figure 1 presents a plot of the obtained mean scores. An examination of this interaction effect indicates that the number of readings was highest when both threat and efficacy were presented in the article titles.

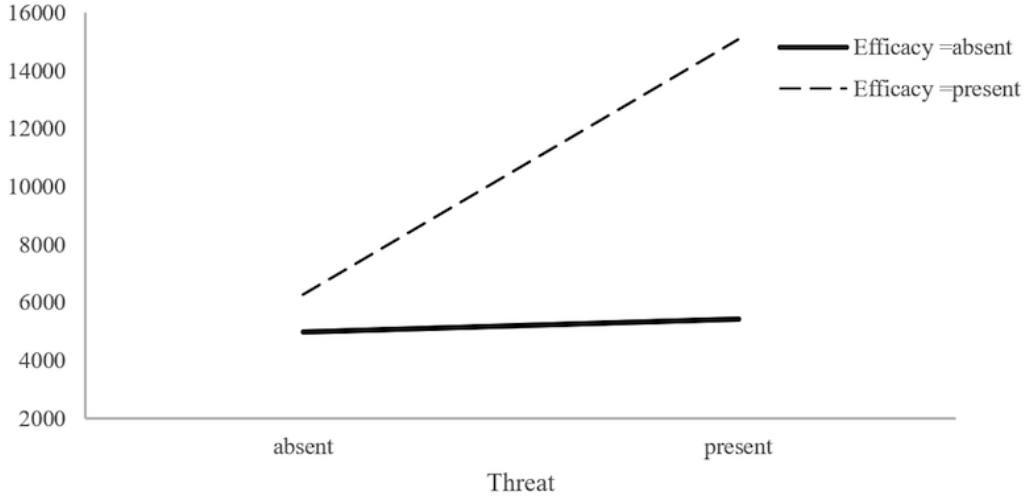
Table 1. Descriptive statistics for the number of readings (N=537).

Presence of threat and efficacy	Mean (SD)	n (%)
Threat absent		
Efficacy absent	4881.70 (3969.35)	96 (17.9)
Efficacy present	6303.09 (5061.33)	299 (55.7)
Total	5957.64 (4852.52)	395 (73.6)
Threat present		
Efficacy absent	5302.03 (6112.63)	98 (18.3)
Efficacy present	15457.27 (8247.67)	44 (8.2)
Total	8448.72 (8286.11)	142 (26.4)
Total		
Efficacy absent	5094.03 (5155.45)	194 (36.1)
Efficacy present	7477.39 (6346.04)	343 (63.9)
Total	6616.36 (6048.11)	537 (100)

Table 2. Threat×efficacy factorial analysis of variance for the amount of reading.

Variable	*F* test (df)	*P* value	Partial η^2
Title length	11.03 (1)	.001	0.02
Threat	62.99 (1)	<.001	0.11
Efficacy	85.87 (1)	<.001	0.14
Threat×Efficacy	51.51 (1)	<.001	0.09

Figure 1. Interaction effect between threat and efficacy upon the number of readings.



Additionally, another ANCOVA was performed to explore how the level of threat and efficacy presented in the main bodies of the articles affected the number of likes, with videos and article length serving as covariates. The results revealed that the main effect for threat was not significant ($F_{2,526}$=.65; P=.52; partial η^2=0.002). However, the main effect for efficacy was significant ($F_{2,526}$=5.45; P=.005; partial η^2=0.02), where articles with a high level of efficacy (mean 46.75, SD 63.18) had a greater number of likes than did those with medium (mean 34.19, SD 47.36) and low levels of efficacy (mean 26.68, SD 32.82; see Tables 3 and 4).

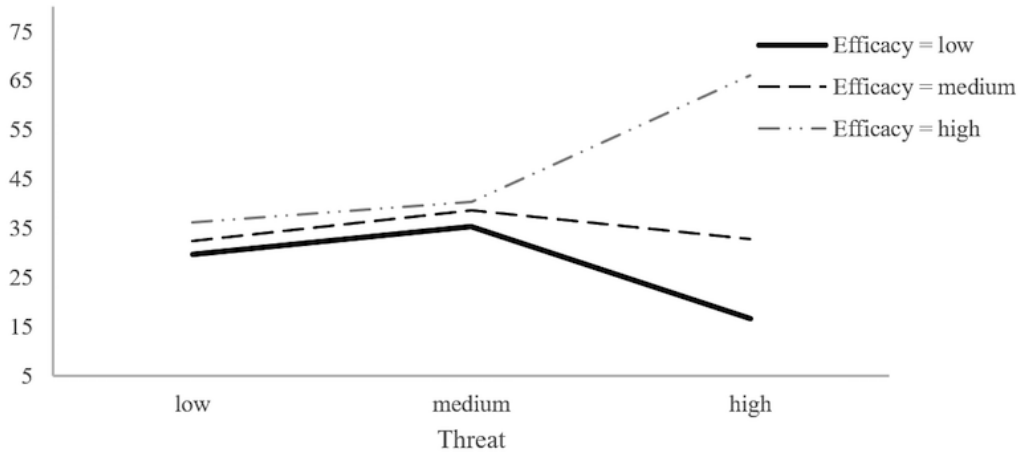
Table 3. Descriptive statistics for the number of likes (N=537).

Level of threat and efficacy	Mean (SD)	n (%)
Low threat		
Low efficacy	25.59 (27.71)	87 (16.2)
Medium efficacy	29.68 (31.61)	60 (11.2)
High efficacy	37.69 (38.18)	123 (22.9)
Total	32.01 (33.97)	270 (50.3)
Medium threat		
Low efficacy	35.63 (43.89)	30 (5.6)
Medium efficacy	39.90 (40.21)	48 (8.9)
High efficacy	41.84 (43.68)	81 (15.1)
Total	40.08 (42.49)	159 (29.6)
High threat		
Low efficacy	15.88 (32.51)	16 (3.0)
Medium efficacy	34.04 (83.64)	23 (4.3)
High efficacy	68.68 (102.19)	69 (12.9)
Total	53.48 (93.08)	108 (20.1)
Total		
Low	26.68 (32.82)	133 (24.8)
Medium	34.19 (47.36)	131 (24.4)
High	46.75 (63.18)	273 (50.8)
Total	38.72 (53.93)	537 (100)

Table 4. Threat×efficacy factorial analysis of variance for intention to the number of likes.

Variable	*F* test (df)	*P* value	Partial η^2
Article length	10.10 (1)	.002	0.02
Video	0.97 (1)	.32	0.00
Threat	0.65 (2)	.52	0.00
Efficacy	5.45 (2)	.005	0.02
Threat×Efficacy	2.53 (4)	.04	0.02

Figure 2. Interaction effect between threat and efficacy upon the number of likes.



The results revealed that the threat×efficacy interaction effect ($F_{4,526}$=2.53; *P*=.04; partial η^2=0.02) was significant. Figure 2 presents a plot of the obtained mean scores, which indicated that for the main bodies of the articles with low or medium levels of threat, regardless of the level of efficacy, the number of likes remained stable and relatively low. Similarly, for the main bodies of the articles with a high level of threat, coupled with low or medium levels of efficacy messages, the number of likes was small. However, for the main bodies of the articles with high levels of threat and efficacy, the number of likes was highest.

Discussion

Principal Findings

This study examined the nature and impact of mobile-based social media messages. Single exposure to information may not lead to immediate behavior change. Media content may stimulate people's information attention and engagement (ie, liking, sharing, and commenting) [8,14,15,19], which promotes actual behaviors [20]. As such, this study explored how mobile-based social media messages affect individuals' engagement with the information based on the EPPM.

First, the WeChat SA article titles with a threat or efficacy component had a greater number of readings than did those without threat and efficacy components. Moreover, the number of readings was highest when both threat and efficacy were present in the article titles. This is consistent with the EPPM model. Titles containing the threat of breast cancer will attract people's attention and arouse their feelings of fear [6,21], which will motivate them to seek preventive solutions. In this regard, once the titles contain the efficacy component, people will perceive that the article will provide effective measures. Thus, they are more likely to read these articles.

In addition, the results revealed that the main bodies of the articles containing a high level of efficacy received a greater number of likes than did those with medium and low levels of efficacy, whereas the level of threat did not significantly affect the number of likes. It is plausible that most people who followed the WeChat SA related to breast cancer were aware of the risk of breast cancer before reading the articles. Moreover, the article titles may have aroused their feelings of fear. In this regard, they prefer to find effective and feasible measures to reduce their fear and anxiety instead of increasing perceived risk about diseases.

More interestingly, the main bodies of the articles that contain high levels of threat and efficacy gained the largest number of likes. However, those with low levels of efficacy but high levels of threat did not lead to public agreement. This is in line with the EPPM. People who experience an overly high perceived threat believe that they are incapable of adopting protective behaviors or their beliefs fall under the ineffectiveness of such behaviors. Thus, they would subsequently enter the fear control process and engage in defensive avoidance or reaction behaviors [7]. In contrast, when people have a higher perception of efficacy than threat, they enter the danger control process and change their perceptions, attitudes, intentions, and even actual behaviors to cope with the threat according to the recommended responses [10].

Implications and Limitations

This study is an initial attempt to examine the nature and impact of mobile-based social media messages on people's engagement with information. First, breast cancer prevention information on the most popular Chinese mobile-based social media platform (ie, WeChat SA) was analyzed and found to be well designed. However, instead of examining the direct effect of social media on individuals' actual behavior, this study investigated how threat and efficacy of breast cancer prevention articles on WeChat SAs affect the number of readings and likes. This provides a new approach or perspective to examine the effects of social media–based health campaigns,

Despite its significant contributions, this study is not without limitations. First, only breast cancer prevention–related WeChat SA articles were examined in this study, which may limit generalization of the findings to other diseases and health conditions. Future studies should examine mobile-based social media content regarding other diseases. Second, only the number of readings and likes were included as indicators of individuals' engagement with the information in this study. Other social cues, such as the number of comments and retweets, may also represent the level of engagement with the information. Thus, the content of comments and the number of retweets could be considered. Finally, this study examined the effects of mobile-based social media only on individuals' engagement with the information rather than their actual behaviors. Although previous studies have indicated the link between engagement with information and actual behavior [8], future research should examine whether and how engagement with information affects actual preventive behavior.

Conflicts of Interest

None declared.

References

1. Zhang XT, Wen D, Liang J, Lei JB. How the public uses social media wechat to obtain health information in china: a survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
2. Wang H, Yu C, Gao X, Welte T, Muscarella AM, Tian L, et al. The osteogenic niche promotes early-stage bone colonization of disseminated breast cancer cells. Cancer Cell 2015 Feb 09;27(2):193-210 [FREE Full text] [doi: 10.1016/j.ccell.2014.11.017] [Medline: 25600338]
3. Jiang S, Beaudoin CE. Smoking prevention in China: a content analysis of an anti-smoking social media campaign. J Health Commun 2016 Dec;21(7):755-764. [doi: 10.1080/10810730.2016.1157653] [Medline: 27232655]
4. Zhang Y, He D, Sang Y. Facebook as a platform for health information and communication: a case study of a diabetes group. J Med Syst 2013 Jun;37(3):9942. [doi: 10.1007/s10916-013-9942-7] [Medline: 23588823]
5. Shi JY, Chen L. Social support on Weibo for people living with HIV/AIDS in China: a quantitative content analysis. Chinese Journal of Communication 2014 Jul 24;7(3):285-298. [doi: 10.1080/17544750.2014.926954]
6. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Commun Monogr 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]
7. Witte K. Fear control and danger control: a test of the extended parallel process model (EPPM). Commun Monogr 2009 Jun 02;61(2):113-134. [doi: 10.1080/03637759409376328]
8. Wakefield M, Loken B, Hornik R. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 09;376(9748):1261-1271 [FREE Full text] [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
9. Cox R. Environmental Communication and the Public Sphere. London: Sage; 2012.
10. Witte K, Cameron KA, McKeon JK, Berkowitz JM. Predicting risk behaviors: development and validation of a diagnostic scale. J Health Commun 1996;1(4):317-341. [doi: 10.1080/108107396127988] [Medline: 10947367]
11. Chen L, Yang X. Using EPPM to evaluate the effectiveness of fear appeal messages across different media outlets to increase the intention of breast self-examination among Chinese women. Health Commun 2018 Aug 06;1:1-8. [doi: 10.1080/10410236.2018.1493416] [Medline: 30080982]
12. Smith SW, Rosenman KD, Kotowski MR, Glazer E, McFeters C, Keesecker NM, et al. Using the EPPM to create and evaluate the effectiveness of brochures to increase the use of hearing protection in farmers and landscape workers. J Appl Commun Res 2008 May;36(2):200-218. [doi: 10.1080/00909880801922862]
13. Shi J, Smith SW. The effects of fear appeal message repetition on perceived threat, perceived efficacy, and behavioral intention in the extended parallel process model. Health Commun 2016;31(3):275-286. [doi: 10.1080/10410236.2014.948145] [Medline: 26305152]
14. Chen L, Guo Y, Shi J. Social support seeking on social media among Chinese gay men living with HIV/AIDS: the role of perceived threat. Telemed J E Health 2018 Aug 14:2018. [doi: 10.1089/tmj.2018.0136] [Medline: 30106664]
15. Rus H, Cameron L. Health communication in social media: message features predicting user engagement on diabetes-related Facebook pages. Ann Behav Med 2016 Dec;50(5):678-689. [doi: 10.1007/s12160-016-9793-9] [Medline: 27059761]
16. Shi J, Poorisat T, Salmon C. The use of social networking sites (SNSs) in health communication campaigns: review and recommendations. Health Commun 2018 Dec;33(1):49-56. [doi: 10.1080/10410236.2016.1242035] [Medline: 27858464]

17. Dixon-Gray L, Mobley A, McFarlane J, Rosenberg K. (Love and Health): a preconception health campaign for second-generation Latinas in Oregon. Am J Health Promot 2013 Jan 02;27(3_suppl):S74-S76. [doi: 10.4278/ajhp.120113-arb-29]
18. Platt T, Platt J, Thiel D, Kardia S. Facebook advertising across an engagement spectrum: a case example for public health communication. JMIR Public Health Surveill 2016 May 30;2(1):e27 [FREE Full text] [doi: 10.2196/publichealth.5623] [Medline: 27244774]
19. Cho M, Schweickart T, Haase A. Public engagement with nonprofit organizations on Facebook. Public Relat Rev 2014 Sep;40(3):565-567. [doi: 10.1016/j.pubrev.2014.01.008]
20. Alhabash S, McAlister A, Lou C, Hagerstrom A. From clicks to behaviors: the mediating effect of intentions to like, share, and comment on the relationship between message evaluations and offline behavioral intentions. J Interact Advert 2015 Jul 15;15(2):82-96. [doi: 10.1080/15252019.2015.1071677]
21. Vogt J, Koster EH, De Houwer J. Safety first: instrumentality for reaching safety determines attention allocation under threat. Emotion 2017 Dec;17(3):528-537. [doi: 10.1037/emo0000251] [Medline: 27869466]

Abbreviations

ANCOVA: analysis of covariance
EPPM: Extended Parallel Process Model
RQ: research question
SA: Subscription Account

Edited by G Eysenbach; submitted 11.03.19; peer-reviewed by G Kolostoumpis, Z Zhou; comments to author 11.05.19; revised version received 22.05.19; accepted 23.05.19; published 24.06.19.

Please cite as:
Chen L, Yang X, Fu L, Liu X, Yuan C
Using the Extended Parallel Process Model to Examine the Nature and Impact of Breast Cancer Prevention Information on Mobile-Based Social Media: Content Analysis
JMIR Mhealth Uhealth 2019;7(6):e13987
URL: http://mhealth.jmir.org/2019/6/e13987/
doi:10.2196/13987
PMID:31237239

Original Paper

A Mobile System to Improve Quality of Life Via Energy Balance in Breast Cancer Survivors (BENECA mHealth): Prospective Test-Retest Quasiexperimental Feasibility Study

Mario Lozano-Lozano[1,2,3,4], MSc, OT; Irene Cantarero-Villanueva[1,2,3,4], PhD; Lydia Martin-Martin[1,2,3,4], PhD; Noelia Galiano-Castillo[1,2,3,4], PhD; Maria-José Sanchez[3,5,6], PhD; Carolina Fernández-Lao[1,2,3,4], PhD; Paula Postigo-Martin[1,2,4], MSc; Manuel Arroyo-Morales[1,2,3,4], MD, PhD

[1]Department of Physical Therapy, Faculty of Health Sciences, University of Granada, Granada, Spain

[2]Sport and Health University Research Institute, Granada, Spain

[3]Biohealth Research Institute in Granada, Granada, Spain

[4]Cuidate-Support Unit for Oncology Patients, Granada, Spain

[5]Andalusian School of Public Health, Granada, Spain

[6]Consortium for Biomedical Research in Epidemiology and Public Health, Madrid, Spain

Corresponding Author:
Irene Cantarero-Villanueva, PhD
Department of Physical Therapy
Faculty of Health Sciences
University of Granada
Avda de la Ilustración, 60
Granada,
Spain
Phone: 34 958248764
Email: irenecantarero@ugr.es

Abstract

Background: Energy balance is defined as the difference between energy expenditure and energy intake. The current state of knowledge supports the need to better integrate mechanistic approaches through effective studies of energy balance in the cancer population because of an observed significant lack of adherence to healthy lifestyle recommendations. To stimulate changes in breast cancer survivors' lifestyles based on energy balance, our group developed the BENECA (Energy Balance on Cancer) mHealth app. BENECA has been previously validated as a reliable energy balance monitoring system.

Objective: Based on our previous results, the goal of this study was to investigate the feasibility of BENECA mHealth in an ecological clinical setting with breast cancer survivors, by studying (1) its feasibility and (2) pretest-posttest differences with regard to breast cancer survivor lifestyles, quality of life (QoL), and physical activity (PA) motivation.

Methods: Eighty breast cancer survivors diagnosed with stage I to IIIA and with a body mass index over 25 kg/m2 were enrolled in this prospective test-retest quasi-experimental study. Patients used BENECA mHealth for 8 weeks and were assessed at baseline and the postintervention period. Feasibility main outcomes included percentage of adoption, usage, and attrition; user app quality perception measured with the Mobile App Rating Scale (MARS); satisfaction with the Net Promoter Score (NPS); and barriers and facilitators of its use. Clinical main outcomes included measuring QoL with the European Organization for Research and Treatment of Cancer QoL Questionnaire Core 30 (EORT QLQ-C30), PA assessment with accelerometry, PA motivation measure with a Spanish self-efficacy scale for physical activity (EAF), and body composition with dual-energy x-ray absorptiometry. Statistical tests (using paired-sample t tests) and Kaplan-Meier survival curves were analyzed.

Results: BENECA was considered feasible by the breast cancer survivors in terms of use (76%, 58/76), adoption (69%, 80/116), and satisfaction (positive NPS). The app quality score did not make it one of the best-rated apps (mean 3.71, SD 0.47 points out of 5). BENECA mHealth improved the QoL of participants (global health mean difference [MD] 12.83, 95% CI 8.95-16.71, P<.001), and EAF score (global MD 36.99, 95% CI 25.52-48.46, P<.001), daily moderate-to-vigorous PA (MD 7.38, 95% CI 0.39-14.37, P=.04), and reduced body weight (MD −1.42, 95% CI −1.97 to −0.87, P<.001).

Conclusions: BENECA mHealth can be considered feasible in a real clinical context to promote behavioral changes in the lifestyles of breast cancer survivors, but it needs to be enhanced to improve user satisfaction with use and functionality. This study highlights the importance of the use of mobile apps based on energy balance and how the QoL of breast cancer survivors can be improved via monitoring.

(JMIR Mhealth Uhealth 2019;7(6):e14136) doi:10.2196/14136

KEYWORDS

mHealth; energy balance; monitoring; breast cancer; survivors; quality of life

Introduction

There is a direct relationship between energy imbalance and an increased risk of not only multiple cancers but also cancer mortality, and a worsening of the effects of the disease [1-3]. Energy balance is defined as the difference between energy expenditure and energy intake [4]. Energy intake that exceeds energy expenditure is the main driver of weight gain; thus, balancing both helps weight maintenance [5].

A panel of experts from the International Agency for Research on Cancer and the World Cancer Research Fund agreed that 16 types of cancer are probably associated with one of the more relevant consequences of energy imbalance, excess fat accumulation in the body, making obesity the second leading cause of cancer worldwide [1,6]. Moreover, since the first decade of the 2000s, the scientific evidence on the benefits of physical activity (PA) in the quality of life (QoL) of cancer survivors (known as "oncological exercise") has grown exponentially, generating dozens of systematic reviews, several international guidelines, and the recommendation to include programs of exercise in cancer survivors care [7]. Dietary and exercise interventions can alter the energy imbalance associated with cancer and potentially decrease the QoL of cancer survivorship [5]. However, the literature shows that despite strong evidence of this association, an insurmountable barrier prevails between "what needs to be done" and "what patients really do," observing a significant lack of adherence to the preceding interventions [1].

In today's progressively technical world, the use of mobile apps in smart devices has become the norm. In the same way, patients increasingly use therapeutic mobile apps related to some form of cancer treatment [8]. More than 2500 mobile apps are defined as apps related to cancer, but this relationship is peripheral or based on unproven claims, such as apps for yoga and naturopathy that claim to help prevent or even cure cancer [9]. In 2017, 15% of studies conducted worldwide were aimed at digital health, with 75% of these studies being conducted in the United States [9]. Recently, 539 apps were considered in a systematic review, which concluded that the effectiveness of most of them had not been validated scientifically [8,10]. Duman-Lubberding and colleagues [11] have developed Oncokompas, an eHealth app to facilitate access to supportive cancer care and monitor cancer patients' QoL [12], specifically in the case of breast cancer [13]. Another study by Gietema and colleagues [14] assessed the feasibility of the Runkeeper app to improve the level of PA of cancer patients. They concluded that there is a need to increase research in the area. Different studies and meta-analyses of cancer patients show the benefits of mHealth, which include reducing fatigue or pain [15,16], distance PA programs with inconclusive results for and against [17-19], the use of social networks by patients of some types of cancer to improve QoL [20], and monitoring of symptoms [21,22]. However, none of these studies refers to monitoring and providing high-quality research feedback to restore the energy balance in cancer patients. The only references found in this field were in healthy populations [23,24], children and adolescents [25], pregnant women [26], hospitalized patients [27], and cardiac surgery [28] and diabetes [29] patients. Furthermore, monitoring using globally extended systems, such as Fitbit wristbands, is being questioned [30]. A recent systematic review of 67 studies concluded that, except for the measurement of steps in adults, there are a limited number of situations in which these devices provide accurate measurement for use in research [30].

In an attempt to stimulate changes in breast cancer survivors' lifestyles based on energy balance, we developed the BENECA mHealth app: Energy Balance on Cancer [31,32]. BENECA mHealth aims to monitor the energy expenditure and energy intake of breast cancer survivors and provide instantaneous, simple, and clear feedback on the users' energy balance, along with recommendations on how to improve it. This strategy was based on a recent systematic review of behavior change techniques for increasing PA in cancer survivors [33], as well as another study carried out by Hillier et al [34], who developed a Web-based program to assess energy balance in healthy adults. The first essential step, to develop and validate our tool, was to ensure the reliability of the BENECA mHealth monitoring system. The results of our previous study showed that it is a direct, rapid, and consistent evaluation system [32]. Based on these results, the goal of this study was to investigate the feasibility of BENECA mHealth in an ecological setting with a population of cancer survivors after they are discharged from their oncology treatment.

This involved studying the adoption of the app, its usage, user app quality perception, and the barriers and facilitators of its use. In addition, we gained insight into pretest-posttest differences with regard to breast cancer survivors' lifestyles, QoL, and PA motivation. This investigation was based on the hypothesis that using the BENECA mHealth app for 8 weeks would help increase the motivation of breast cancer survivors to adhere to healthier lifestyles, thereby improving their QoL.

Methods

Study Design and Patient Recruitment

A prospective test-retest quasi-experimental study was carried out with 80 breast cancer survivors. The breast cancer survivors were selected based on the following eligibility criteria: (1) breast cancer stage I, II, or IIIA, (2) 30 to 75 years old, (3) body mass index (BMI) over 25 kg/m^2, (4) user-level skills for app management, and (5) completed the adjuvant treatment at least 6 months before being included in the study. Eligible participants were excluded if they had mental or physical health conditions that prevented them from walking and/or participating in the assessment or if they did not sign the informed consent form. In addition, participants had to have access to a mobile device or tablet with an internet connection and an Android operating system. The research team loaned out two devices in cases where this was not possible or the operating system was incompatible with the app. All participants were recruited through the oncology unit from the University Hospital Complex of Granada, Spain, after being informed about the study and being referred by their respective oncologist. All eligible participants were contacted via telephone, screened using the inclusion and exclusion criteria, and if they were interested in participating, cited for the baseline assessment.

This study was approved by the ethics committee of the Andalusian Health Service (FIS, PI14-01627; Granada, Spain) and it was performed in accordance with the Helsinki Declaration for biomedical research (14/2017) [35]. Participants completed informed consent forms before the assessment.

BENECA mHealth

The CUIDATE research group developed the Energy Balance on Cancer (BENECA) mHealth app to monitor and provide feedback to breast cancer survivors on healthy eating and PA. A description of the BENECA mHealth System [31,36] and a reliability study for the same [32] were previously published. After the baseline assessment was performed, a member of the research group downloaded the app on a patient's mobile phone and taught them how to use it. The patient then had to prove that she understood the instructions by using the app in the presence of the researcher. Patients had to use BENECA mHealth for 8 weeks during the study. Physical activity (duration and intensity) and diet (food and drink intake) data were recorded via the app (self-recorded). Intake was recorded using a dietary record questionnaire; BENECA is structured with six consumption times. On each day, for each period, users report all food and beverages consumed. For PA, BENECA incorporated the Minnesota Leisure-time Physical Activity Questionnaire. Patients had to record intensity and duration of activities each day; BENECA only recorded those activities with a duration of at least 10 minutes. Using this information, the app sent a notification to the user of their daily energy balance, offering recommendations on diet and PA, which were based on the guidelines of the World Cancer Research Fund International, the strategies for PA and diet in patients with cancer from the American College of Sports Medicine [37], and the recommendations of the American Cancer Society [38]. Users receive a straightforward daily notification if there has been an energy imbalance; any difficulties in handling the app were resolved via calls and text messages between the researcher and patient (Multimedia Appendix 1). BENECA had been developed based on the theories of learning, Goal-Setting Theory, and Social Cognitive Theory to include techniques such as reinforcement, facilitation, self-monitoring, goal setting, feedback on performance, and reviewing goals, which have demonstrated to be promising in increasing PA in different populations [33,39]. A video tutorial was made available to the patients to review the use of the app.

Outcome Measures

Patient demographic and clinical data were obtained at the beginning of the study using a study-specific survey. Baseline data were gathered at the start of the study and again after 8 weeks of using BENECA mHealth. The outcomes measured are presented subsequently.

Feasibility of Main Outcomes

BENECA mHealth was considered feasible for use by breast cancer survivors as long as it met the following criteria, established based on previous studies with eHealth apps [11,13,40,41]: adoption and usage rate over 50%, a positive Net Promoter Score (NPS), and a Mobile App Rating Scale (MARS) score of up to 3.73 out of 5.

Adoption, Usage, and Attrition

The adoption rate was the percentage of the number of breast cancer survivors that agreed to participate in the study and completed the initial assessment, demonstrating the intention to use BENECA mHealth, out of the total number invited to participate in the study. The usage rate is the percentage of breast cancer survivors that used BENECA mHealth, which was determined through the logging data of the app. Both the adoption and usage rates were calculated based on the methods used in a previously published study [13]. The attrition rate is the percentage of breast cancer survivors that stopped using BENECA mHealth and did not use it again, as per Eysenbach's definition [42]. To assess the safety of the process, any adverse effects reported by the patients were recorded through a patient's daily diary.

BENECA mHealth Quality

The MARS was used to assess the quality of BENECA mHealth. The MARS is composed of 23 items grouped into different sections: engagement, functionality, aesthetics, and information quality (with which the overall average score of the scale is obtained). There are also two optional sections: subjective quality (with four items) and app-specific quality (with six items). Each item was assessed independently based on a Likert scale from 1 (inadequate) to 5 (excellent), and the mean score was calculated for each section. This scale has been validated and has proven to be simple, objective, and reliable to assess the quality of mHealth apps [43]. Similarly, the NPS was used to measure satisfaction based on responses to the following question: How likely are you to recommend BENECA mHealth to other breast cancer survivors? The responses were recorded using an 11-point Likert scale in which 0 indicates "not likely" and 10 indicates "very likely." The percentage of *detractors* (those whose scores were from 0 to 6) and *promoters* (those

whose scores were from 9 to 10) were calculated, and each group was given a score between −100 and 100. A positive score is considered good; a negative score is considered bad [44]. This methodology has been used as a predictor of growth and an indicator of customer satisfaction in for-profit industries, and it provides insight into the client experience in nonprofit health care settings [45].

Barriers and Facilitators

After the participants used BENECA mHealth for 8 weeks and completed the corresponding assessment, a trained member of the research team interviewed each participant using a standardized set of interview questions based on a previous study [13]. This interview focused on three main elements: overall experience with BENECA mHealth, congruence between expectations and reality with BENECA mHealth, and the perception and added value of BENECA mHealth. For cases in which the app was no longer used, the participants were asked about their reasons for not using the app and the preferences or needs that would prompt them to use it. Each interview was read several times and transcribed by the same researcher, and the barriers and facilitators reported by the breast cancer survivors were synthesized [46].

Main Clinical Outcomes

Quality of Life

The European Organization for Research and Treatment of Cancer QoL Questionnaire Core 30 (EORTC QLQ-C30) version 3.0 was used to assess the QoL of the participants. This questionnaire is intended to measure general aspects of QoL specific to cancer patients. It contains five functional scales (physical, role, cognitive, emotional, and social functioning), a global health status scale, and symptom scales of fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial problems. It is scored using a four-point Likert scale (from 1="not at all" to 4="very much") and the raw scores are transformed into a 0 to 100 scale. The higher the score on the functional scales, the better the QoL, but the higher the score on the symptom scales, the poorer the QoL [47,48].

Self-Efficacy and Motivation in Relation to Physical Activity

A Spanish self-efficacy scale for physical activity (EAF) was used to measure the self-efficacy and motivation of the participants to engage in PA and incorporate it into their daily activities. It consists of three domains: scheduled physical exercise, PA in daily life activities, and walking, which determine a person's perception of their abilities to engage in PA (self-efficacy for PA) [49].

Physical Activity

Data on PA and the sedentary lifestyle of the breast cancer survivors were collected using accelerometry based on a previously published protocol of use and analysis [50]. A preprogrammed triaxial accelerometer (ActiGraph GT3X+, Pensacola, FL, USA) was used by each patient for eight consecutive days. The participants received a questionnaire diary and an instruction sheet on how to use the device. Only the records of more than 4 days and of at least 10 hours per day were included in the analysis.

Body Composition

Dual-energy x-ray absorptiometry (Discovery DXA densitometer from Hologic, QDR 4500 W) was used for assessing BMI, the percentage of fat mass, and bone mineral density, as previously used for breast cancer patients [51] in accordance with protocol of use [52]. The height and weight of the participants were also measured at baseline as well as hip and waist circumferences.

Statistical Analysis

All analyses were performed using SPSS Statistics version 24 (IBM Corp, Armonk, NY, USA). Statistical significance was assumed when $P<.05$. The logging data from BENECA mHealth were obtained on request from the computer engineers responsible for the development of the app.

First, descriptive measures were used to report the data on adoption, use, attrition, and quality, as well as to report on the clinical and anthropometric variables of the participants. A Kaplan-Meier survival curve was used to visually examine the survival curve of the entire cohort to determine the attrition. In the analysis, an "app survivor" was defined as a breast cancer survivor that maintained logging practices using BENECA mHealth until at least 3 days before the last day of the experimental period. Those defined under "app death" were those who missed five consecutive daily loggings (based on a previous study [53]). A Kaplan-Meier estimator with right-censored data was used. This type of data was used because it best fit our study results. As most of the breast cancer survivors "survived" until the end of the experimental period, we do not know how long they would have continued using BENECA mHealth after this period. Then, a Cox proportional hazard model was used to examine if age, marital status, and employment had any effect on the attrition.

Second, to assess the pretest-posttest differences in the main outcomes, an analysis of paired-sample *t* tests was used and, when appropriate, Wilcoxon signed rank tests were conducted. Moreover, the effect size (ES) estimate was determined and interpreted using Cohen's guidelines of 0.1=small effect, 0.3=medium effect, and 0.5=large effect.

Third, to assess differences between "users" and "nonusers" and the patients' perception of BENECA mHealth quality, a Mann-Whitney test was used for categorizing the breast cancer survivors according to the cut-off used in the survival analysis. A simple linear regression was used to examine the influence of age on the perception of BENECA mHealth quality.

Our data contained a few missing values (5%, 4/80 of the total number of cases), but these can be considered random and inconsequential [54]. Hence, no multiple imputation method was necessary (casewise deletion was used).

Results

Demographic Characteristics

The baseline demographic and clinical characteristics of the participants (mean age 51.80, SD 8.64 years) are presented in Table 1. Of the 80 breast cancer survivors, 50 (62%) were married, 31 (38%) had a higher education, and 40 (50%) were diagnosed with stage II breast cancer, followed in frequency by stage IIIA (28/80, 35%). All participants received instructions on how to use BENECA mHealth to monitor energy intake and expenditure. Four participants were unable to be assessed postintervention (dropouts); three were not assessed due to changes in their health status unrelated to the study, and one decided to discontinue.

Table 1. Participant demographics (N=80).

Variables	Participants
Age (years), mean (SD)	51.80 (8.64)
Body mass index, mean (SD)	29.11 (4.77)
Marital status, n (%)	
Single	16 (20)
Married	50 (63)
Divorced	10 (13)
Other	4 (5)
Education, n (%)	
No education	1 (1)
Primary studies	23 (29)
Secondary studies	25 (31)
Higher education	31 (39)
Employment, n (%)	
Homemaker	18 (22)
Employee	32 (40)
Low	10 (13)
Unemployed by the disease	20 (25)
Cancer stage, n (%)	
I	10 (13)
II	40 (51)
IIIA	28 (36)
Surgery, n (%)	
Lumpectomy	24 (30)
Quadrantectomy	13 (16)
Unilateral mastectomy	27 (34)
Bilateral mastectomy	16 (20)
Medical treatment, n (%)	
None	6 (8)
Radiation therapy alone	10 (13)
Chemotherapy alone	6 (8)
Chemotherapy and radiation therapy	48 (60)
Adjuvant chemotherapy	7 (9)
Neoadjuvant chemotherapy	3 (4)

Feasibility Outcomes

Adoption, Usage, and Attrition Rates

The study design is shown in Figure 1. The adoption rate of BENECA mHealth was 69%; 80 of 116 breast cancer survivors who were invited to participate intended to use BENECA mHealth, filled the informed consent form, and were assessed at baseline. The reasons for not participating in the study included lack of interest (too busy; n=14), incompatibility of the user's mobile operating system with BENECA mHealth (n=11), and failed initial contact (eg, wrong phone number or no answer; n=11).

The usage rate was 73% (58/80) including dropouts and 76% (58/76) excluding dropouts. The reasons for stopping using BENECA mHealth included technical issues, such as difficulty in finding specific foods (n=6), app blocks (n=4), difficulty in calculating proportions of diet registration (n=9), or lack of motivation (n=3). We examined attrition using the Kaplan-Meier survival curve and Cox proportional hazards model. Figure 2 illustrates the attrition curve of the study participants with their respective 95% CIs. The curve is flat at the beginning, begins to get steeper after the first month, and flattens again with time. The Cox proportional hazards model was used to assess the differences in the survival rate using covariables that could affect this rate from the clinical point of view based on a priori knowledge. The results obtained using this model with the covariates were significant at P=.02; the coefficients are shown in Table 2.

Figure 1. Flow diagram of the study design. EAF: self-efficacy scale for physical activity; EORT QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30. *N=75 for accelerometry analyses (one broken device on preassessment).

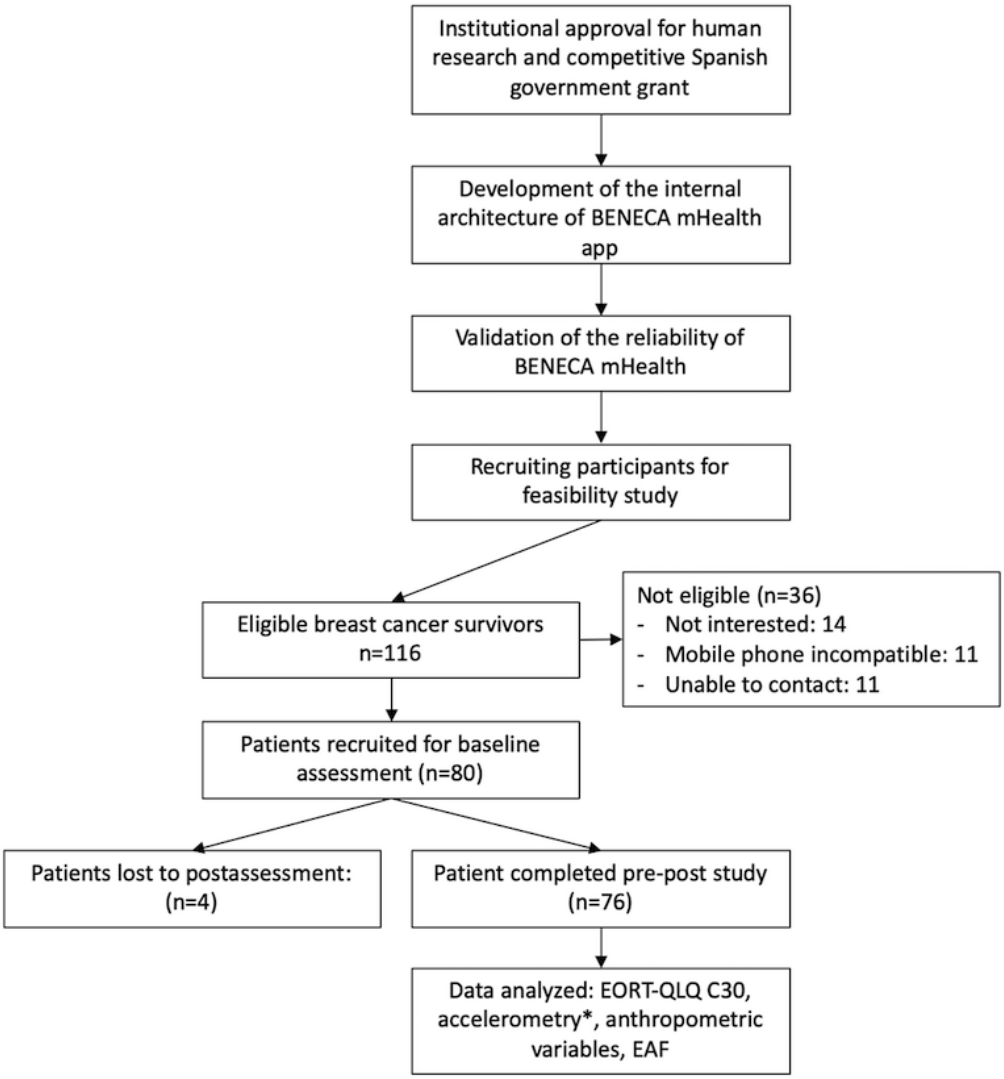


Figure 2. "Survival" of BENECA app participants as shown by a Kaplan-Meier survival curve with 95% CIs (dashed lines).

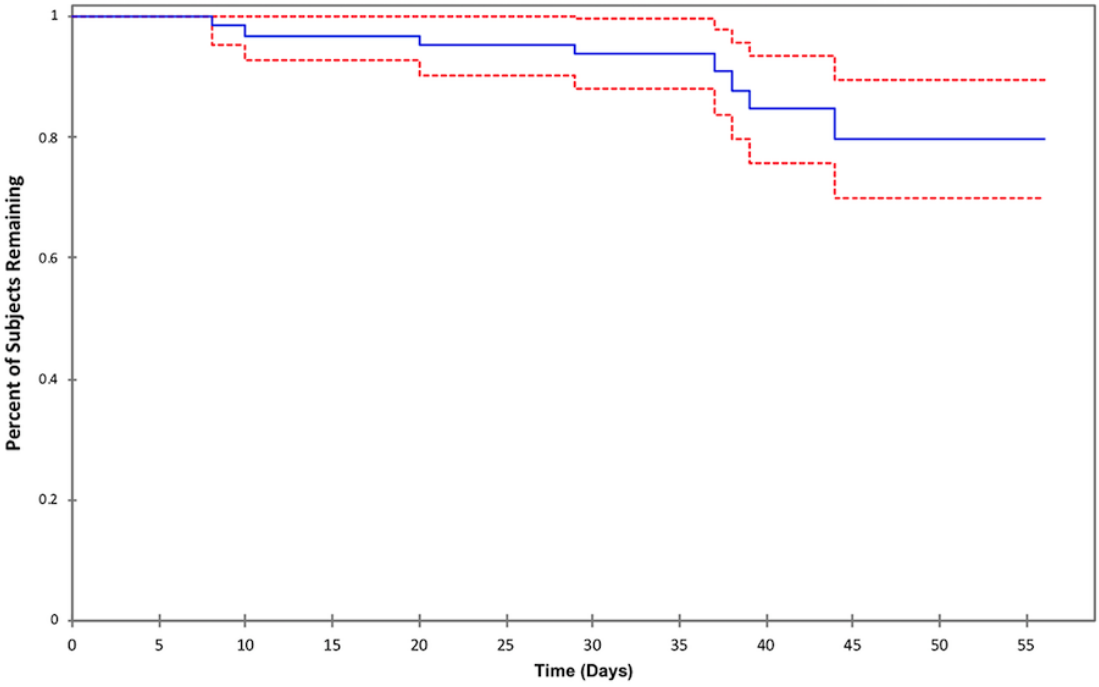


Table 2. Coefficients for the Cox proportional hazards model.

Coefficients	Coefficient estimate (95% CI)	*P* value
Age	1.12 (1.04-1.19)	.001
Marital status[a]		
Married	0.88 (0.25-3.18)	.85
Divorced	0.77 (0.15-4.04)	.76
Other	2.52 (0.35-18.26)	.36
Employment[b]		
Employee	0.46 (0.13-1.59)	.22
Low[c]	1.12 (0.27-4.62)	.87
Unemployed due to the disease	0.46 (0.12-1.67)	.24

[a]Marital status reference category: single.

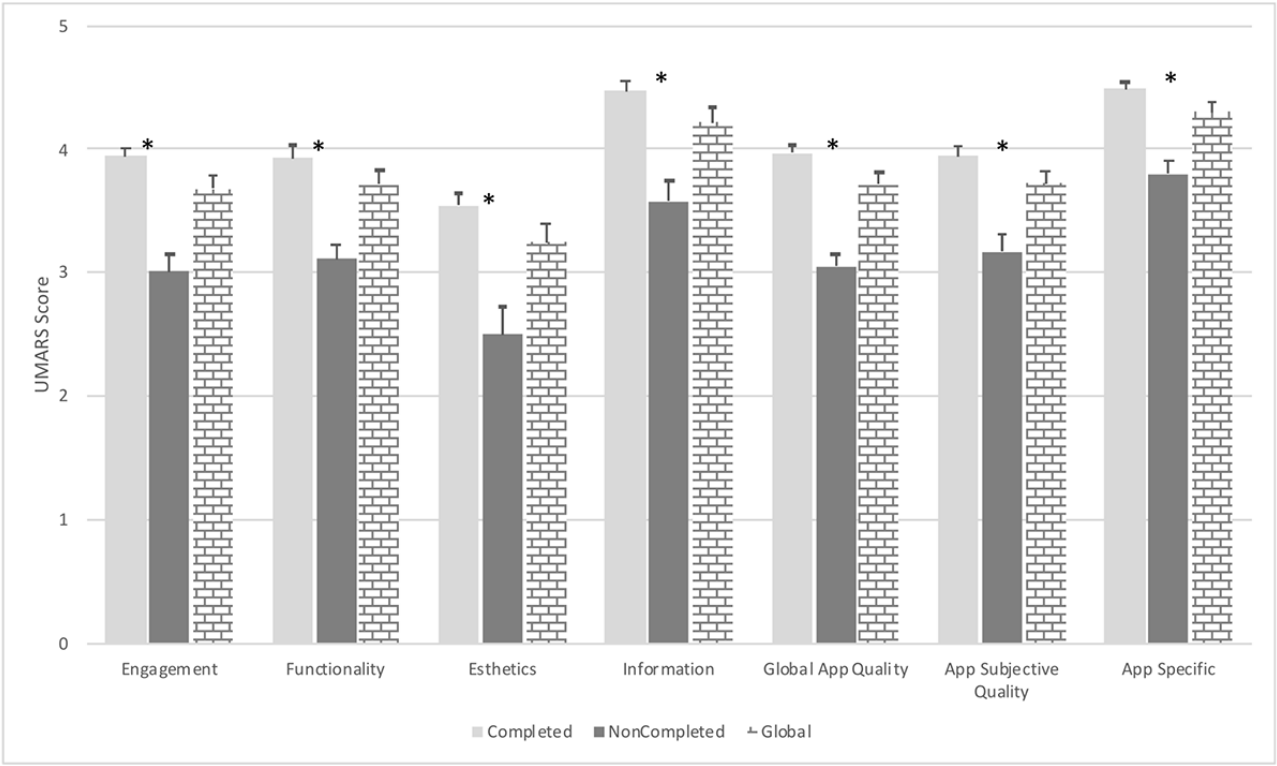
[b]Employment reference category: homemaker.

[c]Unemployed/on leave.

Patients' Perception of BENECA mHealth Quality

The mean MARS quality score for the app was 3.71 (SD 0.47) out of 5, and the NPS was positive (6.58 in range of −100 to 100), consisting of 24% (19/80) detractors, 30% (24/80) promoters, and 46% (37/80) passives. On average, the best-rated MARS category was app-specific change (mean 4.30, SD 0.37), followed by information (mean 4.22, SD 0.51), app subjective quality (mean 3.73, SD 0.46), and functionality (mean 3.71, SD 0.52). The worst-rated section was aesthetics, with a mean of 3.25 (SD 0.63). The specific scores for each section of the MARS are illustrated in Figure 3. The participants were divided according to the cut-off used in the survival analysis. It shows how the participants who used BENECA until the end of the experimental period scored higher and were statistically significant in all sections ($P<.001$). Linear regression showed that the older the patient, the lower the app quality score (beta=−0.29, $t_{75}=-2.64$, $P=.01$).

Figure 3. Mobile App Rating Scale (MARS) mean scoring. Data show differences between completed and noncompleted users and global mean scoring. Completed users are defined as those who used the BENECA mHealth app until study completion (n=58). Noncompleted users are defined as those who stopped using the BENECA mHealth app study completion (n=22). *P<.001. UMARS: User Version of the Mobile Application Rating Scale.



Barriers and Facilitators

In summary, seven barriers and five facilitators were quoted five times or more when the participants were interviewed. Among the barriers, the most common was "BENECA does not have some food items" followed by "difficulty at the time of introducing the intake." Among the facilitators, the most common was "BENECA provides relevant information to the patient" followed by "patient considers it important to know BENECA's feedback on energy balance." Table 3 summarizes the barriers and facilitators mentioned.

Table 3. Barriers and facilitators toward the feasibility of the BENECA mHealth app (N=77).

Barriers and facilitators	n (%)
Barriers	
Extension of BENECA	29 (38)
BENECA does not have an added value for the patient	9 (12)
BENECA does not have some food items	59 (77)
BENECA does not have some physical activities	8 (10)
BENECA feedback is limited	17 (22)
Difficulty at the time of introducing the intake	42 (55)
The patient's perception of BENECA's contribution to her health is negative	2 (3)
Facilitators	
The usefulness of BENECA in general	32 (42)
Ease of introducing physical activity	27 (35)
Patient considers it important to know BENECA's feedback on energy balance	55 (71)
BENECA is easy to use	18 (23)
BENECA provides relevant information to the patient	51 (66)

Main Clinical Outcomes

Quality of Life

The results of the main pre-post analyses of EORT QoL C30 are shown in Table 4. Statistically significant differences were observed after the experimental period with moderate to large effects as follows: general QoL (t_{75}=6.592, P<.001, d=0.87), physical functioning (t_{75}=5.312, P<.001, d=0.63), emotional functioning (t_{75}=2.981, P=.004, d=0.23), cognitive functioning (t_{75}=5.575, P<.001, d=0.75), social functioning (t_{75}=6.619, P<.001, d=0.82), fatigue (t_{75}=−6.003, P<.001, d=0.85), pain (t_{75}=−2.017, P=.047, d=0.23), dyspnea (t_{75}=−5.190, P<.001, d=0.61), and insomnia (t_{75}=−2.905, P=.005, d=0.32). An improvement in the scores of all these items, as well as a reduction in some symptoms, was observed after 2 months of using BENECA mHealth.

Table 4. Within-group pre-post effects on mean quality of life scores on the European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30 (EORTC QLQ-C30).

EORT QLQ-C30 variable	Study group, mean (SD)		Mean difference (95% CI)	*P* value[a]
	Pre (n=76)	Post (n=76)		
Global health	58.54 (14.40)	70.83 (11.26)	12.83 (8.95 to 16.71)	<.001
Physical functioning	75.25 (15.88)	85.35 (13.16)	10.88 (6.80 to 14.96)	<.001
Role functioning	66.45 (26.45)	70.83 (24.36)	5.26 (−1.99 to 12.52)	.15
Emotional functioning	59.06 (19.31)	64.04 (19.82)	5.59 (1.86 to 9.33)	.004
Cognitive functioning	62.5 (22.11)	80.26 (21.38)	17.98 (11.56 to 24.41)	<.001
Social functioning	66.88 (23.94)	86.62 (20.00)	20.17 (14.10 to 26.25)	<.001
Fatigue	42.5 (23.64)	23.68 (15.95)	−19.59 (−26.09 to −13.09)	<.001
Nausea	2.29 (6.35)	2.19 (5.67)	−0.22 (−1.54 to 1.10)	.95
Pain	44.58 (26.22)	38.6 (20.59)	−6.35 (−12.64 to −0.08)	.047
Dyspnea	27.92 (25.13)	12.72 (19.60)	−15.35 (−21.24 to −9.46)	<.001
Insomnia	46.25 (36.16)	35.09 (32.61)	−12.28 (−20.70 to −3.86)	.005
Appetite loss	9.58 (15.18)	7.46 (15.00)	−2.19 (−6.38 to 1.99)	.30
Constipation	21.67 (28.11)	19.74 (29.40)	−1.75 (−8.78 to 5.27)	.62
Diarrhea	10.83 (19.68)	12.72 (18.83)	1.31 (−1.85 to 4.47)	.41
Financial difficulties	19.17 (28.94)	16.67 (24.65)	−2.19 (−5.99 to 1.61)	.25

[a]Paired-sample *t* test or Wilcoxon signed rank test as appropriate. Analyses were performed on only those patients that followed-up.

Self-Efficacy and Motivation for Physical Activity and Accelerometry

The results of the main pre-post analyses using the self-efficacy scale for PA and accelerometry are shown in Table 5. There were significant statistical differences after the experimental period with a moderate ES on the EAF scale as follows: daily PA (t_{75}=5.369, P<.001, d=0.56), walking (t_{75}=6.228, P<.001, d=0.55), and total EAF score (t_{75}=6.423, P<.001, d=0.67). For accelerometry, there were only significant differences in weekday moderate-to-vigorous physical activity (MVPA; t_{75}=2.106, P=.04, d=0.26), observing trend in global MVPA (t_{75}=1.917, P=.06), weekday steps (t_{75}=1.779, P=.08), and global steps (t_{75}=1.647, P=.10). Therefore, after using BENECA mHealth, the users felt more motivated to increase the levels of PA in their daily lives.

Table 5. Within-group pre-post effects on mean scores on the self-efficacy scale for physical activity (EAF) and accelerometry.

Variable	Study group, mean (SD)		Mean difference (95% CI)	*P* value[a]
	Pre (n=76[b])	Post (n=76)		
EAF				
Scheduled PA[c]	81.70 (33.08)	87.71 (19.22)	6.08 (−1.08, 13.24)	.10
Daily live PA	50.06 (22.67)	62.63 (17.64)	12.22 (7.69, 16.76)	<.001
Walking	15.20 (9.03)	20.34 (7.95)	5.12 (3.48, 6.76)	<.001
Total EAF score	146.96 (53.36)	184.61 (48.52)	36.99 (25.52, 48.46)	<.001
Accelerometry				
MVPA[d] weekday	50.68 (25.83)	58.07 (26.05)	7.38 (0.39, 14.37)	.04
MVPA weekend	41.77 (24.55)	42.77 (21.51)	0.99 (−4.62, 6.62)	.73
MVPA global	48.14 (24.31)	53.69 (21.85)	5.55 (−0.22, 11.34)	.06
Steps weekday	7488.97 (3142.34)	8268.41 (3230.87)	779.44 (−94.35, 1653.22)	.08
Steps weekend	6218.50 (3147.26)	6316.87 (2875.87)	98.37 (−678.14, 874.88)	.80
Steps global	7125.97 (2935.94)	7710.82 (2672.78)	584.85 (−123.09, 1292.78)	.10

[a]Paired-sample *t* test or Wilcoxon signed rank test as appropriate. Analyses were performed on only those patients that followed-up.

[b]Accelerometry analyses was perform on 75 participants because there was one more dropout on preassessment (broken device).

[c]PA: physical activity.

[d]MVPA: moderate-to-vigorous physical activity.

Body Composition

The results of the main pre-post analyses of the anthropometric variables are shown in Table 6. Statistically significant differences were observed after the experimental period with a moderate ES as follows: weight (t_{75}=−5.050, *P*<.001, *d*=0.12) and BMI (t_{75}=−4.804, *P*<.001, *d*=0.12). In addition, a trend was observed in waist circumference (t_{74}=−1.900, *P*=.06) and body fat (t_{75}=−1.946, *P*=.06). No differences were observed for hip circumference (t_{74}=−1.007, *P*=.32) and bone mineral density (t_{75}=−1.019, *P*=.31). After 2 months of using BENECA mHealth, a reduction in users' body weight was observed, which could lead to a reduction in the hip circumference and percentage of body fat.

Table 6. Within-group pre-post differences on anthropometric and body composition variables.

Variable	Study group, mean (SD)		Mean difference (95% CI)	*P* value[a]
	Pre (n=76)	Post (n=76)		
Weight (kg)	73.09 (11.14)	71.67 (10.90)	−1.42 (−1.97, −0.86)	<.001
BMI[b] (kg/m^2)	29.11 (4.78)	28.51 (4.73)	−0.57 (−0.81, −0.34)	<.001
Waist circumference (cm)	87.45 (9.26)	86.97 (9.00)	−0.84 (−1.71, 0.04)	.06
Hip circumference (cm)	107.94 (14.23)	107.71 (13.11)	−0.64 (−1.93, 0.63)	.32
Body fat (%)	41.44 (6.23)	39.78 (7.34)	−1.57 (−3.18, 0.04)	.06
Bone mineral density (g/cm^2)	1.02 (0.11)	1.04 (0.14)	0.02 (−0.02, 0.05)	.31

[a]Paired-sample *t* test or Wilcoxon signed rank test as appropriate. Analyses were performed on only those patients that followed-up.

[b]BMI: body mass index.

Discussion

Principal Results

According to our initial hypothesis, after using BENECA mHealth for 8 weeks, the app was considered feasible by the breast cancer survivors in terms of use, adoption, and satisfaction, although the app quality score did not make it one of the best-rated apps. BENECA mHealth was associated with changes in the QoL of breast cancer survivors, as well as their self-perception of effectiveness and motivation for engaging in PA in their daily life.

Comparison With Prior Work

The adoption rate in this study was 69%, and the usage rate was 73% to 76%. These results can be explained by the technical characteristics of BENECA mHealth and its functionality, such as user-friendliness, the use of internationally accepted

measures, and the visual feedback. The results of this study are comparable with those obtained by Melissant et al [13] for a supportive care app for breast cancer survivor, which had an adoption rate of 75% and usage rate of 75% to 84%. Another study of a lifestyle intervention with a mobile app for endometrial and breast cancer survivors recorded a 75% usage rate [55]. However, Duman-Lubberding et al [11] obtained an adoption rate of 64% and a usage rate of 75% to 91% for a similar app for head and neck cancer survivors. The somewhat lower rate of use in our study for the latter may be due to how these data were obtained (ie, by the number of log-ins—objective measure—instead of the self-reported data of those studies—subjective measure). With regard to "app survival," we found that in a study by Springer et al [53] to test an mHealth app targeting healthy eating behavior in the general population, they obtained a survival rate less than 60% using the Kaplan-Meier survival curve. The higher survival rate in our study (over 70%) can be explained by the type of population studied. In general, patients with some type of pathology will be more predisposed to be involved in this type of study than the general population [56]. In addition, experiencing cancer treatment may be a stimulus to use the app, as patients may feel the increased need to learn more about the treatment.

Taking into account the barriers perceived by the participants in the use of the app, the barriers reported by BENECA mHealth were in line with a recently published review on the adherence to online psychological interventions [57] as well as with those in a study by Melissant et al [13] with the Oncokompas app to monitor the QoL of breast cancer survivors (eg, "Oncokompas is too extensive"). The reported mean satisfaction score of the quality of BENECA mHealth, although it may seem not very high, is in line with a recently published study on the quality of 18 mobile apps for pain management using the same MARS quality scale [41]. In addition, the low scores in some sections can be explained by the barriers reported by the patients, such as the difficulties in inputting the intake that makes it very extensive to fill in the app. This barrier was also reported in another feasibility study on head and neck cancer patients [11]. Considering that the minimum score to be considered a best-rated app based on the MARS scale is 3.73 (according to a previous study [41]), BENECA mHealth can be regarded as an app with average ratings. BENECA is currently being improved in an attempt to address the reported barriers.

The benefits of PA for cancer patients have been amply demonstrated [58], although a recent meta-analysis (2013-2018) of distance-based PA behavioral change interventions for cancer survivors concluded that the effects of interventions on PA were small [18]. In addition, although efficacy cannot be discussed in a study such as this, according to the literature, a difference of 8 points between assessments of QoL measured with the EORT QLQ-C30 is the minimum clinically significant difference required to discuss the clinical relevance of the findings [59]. The QoL findings in this study reinforce these preceding conclusions and are consistent with the results of the EAF scale and those observed via accelerometry. Changes are observed in the participants with more motivation to do PA, and it seems that using BENECA mHealth is associated with changes that lead to a positive feedback chain that improves physical and emotional functioning. The significant differences in cognitive functioning can be explained by the actual use of the mobile device, as there is evidence of the cognitive benefits of using electronic devices [60]. Our findings are in agreement with those reported by Pope et al [20], who used a mobile app and social media for 10 weeks to improve the QoL of breast cancer survivors, with a sample size much smaller than ours. However, they differ from the conclusions of McCarroll et al [55], who assessed the effectiveness of a public mobile app (LoseIT) for dietetic intervention for 4 weeks in breast and endometrial cancer survivors. They did not find significant changes in the QoL of the patients. It is possible that the experimental period of 4 weeks and lack of stratification of the type of cancer could explain these differences, despite the use of a powerful questionnaire to assess QoL. Lastly, we only found statistically significant differences in the MVPA of the data obtained via accelerometry, although we observed an improvement in other variables after the use of BENECA mHealth. These results are consistent with those of a clinical trial published in 2018 that used smartwatches and social media PA behavioral change over a 10-week intervention to determine the health outcomes for breast cancer survivors, in which no significant differences in the accelerometry variables were observed [61].

Finally, one of the main challenges not only with cancer patients but with the general population is the maintenance and reduction of body weight [5,62]. Different studies of lifestyle interventions have shown beneficial results, such as the one by von Gurenigen et al [63] in which they evaluated the effectiveness of a face-to-face intervention on diets in obese patients with endometrial cancer, achieving a reduction of approximately 5%. Similarly, McCarroll et al [55] achieved a reduction of approximately 6% from baseline weight. The literature indicates that a weight reduction of 5% is sufficient to reduce medical comorbidities [62]. In our study, an average weight loss of approximately 2% was achieved, which is below the recommendations. This may be because BENECA mHealth is not really a lifestyle intervention mobile app, but rather one that tries to incite behavioral change in users by monitoring their energy balance and making them aware of it. Therefore, we believe that the results obtained can be considered a first step, although future research should corroborate these results. The internal architecture of BENECA mHealth can also be extrapolated to suit patients with other types of cancer.

Strengths and Limitations

It is important to recognize some of the limitations of this study. The main one is its design. It is a nonrandomized, single-arm exploratory study; therefore, the results should be taken with caution. The ideal study would have been a randomized controlled trial (RCT); nevertheless, it was mandatory to develop a feasibility study for this sensitive population before carrying out an RCT. Moreover, due to the nature of the design of this study, the reported results must be confirmed in a larger RCT because the observed changes may not be attributable to the intervention. Secondly, BENECA was only developed for the Android operating system, but we are currently working on the next version of the BENECA app to solve this limitation. Thirdly, BENECA was designed to monitor energy balance and

then propose recommendations based on international guidelines of clinical practice, systematic reviews, and meta-analysis to ensure the recommendations can be generalized. However, we believe that it is a good starting point, especially for very sedentary people. Finally, the generalization of results is limited due to the design of the study, the use of restrictive inclusion and exclusion criteria and the recruitment strategy (the participants were referred by their oncologists), which may involve a bias of the threat of regression to the mean. In addition, another added difficulty could refer to the use of the app by older people in southern Spain, who may not even have mobile phones adapted to the app. Therefore, future studies should be conducted with a larger sample size; a controlled and randomized clinical trial design comparing the use of BENECA with, for example, a face-to-face intervention; and including biomarker measurements such as those for inflammation or development/recurrence of breast cancer.

Despite these limitations, this study also has strengths. These include the wide range of ages of the participants, which makes it possible to generalize the results; the use of energy balance as a means of changing behavior, which has not been studied much; its ease of use; it has high adherence; and it has no adverse effect on the prior validation of BENECA mHealth [32], which guarantees its reliability.

Conclusions

BENECA mHealth can be considered feasible in a real clinical context and has been associated with behavioral changes in the lifestyles of breast cancer survivors, but it needs to be enhanced to improve user satisfaction with use and functionality. Having assumed that BENECA is usable and applicable in a real clinical context, as well as having the first data of its applicability and clinical efficacy, the next step will be to confirm these results through a larger study with a control group. In addition, efforts should focus on overcoming the barriers reported by the participants and developing a new version of BENECA mHealth in which these improvements will be implemented. Finally, future research could focus on its generalization for application to other oncological processes. This study highlights the importance of the use of mobile apps based on energy balance and how the QoL of breast cancer survivors can be improved via monitoring. The results of this study could garner support for the use of this type of strategy in the projected 29.5 million cancer patients in 2040 [64].

Acknowledgments

The study was funded by the Spanish Ministry of Economy and Competitiveness (Plan Estatal de I+D+I 2013-2016), Fondo de Investigación Sanitaria del Instituto de Salud Carlos III (PI14/01627), Fondos Estructurales de la Unión Europea (FEDER), and by the Spanish Ministry of Education (FPU14/01069 and FPU17/00939). This study occurred thanks to additional funding from the University of Granada, Plan Propio de Investigación 2016, Excellence Actions: Units of Excellence; Unit of Excellence on Exercise and Health (UCEES). This work was part of a PhD thesis conducted in the Clinical Medicine and Public Health Doctoral Studies of the University of Granada, Spain.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of feedback messages of the BENECA mHealth app.

[PDF File (Adobe PDF File), 34KB - mhealth_v7i6e14136_app1.pdf]

References

1. Ulrich CM, Himbert C, Holowatyj AN, Hursting SD. Energy balance and gastrointestinal cancer: risk, interventions, outcomes and mechanisms. Nat Rev Gastroenterol Hepatol 2018 Nov;15(11):683-698. [doi: 10.1038/s41575-018-0053-2] [Medline: 30158569]
2. Silvera SA, Jain M, Howe GR, Miller AB, Rohan TE. Energy balance and breast cancer risk: a prospective cohort study. Breast Cancer Res Treat 2006 May;97(1):97-106. [doi: 10.1007/s10549-005-9098-3] [Medline: 16319973]
3. Wang X, Liu W, Xie X. Energy imbalance and cancer: cause or consequence? IUBMB Life 2017 Dec;69(10):776-784 [FREE Full text] [doi: 10.1002/iub.1674] [Medline: 28858429]
4. Hill JO, Wyatt HR, Peters JC. The importance of energy balance. Eur Endocrinol 2013 Aug;9(2):111-115 [FREE Full text] [doi: 10.17925/EE.2013.09.02.111] [Medline: 29922364]
5. Romieu I, Dossus L, Barquera S, Blottière HM, Franks PW, Gunter M, IARC working group on Energy Balance and Obesity. Energy balance and obesity: what are the main drivers? Cancer Causes Control 2017 Dec;28(3):247-258 [FREE Full text] [doi: 10.1007/s10552-017-0869-z] [Medline: 28210884]
6. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K, International Agency for Research on Cancer Handbook Working Group. Body fatness and cancer-viewpoint of the IARC Working Group. N Engl J Med 2016 Aug 25;375(8):794-798. [doi: 10.1056/NEJMsr1606602] [Medline: 27557308]

7. Courneya K. Exercise guidelines for cancer survivors: are fitness and quality-of-life benefits enough to change practice? Curr Oncol 2017 Feb 28;24(1):8. [doi: 10.3747/co.24.3545] [Medline: 28270718]
8. Cannon C. Telehealth, mobile applications, and wearable devices are expanding cancer care beyond walls. Semin Oncol Nurs 2018 Dec;34(2):118-125. [doi: 10.1016/j.soncn.2018.03.002] [Medline: 29627143]
9. Aitken M, Kleinrock M, Simorellis A, Nass D. IQVIA. Global oncology trends 2018 URL: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/global-oncology-trends-2018.pdf
10. Berkowitz C, Zullig L, Koontz B, Smith S. Prescribing an ppp? Oncology providers' views on mobile health apps for cancer care. JCO Clin Cancer Inform 2017 Dec;1:1-7 [FREE Full text] [doi: 10.1200/CCI.17.00107] [Medline: 30657404]
11. Duman-Lubberding S, van Uden-Kraan CF, Jansen F, Witte B, van der Velden LA, Lacko M, et al. Feasibility of an eHealth application. Support Care Cancer 2016 May;24(5):2163-2171 [FREE Full text] [doi: 10.1007/s00520-015-3004-2] [Medline: 26563178]
12. van der Hout A, van Uden-Kraan CF, Witte B, Coupé VM, Jansen F, Leemans C, et al. Efficacy, cost-utility and reach of an eHealth self-management application 'Oncokompas' that helps cancer survivors to obtain optimal supportive care: study protocol for a randomised controlled trial. Trials 2017 Dec 22;18(1):228 [FREE Full text] [doi: 10.1186/s13063-017-1952-1] [Medline: 28532439]
13. Melissant H, Verdonck-de Leeuw IM, Lissenberg-Witte B, Konings I, Cuijpers P, Van Uden-Kraan CF. 'Oncokompas', a web-based self-management application to support patient activation and optimal supportive care: a feasibility study among breast cancer survivors. Acta Oncol 2018 Jul;57(7):924-934. [doi: 10.1080/0284186X.2018.1438654] [Medline: 29451059]
14. Ormel HL, van der Schoot GG, Westerink NL, Sluiter WJ, Gietema JA, Walenkamp AM. Self-monitoring physical activity with a smartphone application in cancer patients: a randomized feasibility study (SMART-trial). Support Care Cancer 2018 Nov;26(11):3915-3923 [FREE Full text] [doi: 10.1007/s00520-018-4263-5] [Medline: 29785635]
15. Hernandez Silva E, 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] [Medline: 30635865]
16. Seiler A, Klaas V, Tröster G, Fagundes C. eHealth and mHealth interventions in the treatment of fatigued cancer survivors: a systematic review and meta-analysis. Psychooncology 2017 Dec;26(9):1239-1253. [doi: 10.1002/pon.4489] [Medline: 28665554]
17. Lee H, Uhm K, Cheong I, Yoo J, Chung S, Park Y, et al. Patient satisfaction with mobile health (mHealth) application for exercise intervention in breast cancer survivors. J Med Syst 2018 Nov 06;42(12):254. [doi: 10.1007/s10916-018-1096-1] [Medline: 30402781]
18. Groen W, van Harten WH, Vallance J. Systematic review and meta-analysis of distance-based physical activity interventions for cancer survivors (2013-2018): We still haven't found what we're looking for. Cancer Treat Rev 2018 Sep;69:188-203. [doi: 10.1016/j.ctrv.2018.07.012] [Medline: 30077954]
19. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile health physical activity intervention preferences in cancer survivors: a qualitative study. JMIR Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
20. Pope Z, Lee JE, Zeng N, Lee HY, Gao Z. Feasibility of smartphone application and social media intervention on breast cancer survivors' health outcomes. Transl Behav Med 2018 Feb 17;9(1):22. [doi: 10.1093/tbm/iby002] [Medline: 29471477]
21. Egbring M, Far E, Roos M, Dietrich M, Brauchbar M, Kullak-Ublick G, et al. A mobile app to stabilize daily functional activity of breast cancer patients in collaboration with the physician: a randomized controlled clinical trial. J Med Internet Res 2016 Dec 06;18(9):e238 [FREE Full text] [doi: 10.2196/jmir.6414] [Medline: 27601354]
22. Cox S, Lane A, Volchenboum S. Use of wearable, mobile, and sensor technology in cancer clinical trials. JCO Clin Cancer Inform 2018 Dec;2:1-11 [FREE Full text] [doi: 10.1200/CCI.17.00147] [Medline: 30652590]
23. Swendeman D, Comulada WS, Koussa M, Worthman CM, Estrin D, Rotheram-Borus MJ, et al. Longitudinal validity and reliability of brief smartphone self-monitoring of diet, stress, and physical activity in a diverse sample of mothers. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e176 [FREE Full text] [doi: 10.2196/mhealth.9378] [Medline: 30249576]
24. Lee M, Lee H, Kim Y, Kim J, Cho M, Jang J, et al. Mobile app-based health promotion programs: a systematic review of the literature. Int J Environ Res Public Health 2018 Dec 13;15(12):2838. [doi: 10.3390/ijerph15122838]
25. 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] [Medline: 28646889]
26. Dodd JM, Louise J, Cramp C, Grivell RM, Moran LJ, Deussen AR. Evaluation of a smartphone nutrition and physical activity application to provide lifestyle advice to pregnant women: The SNAPP randomised trial. Matern Child Nutr 2017 Aug 24;14(1):e12502. [doi: 10.1111/mcn.12502]
27. Paulsen MM, Hagen ML, Frøyen MH, Foss-Pedersen RJ, Bergsager D, Tangvik RJ, et al. A dietary assessment app for hospitalized patients at nutritional risk: development and evaluation of the MyFood app. JMIR Mhealth Uhealth 2018 Sep 07;6(9):e175. [doi: 10.2196/mhealth.9953]

28. Thijs I, Fresiello L, Oosterlinck W, Sinnaeve P, Rega F. Assessment of physical activity by wearable technology during rehabilitation after cardiac surgery: explorative prospective monocentric observational cohort study. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e9865 [FREE Full text] [doi: 10.2196/mhealth.9865] [Medline: 30702433]
29. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
30. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
31. Lozano-Lozano M, Moral-Munoz JA, Galiano-Castillo N, Martín-Martín L, Fernández-Lao C, Arroyo-Morales M, et al. Designing BENECA m-Health app, a mobile health application to monitor diet and physical activity in cancer survivors. In: Rojas I, Joya G, Catala A, editors. Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science. Cham, Switzerland: Springer; 2017:63-70.
32. Lozano-Lozano M, Galiano-Castillo N, Martín-Martín L, Pace-Bedetti N, Fernández-Lao C, Arroyo-Morales M, et al. Monitoring energy balance in breast cancer survivors using a mobile app: reliability study. JMIR Mhealth Uhealth 2018 Mar 27;6(3):e67 [FREE Full text] [doi: 10.2196/mhealth.9669] [Medline: 29588273]
33. Finne E, Glausch M, Exner A, Sauzet O, Stölzel F, Seidel N. Behavior change techniques for increasing physical activity in cancer survivors: a systematic review and meta-analysis of randomized controlled trials. Cancer Manag Res 2018 Oct;10:5125-5143 [FREE Full text] [doi: 10.2147/CMAR.S170064] [Medline: 30464612]
34. Hillier FC, Batterham AM, Crooks S, Moore HJ, Summerbell CD. The development and evaluation of a novel Internet-based computer program to assess previous-day dietary and physical activity behaviours in adults: the Synchronised Nutrition and Activity Program for Adults (SNAPA™). Br J Nutr 2012 Apr;107(8):1221-1231. [doi: 10.1017/S0007114511004090] [Medline: 21861942]
35. Government of Spain. Law 14/2007, of July 3, on Biomedical Research. Inter-University Chair in Law and the Human Genome: Off Bull State, No 159, BOE-A-2007-12945 (Jun 2) 2011:1-48 [FREE Full text]
36. Lozano-Lozano M, Martín-Martín L, Galiano-Castillo N, Álvarez-Salvago F, Cantarero-Villanueva I, Fernández-Lao C, et al. Integral strategy to supportive care in breast cancer survivors through occupational therapy and a m-health system: design of a randomized clinical trial. BMC Med Inform Decis Mak 2016 Nov 25;16(1):150 [FREE Full text] [doi: 10.1186/s12911-016-0394-0] [Medline: 27887610]
37. Schmitz K, Courneya K, Matthews C, Demark-Wahnefried W, Galvão DA, Pinto B, American College of Sports Medicine. American College of Sports Medicine roundtable on exercise guidelines for cancer survivors. Med Sci Sports Exerc 2010 Jul;42(7):1409-1426. [doi: 10.1249/MSS.0b013e3181e0c112] [Medline: 20559064]
38. Kushi L, Doyle C, McCullough M, Rock C, Demark-Wahnefried W, Bandera EV, American Cancer Society 2010 Nutrition and Physical Activity Guidelines Advisory Committee. American Cancer Society Guidelines on nutrition and physical activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2012;62(1):30-67 [FREE Full text] [doi: 10.3322/caac.20140] [Medline: 22237782]
39. Vollmer DD, Fair K, Hong YA, Beaudoin CE, Pulczinski J, Ory MG. Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps. JMIR Mhealth Uhealth 2015;3(1):e31 [FREE Full text] [doi: 10.2196/mhealth.3861] [Medline: 25830810]
40. Cnossen I, van Uden-Kraan CF, Rinkel R, Aalders I, de Goede CJ, de Bree R, et al. Multimodal guided self-help exercise program to prevent speech, swallowing, and shoulder problems among head and neck cancer patients: a feasibility study. J Med Internet Res 2014 Mar 06;16(3):e74 [FREE Full text] [doi: 10.2196/jmir.2990] [Medline: 24610383]
41. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
42. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
43. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
44. Krol M, de Boer D, Delnoij D, Rademakers JJ. The Net Promoter Score--an asset to patient experience surveys? Health Expect 2015 Dec;18(6):3099-3109 [FREE Full text] [doi: 10.1111/hex.12297] [Medline: 25345554]
45. Koladycz R, Fernandez G, Gray K, Marriott H. The Net Promoter Score (NPS) for insight into client experiences in sexual and reproductive health clinics. Glob Health Sci Pract 2018 Dec 03;6(3):413-424 [FREE Full text] [doi: 10.9745/GHSP-D-18-00068] [Medline: 30072372]
46. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 08;320(7227):114-116 [FREE Full text] [doi: 10.1136/bmj.320.7227.114] [Medline: 10625273]
47. Aaronson N, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez N, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 03;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]

48. Nolte S, Liegl G, Petersen MA, Aaronson NK, Costantini A, Fayers PM, EORTC Quality of Life Group. General population normative data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons across 13 European countries, Canada and the Unites States. Eur J Cancer 2019 Jan;107:153-163 [FREE Full text] [doi: 10.1016/j.ejca.2018.11.024] [Medline: 30576971]
49. Fernández Cabrera T, Medina Anzano S, Herrera Sánchez IM, Rueda Méndez S, Fernández Del Olmo A. [Construction and validation of a self-efficacy scale for physical activity]. Rev Esp Salud Publica 2011 Aug;85(4):405-417 [FREE Full text] [doi: 10.1590/S1135-57272011000400009] [Medline: 22392471]
50. Carbonell-Baeza A, Ruiz J, Aparicio V, Ortega F, Munguía-Izquierdo D, Alvarez-Gallardo IC, et al. Land- and water-based exercise intervention in women with fibromyalgia: the al-Andalus physical activity randomised controlled trial. BMC Musculoskelet Disord 2012 Feb 15;13:18 [FREE Full text] [doi: 10.1186/1471-2474-13-18] [Medline: 22336292]
51. Greenspan S, Vujevich K, Brufsky A, Lembersky B, van Londen GJ, Jankowitz R, et al. Prevention of bone loss with risedronate in breast cancer survivors: a randomized, controlled clinical trial. Osteoporos Int 2015 Jun;26(6):1857-1864 [FREE Full text] [doi: 10.1007/s00198-015-3100-7] [Medline: 25792492]
52. Leahy S, Toomey C, McCreesh K, O'Neill C, Jakeman P. Ultrasound measurement of subcutaneous adipose tissue thickness accurately predicts total and segmental body fat of young adults. Ultrasound Med Biol 2012 Jan;38(1):28-34. [doi: 10.1016/j.ultrasmedbio.2011.10.011] [Medline: 22104525]
53. Springer A, Venkatakrishnan A, Mohan S, Nelson L, Silva M, Pirolli P. Leveraging self-affirmation to improve behavior change: a mobile health app experiment. JMIR Mhealth Uhealth 2018 Jul 19;6(7):e157. [doi: 10.2196/mhealth.9151]
54. Schafer J. Multiple imputation: a primer. Stat Methods Med Res 1999 Mar;8(1):3-15. [doi: 10.1177/096228029900800102] [Medline: 10347857]
55. McCarroll M, Armbruster S, Pohle-Krauza R, Lyzen A, Min S, Nash D, et al. Feasibility of a lifestyle intervention for overweight/obese endometrial and breast cancer survivors using an interactive mobile application. Gynecol Oncol 2015 Jun;137(3):508-515. [doi: 10.1016/j.ygyno.2014.12.025] [Medline: 25681782]
56. Anderson A, Borfitz D, Getz K. Global public attitudes about clinical research and patient experiences with clinical trials. JAMA Netw Open 2018 Oct 05;1(6):e182969. [doi: 10.1001/jamanetworkopen.2018.2969]
57. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
58. Segal R, Zwaal C, Green E, Tomasone J, Loblaw A, Petrella T, Exercise for People with Cancer Guideline Development Group. Exercise for people with cancer: a clinical practice guideline. Curr Oncol 2017 Feb;24(1):40-46 [FREE Full text] [doi: 10.3747/co.24.3376] [Medline: 28270724]
59. Kvam A, Fayers P, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. Eur J Haematol 2011 Oct;87(4):330-337. [doi: 10.1111/j.1600-0609.2011.01665.x] [Medline: 21668504]
60. Wilmer H, Sherman L, Chein J. Smartphones and cognition: a review of research exploring the links between mobile technology habits and cognitive functioning. Front Psychol 2017;8:605 [FREE Full text] [doi: 10.3389/fpsyg.2017.00605] [Medline: 28487665]
61. Pope Z, Zeng N, Zhang R, Lee H, Gao Z. Effectiveness of combined smartwatch and social media intervention on breast cancer survivor health outcomes: a 10-week pilot randomized trial. J Clin Med 2018 Jun 07;7(6):140. [doi: 10.3390/jcm7060140]
62. Appel L. UpToDate. Overweight, obesity, and weight reduction in hypertension Internet URL: https://www.uptodate.com/contents/overweight-obesity-and-weight-reduction-in-hypertension#H7 [accessed 2019-03-06]
63. von Gruenigen VE, Waggoner S, Frasure H, Kavanagh M, Janata J, Rose P, et al. Lifestyle challenges in endometrial cancer survivorship. Obstet Gynecol 2011 Jan;117(1):93-100. [doi: 10.1097/AOG.0b013e31820205b3] [Medline: 21173649]
64. International Agency for Research on Cancer. Global Cancer Observatory URL: https://gco.iarc.fr

Abbreviations

BMI: body mass index
EAF: self-efficacy scale for physical activity (Spanish)
EORT QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
ES: effect size
MARS: Mobile App Rating Scale
MVPA: moderate-to-vigorous physical activity
NPS: Net Promoter Score
PA: physical activity
QoL: quality of life
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 26.03.19; peer-reviewed by F Artacho-Cordon, M Robertson; comments to author 18.04.19; revised version received 13.05.19; accepted 16.05.19; published 25.06.19.

Please cite as:
Lozano-Lozano M, Cantarero-Villanueva I, Martin-Martin L, Galiano-Castillo N, Sanchez MJ, Fernández-Lao C, Postigo-Martin P, Arroyo-Morales M
A Mobile System to Improve Quality of Life Via Energy Balance in Breast Cancer Survivors (BENECA mHealth): Prospective Test-Retest Quasiexperimental Feasibility Study
JMIR Mhealth Uhealth 2019;7(6):e14136
URL: http://mhealth.jmir.org/2019/6/e14136/
doi:10.2196/14136
PMID:31237570

Original Paper

Efficacy of the Mindfulness Meditation Mobile App "Calm" to Reduce Stress Among College Students: Randomized Controlled Trial

Jennifer Huberty[1], PhD; Jeni Green[1], MS; Christine Glissmann[2], MS; Linda Larkey[3], PhD; Megan Puzia[4], MS; Chong Lee[1], PhD

[1]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

[2]Department of Student Affairs, University of Redlands, Redlands, CA, United States

[3]College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[4]Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

Corresponding Author:
Jennifer Huberty, PhD
College of Health Solutions
Arizona State University
500 North 3rd Street
Phoenix, AZ, 85004
United States
Phone: 1 602 827 2456
Fax: 1 602 827 2469
Email: jhuberty@asu.edu

Abstract

Background: College students experience high levels of stress. Mindfulness meditation delivered via a mobile app may be an appealing, efficacious way to reduce stress in college students.

Objective: We aimed to test the initial efficacy and sustained effects of an 8-week mindfulness meditation mobile app—Calm—compared to a wait-list control on stress, mindfulness, and self-compassion in college students with elevated stress. We also explored the intervention's effect on health behaviors (ie, sleep disturbance, alcohol consumption [binge drinking], physical activity, and healthy eating [fruit and vegetable consumption]) and the feasibility and acceptability of the app.

Methods: This study was a randomized, wait-list, control trial with assessments at baseline, postintervention (8 weeks), and at follow-up (12 weeks). Participants were eligible if they were current full-time undergraduate students and (1) at least 18 years of age, (2) scored ≥14 points on the Perceived Stress Scale, (3) owned a smartphone, (4) were willing to download the Calm app, (5) were willing to be randomized, and (7) were able to read and understand English. Participants were asked to meditate using Calm at least 10 minutes per day. A P value ≤.05 was considered statistically significant.

Results: A total of 88 participants were included in the analysis. The mean age (SD) was 20.41 (2.31) years for the intervention group and 21.85 (6.3) years for the control group. There were significant differences in all outcomes (stress, mindfulness, and self-compassion) between the intervention and control groups after adjustment for covariates postintervention (all P<.04). These effects persisted at follow-up (all P<.03), except for the nonreacting subscale of mindfulness (P=.08). There was a significant interaction between group and time factors in perceived stress (P=.002), mindfulness (P<.001), and self-compassion (P<.001). Bonferroni posthoc tests showed significant within-group mean differences for perceived stress in the intervention group (P<.001), while there were no significant within-group mean differences in the control group (all P>.19). Similar results were found for mindfulness and self-compassion. Effect sizes ranged from moderate (0.59) to large (1.24) across all outcomes. A significant group×time interaction in models of sleep disturbance was found, but no significant effects were found for other health behaviors. The majority of students in the intervention group reported that Calm was helpful to reduce stress and stated they would use Calm in the future. The majority were satisfied using Calm and likely to recommend it to other college students. The intervention group participated in meditation for an average of 38 minutes/week during the intervention and 20 minutes/week during follow-up.

Conclusions: Calm is an effective modality to deliver mindfulness meditation in order to reduce stress and improve mindfulness and self-compassion in stressed college students. Our findings provide important information that can be applied to the design of future studies or mental health resources in university programs.

Trial Registration: ClinicalTrials.gov NCT03891810; https://clinicaltrials.gov/ct2/show/NCT03891810

(JMIR Mhealth Uhealth 2019;7(6):e14273) doi:10.2196/14273

KEYWORDS

meditation; mental health; mindfulness; smartphone; technology; mobile phone

Introduction

Elevated stress has been reported in more than 75% of college students (ie, aged 18-33 years) [1-3], and college students often report higher levels of stress than people of other age groups [4,5]. Some studies showed that more than 85% of college students reported feeling overwhelmed [4,5]. This is likely due to the unique set of stressors and demands they experience as they gain autonomy from their parents (ie, move away leaving family and friends), learn to manage finances, balance an increased academic workload and extracurricular activities, and make career choices [5-7]. These demands may lead to increased anxiety, loneliness, depression, sleep disturbance, and even suicidal ideation [8-10].

Stress is associated with a greater likelihood of suicide attempts [11], which is the second leading cause of death among teens and young adults (aged 15-24 years) [12]. One report found that 60.8% of college students felt overwhelming anxiety, 38.2% felt so depressed that it was difficult to function, and 10.4% seriously considered suicide [13]. Studies suggest that untreated mental health is highly prevalent among college students. One study reported that <25% of college students received treatment for any type of mental health disorder (eg, depression or anxiety) within the past year [14]. Another study reported that only 36% of students with mental health issues had sought treatment [15]. Many students do not seek treatment partly due to the lack of time, privacy concerns, stigma [16], lack of emotional openness [17], and financial constraints [18-21].

Beyond the potential negative impact that elevated stress may have on mental health, stress may also have a negative impact on the physical health of college students and the health behavior they choose during this time [22]. College students with elevated stress may be more likely to experience poor sleep, increased alcohol consumption, less exercise, and unhealthy eating habits [23-25]. There is an urgent need for effective stress management strategies to address the adverse outcomes of stress and the potential barriers to treatment in college students.

Implementing stress reduction programs on college campuses has become a priority [26,27]. Mindfulness-based interventions, in particular, have become more popular on college campuses recently [28] and may be an effective strategy to reduce stress in college students [29]. Mindfulness is defined as the state of being attentive to and aware of what is taking place in the present moment without judgement [30]. Two of the most popular mindfulness-based interventions are mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy (MBCT) [31,32]. Although effective for stress management, these programs may be rigorous, time consuming, and costly, which may not be ideal for college students.

Mindfulness meditation, a component of MBSR and MBCT, alone has shown to have positive benefits—from reducing stress and anxiety to improving overall well-being [29,33]. The mechanisms of mindfulness meditation are unclear but constructs such as increasing self-compassion and mindfulness have been shown to mediate the effect of meditation programs on stress in a variety of populations [34-36]. Meta-analyses have suggested that self-compassion is strongly related to mental health [37] and well-being [38]. However, the stress-reducing effects in college students is less clear. More research is needed to understand what aspects of mindfulness meditation interventions may be working to reduce stress in college students.

Mindfulness meditation may be a practical approach to reduce stress in college students. In a recent narrative synthesis of studies on the effects of mindfulness meditation on stress in college students, 16 of the 22 studies reported significant decreases in self-reported stress [29]. Significant findings were reported in interventions ranging from 2.25 hours to as much as 30 hours of total required participation in meditation over the course of the intervention (eg, 3-8 weeks) [39-44]. Although these findings are promising, several limitations must be noted, including low adherence rates, lack of control groups, and small sample sizes [29]. More studies are needed to address the existing limitations of mindfulness meditation interventions and explore more convenient ways to deliver mindfulness meditation in order to encourage participation and increase accessibility among college students.

Approximately 98% of college-aged students use the internet [45] and 85% own a smartphone [46]. Many college students have expressed being receptive to online mental health treatments, with one study reporting that college students were more likely to seek help online than face-to-face [47]. The development of mobile health apps has been increasing exponentially, and the use of apps has been reported to improve the efficiency of health care delivery and the effectiveness of treatment [48]. It is also important to note that while apps may offer the ability for increased access to health information, remote care, and user autonomy [49], there is a need to ensure safety of users, as there is no regulatory oversight. App users should consider the quality and accuracy of evidence-based information, qualifications of content developers, medical/health claims [50], information sharing or data security [51], informed consent, and user competence (eg, training required to use techniques) [49]. Independent testing of mobile apps is highly warranted.

Mindfulness meditation mobile apps may be a promising approach to reduce stress and address the barriers for stress management reported by college students. A recent study conducted a search for “mindfulness” mobile apps in Apple iTunes and Google Play Store and found a total of 560 apps

that were accessible and in English [52]. The top mindfulness apps listed in Apple iTunes (in a current search as of April 2019) was Calm (named "2017 app of the year" by Apple, #2 Health & Fitness app, 4.8/5 star reviews), followed by Aura (#6 Health & Fitness app, 4.7/5 star reviews) and Headspace (#9 Health & Fitness app, 4.9/5 star reviews). Similar results were found in the Google Play Store, with the top mindfulness apps being Headspace (#6 Health & Fitness app, 4.5/5 star reviews), followed by Calm (#8 Health & Fitness app, 4.5/5 star reviews) and Deep Calm (#44 Health & Fitness app, 4.5/5 star reviews). Although the popularity of mindfulness apps is increasing, only a few studies have examined the feasibility and efficacy of such mindfulness-based mobile apps for reducing stress among college students.

The purpose of this study was to test the efficacy [53] of an 8-week mindfulness meditation intervention delivered via a consumer-based mobile app (ie, Calm) as compared to a wait-list control group on stress in college students with elevated stress (≥14 on the Perceived Stress Scale [PSS]). We also explored the feasibility (ie, acceptability and demand) of the intervention delivered via the mobile app and examined the sustained effects (at 12 weeks from baseline) of the intervention on stress, mindfulness, and self-compassion. Additionally, exploratory analyses examined the potential effects of the mindfulness intervention on health behaviors (ie, sleep disturbance, alcohol consumption [binge drinking], physical activity, and healthy eating [fruit and vegetable consumption]).

We hypothesized that college students in the intervention group would have significant improvements in perceived stress as compared to the wait-list control. We also hypothesized that stress, mindfulness, and self-compassion would have sustained effects in the intervention group as compared to the wait-list control. Finally, we hypothesized that the intervention may improve health behaviors. The findings from this study will provide important insights on the potential stress-reducing effects of a consumer-based mindfulness meditation mobile app (ie, Calm) and could be applied to the design of future mindfulness meditation interventions in stressed college students.

Methods

Ethics Approval

This study was approved by an Institutional Review Board at a large university in the Southwestern United States (STUDY00006896). All participants provided electronic consent prior to participation in the study. The datasets generated or analyzed during the study are available from the corresponding author upon request.

Study Design

This study was a randomized, wait-list control trial (trial registration: ClinicalTrials.gov NCT03891810) with assessments conducted at baseline, postintervention (8 weeks), and at follow-up (12 weeks). Participants randomized to the intervention group participated in an 8-week mindfulness meditation mobile app intervention of at least 10 minutes per day. Those randomized to the wait-list control group received the intervention after 12 weeks.

Sample Size

The sample size was determined to detect significant change in perceived stress between the intervention and control groups (effect size=0.76) based on previous studies [28,54]. The power analysis showed that 93 participants were sufficient to have 85% statistical power at a two-sided α of 0.05 for significance. Accounting for a 10% dropout rate, the total sample size was planned to be 104 participants.

Recruitment and Selection

Participants were recruited between January and April 2018 via social media (ie, Facebook and Instagram), email listservs, and flyers and by emailing university professors. Interested students were sent an eligibility survey via Qualtrics (ie, online survey database). Participants were eligible for the study if they were a current full-time undergraduate student attending a university in the Southwestern United States and (1) at least 18 years old, (2) scored ≥14 points on the PSS, (4) owned a smartphone, (5) were willing to download the Calm app, (6) were willing to be randomized, and (7) were able to read and understand English. Participants were excluded if they had a current mindfulness practice (ie, practice ≥15 min per day of meditation, yoga, and body scan within the past 6 months) [55] or were currently using the Calm app or another meditation app. Eligible participants were sent a link via Qualtrics to an online intake video that was approximately 5 minutes long. Participants were required to view the video in its entirety and correctly answer three follow-up questions regarding information presented in the video (eg, how long will you be asked to participate in mediation?). If participants missed a question, they were redirected to watch the video again and answer the questions until all three were answered correctly. Once the questions were answered correctly, participants were sent the informed consent and baseline questionnaire via Qualtrics software (Qualtrics, Provo, UT). After the informed consent and baseline questionnaire was completed, participants were randomized into the intervention group or the wait-list control group.

Randomization and Blinding

Participants were randomized using a list created from an online randomization tool (randomizer.org) that was set to randomize in a ratio of 1:1. After participants completed baseline assessments, a research team member (unblinded) allocated participants to a group using the randomization list. Study staff and participants were unblinded to group allocation.

Intervention Group

After randomization, the intervention group was emailed instructions on how to download the Calm mobile app. Calm is a consumer-based mindfulness meditation mobile app that offers a range of mindfulness meditation practice guide modules that vary in length, instruction, and content. Mindfulness meditation is the practice of moment-to-moment awareness in which the person purposefully focuses on the present without judgement. Vipassana is a technique of mindfulness that explores how the mind influences the body and how the body influences the mind (ie, objective observation of physical

sensations in the body) [56]. The goal is to sharpen concentration, maintain awareness, and develop equanimity by releasing habitual tendencies toward craving and aversion. Calm also integrates some cognitive behavioral therapy (CBT) techniques into the meditation sessions on occasion. The CBT-influenced sessions encourage users to develop awareness of their thoughts, interpretations, and emotional and physiological responses in order to alter their perception of a situation or create a new, more balanced thought process [57].

Participants can meditate using the "daily Calm" set of guided meditations or may choose from a number of programs offering multiday meditations specific to goals (ie, happiness or self-esteem). Calm also offers other individual guided and unguided (eg, a brief introductory guidance followed by a chosen period of silence or sounds from nature) meditations.

Once downloaded, participants were asked to complete the "7 Days of Calm" program for the first week of the intervention to familiarize themselves with the principles of meditation and to standardize the introductory teaching content. Each day, the "7 Days of Calm" program started with an educational component on a principle of mindfulness meditation (eg, being present, returning to the here and now, and pulling out of autopilot). For each 10-minute session, after a principle was discussed, a related mindfulness meditation exercise was introduced and guided (eg, body scan, breath focus, and loving kindness). After participants completed the "7 Days of Calm" program, they were allowed to choose any meditation they desired. They could choose meditations from the College Collection, which addresses topics such as stress, sleep, self-compassion, and concentration or they could choose another series such as "7 Days of Managing Stress." During the 8-week intervention, participants were asked to complete at least 10-minutes daily of meditation and could exceed that time limit by choosing additional meditations. If participants were not achieving 30 minutes of meditation per week, they were sent a text reminder to meditate. After the 8-week intervention, participants still had access to calm and could use it at their own leisure for 1 additional month (12 weeks from baseline).

Wait-List Control Group

Participants randomized to the wait-list control group received an email with their group assignment and stating that they would receive access to the Calm app after 12 weeks. They were also asked not to participate in any mindfulness activities (eg, yoga, meditation, and qigong) during this time. After 8 weeks, participants received a Qualtrics link to the postassessment (same surveys that the intervention group received). After 12 weeks, participants were sent a Qualtrics link to the follow-up assessment (same surveys that the intervention group received) and an email with instructions on how to download Calm and the assigned username and password.

Measures and Incentives

Both the intervention and wait-list control groups were administered three surveys—at baseline (week 0), postintervention (week 8), and at follow-up (week 12)—to assess perceived stress, mindfulness, self-compassion, health behaviors, and feasibility outcomes via online surveys (Qualtrics). Data on demographics, mental health history, medication use, and counseling activities were collected at baseline. Participants were incentivized with a US $5 gift card for completing baseline questionnaires, US $10 gift card for completing the postintervention questionnaires, and US $15 gift card for completing follow-up questionnaires. Participants could choose to receive their gift card from Starbucks, Target, or Amazon. Calm memberships were provided to participants for free during the study.

Perceived Stress

Stress was measured using the PSS [58,59]. This scale is a 10-item inventory used for the assessment of perceived stress. The scale measures the degree to which situations are appraised as stressful. The response items are rated on a 5-point Likert scale ranging from 0 (never) to 4 (very often). Scores range from 0 to 40, with higher scores indicating higher levels of perceived stress. The PSS has been shown to be reliable in undergraduate college samples [58,59]; in this study, the alpha coefficient for baseline PSS scores was 0.82.

Mindfulness

Mindfulness was measured using the Five Factor Mindfulness Questionnaire (FFMQ) [60]. The FFMQ is a 39-item self-report inventory used for the assessment of multiple constructs of mindfulness skills. The inventory assesses five subscales: observing, describing, acting with awareness, nonjudgment of inner experience, and nonreactivity to inner experience. The response items are rated on a 5-point Likert scale ranging from 1 (never or very rarely true) to 5 (very often or always true). The facet scores range from 8 to 40, with the exception of nonreactivity to inner experience, which ranges from 7 to 35. Higher scores indicate higher levels of mindfulness. The five FFMQ subscales have high internal reliability, test-retest reliability, and validity in undergraduate samples [60]. Consistent with this research, the FFMQ subscale reliability in this study was found to be high (at baseline, alpha coefficients were between 0.80 and 0.92).

Self-Compassion

Self-compassion was measured using the Self-Compassion Survey Short-Form (SCS-SF) [61]. The SCS-SF is a 12-item survey assessing three subscales: self-kindness versus self-judgment, common humanity versus isolation, and mindfulness versus over-identification. The response items are rated on a 5-point Likert scale ranging from 1 (almost never) to 5 (almost always). Higher scores indicate higher levels of self-compassion. The SCS-SF is a valid and reliable measure to assess self-compassion in college students (α=0.87 in our sample), and previous research has demonstrated near-perfect correlations (r=0.98) with the long-form version [61].

Secondary Outcomes

Health behaviors measured included sleep disturbance, alcohol consumption (ie, binge drinking), physical activity, and fruit and vegetable consumption. Sleep disturbance (ie, sleep quality) was measured using the Patient-Reported Outcomes Measurement Information System (PROMIS) (α=0.97 in our sample) [62]. The Youth Risk Behavior Surveillance (YRBS) survey was used to measure binge drinking, physical activity

participation, and healthy eating [63]. Specifically, items from the YRBS survey inquired about whether participants had consumed alcohol (ie, engaged in binge drinking; for women, consuming four consecutive alcoholic beverages in a 2-hour period; for men, five alcoholic beverages) at any point during the past 7 days, engaged in at least 150 minutes of physical activity during the past 7 days, and eaten five servings of fruits or vegetables on most of the past 7 days.

Feasibility

Feasibility measures included acceptability and demand. Acceptability was measured with a satisfaction survey postintervention (week 8). Demand was measured using adherence (minutes/week) to the meditation intervention. Adherence to meditation was recorded in weekly reports from the Calm informatics team. Reports included the date and time of each meditation participated in, the title of the meditation, and the duration of participation (ie, the time spent viewing the meditation) for each participant.

Statistical Analyses

General linear models (GLMs) were used to examine differences in means between groups at baseline and to test the initial efficacy of the intervention, examining mean differences of change (ie, change scores) for perceived-stress, mindfulness, and self-compassion after 8 weeks of mindfulness meditation between the intervention and control groups after adjustment for covariates (age, gender, and race). To examine the sustained effects, linear mixed models were used to test mean differences for perceived stress, mindfulness, and self-compassion between groups, time, and group×time interaction factors, with adjustment for covariates (age, gender, and race) performed using Bonferroni posthoc tests. GLMs were also used to evaluate changes in sleep disturbances; however, because sleep was exploratory, covariates were not included in these models. The McNemar tests were used to examine changes in other exploratory outcome variables (ie, binge drinking, physical activity, and healthy eating), which were coded dichotomously.

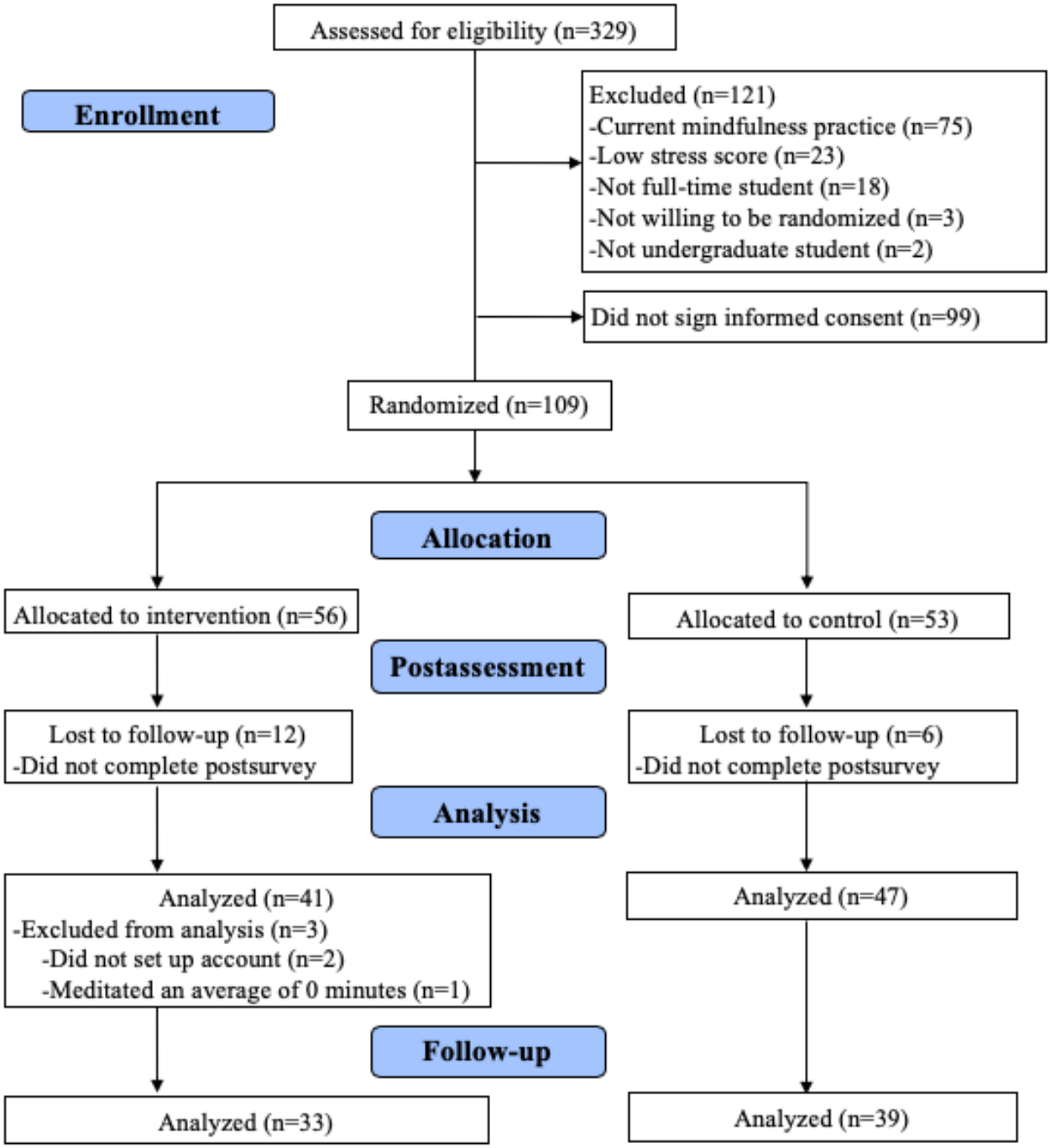
Feasibility potential was examined using adherence to the intervention. Adherence was calculated by averaging the weekly meditation minutes provided by the reports from the Calm informatics team. Acceptability measures and type of meditation were summarized using frequency with percentages. The Cohen *d* statistic was used to compute effect size. A two-sided α of 0.05 was used to determine statistical significance between groups, and a Bonferroni correction α of 0.016 was used for within-group comparison. All statistical procedures were performed using the IBM Statistical Package for Social Sciences (SPSS) version 24.0 (IBM Corp, Armonk, NY) and Statistical Analysis Systems Software, version 9.4 (SAS Institute, Cary, NC).

Results

Participant Enrollment

A total of 329 Arizona State University students completed the eligibility questionnaire. Of those who met the study eligibility, 37% (n=121) were ineligible and 30% (n=99) did not sign the informed consent. Subsequently, 56 participants were randomized to the intervention group and 53, to the wait-list group, yielding a total of 109 consented/randomized participants. We excluded participants who did not complete the postsurvey (n=18), set up a Calm account (n=2), or meditate at all (n=1), resulting in a total of 88 participants to be included in the analysis (Figure 1).

Figure 1. Enrollment flow diagram.



Participant Demographics

Baseline characteristics of the study participants are presented in Table 1. The mean age (SD), adjusted for gender and race, was 20.41 (2.31) years for the intervention group and 21.85 (6.3) years for the control group. The majority of participants were female (79/88, 88%), freshman (27/88, 31%), non-Hispanic (66/88, 77%), and of white race (48/88, 59%). There were no differences in demographic characteristics, mental diagnosis, medication use, and counseling activities between the intervention and control groups (using Chi-square tests, all $P>.17$).

Table 1. Baseline characteristics of the study participants (N=88).

Characteristics	Intervention (n=41)	Control (n=47)	*P* value
Age (years)[a], mean	20.41	21.85	.18
Gender, n (%)			.57
Male	5 (6)	4 (5)	
Female	36 (41)	43 (49)	
Class, n (%)			.61
Freshman	10 (11)	17 (19)	
Sophomore	12 (14)	10 (11)	
Junior	10 (11)	12 (14)	
Senior	9 (10)	8 (9)	
Ethnicity, n (%)			.17
Hispanic	13 (15)	7 (8)	
Non-Hispanic	27 (31)	39 (44)	
Prefer not to respond	1 (1)	1 (1)	
Race, n (%)			.54
White/Caucasian	25 (28)	23 (26)	
Asian/Asian American	6 (7)	9 (10)	
Black	1 (1)	4 (5)	
Biracial/multiracial	3 (3)	7 (8)	
Other	3 (3)	2 (2)	
Prefer not to respond	3 (3)	2 (2)	
Mental diagnosis, n (%)			.53
Yes	12 (14)	11 (13)	
No	29 (33)	36 (41)	
Medications, n (%)			.94
Yes	5 (6)	6 (7)	
No	36 (41)	41 (47)	
Counseling, n (%)			.20
Yes	2 (2)	6 (7)	
No	39 (44)	41 (47)	

[a]Adjusted for gender and race.

Primary Outcomes (Perceived Stress, Mindfulness, and Self-Compassion)

Table 2 describes the means of perceived stress, mindfulness, and self-compassion between the intervention and control groups at baseline, postintervention, and at follow-up. The baseline mean (SD) scores of self-reported stress (intervention: 23.11 [4.93] vs control: 21.88 [4.94]; *P*=.25), mindfulness (intervention: 109.41 [17.36] vs control: 113.88 [17.35]; *P*=.25), and self-compassion (intervention: 29.81 [8.13] vs control: 31.85 [8.2]; *P*=.25) showed no significant differences between study groups.

Table 3 summarizes the results of 8 weeks of meditation on changes in stress, mindfulness, and self-compassion, testing the initial efficacy of the intervention compared to wait-list control participants. Participants in the intervention group had a significant reduction in perceived stress compared with the control group (∇=–7.13; *P*<.001; effect size=1.24). Participants in the intervention group had significant improvements in total mindfulness (Δ=19.23; *P*<.001; effect size=1.11) and all five constructs: observe (Δ=3.735; *P*<.001; effect size=0.67), describe (Δ=3.528; *P*<.001, effect size=0.59), acting with awareness (Δ=4.737; *P*<.001; effect size=0.83), nonjudgment of inner experience (Δ=4.938; *P*<.001; effect size=0.76), nonreactivity to inner experience (Δ=3.781; *P*<.001; effect size=0.92). Participants in the intervention group also had significant improvements in self-compassion (Δ=8.223; *P*<.0001; effect size=0.84) compared to participants in the control group.

Table 2. Mean scores per study group for perceived stress, mindfulness, and self-compassion.

Variable	Baseline (n=88)				Postintervention (n=88)				Follow-up (n=71)			
	Intervention		Control		Intervention		Control		Intervention		Control	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
PSS[a]	41	23.11 (4.93)	47	21.88 (4.94)	41	16.15 (6.16)	47	20.02 (6.16)[b]	32	15.89 (6.71)	39	19.86 (6.70)[b]
FFMQ[c]	38	109.41 (17.36)	45	113.88 (17.35)	40	129.20 (18.32)	44	111.07 (18.31)[d]	32	132.50 (20.83)	39	114.82 (20.81)[d]
Observe	40	23.96 (5.83)	47	24.40 (5.84)	41	27.80 (5.60)	47	23.17 (5.60)[d]	32	28.07 (5.75)	39	24.94 (5.74)[e]
Describe	40	23.22 (6.08)	46	23.75 (6.08)	41	26.75 (5.85)	46	23.48 (5.90)[b]	32	28.11 (7.68)	39	24.03 (6.88)[b]
Act aware	39	21.62 (6.13)	47	21.80 (6.10)	40	26.42 (5.38)	46	21.64 (5.41)[d]	32	26.77 (5.66)	39	22.22 (5.71)[b]
Nonjudgment	39	20.95 (6.75)	46	24.02 (6.79)[e]	41	25.88 (6.21)	46	23.04 (6.28)[e]	32	27.29 (6.72)	39	23.15 (6.71)[b]
Nonreactivity	40	18.89 (4.39)	47	19.65 (4.39)	41	22.72 (3.97)	47	19.76 (3.95)[d]	32	22.25 (4.12)	39	20.49 (4.12)
SCS-SF[f]	39	29.81 (8.13)	46	31.85 (8.20)	40	37.06 (9.02)	47	31.91 (8.92)[d]	32	39.16 (8.97)	39	33.66 (8.96)[b]

[a]PSS: Perceived Stress Scale.

[b]*P*<.01.

[c]FFMQ: Five Facet Mindfulness Questionnaire.

[d]*P*<.0001

[e]*P*<.05.

[f]SCS-SF: Self-Compassion Scale - Short Form.

Table 3. Pre-post change scores in perceived stress, mindfulness, and self-compassion between groups (N=88; all *P*<.001). Scores are adjusted for age, gender, and race.

Variable	Change score	
	Intervention (n=41)	Control (n=47)
PSS[a]	–7.137	–2.013
FFMQ[b]	19.23	–2.408
Observe	3.735	–1.272
Describe	3.528	–0.257
Act aware	4.737	–0.260
Nonjudgment	4.938	–0.838
Nonreactivity	3.781	0.138
SCS-SF[c]	8.223	0.294

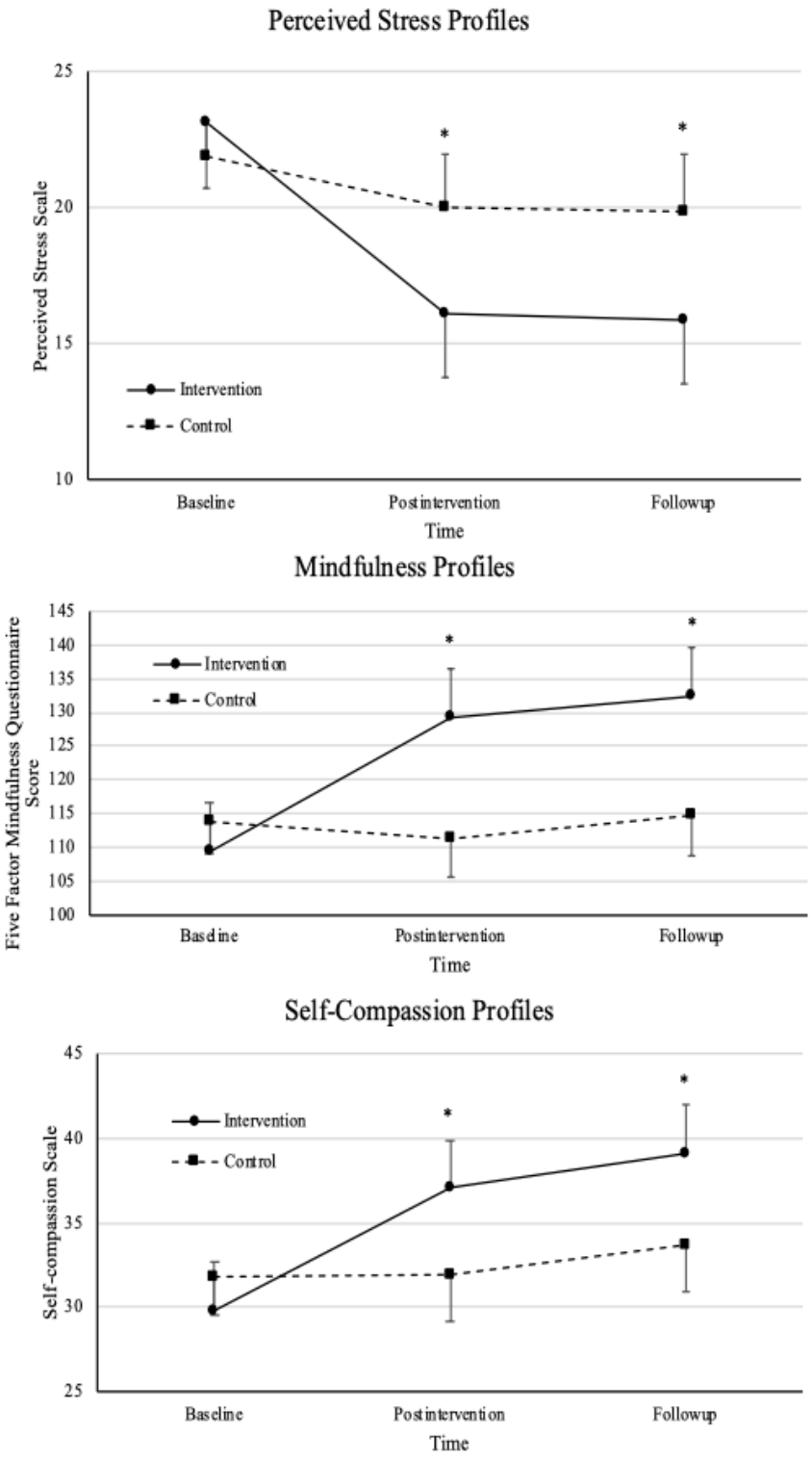
[a]PSS: Perceived Stress Scale.

[b]FFMQ: Five Facet Mindfulness Questionnaire.

[c]SCS-SF: Self-Compassion Scale - Short Form

Additional analyses using linear mixed models were used to examine effects of the intervention compared to control over time, including initial and sustained effects. Figure 2 illustrates the means of perceived stress, mindfulness, and self-compassion and the 95% CIs between intervention and control groups across baseline, postintervention, and at follow-up. On an average, participants in the intervention group had significantly decreased perceived stress scores postintervention (8 weeks) as compared to those in the control group, and these changes persisted through follow-up (12 weeks; Figure 2 and Table 2). Compared to those in the control group, participants in the intervention group showed significantly increased mindfulness and self-compassion scores postintervention (8 weeks), and these changes persisted through follow-up (12 weeks).

Figure 2. Effects of mindfulness meditation on perceived stress, mindfulness, and self-compassion.



There was a significant interaction between group and time factors in perceived stress (P=.002), mindfulness (P<.001), and self-compassion (P<.001). Bonferroni posthoc tests showed significant within-group mean differences for perceived stress in the intervention group (baseline and postintervention: P<.001; baseline and follow-up: P<.001), while there were no significant within-group mean differences for perceived stress in the control group (all P>.19). For mindfulness and self-compassion, there were significant within-group mean differences in the intervention group (baseline and postintervention: P<.001; baseline and follow-up: P<.001), while there were no significant within-group mean differences in the control group (all P>.40). In the intervention group, there were no statistical differences in perceived stress, mindfulness, and self-compassion between the postintervention phase and follow-up (all P>.05).

Secondary Outcomes (Sleep Quality, Binge Drinking, Physical Activity Participation, and Healthy Eating)

Linear models indicated that among participants using the mindfulness meditation app, there was a significant decrease in sleep disturbance, as indicated by *t* scores on PROMIS (*P*=.02; effect size=0.79), while there were no changes among control participants (*P*=.95; effect size=0.02). However, results from linear mixed models showed that the interaction between the time and intervention groups was nonsignificant, suggesting that sleep improvements did not differ across the two groups (*P*=.11; effect size=0.34).

Data from the three YRBS items were used to create dichotomous variables reflecting whether in the past 7 days, participants engaged in binge drinking (ie, consumed four or more [for women] or five or more [for men] alcoholic beverages in a row on at least one day), engaged in 150 minutes of physical activity, and ate five serving of fruits and vegetables per day on most days. Results from the McNemar tests indicate that there were no significant changes in binge drinking for students in either study group (all *P*>.51) for physical activity or eating fruits and vegetables.

Acceptability (Satisfaction, Enjoyment, Intent to Continue Use)

We assessed acceptability via a satisfaction survey postintervention (week 8) in the intervention group. Approximately, 51% (21/41) of the students indicated that Calm was helpful to very helpful in reducing their stress in the short term, and these findings were similar for the long term (20/41, 49%). Over half of the students said they would use Calm to help reduce stress in the future. Approximately 85% (35/41) of the sample was somewhat to very satisfied using Calm, and 85% (35/41) indicated that they somewhat to very much enjoyed using Calm. About 68% (28/41) of the sample said they were likely to extremely likely to use Calm in the future, and 76% (31/41) would be likely to extremely likely to recommend Calm to other college students.

Demand (Adherence)

Participants in the intervention group engaged in an average of 37.9 (SD 30.5) minutes of meditation per week over the course of the 8-week study. Over half of the participants (23/41, 56%) completed more than 30 minutes per week using the Calm app, with 22% (9/41) of participants meditating more than 60 minutes a week. Approximately one-third (14/41, 34%) of intervention participants continued to meditate during the follow-up period (12 weeks from baseline) and spent an average of 20.4 (SD 23.9) minutes meditating.

Discussion

Principal Findings

The purpose of this study was to test the efficacy and sustained effects of an 8-week mindfulness meditation intervention delivered using a consumer-based mobile app (ie, Calm) to reduce stress in college students and to examine feasibility components germane to planning the next steps in research. This was one of the first studies to test the efficacy of such a product in undergraduate college students. Our findings demonstrate significant between-group differences on all main outcomes variables including perceived stress, all five factors of mindfulness (observing, describing, acting with awareness, nonjudgment of inner experience, and nonreactivity to inner experience), and self-compassion postintervention. These effects were sustained at follow-up, and effect sizes ranged from moderate (0.59) to large (1.24) across all outcomes. Additionally, Calm was perceived as helpful; many participants enjoyed Calm and were likely to continue using the app. These findings highlight the promise of Calm to reduce stress in college students.

Perceived Stress

To our knowledge, this study was one of the first to demonstrate the stress-reducing effects of a mindfulness meditation mobile app (ie, Calm) in college students. The magnitude of change in perceived stress was greater in the intervention group than in the control group postintervention and at follow-up. Our findings are similar to those of a recent study that tested whether two mindfulness meditation apps (Headspace and Smiling Mind) led to improvements in mental health as compared to a control group in undergraduate students (N=208) [64]. Students were asked to meditate for 10 minutes daily for 40 days (10 days of requested use followed by 30 days of discretionary use). Students using Headspace (not Smiling Mind) reported significant reductions in stress and were maintained at follow-up. However, unlike our study, the app usage was measured by self-report, nearly half of all participants never used the app during the open-access period, students did not have elevated levels of stress as part of eligibility criteria, and the effect size was small (*d*=0.26). Another study testing 10-20 minutes/day of Headspace use for 30 days among medical students (N=88) reported similar findings on stress compared to a wait-list control group both postintervention (30 days) and at follow-up (60 days) [65]. App usage in this study was self-reported, and participants were also instructed to send a screenshot to the research team of the minutes meditated report in Headspace. However, this study only reported on the days of app usage rather than minutes meditated and did not report effect sizes. Reporting days of app usage does not reflect the total amount of minutes meditated and does not illustrate how minutes of meditation may vary across days. Our findings are also similar to those of a study testing a newly developed mindfulness app (Wildflowers) on stress compared to a cognitive training control group (app called 2048) in college students (N=86) [66]. Students were asked to use the Wildflowers app daily for 3 weeks and could choose lessons on meditation or choose from a library of guided meditations. Students using the Wildflowers app had improved mood and reduced stress following each training session. However, the study did not have adequate power, did not report on the number of minutes students had meditated (only sessions completed), and had small effect sizes (*r*=–0.08 to –0.37). Although all the aforementioned studies demonstrated a reduction in the number of minutes (or sessions) meditated, ours was the longest intervention (8 weeks) and showed the strongest effect sizes for perceived stress (effect size=1.24).

Although the average minutes of meditation in our study significantly decreased at follow-up, impressively, the reduction

in perceived stress remained significant at follow-up (12 weeks from baseline). Similar to findings of other studies, our findings suggest that brief mindfulness meditation interventions delivered via mobile apps show promise in reducing stress in college students, and there *may be* added benefit for weeks following the intervention period, regardless of the time spent in meditation. Delivering mindfulness meditation may be a more practical approach to reducing stress, as it requires fewer resources (eg, cost, staff, or brick and mortar building), students can participate remotely, and there are fewer time constraints. However, there is a need to independently test and determine the efficacy of mindfulness meditation using mobile apps.

Mindfulness

We also demonstrated significant improvements with large effects (effect size=1.11) in overall mindfulness scores and significant improvements with moderate-to-large effects (effect size=0.59-0.92) in all five factors of the FFMQ (observing, describing, acting with awareness, nonjudgment of inner experience, and nonreactivity to inner experience). Yamada and colleagues [67] implemented a brief 10-minute mindfulness meditation intervention (led by an instructor twice a week) over a 15-week semester in college students. Those participating in the mindfulness group, on an average, increased their mindfulness scores by nearly 9% as compared to the control group (no change). However, investigators did not provide the mindfulness factor scores. Greeson and colleagues [42] implemented a mindfulness meditation intervention (Koru) in college students and found significant increases in mindfulness and large effect sizes at the end of the 4-week intervention compared to a wait-list control group [42]. The Koru program includes 75-minute sessions taught in a small group format and also requires 10-minute daily meditation practice. Both the aforementioned studies, and our current study, demonstrate that brief mindfulness meditation interventions have the potential to improve mindfulness scores; however, the aforementioned studies did not include a follow-up period to assess the sustained effects of the intervention. In our study, mindfulness was greater even at the follow-up assessment. More studies are needed to test the sustained effects of brief mindfulness meditation interventions.

Self-Compassion

We found significant improvements in self-compassion in the intervention group as compared to the wait-list control group after the 8-week intervention with a large effect size (effect size=0.84). Improvements in self-compassion were sustained at follow-up. Greeson and colleagues (2014) also found significant increases with a large effect size in self-compassion, measured by the 26-item Self-Compassion Scale [68], as compared to a wait-list control group [42], but did not include a follow-up assessment. Another study implemented a brief mindfulness intervention (based on MBSR) that included 10-minute periods of mindfulness training followed by 5 minutes of discussion across 28 clinical interviewing classes (7 hours of instruction over 10 weeks) in social work students [69]. In contrast to our findings, this study did not find significant improvements in self-compassion. However, this study was limited by a lack of randomization and a small sample size. Despite these findings, our study and others demonstrate that self-compassion often improves with increasing mindfulness in students [29,36,70], but the sustained effects after the intervention period are unclear. Self-compassion has been considered a core quality of mindfulness [71] and is associated with lower stress and psychological symptoms in addition to greater well-being [72,73]. In one study, mindfulness and self-compassion were mediators in the pathway to emotional well-being in adolescents [36]. More research investigating the impact of brief mindfulness meditation interventions, especially those delivered via mobile apps, on self-compassion is needed to determine the efficacy of such apps and potential to sustain effects after the intervention period.

Sleep Quality and Health Behaviors

Our findings suggest that Calm may be a promising approach to help improve sleep quality (ie, reduce sleep disturbance) in college students. Similar to our findings, Greeson and colleagues [42] reported a significant reduction in sleep problems as compared to a wait-list control group after participation in a 4-week mindfulness meditation intervention. However, the intervention was only for 4 weeks and delivered in person. Another study tested the impact of an 8-week internet-based mindfulness training program to improve mental health and sleep in college students and young adults compared to an internet-based cognitive behavioral training program [74]. Each session lasted 30-45 minutes and both groups experienced significant reductions in sleep disturbance. This study also did not deliver the intervention with a mobile app and had a high attrition rate. One study tested a mindfulness-based health app on mental health and sleep behavior in college students [75]. Students were asked to use the app 5 days per week for 4 weeks, but there were no significant changes in sleep. There are limited and mixed data available on the impact of mobile meditation apps and sleep outcomes in college students, and there is a need for further exploration of meditation delivered via an app and sleep outcomes in college students.

We did not find an association between mindfulness meditation and changes in alcohol consumption (ie, binge drinking), physical activity, or healthy eating (ie, fruit and vegetable consumption). There are a lack of studies testing the impact of mindfulness meditation mobile apps on health behavior outcomes in college students. One study testing the effects of a mindfulness app on weight and weight-related behaviors reported reductions in emotional eating and uncontrolled eating after engagement in an 11-week intervention compared to a self-monitoring electronic diary control group [76]. However, this study had low adherence and did not report on fruit and vegetable consumption. A cross-sectional survey administered to college women reported that higher levels of mindfulness were related to healthy eating practices, better sleep quality, and better physical health [77]. Another cross-sectional survey implemented in adults aged 18-25 years suggested that self-compassion (a potential mechanism of mindfulness) was positively correlated with intentions to engage in health-promoting behaviors (eg, healthy eating, physical activity, good sleep hygiene, and stress management) [78]. Papies (2017) suggests that mindfulness interventions may be most effective if targeted at a specific behavior (ie, reducing alcohol

consumption, healthy eating, and reducing smoking) [79]. Although mindfulness-based interventions show promise in improving health behaviors, more research in this area is warranted.

Feasibility

Approximately 85% of the sample was satisfied with and found Calm enjoyable. None of the students in our sample indicated that they were dissatisfied with using Calm. A majority said they would use the app in the future or refer it to other college students. There is limited evidence regarding the acceptability of mindfulness meditation mobile apps, specifically in college students. Similar to our data, a study by Donovan and colleagues [80] suggested high satisfaction with a mobile app (BodiMojo) aiming to teach adolescents (13-22 years old) about mindfulness and self-compassion, with 92% of the sample rating their enjoyment of the program a score of 6 or 7 (1=not at all, 7=very much) [80]. Another study conducted in the Netherlands by Emmerik and colleagues [81] implemented a mindfulness-based mobile app (VGZ Mindfulness Coach) to improve mindfulness, general psychiatric symptoms, and quality of life in adults aged ≥18 years [81]. Participants in the study reported high satisfaction with an average satisfaction score of 4.18 (SD 0.84) on a scale of 1-5 (higher scores indicate higher satisfaction) [81]. The aforementioned research is similar to our findings and demonstrates high satisfaction among mindfulness-based mobile apps. However, it is difficult to compare the mobile apps against each other because of the various mindfulness components used, differences in application interfaces, usability, and length of programs. Future studies may assess the satisfaction of the various components of behavioral/mental health mobile apps, as studies have reported that satisfaction of these apps may be dependent upon components such as accountability, tracking progress, reminders/push notifications, and convenience [82,83]. This information could be particularly useful for researchers to design/modify new or existing mindfulness meditation mobile apps to be more relatable to the user and potentially attenuate the decline in usage over time.

The average participation in meditation across the 8-week study was 38 minutes/week, which was sufficient to see effects in perceived stress in our sample. The decrease in minutes of meditation practiced by college students over the length of the intervention is not surprising, as college students have a number of barriers that may keep them from participating in daily meditation (eg, time constraints, financial pressure, educational expectations, and interpersonal relationships). In a formative analysis by Lauricella [84], a majority of undergraduate students who participated in a face-to-face mindfulness meditation intervention followed by a digital session of the same technique preferred the in-person version as compared to the digital practice [84]. The minority (~25%) that preferred the digital practice found the solitude and confidentiality of this format to be favorable. It is important to understand more about why college students do or do not participate in meditation delivered digitally. Gender differences regarding use of mindfulness-based mobile apps may be important considerations in future studies. Men are often underrepresented in mindfulness interventions, and women may be more engaged in the intervention [85] and experience greater benefit [86]. We also provided students with incentives, which may have impacted intervention results and adherence. Future studies may consider replicating this study without the use of incentives to observe real-world participation. Furthermore, many consumer-based mobile apps do not include theory-based strategies for behavior change, which may impact adherence rates [87]. For example, self-monitoring (ie, the ability to track minutes) and social support are strategies highly associated with adherence to behavior. Calm is a consumer-based app that does include tracking and reminders for participants, but other strategies may be necessary. Future studies could incorporate social support with an app and examine the effects of such an addition on adherence.

It is also important to note that mobile app usage rapidly declines within the first 90 days of downloading the app [88]. According to a 2018 report by Localytics, 21% of users abandon using an app after only one use, 57% of users stop using the app within the first 4 weeks, and average retention by 90 days is only 29% across all industries (media and entertainment, electronic commerce/retail, travel and lifestyle, business, and technology) [88,89]. Several studies that have investigated barriers to health-related mobile app usage have identified a lack of regulatory oversight, limited evidence-based literature, privacy/security concerns [87], poor functionality or incompatibility, lack of desired features, abandonment of a health goal [90], loss of interest, hidden costs, and data entry burden [91]. Mobile apps have the potential to provide affordable, convenient access, and evidence-based information to a large and broad audience. However, as noted previously, there are several limitations that need to be addressed to improve the quality and efficacy of mobile apps.

Limitations

Although the findings in this study are promising regarding the efficacy of a mindfulness-based mobile app to reduce stress and improve mindfulness and self-compassion among college students, with moderate-to-large effect sizes, several limitations must be noted. First, the sample mostly comprised white, young female adults; therefore, the generalizability of our findings may be limited. In future interventions, it is important to explore efficacy in more diverse samples (eg, race/ethnicity, education level, and chronic disease status) [44,54,92]. At the time of the intervention, Calm was only offered in English. Exploration of the efficacy of Calm with non-English speaking populations (and offered in other languages) is warranted. Second, all data collected in the study relied on self-report, which may lead to social desirability or response bias where participants respond in what they believe is a more favorable manner (although the wait-list control condition assists partially in controlling for this bias). We had a wait-list control group and were not able to control for maturation or expectation effects. Future studies warrant determining the best control condition. In our study, we tracked the minutes of meditation and not the frequency of sessions. Other studies have tracked both but only conducted analysis with one or the other [65,66]. Future studies should consider measuring and reporting both minutes and frequency (eg, days/sessions) related to participation. Our estimated sample size of 104 was not achieved (88 included in data analysis), but we did have adequate power because our effect sizes were strong. However, analyses of all secondary outcomes were

exploratory and not sufficiently powered to detect an effect, and results for these outcomes should be interpreted with caution. Finally, we did not collect information regarding students' school schedule, homework load, or testing schedule, and therefore, it is unclear how these factors may have played a role in students' participation or stress scores.

Conclusions

The findings in this study demonstrate the efficacy of Calm to reduce stress and improve mindfulness and self-compassion in short-term contexts in stressed college students. Although direct comparisons are not possible, it appears that the degree of improvement in response to a smartphone-based mindfulness app may be similar to programs that require in-person attendance with even greater opportunity for convenient compliance and continued use. Students also reported high satisfaction using Calm to reduce stress. Currently, there is limited literature exploring the effectiveness and acceptability of mobile apps for delivering mindfulness meditation. More research in this area is needed to establish efficacy and to explore the degree to which effects are sustained in the short and long term. Particularly, we suggest future independent replications of this work to establish efficacy. The findings presented here provide important information that can be applied to the design of future studies or mental health resources in university programs. Calm may be a cost-effective, convenient, easily disseminated, and enjoyable way to manage stress among college students.

Acknowledgments

This study was supported by Arizona State University Counseling and Health Services.

Conflicts of Interest

JH is currently the Director of Science at Calm (although she was engaged in this role almost 1 year after the design, collection of data, and analysis of results of the study presented in this paper). None of the authors declare any conflicts of interest.

Multimedia Appendix 1
CONSORT - EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 3736 KB - mhealth_v7i6e14273_app1.pdf]

References

1. Pierceall EA, Keim MC. Stress and Coping Strategies Among Community College Students. Community College Journal of Research and Practice 2007 Sep 06;31(9):703-712. [doi: 10.1080/10668920600866579]
2. Almojali AI, Almalki SA, Alothman AS, Masuadi EM, Alaqeel MK. The prevalence and association of stress with sleep quality among medical students. J Epidemiol Glob Health 2017 Dec;7(3):169-174 [FREE Full text] [doi: 10.1016/j.jegh.2017.04.005] [Medline: 28756825]
3. American College Health Association. Hanover, MD American College Health Association-National College Health Assessment II: Spring 2018 Reference Group Executive Summary URL: https://www.acha.org/documents/ncha/NCHA-II_Spring_2018_Reference_Group_Executive_Summary.pdf [accessed 2019-06-18]
4. American Psychological Association. 2012. Stress by Generation URL: https://www.apa.org/news/press/releases/stress/2012/generations [accessed 2019-06-18] [WebCite Cache ID 77OErQthM]
5. Neely ME, Schallert DL, Mohammed SS, Roberts RM, Chen Y. Self-kindness when facing stress: The role of self-compassion, goal regulation, and support in college students' well-being. Motiv Emot 2009 Jan 6;33(1):88-97. [doi: 10.1007/s11031-008-9119-8]
6. Anastasiades MH, Kapoor S, Wootten J, Lamis DA. Perceived stress, depressive symptoms, and suicidal ideation in undergraduate women with varying levels of mindfulness. Arch Womens Ment Health 2017 Dec;20(1):129-138. [doi: 10.1007/s00737-016-0686-5] [Medline: 27778150]
7. Newman B, Newman P. Development Through Life: A Psychosocial Approach. Stamford, CT: Wadsworth Publishing; 2019.
8. McIndoo CC, File AA, Preddy T, Clark CG, Hopko DR. Mindfulness-based therapy and behavioral activation: A randomized controlled trial with depressed college students. Behav Res Ther 2016 Feb;77:118-128. [doi: 10.1016/j.brat.2015.12.012] [Medline: 26745622]
9. Roberts KC, Danoff-Burg S. Mindfulness and health behaviors: is paying attention good for you? J Am Coll Health 2010;59(3):165-173. [doi: 10.1080/07448481.2010.484452] [Medline: 21186446]
10. O'Donovan A, Hughes B. Social support and loneliness in college students: effects on pulse pressure reactivity to acute stress. Int J Adolesc Med Health 2007;19(4):523-528. [Medline: 18348427]
11. Liu CH, Stevens C, Wong SHM, Yasui M, Chen JA. The prevalence and predictors of mental health diagnoses and suicide among U.S. college students: Implications for addressing disparities in service use. Depress Anxiety 2019 Dec;36(1):8-17. [doi: 10.1002/da.22830] [Medline: 30188598]

12. Centers for Disease Control Prevention. 2016. Leading causes of death by age group, United States - 2016 URL: https://www.cdc.gov/injury/wisqars/pdf/leading_causes_of_death_by_age_group_2016-508.pdf [accessed 2019-06-18] [WebCite Cache ID 77OFQIpzT]
13. American College Health Association. American College Health Association-National College Health Assessment II: Fall 2010 Reference Group Executive Summary URL: https://www.acha.org/documents/ncha/ACHA-NCHA-II_ReferenceGroup_ExecutiveSummary_Fall2010.pdf [accessed 2019-06-18]
14. Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu S, et al. Mental health of college students and their non-college-attending peers: results from the National Epidemiologic Study on Alcohol and Related Conditions. Arch Gen Psychiatry 2008 Dec;65(12):1429-1437 [FREE Full text] [doi: 10.1001/archpsyc.65.12.1429] [Medline: 19047530]
15. Eisenberg D, Hunt J, Speer N, Zivin K. Mental health service utilization among college students in the United States. J Nerv Ment Dis 2011 May;199(5):301-308. [doi: 10.1097/NMD.0b013e3182175123] [Medline: 21543948]
16. Eisenberg D, Downs MF, Golberstein E, Zivin K. Stigma and help seeking for mental health among college students. Med Care Res Rev 2009 Oct;66(5):522-541. [doi: 10.1177/1077558709335173] [Medline: 19454625]
17. Komiya N, Good GE, Sherrod NB. Emotional openness as a predictor of college students' attitudes toward seeking psychological help. Journal of Counseling Psychology 2000;47(1):138-143. [doi: 10.1037/0022-0167.47.1.138]
18. Givens JL, Tjia J. Depressed medical students' use of mental health services and barriers to use. Acad Med 2002 Sep;77(9):918-921. [Medline: 12228091]
19. Megivern D, Pellerito S, Mowbray C. Barriers to higher education for individuals with psychiatric disabilities. Psychiatr Rehabil J 2003;26(3):217-231. [Medline: 12653444]
20. Mowbray CT, Megivern D, Mandiberg JM, Strauss S, Stein CH, Collins K, et al. Campus mental health services: recommendations for change. Am J Orthopsychiatry 2006 Apr;76(2):226-237. [doi: 10.1037/0002-9432.76.2.226] [Medline: 16719642]
21. Tjia J, Givens JL, Shea JA. Factors associated with undertreatment of medical student depression. J Am Coll Health 2005;53(5):219-224. [doi: 10.3200/JACH.53.5.219-224] [Medline: 15813232]
22. Nelson MC, Story M, Larson NI, Neumark-Sztainer D, Lytle LA. Emerging adulthood and college-aged youth: an overlooked age for weight-related behavior change. Obesity (Silver Spring) 2008 Oct;16(10):2205-2211. [doi: 10.1038/oby.2008.365] [Medline: 18719665]
23. Leppink EW, Odlaug BL, Lust K, Christenson G, Grant JE. The Young and the Stressed: Stress, Impulse Control, and Health in College Students. J Nerv Ment Dis 2016 Dec;204(12):931-938. [doi: 10.1097/NMD.0000000000000586] [Medline: 27575792]
24. Hudd S, Dumlao J, Erdmann-Sager D. Stress at college: Effects on health habits, health status and self-esteem. College Student Journal 2000;34(2).
25. Weidner G, Kohlmann C, Dotzauer E, Burns LR. The effects of academic stress on health behaviors in young adults. Anxiety, Stress & Coping 1996 Jan;9(2):123-133. [doi: 10.1080/10615809608249396] [Medline: 26281194]
26. Kitzrow M. The mental health needs of today's college students: Challenges and recommendations. NASPA Journal 2009;46(4):646-660. [doi: 10.2202/1949-6605.1310]
27. Chiauzzi E, Brevard J, Thum C, Thurn C, Decembrele S, Lord S. MyStudentBody-Stress: an online stress management intervention for college students. J Health Commun 2008 Sep;13(6):555-572. [doi: 10.1080/10810730802281668] [Medline: 18726812]
28. Regehr C, Glancy D, Pitts A. Interventions to reduce stress in university students: a review and meta-analysis. J Affect Disord 2013 May 15;148(1):1-11. [doi: 10.1016/j.jad.2012.11.026] [Medline: 23246209]
29. Bamber MD, Kraenzle Schneider J. Mindfulness-based meditation to decrease stress and anxiety in college students: A narrative synthesis of the research. Educational Research Review 2016 May;18:1-32. [doi: 10.1016/j.edurev.2015.12.004]
30. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [Medline: 12703651]
31. Kabat-Zinn J. Mindfulness-based stress reduction (MBSR). Constructivism in the Human Sciences 2003;8(2):73.
32. Kazantzis N, Freeman A, Reinecke M, Dattilio F. Mindfulness-based cognitive therapy. In: Cognitive and Behavioral Theories in Clinical Practice. New York: Guilford Publications, Inc; 2009:307-331.
33. Eberth J, Sedlmeier P. The Effects of Mindfulness Meditation: A Meta-Analysis. Mindfulness 2012 May 2;3(3):174-189. [doi: 10.1007/s12671-012-0101-x]
34. Duarte J, Pinto-Gouveia J. Mindfulness, self-compassion and psychological inflexibility mediate the effects of a mindfulness-based intervention in a sample of oncology nurses. Journal of Contextual Behavioral Science 2017 Apr;6(2):125-133. [doi: 10.1016/j.jcbs.2017.03.002]
35. Baer RA, Lykins EL, Peters JR. Mindfulness and self-compassion as predictors of psychological wellbeing in long-term meditators and matched nonmeditators. The Journal of Positive Psychology 2012 May;7(3):230-238. [doi: 10.1080/17439760.2012.674548]
36. Bluth K, Blanton P. Mindfulness and Self-Compassion: Exploring Pathways to Adolescent Emotional Well-Being. J Child Fam Stud 2014 Oct 01;23(7):1298-1309 [FREE Full text] [doi: 10.1007/s10826-013-9830-2] [Medline: 25328369]

37. MacBeth A, Gumley A. Exploring compassion: a meta-analysis of the association between self-compassion and psychopathology. Clin Psychol Rev 2012 Aug;32(6):545-552. [doi: 10.1016/j.cpr.2012.06.003] [Medline: 22796446]
38. Zessin U, Dickhäuser O, Garbade S. The Relationship Between Self-Compassion and Well-Being: A Meta-Analysis. Appl Psychol Health Well Being 2015 Nov;7(3):340-364. [doi: 10.1111/aphw.12051] [Medline: 26311196]
39. Lynch S, Gander M, Kohls N, Kudielka B, Walach H. Mindfulness-based Coping with University Life: A Non-randomized Wait-list-controlled Pilot Evaluation. Stress and Health 2011 Mar 21;27(5):365-375. [doi: 10.1002/smi.1382]
40. Van Gordon W, Shonin E, Sumich A, Sundin EC, Griffiths MD. Meditation Awareness Training (MAT) for Psychological Well-Being in a Sub-Clinical Sample of University Students: A Controlled Pilot Study. Mindfulness 2013 Feb 7. [doi: 10.1007/s12671-012-0191-5]
41. Hindman R. The Catholic University of America. A comparison of mindfulness-based programs for stress in university students URL: https://cuislandora.wrlc.org/islandora/object/etd%3A400 [accessed 2019-06-18]
42. Greeson JM, Juberg MK, Maytan M, James K, Rogers H. A randomized controlled trial of Koru: a mindfulness program for college students and other emerging adults. J Am Coll Health 2014;62(4):222-233. [doi: 10.1080/07448481.2014.887571] [Medline: 24499130]
43. Hoffmann GA. Marquette University. Mindfulness meditation for college students: A study of its utility and promotion of its practice post treatment URL: https://epublications.marquette.edu/dissertations/AAI3201922 [accessed 2019-06-18]
44. Lever Taylor B, Strauss C, Cavanagh K, Jones F. The effectiveness of self-help mindfulness-based cognitive therapy in a student sample: a randomised controlled trial. Behav Res Ther 2014 Dec;63:63-69. [doi: 10.1016/j.brat.2014.09.007] [Medline: 25302763]
45. Pew Research Center: Internet & Technology. 2017. Internet use by age URL: https://www.pewinternet.org/chart/internet-use-by-age/ [accessed 2019-06-18] [WebCite Cache ID 77OFz1jRI]
46. Emanuel R. The American college student cell phone survey. College Student Journal 2013;47(1):75-82.
47. Lungu A, Sun M. Time for a Change: College Students' Preference for Technology-Mediated Versus Face-to-Face Help for Emotional Distress. Telemed J E Health 2016 Dec;22(12):991-1000. [doi: 10.1089/tmj.2015.0214] [Medline: 27223266]
48. Vital WC. UN Foundation-Vodafone Foundation Partnership. 2009. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World URL: http://www.globalproblems-globalsolutions-files.org/unf_website/assets/publications/technology/mhealth/mHealth_for_Development_full.pdf [accessed 2019-06-18]
49. Prentice J, Dobson K. A review of the risks and benefits associated with mobile phone applications for psychological interventions. Canadian Psychology/Psychologie canadienne 2014 Nov;55(4):282-290. [doi: 10.1037/a0038113]
50. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
51. Carter A, Liddle J, Hall W, Chenery H. Mobile Phones in Research and Treatment: Ethical Guidelines and Future Directions. JMIR Mhealth Uhealth 2015;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
52. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. JMIR Mhealth Uhealth 2015;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
53. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
54. Warnecke E, Quinn S, Ogden K, Towle N, Nelson MR. A randomised controlled trial of the effects of mindfulness practice on medical student stress levels. Med Educ 2011 Apr;45(4):381-388. [doi: 10.1111/j.1365-2923.2010.03877.x] [Medline: 21401686]
55. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]
56. Gethin R. The Foundations Of Buddhism (opus). New York: Oxford University Press; 1998.
57. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
58. Spacapan S, Oskamp S. The Social Psychology Of Health: The Claremont Symposium On Applied Social Psychology. New York: Sage Publications, Inc; 1989.
59. Roberti J, Harrington L, Storch E. Further psychometric support for the 10 item version of the perceived stress scale. Journal of College Counseling 2006;9(2):135-147. [doi: 10.1002/j.2161-1882.2006.tb00100.x]
60. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
61. Raes F, Pommier E, Neff KD, Van Gucht D. Construction and factorial validation of a short form of the Self-Compassion Scale. Clin Psychol Psychother 2011 Jun;18(3):250-255. [doi: 10.1002/cpp.702] [Medline: 21584907]
62. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]

63. Kolbe L, Kann L, Collins J. Overview of the Youth Risk Behavior Surveillance System. Public Health Rep 1993;108 Suppl 1:2-10 [FREE Full text] [Medline: 8210269]
64. 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 2018 Oct 31;10(5):863-876. [doi: 10.1007/s12671-018-1050-9]
65. Yang E, Schamber E, Meyer R, Gold JI. Happier Healers: Randomized Controlled Trial of Mobile Mindfulness for Stress Management. J Altern Complement Med 2018 May;24(5):505-513. [doi: 10.1089/acm.2015.0301] [Medline: 29420050]
66. Walsh K, Saab B, Farb N. Effects of a Mindfulness Meditation App on Subjective Well-Being: Active Randomized Controlled Trial and Experience Sampling Study. JMIR Ment Health 2019 Jan 08;6(1):e10844 [FREE Full text] [doi: 10.2196/10844] [Medline: 30622094]
67. Yamada K, Victor TL. The Impact of Mindful Awareness Practices on College Student Health, Well-Being, and Capacity for Learning: A Pilot Study. Psychology Learning & Teaching 2012 Jan;11(2):139-145. [doi: 10.2304/plat.2012.11.2.139]
68. NEFF KD. The Development and Validation of a Scale to Measure Self-Compassion. Self and Identity 2003 Jul;2(3):223-250. [doi: 10.1080/15298860309027]
69. Gockel A, Burton D, James S, Bryer E. Introducing Mindfulness as a Self-Care and Clinical Training Strategy for Beginning Social Work Students. Mindfulness 2012 Aug 3;4(4):343-353. [doi: 10.1007/s12671-012-0134-1]
70. Birnie K, Speca M, Carlson LE. Exploring self-compassion and empathy in the context of mindfulness-based stress reduction (MBSR). Stress and Health 2010 Nov 29;26(5):359-371. [doi: 10.1002/smi.1305]
71. Snyder C, Lopez S. Meditation and positive psychology. In: Handbook Of Positive Psychology. New York: Oxford University Press; 2019:632-645.
72. Neff KD, Rude SS, Kirkpatrick KL. An examination of self-compassion in relation to positive psychological functioning and personality traits. Journal of Research in Personality 2007 Aug;41(4):908-916. [doi: 10.1016/j.jrp.2006.08.002]
73. Fredrickson BL, Cohn MA, Coffey KA, Pek J, Finkel SM. Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. J Pers Soc Psychol 2008 Nov;95(5):1045-1062 [FREE Full text] [doi: 10.1037/a0013262] [Medline: 18954193]
74. Mak W, Chio F, Chan A, Lui WW, Wu EK. The Efficacy of Internet-Based Mindfulness Training and Cognitive-Behavioral Training With Telephone Support in the Enhancement of Mental Health Among College Students and Young Working Adults: Randomized Controlled Trial. J Med Internet Res 2017 Dec 22;19(3):e84 [FREE Full text] [doi: 10.2196/jmir.6737] [Medline: 28330831]
75. Lee RA, Jung ME. Evaluation of an mHealth App (DeStressify) on University Students' Mental Health: Pilot Trial. JMIR Ment Health 2018 Jan 23;5(1):e2 [FREE Full text] [doi: 10.2196/mental.8324] [Medline: 29362209]
76. Lyzwinski L, Caffery L, Bambling M, Edirippulige S. The Mindfulness App Trial for Weight, Weight-Related Behaviors, and Stress in University Students: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Apr 10;7(4):e12210 [FREE Full text] [doi: 10.2196/12210] [Medline: 30969174]
77. Murphy M, Mermelstein L, Edwards K, Gidycz CA. The benefits of dispositional mindfulness in physical health: a longitudinal study of female college students. J Am Coll Health 2012;60(5):341-348. [doi: 10.1080/07448481.2011.629260] [Medline: 22686356]
78. Sirois FM. A self-regulation resource model of self-compassion and health behavior intentions in emerging adults. Prev Med Rep 2015;2:218-222 [FREE Full text] [doi: 10.1016/j.pmedr.2015.03.006] [Medline: 26844074]
79. Karremans J, Papies E. Mindfulness and health behaviour: Examining the roles of attention regulation and decentering. In: Mindfulness In Social Psychology (current Issues In Social Psychology). New York: Routledge; 2019:102-116.
80. Donovan E, Rodgers R, Cousineau T, McGowan KM, Luk S, Yates K, et al. Brief report: Feasibility of a mindfulness and self-compassion based mobile intervention for adolescents. J Adolesc 2016 Dec;53:217-221. [doi: 10.1016/j.adolescence.2016.09.009] [Medline: 27814499]
81. van Emmerik AAP, Berings F, Lancee J. Efficacy of a Mindfulness-Based Mobile Application: a Randomized Waiting-List Controlled Trial. Mindfulness (N Y) 2018;9(1):187-198 [FREE Full text] [doi: 10.1007/s12671-017-0761-7] [Medline: 29387266]
82. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
83. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
84. Lauricella S. Mindfulness Meditation with Undergraduates in Face-to-Face and Digital Practice: a Formative Analysis. Mindfulness 2013 May 21;5(6):682-688. [doi: 10.1007/s12671-013-0222-x]
85. Bluth K, Roberson PNE, Girdler SS. Adolescent Sex Differences in Response to a Mindfulness Intervention: A Call for Research. J Child Fam Stud 2017 Jul;26(7):1900-1914 [FREE Full text] [doi: 10.1007/s10826-017-0696-6] [Medline: 29051700]

86. de Vibe M, Solhaug I, Tyssen R, Friborg O, Rosenvinge JH, Sørlie T, et al. Mindfulness training for stress management: a randomised controlled study of medical and psychology students. BMC Med Educ 2013 Aug 13;13:107 [FREE Full text] [doi: 10.1186/1472-6920-13-107] [Medline: 23941053]
87. Kao C, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
88. Pero J. Localytics. Mobile apps: What's a good retention rate? URL: http://info.localytics.com/blog/mobile-apps-whats-a-good-retention-rate [accessed 2019-06-18] [WebCite Cache ID 77OGR3QjO]
89. Rodde T. Localytics. 2018. 21% of users abandon an app after one use URL: http://info.localytics.com/blog/21-percent-of-users-abandon-apps-after-one-use [accessed 2019-06-18] [WebCite Cache ID 77OGZKpV2]
90. Murnane E, Huffaker D, Kossinets G. Mobile health apps: adoption, adherence, and abandonment. In: UbiComp/ISWC'15 Adjunct. New York: ACM; 2015 Presented at: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers; September 7-11, 2015; Osaka p. 261-264 URL: http://stanford.edu/~emurnane/files/UbiComp15_Mobile.pdf [doi: 10.1145/2800835.2800943]
91. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
92. Deckro G, Ballinger K, Hoyt M, Wilcher M, Dusek J, Myers P, et al. The evaluation of a mind/body intervention to reduce psychological distress and perceived stress in college students. J Am Coll Health 2002 May;50(6):281-287. [doi: 10.1080/07448480209603446] [Medline: 12701653]

Abbreviations

CBT: cognitive behavioral therapy
FFMQ: Five-Factor Mindfulness Questionnaire
GLM: general linear model
MBCT: Mindfulness-Based Cognitive Therapy
MBSR: Mindfulness-Based Stress Reduction
PROMIS: Patient-Reported Outcomes Measurement Information System
PSS: Perceived Stress Scale
SCS-SF: Self-compassion Scale Short-Form
SPSS: Statistical Package for Social Sciences
YRBS: Youth Risk Behavior Surveillance

Edited by G Eysenbach; submitted 04.04.19; peer-reviewed by P Giacobbi, Jr., Z Reis, J Flett; comments to author 25.04.19; revised version received 03.05.19; accepted 09.06.19; published 25.06.19.

Please cite as:
Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C
Efficacy of the Mindfulness Meditation Mobile App "Calm" to Reduce Stress Among College Students: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(6):e14273
URL: http://mhealth.jmir.org/2019/6/e14273/
doi:10.2196/14273
PMID:31237569

Original Paper

Evaluation of the Effectiveness of a Musical Cognitive Restructuring App for Black Inner-City Girls: Survey, Usage, and Focus Group Evaluation

Angela Neal-Barnett[1], PhD; Robert Stadulis[1], EdD; Delilah Ellzey[1], MA; Elizabeth Jean[1], MA; Tiffany Rowell[1], BA; Keaton Somerville[1], MPS; Kallie Petitti[1], EdS; Benjamin Siglow[1], BA; Arden Ruttan[2], PhD; Mary Hogue[3], PhD

[1]Kent State University, Department of Psychological Sciences, Program for Research on Anxiety Disorders among African Americans, Kent, OH, United States

[2]Kent State University, Department of Computer Science, Kent, OH, United States

[3]Kent State University, College of Business, Kent, OH, United States

Corresponding Author:
Angela Neal-Barnett, PhD
Kent State University
Department of Psychological Sciences
Program for Research on Anxiety Disorders among African Americans
600 Hilltop Dr
Kent, OH, 44242-0001
United States
Phone: 1 330 672 2166
Email: aneal@kent.edu

Abstract

Background: Research on mobile health (mHealth) app use during adolescence is growing; however, little attention has been paid to black adolescents, particularly black girls, who are generally underresearched and underserved in psychological intervention research. Cognitive restructuring is an important tool in anxiety and fear management and involves two parts: (1) recognizing and deconstructing erroneous thoughts and (2) replacing negative anxiety and stress-provoking thoughts with positive thoughts. In our work with black adolescent females, we found that cognitive restructuring is a difficult skill to practice on one's own. Thus, drawing upon the importance of music in the black community, we developed the Build Your Own Theme Song (BYOTS) app to deliver a musical form of the technique to middle-school black girls.

Objective: Our aim in this mixed methods study is to evaluate the effectiveness of the BYOTS app. We hypothesize that participants will expect the app to be effective in reducing negative thoughts and that the app will meet their expectations and data generated from the app will demonstrate a reduction in negative thinking and anxiety.

Methods: A total of 72 black or biracial seventh- and eighth-grade adolescent females were enrolled in Sisters United Now (SUN), an eight-session culturally infused and app-augmented stress and anxiety sister circle intervention. Before using the BYOTS app, girls completed the Multidimensional Anxiety Scale for Children 2 and the App Expectations Survey. Usage data collected from the app included an assessment of negative thinking before and after listening to their song. After completion of the intervention, focus groups were held to gather qualitative data on participants' app experience.

Results: Results using paired sample *t* tests indicated negative thinking was significantly lower at day 7 than day 1 (t_{31}=1.69, *P*=.05). Anxiety from preuse to postuse of the app was also reduced (t_{38}=2.82, *P*=.004). Four effectiveness themes emerged from the focus groups: difference in behavior and temperament, promoted calmness, helpfulness in stressful home situations, and focused thinking via the SUN theme song.

Conclusions: The BYOTS app is a useful tool for delivering musical cognitive restructuring to reduce negative thinking and anxiety in an underserved urban population. Changes were supported both quantitatively and qualitatively. Participants, their peers, and their family noted the difference. Findings support expanding the research to black girls of various socioeconomic statuses and geographic diversity. Currently, the app augments SUN, a culturally relevant intervention. Future research will explore BYOTS as a stand-alone app.

(JMIR Mhealth Uhealth 2019;7(6):e11310) doi:10.2196/11310

KEYWORDS

black girls; musical cognitive restructuring; mHealth; negative thinking

Introduction

Background

Anxiety is a common problem among adolescents and has an adverse impact on their lives. Anxiety symptoms are associated with poor academic performance [1,2], difficulty with social skills and peer relationships [3,4], and behavior difficulties [5]. Several studies have suggested racial differences in symptoms and intensity of anxiety in youth. Compton and colleagues [6] found that black youth reported significantly more harm avoidance and physical avoidance than their white peers. Latzman and colleagues' [7] study of depression and anxiety across three cohorts of youth—elementary, middle school, and high school—found higher anxiety symptoms in middle school.

Anxiety and Black Females

Anxiety symptoms for black females are chronic and more intense [8]. Regularly elevated anxiety levels are reported for black girls residing in urban areas across the United States [9-12]. Research indicates that, if left unaddressed, anxiety difficulties in adolescence are associated with a higher risk for anxiety disorders as adults [13,14]. Black women with higher levels of anxiety also report poor emotional regulation [13]. Given the elevated anxiety levels among urban black girls [9-12], the intensity and chronicity of anxiety symptoms [8], and the implications for adulthood [13,14], early intervention with this population is appropriate. Yet the literature is clear as it relates to anxiety and other mental health issues: black girls are underresearched and underserved [15,16]. Reasons for this include, but are not limited to, cultural mistrust, access to mental health care, access to same-race mental health care, and stigma [15,16].

mHealth and Adolescents

A plethora of mHealth apps exist for anxiety, but few are evidence-based, and even fewer are specifically designed for adolescents [17,18]. Nonetheless, the available literature suggests that given adolescents' frequent use of mobile phones, mHealth apps can serve as effective forms of early intervention or augmentations to more traditional forms of intervention [19]. This technology may be particularly helpful for underserved black adolescents as an early anxiety intervention because 81% of this population reported using a mobile phone on a daily basis [20,21].

Cognitive Restructuring

Cognitive restructuring is an important tool in anxiety management [22]. Operationally defined, cognitive restructuring involves two parts: (1) recognizing and deconstructing erroneous thoughts and (2) replacing the negative anxiety and stress-provoking thoughts with positive thoughts [22]. In our work with black females, we have found at times that cognitive restructuring is a difficult skill to practice on one's own [23]. Outside of the therapy setting, our clients reported that they found it "too hard" and that they "couldn't do it," suggesting there may be a need for a culturally informed technique to assist in the steps necessary to teach and execute cognitive restructuring. One way to do so is to merge the technique with music.

Black Adolescents and Music

Music is an important part of almost every adolescent's life [24,25]. The tempo, genre, rhythm, and lyrics combine to allow youth to express thoughts and feelings they may not be able to verbalize in other ways [25]. Female adolescents are more likely to use music to express their emotional state than their male counterparts [26]. Unlike adolescents of previous eras, mobile phone technology allows today's youth to take their music everywhere they go [20].

For black youth, music is a powerful tool for expressing and overcoming emotion [23,27-32]. Over the last 70 years, songs have been an integral part of the black youth experience, providing hope, empowerment, faith, and encouragement [23,27-32]. Reflecting on the importance of spirituals and gospel music to black adolescents and emerging adults during the civil rights era, icon Euvester Simpson said, "The only thing that got us through that...we sang. We sang all night. I mean songs got us through so many things, and without that music, I think many of us would have just lost our minds or lost our way completely" [30]. The importance of gospel music in stressful situations still resonates today. A recent study with black young adults found many used gospel songs with instructive lyrics when stressed [31].

The expression of emotion and feelings is not limited to religious music. Hip-hop emerged from low-income urban areas to give a voice to black youth who were disenfranchised and marginalized [23]. As a result, for the first time, their experiences and concerns were validated and heard [23,32]. Although the commercial success of the genre co-opted the topics of the rhymes, a recent resurgence has resulted in a return to lyrics reflecting social justice, racism, and inequality issues of concern to black youth [23]. Indeed, recent research suggests that a component of hip-hop, rap, is the perfect form of music therapy for black adolescents [32].

Building on the evidence that music provides a voice to the emotions and experience of females [1,2,33] and black adolescents [23,27-32], and the potential impact of mHealth apps on the emotional lives of adolescents [19,21], we developed the Build Your Own Theme Song (BYOTS) app [33,34]. Whereas other researchers have used cognitive restructuring and music separately within the same intervention with black African couples [35], in our work we have integrated music directly into the cognitive restructuring process. Rather than developing positive statements to replace negative thoughts, our participants are taught to build their own theme song with positive lyrics to counteract negative thoughts. In our intervention studies, when a black female adolescent pays

attention to her thoughts and recognizes that she is thinking negatively, she uses her theme song to counteract and replace those thoughts with positive thoughts.

Build Your Own Theme Song App

The BYOTS mHealth app consists of a self-penned, self-recorded theme song based on the girl's favorite song. Consistent with the research that shows mHealth apps for anxiety have the greatest impact when augmenting a face-to-face intervention [17], the BYOTS mHealth app is integrated into an eight-session culturally infused Sisters United Now (SUN) sister circle that takes place during the school day [11,33,34]. SUN and the app were developed to address the neglect of black adolescent girls in anxiety-related interventions [11,15]. During the intervention, girls are introduced to the negative and positive thought cycles. They learn how a theme song can interrupt the negative thought cycle by replacing the negative thoughts with positive affirming thoughts. Within our work, we have labeled this process as musical cognitive restructuring. The girls first choose their favorite song. As the intervention takes place in schools, in accordance with school rules, the song must not contain profanity. They then use a vision statement that was generated earlier in the intervention and a positive word bank to rewrite their favorite song into their own personal theme song. After the lyrics are revised, the new theme song is recorded on the BYOTS app.

An additional feature of the app is the pre-post app survey. When participants open the app, the survey is presented. Based on the survey score, the girl is instructed to play her theme song. After the song is completed, the survey appears again. Based on the survey score, the girl either receives a positive message praising her for changing her thinking or instructing her to play her theme song again. Three push alerts per day (before school, after school, and at night) are built into the app to prompt girls to use it multiple times per day. The girls are also instructed to use the app without prompts when they experience negative thinking.

In this mixed methods study, we present data on the effectiveness of the BYOTS app. Although a repeated measures approach to measure the participants' evaluation of the app's effectiveness would be desirable, a preassessment of their app expectations was only possible. Specifically, we predict that participants would expect the app to be helpful in reducing negative thoughts and that, after working with the app, their positive expectations will be met. To test this prediction, we examine girls' expectations of the app and compare these to their experience after using their app. Specifically, we examine the ease of using the app, including recording their theme song. We also looked at their comfort using the app in various settings and with others (family and friends) present. We also hypothesize that using the app will reduce negative thinking. To test this hypothesis, evidence of negative thinking was examined on day 1 and day 7 of the 1-week app experience; we predict that the presence of negative thinking will be lower at the end of day 7. Anxiety was also assessed before and after the app-augmented SUN experience; we predict that anxiety will be reduced after the SUN intervention.

Methods

Materials

The Multidimensional Anxiety Scale for Children 2 (MASC-2) [35] is a 50-item instrument that assesses anxiety symptoms using a four-point Likert-like scale from 0 to 3 (0=never, 1=rarely, 2=sometimes, and 3=often) in children. It contains items such as "I worry about other people laughing at me" and "I have trouble asking other kids to play with me." Total anxiety score was used to assess overall anxiety symptom level. This measure has been found to have good reliability (Cronbach alpha=.90) and good convergent validity [11,12]. Additionally, this measure has been found to be reliable in our sample (Cronbach alpha=.87) [11,36].

Pre and Post In-App Survey

This survey consists of five items. The first statement assesses general stress level, and the remaining four questions assess negative thought. Questions about negative thought included (1) I have little control over important things in my life, (2) I am confident in my ability to handle personal problems, (3) things are not going my way, and (4) only good things lie before me. Two items are reverse coded. The survey uses a seven-point Likert scale (1=strongly disagree, 2=disagree, 3=slightly disagree, 4=neither, 5=slightly agree, 6=agree, 7=strongly agree). All app responses were uploaded to a secure online server. For each entry, responses were combined for negative thought total score. A score greater than 16 on the four negative thinking questions (the median score on the scale that could range from 4 to 28) resulted in a prompt to listen to their theme song. On both days 1 and 7, the first survey response of the day was used in the analysis. If there was only one survey response for the day, this entry was retained for the analysis. Ellzey's recent results supported the validity of the In-App Survey [33].

The Expectations of App Survey

This measure consisted of 14 questions on a seven-point Likert scale and was used specifically to measure girls' expectations of the BYOTS mobile app prior to the start of the intervention. Sample questions included, "It will be practical to pull out my phone and use the app when I feel stressed" and "I believe the app will be useful in calming my nerves for specific situations when I feel stressed." Higher scores indicated more positive expectations. A Cronbach alpha coefficient was calculated (Cronbach alpha=.893) indicating the survey was reliable for this sample.

Theme Song Building Session Evaluation

Administered after the completion of the theme song lyrics and recording, this five-item evaluation examined girls' satisfaction with using the app to build and record their theme song. Sample statements included "I enjoyed recording my own theme song" and "Practicing using the app during the session helped me learn." Each item was rated on a five-point scale from strongly disagree (1) to strongly agree (5).

Guided Discussion Focus Group Transcripts

At the end of the intervention, guided discussions to gain insight into participants' experiences with the app were held. These

discussions were audio recorded and transcribed. Open-ended questions were posed, such as "What made it easy to use? What made it hard?" and "Has using the app changed the way you think about yourself?" All participants were encouraged to provide their perceptions on how helpful the app was in changing their negative thinking. Audio recordings of the six focus discussions were transcribed.

Participants

Participants were 72 black or biracial seventh- and eighth-grade adolescent females. However, the sample size for each analysis varied due to initial connectivity issues between the app and the secure server in which data were unable to be matched to the participant. Participants were between the ages of 12 and 15 years and enrolled in the SUN program. Participants attended one of two middle schools located in a large midwestern, low-income, urban school district in the United States. All students within the district were enrolled in the federal free breakfast and lunch program. Informed consent was obtained from both the girls and their parent(s) or guardian(s). Participants either self-selected to participate or were recommended for SUN by their school counselor. The study was approved by the Kent State University Institutional Review Board, Kent, Ohio.

Procedure

As part of the SUN intervention, girls recorded their theme songs into the app and practiced using the app in-session. The girls were then prompted three times throughout the day for a 1-week period to use the app. Data on app usage were collected in real time. On their return to SUN, a focus group was led by a third-party individual to assess the girls' experiences with the app.

Statistical Analysis

Quantitative data was analyzed using SPSS. The following App Expectations Survey items were chosen to test our predictions about the app's effectiveness: "Girls will find the app helpful" (item 8), "I expect that after using the app I will feel less anxious in general" (item 9), "The app makes me feel powerful" (item 10), and "I want to use the theme song app" (item 14).

At the end of the intervention, girls completed an evaluation, which included an item to assess their perception of the effectiveness of listening to their song to help reduce their negative thinking. To make this five-point scale comparable to the expectation survey, scores were converted to a seven-point scale to match the expectation scale (maximum score of 5 on the evaluation survey converts to a score of 7, midpoint 3=4, and minimum 1=1, etc). This transformation of the data, a standard practice to enable the direct comparison of scales, was necessary to determine if there was a significant difference between expectations and the app experience using an independent *t* test [37-39].

To determine whether negative thinking had been reduced because of the app experience, days 1 and 7 app scores were compared with the exception of three participants who began using the app on day 2. A paired sample *t* test was then used to examine whether day 7 negative thinking scores as measured by the app survey were lower than day 1 scores. Pre and post MASC-2 total anxiety scores also were compared using a paired sample *t* test.

Qualitative data was analyzed using content analysis. Three black women between the ages of 21 and 25 years who were familiar with the culture and idiom of urban black girls transcribed the focus group audio digitals. A coding form was developed to assess each focus group question for each of the SUN sister circles. Coders read the transcripts, identified major themes, and included quotes that reflected the identified themes. In addition, coders determined if the themes identified related to the primary research questions of app feasibility and effectiveness. A consensus meeting was held with the data supervisor after coding was completed. Coders reviewed the themes and discussed any differences in their coding. Ultimately, the themes that emerged were the consensus of the three coders.

Results

Quantitative Evidence of Effectiveness

Descriptive statistics were calculated for relevant app expectation items that addressed the potential effectiveness of the app and are presented in Table 1. Scores ranged from 1 to 7 (7=strongly agree). The results for all expectation items were well above the median score (4) of the App Expectation Survey Likert scale. Also, the descriptive analysis indicated that the modes were all high (all scores of 7 or 6) and skewness was strongly negative for each item, indicating that girls' expectations for the app were very high.

Descriptive statistics were calculated for each app expectation (mean and standard deviation) where scores could range from 1 to 7 (7=strongly agree): item 6, "It will be practical to pull out my phone and use the app when I feel stressed" (mean 6.2, SD 1.2); item 8, "I believe the app will be useful in calming my nerves for specific situations when I feel anxious" (mean 5.9, SD 1.4); item 9, "I expect that after using the app I will feel less anxious in general" (mean 5.9, SD 1.3), item 10, "I expect my theme song will make me feel powerful" (mean 6.1, SD 1.2)" and item 14, "I want to use the theme song app" (mean 5.7, SD 1.5). The results for all expectation items were well above the median score (4 of 7) of the Application Expectation Survey Likert scale. Also, the descriptive analysis indicated that the modes were all high (all scores of 7 or 6) and that skewness was strongly negative for each item further indicating that girls' expectations for the app were very high.

The next question was to determine if these high expectations were maintained after experiencing the app. Independent *t* tests compared app expectation items 6 and 8 (these items were judged to be the best indicators of expectations of use and usefulness) with the session 7 evaluation question "I enjoyed recording my theme song" (mean 5.8, SD 1.5). The resulting *t* values were not significant. Thus, participants' expectations of the app were high, and these expectations were maintained after actually experiencing the app.

Table 1. Descriptive statistics for selected App Expectation Survey items.

App Expectation Survey items	N	Participant responses (7=strongly agree)	
		Mean (SD)	Median
App will be practical to use	60	6.2 (1.2)	7.0
App will useful in calming nerves	59	5.9 (1.4)	6.0
Will feel less anxious using the app	59	5.9 (1.3)	6.0
The theme song will make me feel more powerful	60	6.1 (1.2)	6.0
Want to use the theme song app	59	5.7 (1.5)	6.0

A paired samples *t* test was used to determine whether negative thought was significantly lower at day 7 than day 1. Average negative thought scores on day 7 (mean 12.81, SD 4.22) were significantly lower than average negative thought scores on day 1 (mean 14.20, SD 4.10; t_{31}=1.69, *P*=.05, Cohen *d*=0.30). On day 7, 42% of participants reported lower average negative thought scores versus day 1.

A paired sample *t* test compared girls' pre-app to post-app total MASC anxiety T-scores to determine if a change in anxiety was evident. Results showed there was a significant decrease (t_{38}=2.82, *P*=.004) in anxiety scores from preintervention (mean 56.28, SD 11.18) to postintervention (mean 53.21, SD 11.31) with a medium effect size (Cohen *d*=0.517).

Qualitative Evidence of Effectiveness

During the guided discussions, four themes emerged related to the effectiveness of the BYOTS app: (1) differences in behavior and temperament, (2) promotes calmness, (3) helpfulness in stressful home situations, and (4) focused thinking via the SUN song

The theme "differences in behavior and temperament" relates to behavioral and attitude changes observed in participants by others as well as their own self-awareness of that change. These changes are illustrated in the following sample statements:

> *My mother has seen changes because I always use to catch an attitude with my sister and brother, but now I worries [sic] about myself and not everyone else.*

> *I don't cuss people out anymore.*

A second theme identified reflects girls' perceptions that the app "promoted calmness." The following statement is an example of this theme:

> *When it tells you 'you're stressed' at the end, it helps you think to yourself that you need to calm down.*

The third theme that emerged was that girls found the app "helpful in home and familial situations." The following statement exemplifies this theme:

> *I was arguing with my mom and sister. I ran up to my room, slammed the door. Then I thought, wait a minute, I can use my app. I listened to my song, went downstairs, and we worked it out.*

The final theme that emerged centered on the ability of the app to focus one's thinking. As noted earlier in this paper, when girls opened the app, it played the first seven bars of their theme song. The familiarity of the song allowed girls to focus on their thoughts. This theme is reflected in the statement:

> *When the S.U.N. song began to play, it calmed me down, and I could focus on my thinking.*

Discussion

Principal Findings

Despite having higher levels of anxiety, and the consequences of that anxiety on adulthood, urban black girls are conspicuously absent from the mHealth app literature. To the best of our knowledge, this study is the first to develop an anxiety-related app for this population and evaluate its effectiveness. Our mixed methods results indicate that musical cognitive restructuring reduced their negative thinking. Indeed, negative thoughts decreased from day 1 to day 7. Changes in thinking and behavior were noted by the girls themselves, their friends and family, and by the data collected within the app. Anxiety was also reduced pre-post intervention.

Our findings suggest that an mHealth app can be an important tool in anxiety intervention with this population. It should be noted that BYOTS is not a stand-alone app but is used in conjunction with SUN culturally relevant intervention. Therefore, although we are able to say that the app played a role in the reduction of anxiety, we cannot say that it was solely responsible for the reduction in anxiety. We have designed a study in which we compare three groups (app only, intervention only, app-augmented intervention); this will allow us to further understand the app's effectiveness in reducing anxiety.

Not only the girls themselves, but their families and friends, noted the change in thinking and attributed that change to using the app. This attribution suggests that we take a closer look at the mechanisms underlying musical cognitive restructuring to determine what components are facilitating the change. Is it the lyrics, the rhythm, the tempo, the fact that it is contained within an app, the self-monitoring aspect, or some combination of these factors? A closer look at key elements would allow us to gain further understanding as to why BYOTS works. To this end, we recently completed a study examining the lyric, rhythm, and tempo component of participants' theme songs. Findings should be available soon.

As a result of the observed changes, family members—particularly mothers—expressed interest in using the app. Although BYOTS is designed for adolescent females, the possibility exists that the app may also be an effective tool for adult black women.

The emergent theme "focused thinking via the song" was an unexpected finding. The SUN intervention has a theme song that is used to close each session. When the BYOTS app is opened, the first seven bars of the song are heard. It appears that for some participants, hearing the intervention theme song was a cue to self-monitor thoughts [40], the first step in cognitive restructuring. In subsequent planned studies, we will further explore this finding.

Girls' expectations of the app matched their experience with the app. Comparing expectations and experience is a standard procedure in intervention research; however, this result holds added meaning for our sample. As part of the focus groups, girls shared their skepticism as to whether the app would meet their expectations. Participants indicated that they had worked with other researchers where what was expected had not matched the actual experience. They expressed great delight "it really worked" regarding their app experiences.

A frequent question that arises in our work is "what about adolescent black boys?" Clearly, a need exists to develop an anxiety mHealth app for this group. Given the robust finding that boys respond to the tempo and rhythm to elevate mood [26], rather than the lyrics of a song, in its current iteration BYOTS may not be a good fit for black males. We plan to undertake a series of focus group studies to determine if and how the app should be modified for this population.

Limitations

This study represents an open trial and a first step in empirically evaluating the BYOTS app. The study was also limited to an urban, low-income population. However, black girls are heterogenous and encompass all socioeconomic statuses and reside in various types of locales. Subsequent studies would extend the work to black girls of various socioeconomic statuses residing in urban, rural, and suburban locales. Already, we are conducting an open trial with girls residing in a small town, and developing and laying the groundwork for a multisite randomized controlled trial.

For research purposes, use of the app was limited to a 1-week period. In subsequent studies, we plan to extend app use to 1 month. This will allow us to gather further data, including frequency of unprompted use, time of day most used, and so on.

Conclusions

The BYOTS app appears to be a useful tool in reducing negative thinking among urban, low-income, middle-school black girls. Our findings add to the knowledge base as to how mHealth can be used with underserved and underresearched adolescents. As elucidated in our Discussion, these results provide direction for subsequent research with this mHealth app.

Acknowledgments

This research was supported in part by grants from the Akron Community Foundation's Women's Endowment Fund, Kent State Research and Sponsored Programs, and the Kent State Applied Psychology Center.

Conflicts of Interest

AN-B is the VP Workforce Diversity and Inclusion for A3B, LLC; she owns the copyright Build Your Own Theme Song (BYOTS) app. None of the other authors have a conflict of interest to declare.

References

1. Mazzone L, Ducci F, Scoto MC, Passaniti E, D'Arrigo VG, Vitiello B. The role of anxiety symptoms in school performance in a community sample of children and adolescents. BMC Public Health 2007 Dec 5;7:347 [FREE Full text] [doi: 10.1186/1471-2458-7-347] [Medline: 18053257]
2. Wood JJ. Effect of anxiety reduction on children's school performance and social adjustment. Dev Psychol 2006 Mar;42(2):345-349. [Medline: 16569172]
3. Fordham K, Stevenson-Hinde J. Shyness, friendship quality, and adjustment during middle childhood. J Child Psychol Psychiatry 1999 Jul;40(5):757-768. [Medline: 10433409]
4. La Greca AM, Lopez N. Social anxiety among adolescents: linkages with peer relations and friendships. J Abnorm Child Psychol 1998 Apr;26(2):83-94. [Medline: 9634131]
5. Barrett S, Heubeck BG. Relationships between school hassles and uplifts and anxiety and conduct problems in grades 3 and 4. J Appl Dev Psychol 2000 Sep;21(5):537-554. [doi: 10.1016/S0193-3973(00)00053-8]
6. Compton SN, Nelson AH, March JS. Social phobia and separation anxiety symptoms in community and clinical samples of children and adolescents. J Am Acad Child Adolesc Psychiatry 2000 Aug;39(8):1040-1046. [doi: 10.1097/00004583-200008000-00020] [Medline: 10939233]
7. Latzman RD, Naifeh JA, Watson D, Vaidya JG, Heiden LJ, Damon JD, et al. Racial differences in symptoms of anxiety and depression among three cohorts of students in the southern United States. Psychiatry 2011;74(4):332-348. [doi: 10.1521/psyc.2011.74.4.332] [Medline: 22168294]
8. Breslau J, Aguilar-Gaxiola S, Kendler KS, Su M, Williams D, Kessler RC. Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. Psychol Med 2006 Jan;36(1):57-68 [FREE Full text] [doi: 10.1017/S0033291705006161] [Medline: 16202191]

9. Frazier F, Belliston L, Brower L, Knudsen K. Placing Black Girls at Promise: Report of the Rise Sister Rise Study. Executive Summary. Columbus, OH: Ohio Department of Mental Health; 2011. URL: http://www.risesisterrise.net/uploads/3/4/5/5/34555752/rsr_columbus_executive_summary.pdf [accessed 2019-05-17] [WebCite Cache ID 78Qdht0HM]
10. Jones-DeWeever A. Black Girls in New York City: Untold Strength and Resilience. Washington, DC: Institute for Women's Policy Research; 2009 Feb 28. URL: https://iwpr.org/publications/black-girls-in-new-york-city-untold-strength-and-resilience/ [accessed 2019-05-13] [WebCite Cache ID 78LpQDGt6]
11. Pettiti K, Ellzey D, Davis M, Pugh D, Stadulis R, Neal-Barnett A. Anxiety and stress in inner city Black American middle school girls: levels and relationships. J Psychiatry Mental Disorder 2017 May 1:2-1011 [FREE Full text]
12. Cooley-Strickland M, Griffin R, Darney D, Otte K, Ko J. Urban African American youth exposed to community violence: a school-based anxiety preventive intervention efficacy study. J Prev Interv Community 2011;39(2):149-166 [FREE Full text] [doi: 10.1080/10852352.2011.556573] [Medline: 21480032]
13. Graham J, Calloway A, Roemer L. The buffering effects of emotion regulation in the relationship between experiences of racism and anxiety in a Black American sample. Cog Ther Res 2015;39(5):553-563. [doi: 10.1007/s10608-015-9682-8]
14. Ginsburg G, Becker-Haimes E, Keeton C, Kendall P, Iyengar S, Sakolsky D, et al. Results from the Child/Adolescent Anxiety Multimodal Extended Long-Term Study (CAMELS): primary anxiety outcomes. J Am Acad Child Adolesc Psychiatry 2018;57(7):471-480. [Medline: 29960692]
15. Neal-Barnett A, Stadulis R, Murray M, Payne MR, Thomas A, Salley BB. Sister Circles as a culturally relevant intervention for anxious African American women. Clin Psychol (New York) 2011 Sep;18(3):266-273 [FREE Full text] [doi: 10.1111/j.1468-2850.2011.01258.x] [Medline: 22081747]
16. Watson Coleman B. Congressional Caucas on Black Women and Girls. URL: https://watsoncoleman.house.gov/cbwgcaucus/ [accessed 2019-05-13] [WebCite Cache ID 78Lpxn67L]
17. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
18. Pramana G, Parmanto B, Kendall PC, Silk JS. The SmartCAT: an m-health platform for ecological momentary intervention in child anxiety treatment. Telemed J E Health 2014 May;20(5):419-427 [FREE Full text] [doi: 10.1089/tmj.2013.0214] [Medline: 24579913]
19. Reid SC, Kauer SD, Hearps SJ, Crooke AH, Khor AS, Sanci LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomized controlled trial. BMC Fam Pract 2011;12:131 [FREE Full text] [doi: 10.1186/1471-2296-12-131] [Medline: 22123031]
20. Rideout VJ, Scott KA, Clark D. The Digital Lives of African American Tweens, Teens, and Parents: Innovating and Learning With Technology. Executive Summary. Tempe, AZ: Center for Gender Equity in Science and Technology; 2016. URL: https://cgest.asu.edu/sites/default/files/digital_lives_executive_summary_single.pdf [accessed 2019-05-17] [WebCite Cache ID 78QeMQOs8]
21. Chan A, Kow R, Cheng J. Adolescents' perceptions on smartphone applications (apps) for health management. JournalMTM 2017 Aug;6(2):47-55. [doi: 10.7309/jmtm.6.2.6]
22. Hope DA, Burns JA, Hayes S, Herbert JD, Warner MD. Automatic thoughts and cognitive restructuring in cognitive behavioral group therapy for social anxiety disorde. Cog Ther Res 2010 Feb 1;34(1):1-12 [FREE Full text]
23. Alridge D, Steward J. Introduction: hip-hop in history: past, present, and future. J Afr Am Hist 2005 Jun 1;90(3):190-195. [doi: 10.1086/JAAHv90n3p190]
24. Gantz W, Gartenberg H. Pop music and adolescent socialization: an information perspective. 1979 Jun 1 Presented at: Annual Meeting of the International Communication Association; May 1-5, 1979; Philadelphia, PA URL: https://archive.org/details/ERIC_ED178966
25. Saarikallio S, Erkkilä J. The role of music in adolescents' mood regulation. Psychol Music 2007 Jan 1;35(1):88-109. [doi: 10.1177/0305735607068889]
26. Council on Communications and Media. American Academy of Pediatrics Policy Statement: impact of music, music lyrics, and music videos on children. Pediatrics 2009 Nov 1;124(5):1488-1494. [Medline: 19841124]
27. Southern E. The Music Of Black Americans: A History (Third Edition). New York: WW Norton & Company; 1997.
28. Jones F, Jones AC. The Triumph of the Soul: Cultural and Psychological Aspects of African American Music. Westport, CT: Greenwood Publishing Group; 2000.
29. Smithsonian. African American music URL: https://www.si.edu/spotlight/african-american-music [accessed 2019-05-13] [WebCite Cache ID 78Lrh7Pwl]
30. Dittmer J, Simpson E. Library of Congress. Euvester Simpson oral history interview conducted by John Dittmer in Jackson, Mississippi, 2013-03-12 URL: https://www.loc.gov/item/afc2010039_crhp0072/ [accessed 2019-05-13] [WebCite Cache ID 78Lrmg6yK]
31. Hamilton JB, Stewart JM, Thompson K, Alvarez C, Best NC, Amoah K, et al. Younger African American adults' use of religious songs to manage stressful life events. J Relig Health 2017 Feb;56(1):329-344. [doi: 10.1007/s10943-016-0288-6] [Medline: 27464642]
32. Hadley S, Yancy G. Therapeutic Uses of Rap and Hip-hop. New York: Routledge; 2012.

33. Ellzey D. OhioLINK. Kent, OH: Kent State University; 2018 Apr 1. Musical cognitive restructuring based app for black females' negative thoughts and anxiety [electronic thesis or dissertation] URL: https://etd.ohiolink.edu/pg_10?::NO:10:P10_ETD_SUBID:164420 [accessed 2019-05-17] [WebCite Cache ID 78QuvnSc0]
34. Ezegbe B, Ede M, Eseadi C, Nwaubani O, Akaneme I, Aye E, et al. Effect of music therapy combined with cognitive restructuring therapy on emotional distress in a sample of Nigerian married couples. Medicine (Baltimore) 2018 Aug;97(34):e11637 [FREE Full text] [doi: 10.1097/MD.0000000000011637] [Medline: 30142755]
35. March J. MASC 2: Multidimensional Anxiety Scale for Children. Second Edition. Cheektowaga, NY: Multi-Health Systems; 2013.
36. Davis M, Stadulis R, Neal-Barnet AM. Assessing the effects of the acting white accusation among black girl: social anxiety and bullying victimization. J Natl Med Assoc 2018 Feb 1;110(1):23-28. [Medline: 29510839]
37. Knapp T. Treating ordinal scales as interval scales: an attempt to resolve the controversy. Nurs Res 1990;39(2):121-123. [Medline: 2315066]
38. Bishop PA, Herron RL. Use and misuse of the Likert item responses and other ordinal measures. Int J Exerc Sci 2015;8(3):297-302 [FREE Full text] [Medline: 27182418]
39. IBM Support. 2019. Technote URL: http://www-01.ibm.com/support/docview.wss?uid=swg21482329 [accessed 2019-04-01] [WebCite Cache ID 77JtJJnP6]
40. Snyder M. Public Appearances-Private Realities: The Psychology of Self-Monitoring. New York: WH Freeman & Co; 1984.

Abbreviations

BYOTS: Build Your Own Theme Song
MASC-2: Multidimensional Anxiety Scale-2
SUN: Sisters United Now

Edited by G Eysenbach; submitted 15.06.18; peer-reviewed by K Aguirre, M Birk; comments to author 06.10.18; revised version received 18.01.19; accepted 15.04.19; published 27.06.19.

Please cite as:
Neal-Barnett A, Stadulis R, Ellzey D, Jean E, Rowell T, Somerville K, Petitti K, Siglow B, Ruttan A, Hogue M
Evaluation of the Effectiveness of a Musical Cognitive Restructuring App for Black Inner-City Girls: Survey, Usage, and Focus Group Evaluation
JMIR Mhealth Uhealth 2019;7(6):e11310
URL: http://mhealth.jmir.org/2019/6/e11310/
doi:10.2196/11310
PMID:31188130

Original Paper

Understanding the Role of Healthy Eating and Fitness Mobile Apps in the Formation of Maladaptive Eating and Exercise Behaviors in Young People

Mahsa Honary[1*], PhD; Beth T Bell[2*], PhD; Sarah Clinch[3*], PhD; Sarah E Wild[2*], BSc; Roisin McNaney[4*], PhD

[1]School of Computing and Communications, Lancaster University, Lancaster, United Kingdom

[2]School of Psychological and Social Sciences, York St John University, York, United Kingdom

[3]School of Computer Science, University of Manchester, Manchester, United Kingdom

[4]School of Computer Science, Electrical Engineering and Engineering Maths, University of Bristol, Bristol, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Mahsa Honary, PhD
School of Computing and Communications
Lancaster University
InfoLab21
Lancaster, LA1 4WA
United Kingdom
Phone: 44 1524 593566
Email: m.honary@lancaster.ac.uk

Abstract

Background: Healthy eating and fitness mobile apps are designed to promote healthier living. However, for young people, body dissatisfaction is commonplace, and these types of apps can become a source of maladaptive eating and exercise behaviors. Furthermore, such apps are designed to promote continuous engagement, potentially fostering compulsive behaviors.

Objective: The aim of this study was to identify potential risks around healthy eating and fitness app use and negative experience and behavior formation among young people and to inform the understanding around how current commercial healthy eating and fitness apps on the market may, or may not, be exasperating such behaviors.

Methods: Our research was conducted in 2 phases. Through a survey (n=106) and 2 workshops (n=8), we gained an understanding of young people's perceptions of healthy eating and fitness apps and any potential harm that their use might have; we then explored these further through interviews with experts (n=3) in eating disorder and body image. Using insights drawn from this initial phase, we then explored the degree to which leading apps are preventing, or indeed contributing to, the formation of maladaptive eating and exercise behaviors. We conducted a review of the top 100 healthy eating and fitness apps on the Google Play Store to find out whether or not apps on the market have the potential to elicit maladaptive eating and exercise behaviors.

Results: Participants were aged between 18 and 25 years and had current or past experience of using healthy eating and fitness apps. Almost half of our survey participants indicated that they had experienced some form of negative experiences and behaviors through their app use. Our findings indicate a wide range of concerns around the wider impact of healthy eating and fitness apps on individuals at risk of maladaptive eating and exercise behavior, including (1) guilt formation because of the nature of persuasive models, (2) social isolation as a result of personal regimens around diet and fitness goals, (3) fear of receiving negative responses when targets are not achieved, and (4) feelings of being controlled by the app. The app review identified logging functionalities available across the apps that are used to promote the sustained use of the app. However, a significant number of these functionalities were seen to have the potential to cause negative experiences and behaviors.

Conclusions: In this study, we offer a set of responsibility guidelines for future researchers, designers, and developers of digital technologies aiming to support healthy eating and fitness behaviors. Our study highlights the necessity for careful considerations around the design of apps that promote weight loss or body modification through fitness training, especially when they are used by young people who are vulnerable to the development of poor body image and maladaptive eating and exercise behaviors.

(JMIR Mhealth Uhealth 2019;7(6):e14239) doi:10.2196/14239

KEYWORDS

weight loss; mobile apps; eating disorders; body image diet; exercise; mental health

Introduction

Background and Related Work

Body dissatisfaction, the subjective experience of negative thoughts and feelings toward one's own body [1], is so prevalent among young people (defined by the United Nations as those aged 15 to 24 years [2]) in modern Western societies that it is regarded as *normative discontent* [3,4]. Body dissatisfaction has been linked with a number of maladaptive eating and exercise behaviors, including restrained eating practices, consuming less fruit and vegetables, low levels of physical activity, excessive exercise, binge-purge cycles, and anabolic steroid use [5,6]. Furthermore, body dissatisfaction is regarded as both an important risk factor for, and is symptomatic of, clinical eating disorders, such as anorexia and bulimia [7,8], the majority of which develop during adolescence and early adulthood [9].

The causes of body dissatisfaction and associated maladaptive eating and exercise behaviors are diverse, with research implicating a combination of biological, psychological, and sociocultural factors [7,10,11]. Sociocultural theories emphasize the role of specific agents, such as parents, peers, and the media, in shaping negative attitudes toward the body [12], with body dissatisfaction arising because of perceived pressure from sociocultural agents to conform to an unrealistic, culturally defined body and beauty ideal. For women, this has been described as thin and toned, yet curvaceous with pert breasts and buttocks, whereas for men it is muscular yet lean with little body fat [13]. The complex and unrealistic nature of this ideal makes it impossible for the majority of young people to achieve, leading to negative feelings around their own bodies [7,12]. In turn, these feelings can motivate an engagement in maladaptive eating and exercise behaviors, aimed at changing the body [7,12].

Perpetuating these social and emotional pressures is the fact that many of these behaviors (eg, clean eating, over-exercising, and cutting out food groups) have become the cultural norm, with magazines and celebrities on social media advocating calorie restriction as an everyday part of how we think about food [14]. Following this longstanding cycle of *diet culture*, parents, who themselves engage in dieting behaviors, can be the ones who convey messages on calorie restriction and *good* versus *bad* foods to children from a young age [15]. Thus, to many young people growing up in this environment, these ways of thinking about food and exercise are seen as the norm and are often engaged with regardless of whether the young person is overweight or not [16]. Ironically, calorie restriction has been demonstrated to lead to weight gain and eating disorders over time in young people [17].

In recent years, the emergence and increasing availability of new digital media and technologies has drastically changed the social landscape. As a consequence, theories of body dissatisfaction and maladaptive eating and exercise behaviors have needed to be adapted. Research in this field has typically focused on how social media influences how young people think, feel, and behave with regard to their body. For example, research has highlighted the role of social media image sharing practices in body dissatisfaction [18]; the further normalization of maladaptive body shaping strategies through user-generated social media content [13]; and the use of social media spaces to create communities centered around maladaptive eating and exercise behaviors [19].

Healthy Eating and Fitness Apps

The rise of healthy eating and fitness apps is considered to be yet another type of media-influencing body dissatisfaction [20]. The nature of personal mobile engagement itself poses a significant challenge, with vulnerable populations freely accessing content without the extent of their engagement becoming visible to others (eg, a parent who might be unaware of their child's engagement in calorie-counting practices). However, although there is a wealth of research exploring the dangers of social media use in the development of maladaptive behaviors in young people, comparatively little attention has been paid to healthy eating and fitness mobile apps, despite their ability to become a tool for supporting restrictive eating behaviors (eg, following a 1200-calorie-a-day diet regardless of hunger levels).

The Health and Fitness category accounts for a large proportion of apps in both the Android and Apple app stores (recent statistics place them as the 9th and 8th largest categories of apps, respectively) [21,22]. Although this category includes a variety of health-related domains (including mental health, smoking cessation, and women's health), a significant portion relates to diet, healthy eating, and fitness. In this paper, healthy eating and fitness apps refer to mobile apps that aim to promote healthy eating and fitness (exercise) activities supporting users in monitoring or attending to food and exercise, typically by providing users with additional information about these activities such as calories consumed or burned as well as projections on weight loss goals. The top 10 lists for this category, in both stores, are dominated by apps such as MyFitnessPal [23], Fitbit [24], Strava [25], and Sweatcoin [26], all of which are focused on self-tracking of food or exercise in similar ways.

Healthy eating and fitness apps incorporate a variety of strategies to promote health and fitness behaviors including behavior tracking, goal setting, feedback, rewards, social connectivity, and remote coaching [27]. Adolescent users of Fitbit (a wearable activity monitor with an associated mobile phone app) found that the technology encouraged them to engage in physical activity and increase knowledge of their own health behaviors in the short term [28]. However, there may be a dark side to healthy eating and fitness app use. Many apps are centered on weight loss, encouraging users to set weight-related goals that are achieved through calorie restriction and exercise [28,29]. This emphasis is problematic as the benefits of weight loss to physical health remain a controversial topic within the medical research literature [30,31]. In particular, weight loss is generally only advocated for individuals with high body weight [30],

raising concerns about how the cultural preoccupation with weight loss stigmatizes overweight bodies, which may actually demotivate such individuals from engaging in healthy practices in the first place. The growing body of research also shows that body size is not necessarily indicative of good health [31,32]. In addition, calorie restriction has been widely recognized as an ineffective method of weight loss in the long term and an inefficient way of promoting more healthy eating behaviors. Instead, approaches focused on mindfulness and self-awareness, such as intuitive eating, have been advocated more recently [33,34].

The weight loss frameworks that underpin many healthy eating and fitness apps may not be conducive to the formation of positive health and fitness behaviors. In support of this, Eikey et al found that around 7% (1261/18,601) of female app users set weight goals that are under what is considered to be healthy [29], indicative of a desire to achieve unrealistic appearance goals, rather than improved health. The same study found that users felt encouraged by apps to engage in, what the authors considered to be, maladaptive eating and exercise behaviors [29]. For example, some participants described how app feedback following the logging of calories over their set daily target led them to engage in purging behaviors, whereas others described feeling obsessed with thoughts of food and calorie content. Thus, apps focusing on weight loss may inadvertently promote maladaptive eating and exercise strategies.

Research suggests that healthy eating and fitness apps in the personal health domain employ behavior change models because of their potential to support individuals in the process of adapting and maintaining a new healthy behavior [35,36]. The majority focus on habit formation, where habits are defined as automatic responses to contextual cues and are formed as the behavior is repeated in a stable context [37] (eg, using reminders, goal setting, and positive reinforcement). In addition, some apps (eg, Fitbit's companion app) use competitive techniques such as leaderboards to persuade individuals to be healthier in a playful sense [38,39]. The vast majority of healthy eating and fitness apps implement functionality for the user to self-track, which refers to recording and analyzing data about oneself on a regular basis to monitor a behavior [40]. Self-tracking is largely believed to only be effective if the monitoring continues [41]. Hence, self-tracking can become a system of constant recordkeeping and monitoring, requiring repetitive behaviors [42]. The literature has shown that self-tracking practices, promoted by healthy eating and fitness apps, can have a negative impact on the well-being of young people, by encouraging addictive tracking behaviors and negatively influencing body image [40]. Theoretical work [43] argues that long-term use of digital technologies to monitor health behaviors may objectify the body, encouraging a mind-body dissociation that results in negative affect and the *anaesthetization* of the human experience. Such experiences may prevent the active adoption and ownership of positive health and fitness behaviors [43].

This Research: Aims and Research Questions

Despite the popularity of self-tracking apps, and the fact that they are largely marketed at younger demographics (with those aged 18 to 29 years being the most regular users [44]), there is a distinct lack of research on young people's experiences of healthy eating and fitness apps and self-tracking practices [45] and even less so on the potential harm that these self-tracking behaviors might have [46]. Our study attempts to address this gap in the literature by highlighting potential areas for concern. By drawing attention to this issue, we hope to encourage the development of future technologies that are sensitive to the sociocultural pressures surrounding diet and exercise that young people are already contending with.

In this paper, we describe a series of engagements conducted with 106 participants over the course of approximately 12 months. First, we conducted an anonymous Web-based survey (63 male and 32 female) to gather an overview of the types of healthy eating and fitness apps young people are using, their reasons for nonuse, and any maladaptive eating and exercise behaviors that they had noticed in response to app use. We then conducted 2 workshops with 8 young people to gain an insight into their overarching positive and negative experiences of engaging with healthy eating and fitness apps. We conducted a further 3 interviews with experts in the domain of eating disorder and body image to gain a deeper understanding of clinical presentations of maladaptive behaviors and how the use of healthy eating and fitness apps might exasperate such behaviors. Finally, drawing from insights gathered during Phase 1, we conducted a review of the top 100 healthy eating and fitness apps on the Google Play Store (ie, for Android apps) to explore the extent to which they had the potential to contribute to maladaptive eating and exercise behaviors highlighted by participants.

Through this body of research, we aimed to address 4 research questions: (1) What are the experiences of both female and, yet underrepresented, male young people around using healthy eating and fitness apps?; (2) How do experts in body image and eating disorder feel about healthy eating and fitness app use among young people?; (3) To what extent do current healthy eating and fitness apps on the market address, or indeed confound, the issues identified by young people; (4) How might the designers of future tools support young people in their healthy eating and fitness goals to ensure that they are developing these responsibly, without them becoming a potential source of issue, especially for those at risk of maladaptive eating and exercise behaviors.

Methods

Phase 1 Methods: Understanding Perceptions and Potential Harm

The research presented in this paper received ethical approval from the research ethics committee of Lancaster University. All participants consented to take part in the study, with survey participants being presented with the study information on a welcome page and workshop and interview participants providing written informed consent.

Survey

To begin our study, we first wanted to gain some broad insights into the types of healthy eating and fitness apps that young

people currently use, their motivations for using healthy eating and fitness apps, and the reasons they might have for discontinuing use of these types of apps. In addition, we wanted to explore if young people currently reported any maladaptive eating and exercise behaviors associated with their healthy eating and fitness app use, to provide an overarching understanding of the possible issues they might have and to better inform the design of our in-depth workshops.

We created an anonymous Web-based survey, distributed through social media using a snowball sampling method. With acknowledgment that the existing literature around body image and maladaptive eating and exercise behavior centers mainly around cisgender female populations [47], we wanted to, where possible, gather a gender-balanced perspective during recruitment. As such, when we released our call for participation, we noted that we were interested in the views of people identifying as both female and male.

Participating respondents (n=95, 63 male) were asked to follow a link to the survey where they were then asked the following: the gender they identified with, their nationality, whether or not they currently use a healthy eating and fitness app (and if not, why), which apps they used now or in the past, their primary motivation for using these apps, and whether they felt the app(s) caused any negative experiences and behaviors. Responses were collected in both free text and tick box format. After a 4-week period, the survey was closed and responses were collated.

Workshops

We next ran 2 workshops to better understand how young people engage with healthy eating and fitness apps and what they perceive to be the dangers of the use of these apps. We wanted to balance the perspectives of both (1) typical users and (2) those advocating for positive body image, who we felt represented current body positive activism taking place in Web-based communities such that that young people may come across in their own social media feed, thus representing a possible positive influence.

Workshop Participants

Despite opening our workshops to both genders, the workshop participants were all female. Participants were recruited via an email advertisement. The first workshop was held with 6 female students recruited from York St. John University: 3 studying psychology (W1-P1, W1-P2, and W1-P3) and 3 studying sports sciences (W1-P4, W1-P5, and W1-P6). The workshop was facilitated by 3 researchers. The second was held with 2 *peer educators* from a local youth organization who had undertaken training explicitly around promoting a positive body image in young people (W2-P1 and W2-P2). Peer educators are trained 14- to 25-year-olds who help youth groups explore important topics such as body confidence and self-esteem, using their own experiences to bring the subject to life [48]. The second workshop was facilitated by 2 researchers who were also present in the first workshop.

All participants were aged between 18 and 25 years and had current or past experience of using healthy eating and fitness apps. Participants were provided with a £10 Amazon voucher as a thank you gift for their time. Each workshop lasted approximately 2 hours. Both workshops were audio-recorded and transcribed with the participants' knowledge and consent for qualitative thematic analysis at a later date.

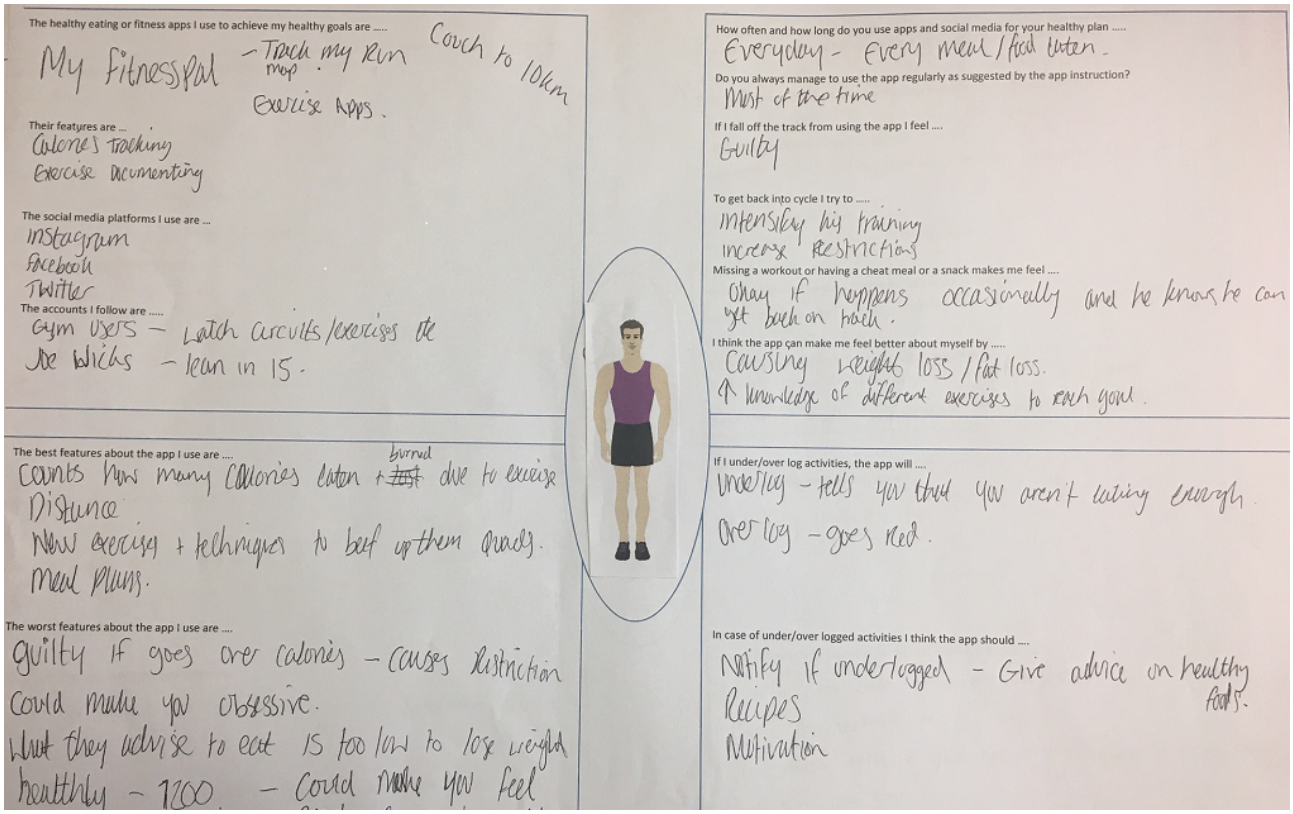
Workshop Activities

Workshops 1 and 2 both followed the same format. We first facilitated a series of activities designed to encourage participants to talk about their experiences of healthy eating and fitness apps. Owing to the potential sensitive nature of the topic area (ie, maladaptive eating and exercise behavior formation), we asked participants to work with personas to allow for a level of dissociation from their own personal experiences [49,50]. For the first activity, using a provided storyboard sheet, (see Figure 1), participants each created a persona that, to them, represented a person who was engaging with healthy eating and fitness apps. They were prompted to think about their persona's health experiences, including their health goals, what they considered to be a healthy body, and their strategies for achieving their health goals.

We then randomly split the group into 2 and asked each group to select a persona to work with for the next activity, which aimed to better understand their relationship with apps and their patterns of use (note that workshop 2 participants each worked with their own personas for this activity). Participants were asked to identify the healthy eating and fitness apps they use to achieve their goals and these apps' features; the duration and frequency of engaging with the apps; whether or not they followed the app's instruction on usage; how they felt when falling off track when using the app; what they do to get back on track; and how the app might make them feel about themselves. We then wanted to know what happens when they under or over log their activities in the app; the app's response in this situation; and what they think the apps should do in this scenario.

Following this, the final activity revolved around having a discussion across the whole group about the effect of following a healthy eating and fitness goal and the use of technology to achieve this goal on the individual's social life. This included the effect on social activities and any impact on friends and families.

Figure 1. Example of persona activities.



Expert Interviews

Following our workshops that allowed us to gather young people's experiences around healthy eating and fitness app usage, we wanted to hear experts' views on the potential impact of healthy eating and fitness app use on body image and self-esteem within young people. This aimed to allow us to map workshop participants' experiences to possible negative experiences and behaviors (ie, what *could* happen if unhealthy eating behaviors become more extreme). We carried out 3 in-depth interviews with experts exhibiting a range of eating disorder and body image experiences (from clinical practice, personal experience, and academia): (1) a clinical psychologist with expertise in child and adolescent mental health (E-1); (2) an eating disorders campaigner, activist, and writer with personal experience of eating disorder who advocates for greater recognition of eating disorder service delivery and the male experience (E-2); and (3) an academic researcher in social psychology with expertise in body image, exercise psychology, motivation, and well-being (E-3).

Each interview was conducted in a semistructured manner, with the topic guide centering around the interviewee's personal or professional experiences with maladaptive eating and exercise behaviors and negative body image; their experience of healthy eating and fitness apps; their views on the impact of healthy eating and fitness apps on young people, including any negative experience and behavior that app use might incite; and their opinions on how future apps might be improved to make them safer to those at risk of maladaptive eating and exercise behaviors. Interviews were all conducted remotely via WebEx and lasted between 30 and 45 min. Each interview was audio-recorded and transcribed for later analysis. Experts were not paid for their time.

Phase 2 Methods: Reviewing Current Apps

The first stage of our study allowed us to identify any negative experiences and behaviors that young people might have when using healthy eating and fitness apps, and understand how these behaviors might be associated with more extreme maladaptive eating and exercise behaviors. Phase 2 of our research drew upon the increasingly common practice of surveying and evaluating mobile apps in targeted health domains [51-56]. The majority of these previous mobile app reviews have focused on engagement, functionality, aesthetics, and perceived quality, using tools such as the mobile app rating scale [57]. This scale facilitates structured scoring on not only user experience but also on the likely impact of an app on user knowledge, attitudes, intention, and action in the target health behavior. However, as our app review is focusing on the potential issues with such features that, for example, promote engagement or change attitudes toward those deemed as *healthy*, we did not feel this was an appropriate tool for our own review. To the best of our knowledge, the app review conducted in this paper is the first to have the explicit goal of identifying the risks for apps targeting healthy eating and fitness.

We conducted a systematic analysis of the top 100 healthy eating and fitness apps that are currently available on the market for young people to freely download and use. We used content analysis to examine the potential of these apps to elicit the negative experiences and behaviors identified in Phase 1. A breakdown of the review methods is given below.

Identifying the Top 100 Apps

We focused our attention on Google Play, the Android app store, because of its market dominance when compared with other operating systems [58]. We first developed a simple tool in Python to allow us to scrape the app store and return the top healthy eating and fitness apps based on a series of filter terms linked to healthy eating and fitness. For eating, there were 20 filter terms in total (eg, calorie, weight, nutrition, and diet), and for fitness, there were 18 terms in total (eg, exercise, fit, flexibility, and training). We focused only on freely available apps as findings have shown that, although young people (aged 18 to 34 years) are more likely to pay for an app than other age groups, the majority still do not pay for any mobile apps at all [59]. Our initial search yielded 540 free healthy eating and fitness apps. We re-ordered the list based on popularity (ie, number of downloads) and then manually reduced the list by disregarding apps that did not match our definition of healthy eating and fitness apps. This included removing apps that were explicitly (1) *wellness*-related (eg, containing the terms yoga, calm, and meditation); (2) route planners for cycling, running, and walking; (3) those that monitor physiological measures (eg, heart rate and blood pressure); and (4) specific to an aspect of health that was not diet or exercise related (eg, menstrual tracking, sleep, drink water reminders, and quit smoking). We then selected the top 100 of the remaining apps for our review.

Review Methods

An app review criteria list was developed based on the themes that emerged from our Phase 1 analysis, reflecting the concerns expressed by experts and users regarding the specific types of logging behaviors facilitated by apps and the way in which apps respond to these behaviors. More specifically, we created a series of reviewing questions that aimed to identify the following:

- General Purpose (Qualitative)
- Data Logging including the types of data entry and behavior logging facilitated by the app (Qualitative)
- Specific Behaviors including the ability to set underweight goals and ability to report low or high calorie consumption (both quantitative: Yes or No) as well as the app responses to these behaviors and other behaviors of interest, including continuous use, not meeting exercise goals, and over and under logging calories (Qualitative)
- Feedback including whether the reviewer perceived the app to rely on mainly positive or negative feedback to motivate users (Qualitative) and whether the app was more focused on what the user had or had not achieved (Qualitative).
- Data Sharing including ability to share data with other users (Quantitative: Yes or No) and across social media (Quantitative: Yes or No) and ability to rank self among other users (Quantitative: Yes or No).

Each app was reviewed by one of the authors who downloaded the app to their personal handset, engaged with the app for a week, and responded to the review questions via a Web-based review form. Apps were downloaded in batches of 10 so that the reviewer was assessing no more than 10 apps at any one time. To assess for interrater reliability of the reviews, a 20% subset of the apps were coded according to the same questions by a research intern, blind to the aims of the study. Responses by the 2 reviewers were then scrutinized by the research team for agreement. For items that generated quantitative data, we followed the guidance of West et al [36] and divided the number of agreements by the number of disagreements. The concordance rate of 88% was comparable with that reported by [36]. For items that generated qualitative responses, these responses were compared for any substantial differences, but none were found. Thus, an agreement between reviewers was found to be acceptable across all review criteria, both qualitative and quantitative.

Results

Data Analysis: Phase 1

A content analysis [60] was conducted on the survey data to look for emergent themes relating to each specific question. The qualitative data collected from the 2 workshops and the expert interviews were coded by 2 members of the research team using Braun and Clarke thematic analysis [61] to identify overarching themes from discussions, capturing the core topics and concerns coming from the data. The analysis was inductive and, therefore, data driven. Coders then worked together and discussed any discrepancies before agreeing on a final set of themes. The identified themes were discussed and refined with input from another researcher to ensure that they were representative of the data and could inform the app review process.

Phase 1 Findings

Survey Findings

In total, we had 95 respondents who took part in the survey, 63 identifying as male and 32 as female. The majority of respondents were based in the United Kingdom (75), but we also had responses from the rest of Europe (9), United States (6), Canada (3), and China (2). Respondents reported using a range of healthy eating and fitness apps (see Figure 2 for details), but MyFitnessPal, a calorie-tracking diet-focused app, was by far the most widely used (n=46). There were also a wide range of other lesser known apps that individuals (1 to 2 people in each instance) reported using (n=22). Respondents also reported a variety of reasons for engaging with apps (see Table 1 for details). Calorie-tracking (31%), weight loss (20%), and exercise (21%) were the 3 most common reasons. Please note that this was a free text box and, as such, participants could select multiple motivations.

Figure 2. Healthy eating and fitness app use as reported by survey participants. The 7 most popular apps are used by between 5.2% and 48.4% of respondents. A further 23.2% listed other apps (not itemized here because of their limited popularity).

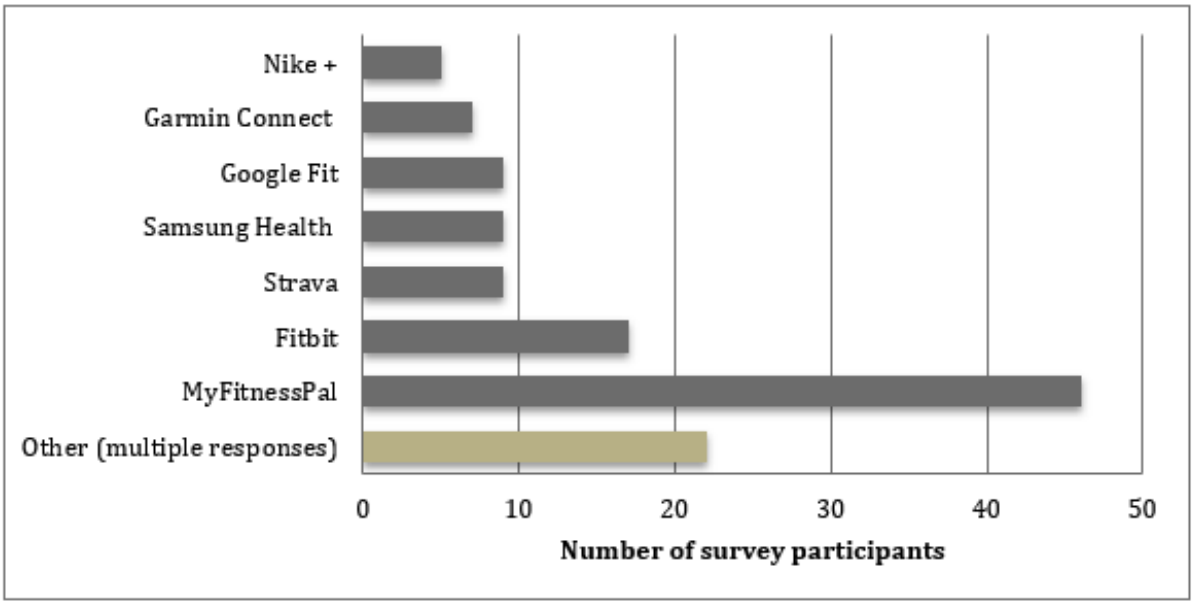


Table 1. Reported motivation to engage with healthy eating and fitness apps.

Reason provided	Female, n	Male, n	Total (N=95), n (%)
Track calories	13	18	31 (33)
Track macro nutrient values	4	5	9 (9)
Exercise	10	10	20 (21)
Weight loss	11	8	19 (20)
Track routes or time for specific activities, for example, cycling or running	0	10	10 (10)
Self-improvement (eg, improved body image)	1	10	11 (12)

App Nonuse

A total of 32 participants (21 male and 11 female) reported that they no longer used their healthy eating and fitness app (see Table 2 for details—as above, participants could provide multiple responses). Respondents' primary reasons for discontinuing use of an app were that it was (1) too demanding on their lives, reporting that they were "too busy" or that it took "too much time to track" and (2) that they lacked motivation, discussing *laziness*, *apathy*, and a *lack of discipline* as contributing factors. Males were more likely to report that they had met health goals, for example, "no longer needed after my weight loss" and "my goal was to get in shape". Other responses for male respondents focused mainly on challenges with the technology, for example, "Background processes making phone run slower", where for 2 of the females, it was an eating disorder that caused them to halt their use: "I have an eating disorder and it was toxic"; "Eating disorder recovery".

Reports of Negative Experiences and Behaviors

A total of 41 respondents indicated that they had experienced some form of negative experiences and behaviors through their app use (see Table 3 for details—as the above participants could provide multiple responses). Of those, the most commonly reported behavior across both males and females (n=15) was obsessive calorie counting or logging, with comments such as "I didn't like MyFitnessPal as this was about counting calories and I became obsessive. It wasn't manageable"; "I got obsessive over calorie counting"; and "Want to log everything I eat". There were also reports relating to guilt and food restriction: "constant guilt"; "MyFitnessPal can cause a self-judgment based on bad diet days"; "yes. Starvation"; and around impacts to social life; "Purposefully taking time away from studying and social activities to exercise"; and "stopped socializing (drinking) as much, avoided eating out". Several respondents discussed general impacts to mood, indicating that they became "miserable", "anxious", and "demoralized". The males in particular also discussed feelings of *disappointment* and *failure* when not achieving their goals.

Table 2. Reasons provided for app nonuse.

Reason provided	Female, n	Male, n	Total (N=32), n (%)
App use too demanding on their life	4	5	9 (28)
Lack of motivation	4	5	9 (28)
Health goals were met	0	5	5 (16)
Other	3	3	6 (19)
No response	0	3	3 (9)

Table 3. Reports of negative or obsessive app use behaviors from respondents.

Behavior reported	Female, n	Male, n	Total (N=41), n (%)
Obsessive counting or logging	8	7	15 (36)
Guilt and food restriction	3	2	5 (12)
Impact on social life	3	2	5 (12)
Impact on mood	2	5	7 (17)
Cheating the technology	2	0	2 (5)
Obsessive, but the push needed	2	2	4 (10)
Non applicable response	0	3	3 (7)

However, for a handful of respondents, there was a feeling that, although the app led to obsessive exercise behaviors, these were the push that they needed to succeed, forcing them to do exercise when tired: “You have to be a bit obsessive to get good results with a calorie counting app”; “Negative, no. Obsessive, possibly. In the case of exercise, there is aspects of obsession anyway”.

For 2 of the females, there was a discussion of cheating the technology, for example, “I wouldn't enter ‘bad’ foods into the app”, where for 3 of the males the negative feelings they harbored were relating to the technology itself and annoyance that was felt when their effort was not recorded: “Yes - if steps aren’t recorded it feels like they haven't happened”.

Workshop and Interview Findings

The following sections describe a synthesis of findings from workshops 1 and 2 and the expert interviews. A total of 4 overarching themes emerged from our thematic analysis: (1) factors relating to motivation and guilt surrounding goal attainment; (2) impact of healthy eating and fitness behaviors on social relationships; (3) app response to logging behaviors, in particular when these might be seen as negative; and (4) a need to reconnect to the body and build self-awareness around healthy versus unhealthy behaviors. We append quotes provided with participant numbers and an indication of whether quotes are from workshop participants (eg, W1-P2) or our experts (eg, E-1).

Goal Attainment: Positive Motivation or Guilt Inducing?

Participants discussed a range of healthy eating and fitness apps that they used; however, as can be seen in the data presented, the ones that massively dominated conversation centered around MyFitnessPal and FitBit. There was much discussion around the qualities that these apps had and differences between the active and passive tracking functions they exhibited. All of the participants, in different ways, highlighted how they favored apps that passively track physical activities, compared with ones requiring users to actively log exercise or food intake data. For example, W2-P2 described:

> *Myfitnesspal is a lot about weight, it’s mainly about logging your food and your calories, whereas fitbit I think is a really positive tool [...] people use them to just increase your daily step count, make yourself just more generally physically active.*

Continuing the discussion over active versus more passive tracking, W2-P2 noted how the burden of inputting individual food items, particularly in home-cooked meals, could actually lead to a somewhat counterintuitive shift toward readymade food: “you’d definitely eat more processed things, ‘cos they’re easier to log ‘cos it’s got a barcode” (W2-P2).

Apps that implemented a positive approach to encourage users for their effort were seen to be more appealing. Participants noted that receiving “little badges” (W2-P2) or a simple message such as “nice one, you’ve done that well” (W1-P5) could be *rewarding* or *motivating*. There was a sense that apps should send positive notifications regardless of whether the daily goal was achieved:

> *[On Fitbit] at the end of the day if you don’t hit your daily step target [...] it just goes back to the next day, but then if you do hit it, it goes green and it’s like “woohoo”. But MyFitnessPal, I think it’s like addictive methods of trying to get you to use it to reach your goals.* [W2-P2]

However, this concept of achieving set goals was not always seen to be a positive thing. Participants described how fixating around the attainment of a goal could lead to obsessive behaviors: “You don’t want to do badly, so it makes you more obsessive of reaching your goals” (W1-P1) and how not reaching a specific goal could lead to feelings of distress and guilt: “The

worst features of the apps are that he [the persona character] feels guilty if he goes over calories and that can cause further restriction and it's made him obsessive" (W1-P4), which could in turn lead to negative counter-behaviors, such as meal skipping: "you could reach your calories at like 3pm and then you don't feel like you could eat anything else for that day" (W1-P1).

E-3 described how the prescription of goals—such as hitting 10,000 steps a day—may overshadow any intrinsic motivations for exercising (ie, with focus being placed on reaching a target number rather than factors such as improving fitness or mental clarity):

> *The monitoring apps can make you think about physical activity in quite a sad way, or quite a compulsive way. If you have a step counter actually you think "I'm going to go for a walk to get my step count up" and your motivation behind doing it is quite external, it's not about enjoying the walk for example, it's not about getting some fresh air or just clearing your head for a minute, actually suddenly that walk is all about "well this is going to get me about 3000 steps if I do this 20 minute walk at lunch time" for example.* [E-3]

E-1 discussed how a focus on these aspects of quantification and goal attainment could become a potential risk for the emergence of maladaptive eating and exercise behavior, if paired with obsessive personality traits or perfectionism:

> *People who tend to have an obsessive personality and also high levels of perfectionism will sustain their use of fitness apps and count every calorie, count every step, count every episode that they go to the gym and all their exercise and use a mobile platform to collect that data.*

He noted how, in clinical care, "we actively discouraged [people with eating disorders] from using those apps and various wrist monitors that count the steps and activity levels" as they can become "triggering" as well as sustaining "obsessional and restrictive and rigid behaviors". E-2 further highlighted this point by describing his own personal experience of using healthy eating and fitness apps to track his food intake and exercise levels:

> *I got to this point where I felt like I was really in control and I found the counting of the calories wasreally empowering...especially running and walking [...] you can put it on and it tracks you and it gives you your speed, and I was really fixated on how far I was going and how fast I was going.* [E-2]

E-2 also highlighted how this eventually led to him losing a sense of control:

> *I no longer felt in control, it sort of started spiraling and I was losing like five pounds, six pounds.*

At this point, he realized that he has established an unhealthy relationship with healthy eating and fitness apps:

> *It all had to sort of go because I just realized how unhealthy it was and I know that I can't really use those kind of apps in a healthy way.*

Impact on Relationships

Participants discussed how heavy dieting and exercise regimes could have a negative impact on maintaining a good social life; many social gatherings focus on food, and often, a lot of time is required to maintain a heavy gym regime. W2-P1 described how, for some people, socializing becomes challenging because of the need to make dietary sacrifices:

> *If you're very religiously logging it and then you think oh I can't go out to that restaurant or I don't want to go out for drinks.*

The concept of religiously monitoring caloric intake in an attempt to lose weight was highlighted by E-1 as a key challenge that could ultimately lead to social isolation and impact of the psycho-social development of a young person:

> *For a lot of people it becomes an obsessive, all-consuming pursuit and therefore all energy is invested in that and there is no space left to do other things like invest time in social relationships, in social gatherings, education, family, friends and other sorts of developmental aspects of growing up.*

It was noted how engagement with apps on a mobile phone is largely, by design, a personal and private experience, which could in fact intensify negative experiences and behaviors, as it is easy to hide from family and friends:

> *...by the time you've noticed it they're probably already obsessive, because apps are actually very easy to hide that you're using it* [W1-P1]

W2-P2 also described how easy it is for young people to engage in secretive behaviors without anyone finding out:

> *It's on a phone where you've got a password. Sure, people can have visibility of it but they may not even know you have it. So say if it's a young, like a teenager, their parents may not even know they're using it, why would they, it's on a locked phone [...] I think you can quite easily start to like manipulate your calories in a negative way and nobody would know you're doing it.*

App Response to Logging Behaviors

Participants discussed the issue of under logging calorie intake and how an app may respond in such cases:

> *It comes up in red basically giving you a warning that you're not eating as much, or, if you carry on like this you'll be losing two pounds in a week.* [W1-P2]

However, they discussed that providing such weight loss projection could motivate users to eat less to lose weight faster within the estimated time frame. Participants explained how, when the user underlogs, apps currently only respond with a warning notification, instead of taking any actions to prevent this from happening. For example, W2-P2 described how she restricted her calorie intake to 1200 calories per day for 7 months without any guidance from the app to re-evaluate her goals:

> *It didn't come up like "you're doing this and you should not be doing it" kind of thing, so I think it is negative, it should prompt you to re-evaluate your weight and your goals but it doesn't, it automatically recalculates your weight, but you can keep changing your end goal and then I'm not sure it brings up like well this is in an underweight category of [Body Mass Index] BMI.*

Our health experts also envisioned apps taking more responsibility to protect users, particularly those at risk, by not just focusing on sending notifications to encourage achieving the daily goal but also by notifying when overuse happens:

> *I think apps could be more intelligent and could give feedback about over-use [...] if you spend too much time on your fitness app or on your diet app or too much time in the gym if you are monitoring all your activities, one could have a warning come up that this might not be the most healthy thing to do.* [E-1]

On the contrary, the need for a *gentle* approach to encourage engagement was also seen as beneficial:

> *If someone's not doing as much physical activity, having an app which kind of brings you back in gently rather than saying "you haven't exercised for five days, what are you doing," actually that could be really helpful, so trying to think about encouraging people to do those healthy activities rather than worrying too much about what those outcomes are going to be.* [E-3]

A Need to Reconnect to the Body

Each of the experts highlighted the importance of being able to *listen to your body* and gain the power to *switch off* when needed. They discussed the notion of feeling controlled, or being driven, by apps and how one should have a period of detox if this happens. E-1 highlighted the need for a psycho-educational component about the harms and benefits of healthy eating and fitness apps, explaining that young people in recovery learn to disconnect themselves from their phones:

> *...trying to live with the anxiety of not knowing the detail of their activity levels and input and outputs; that is something pretty much promoted in recovery.*

He further explained how they aim to teach young people to trust their own bodies and themselves, to look after themselves:

> *It's natural for one to allow your body to find its own equilibrium and its own input and output in terms of food and activity levels, and to rest when you feel the need to rest and to exercise when you feel the need to exercise.* [E-1]

This view was also shared by E-2 who explained the need to be mindful around logging behaviors:

> *Am I checking this because of anxiety and habit or because I want to use it? How autonomous am I in doing this? [...] whether you feel compelled to use it, I think people need to work out whether that's the case for them or not, like having that awareness is really important and that's what gives people the power.* [E-2]

This was echoed by E-3:

> *I try to think about what am I doing, why am I doing this, what's this app encouraging me to do, is it actually doing what I want it to be making me do or is it making me do something negative.* [E-3]

It was acknowledged by E-2 that this can be counter-intuitive to the way healthy eating and fitness apps are currently designed, with users often being actively encouraged to log and monitor themselves constantly: "if you don't [log] then you mess up your statistics". He explained how, on one occasion, forgetting his phone at home and not having the ability to track his steps made him feel anxious:

> *I felt more anxious because I didn't know how much I'd done and I needed the steps, like the app to tell me that I'd done enough exercise, rather than listening to my body, which I think is a big thing with these apps and social media—they're prescribed—I think that sort of disconnects people from their bodies a little bit.* [E-2]

In summary, survey results showed that almost half of our respondents had some form of negative experience using healthy eating and fitness apps. Approximately one-third of respondents reported that they had stopped engaging with healthy eating and fitness apps, typically citing that they found them too demanding or that they lacked the motivation for long-term use. Workshops and interviews further explored personal and professional perspectives on the impact of healthy eating and fitness apps on young people. Participants described how the pressure of attaining health goals, set through the app, could lead to negative experiences and behaviors. There was much consensus that healthy eating and fitness apps can bring feelings of guilt leading to negative experiences and behaviors such as meal skipping. The quantification element of the healthy eating and fitness apps was also seen as a contributing factor for development of maladaptive eating and exercise behaviors, particularly for those with obsessive personalities. Participants concerningly discussed the possible impact of obsession with logging behaviors on social isolation. This was backed up by experts highlighting the importance of being mindful around logging behaviors.

The next stage of our research then focused on exploring the specific features of healthy eating and fitness apps that are currently available on the market, including the ability to log behaviors of interest and how the app responds to user engagement and behavior logging. We wanted to understand the potential of the top 100 healthy eating and fitness apps to elicit the negative experiences and behaviors reported in the first stage.

Table 4. Coding criteria and frequency among apps.

Coding criteria	n
General purpose	
Exercise promotion	62
Diet	10
Exercise and diet	25
Appearance goals	65
Data logged	
Exercise completed	84
Food consumed	23
Weight	56
Specific behaviors	
Allows underweight goal setting	21
Allows reporting of high or low calorie intake	16
Response to low calorie intake	8
Response to high calorie intake	8
Response to not meeting exercise goals	20
Rewards continuous use	60
Feedback	
Overall positive	72
Overall negative	19
Achievement-focused	49
Failure-focused	26
Data sharing	
Other users	32
Social media	59
Automatic ranking	6

Data Analysis: Phase 2

Data from the reviews were collated into a spreadsheet. For review responses that resulted in qualitative data, we performed a qualitative content analysis [62]. As follows, responses were reviewed by 2 members of the research team to identify key themes. Then, the frequency of occurrence of these themes among the apps was coded by these 2 members of the research team individually and summed to produce a score signifying the frequency of that theme with the data set. If there was insufficient information in the text provided in the initial review data to determine whether themes were present in apps, then the coders downloaded the app to clarify whether the app met the coding requirements. Interrater reliability between the 2 coders was again assessed by dividing the number of agreements by the number of disagreements, resulting in an acceptable concordance rate of 90% [36]. The coding criteria developed through this process and frequency of occurrence within apps are shown in Table 4.

For reviews that resulted in quantitative data, responses were summed to produce scores that reflect the number of apps belonging to each possible category. Interrater reliability of this coding had already been determined in the previous review phase (see Table 4).

Phase 2 Findings: App Review Findings

General App Features

Descriptions of apps were coded according to the health behavior targeted by the app. The vast majority of the apps focused on exercise promotion (n=62) and exercise and diet combined (n=24). A total of 3 apps were found to not focus on either diet or exercise. This included 1 app that focused on editing the body to assess how one would look like with various bodily enhancements (Body Editor–Breast Enlarger, Body Shape Editor) and 2 that focused on weight monitoring (eg, Monitor your Weight). Around two-thirds of the apps included the ability to set appearance-related goals, such as weight loss and enhancement in muscle tone (n=65).

The majority of apps collected some form of data pertaining to physical activity (n=84). This included data such as daily steps, cycling, and the completion of workouts specified by the app. A much smaller percentage of apps collected dietary data

pertaining to food consumption (n=23). More than half requested user data pertaining to weight (n=56).

Specific Behaviors

In total, 21 apps allowed the setting of underweight body goals (corresponding to a BMI<18.5). Qualitative responses to other questions highlighted how this could include setting extremely low BMI targets in some apps: “I still can set a BMI goal<13”.

Around two-thirds of the apps that facilitated the recording of dietary intake allowed the reporting of calorie consumption that was 50% more or 50% less than the recommended daily amount for the average individual (n=16). Some apps responded differently to the entering of low-calorie consumption (n=8), for example, highlighting that a goal had not been met or not allowing the calories to be logged. Others had a minimum calorie entry threshold and would not allow calories to be logged under this target. Some apps treated the under logging of calories the same as logging calorie content in line with their goals (n=8), including those that praised under logging calories in the same way that they praised calorie entry consistent with goals (eg, “You made that look easy”—Lose it!). A small proportion of apps also responded differently to logging calories that exceeded the daily target (n=8). Example responses to exceeding calorie goals included numeric feedback, reminders about goals, and visual feedback such as turning the calorie counter red or sad face emoji. Furthermore, some apps had a more questioning approach, for example, asking if the calorie entry is correct.

A fifth of the apps responded to failure to meet exercise goals (n=20). The nature of these responses varied and included reminders, numerical feedback. Similar to when logging a high calorie intake, some apps adopted a questioning approach (“Wanna give up? Think about why you started?”). Two-thirds of the apps offered praise and rewards for continuous use (n=60). Qualitative responses as to the nature of this praise varied. Many offered direct messages of praise such as congratulations messages or fireworks. Elements of gamification were present in some apps, with trophies, badges, and the unlocking of levels being used to reward continuous use.

Feedback

Apps were coded according to whether positive or negative feedback was used to motivate users’ engagement with the app. The majority focused on positive feedback (n=72), whereas a minority of apps were felt to focus on negative feedback (n=19). The remainder of apps either provided neutral feedback (n=8) or provided both positive and negative feedback to the extent that the reviewers could not differentiate between which was being used more in the app (n=1).

Examples of both positive and negative feedback were provided in qualitative responses to other questions (eg, how the app responds to specific user behaviors). Some apps sent positive feedback to encourage users to keep up with their initial goal “Hey [user name]! You’ll feel great after a workout. Strive for progress, not perfection” or suggested that the user adjust their personal goal so that it was more in line with their lifestyle “You haven't been active lately, do you want to adjust your goal”. An example of negative feedback included users being punished for not reaching their personal target “if you miss a day you'll get punished by losing a heart”.

Furthermore, apps were coded as to whether they focus on what the user has achieved or what they have not achieved. Just under half of the apps were positively focused on achievements (n=49), highlighting what the user had done, whereas around one quarter were negatively focused on achievements (n=26), highlighting what the user had not done. The remainder did not focus on achievements (n=25) and included apps that did not involve goal setting or targets and those with predetermined reminders not linked to user behavior.

Data Sharing

Around one-third of apps facilitated the sharing of data with other app users (n=32). The remaining apps either did not facilitate this (n=66) or did not collect any user data (n=2). A much higher proportion of apps facilitated the sharing of app-related achievements through social media sites (n=59), whereas the remainder did not (n=39) or did not collect any data (n=2).

A small proportion of apps automatically ranked users among others (n=6) but all of these apps allowed this feature to be deactivated. This ranking approach was evident in the qualitative responses about app feedback. For example, a push-up workout app compared user reports about their workouts with other users in similar categories (based on BMI, gender, and age): “you did X push-ups, this is better than XX% of users today”.

In summary, we wanted to identify the logging functionalities available in the top 100 healthy eating and fitness apps that have the potential to elicit negative experiences and behaviors captured in the first stage. We found that apps responded differently to data-logging behaviors, for example, some allowing users to under log their calorie consumption whereas others using guilt-inducing techniques when calories exceed the daily goal. Alarmingly, 21% of the apps we reviewed allowed underweight goal setting (BMI<18.5). In addition, one quarter of the apps focused on negative reinforcement for maintaining health goals. A substantial number of apps facilitate sharing data with other users (n=32) and across social media (n=59).

Discussion

Principal Findings

The aim of this research was to understand the potential role of healthy eating and fitness apps in the development of maladaptive body-related attitudes and eating and exercise behaviors in young people. Drawing on the findings from Phase 1 and Phase 2, we have identified 5 important ways in which healthy eating and fitness apps may potentially exert negative impacts on users. In this section, we reflect on these themes and how future research may mitigate against these issues.

The Attraction of the Number Game

Calorie counting is a prominent feature of many currently available healthy eating and fitness apps and a key reason why many of the survey respondents, both male and female, reported using healthy eating and fitness apps. However, participants in

both the survey and workshops also described how obsessive thoughts and behaviors around calorie counting and logging could emerge. For example, not achieving a calorie count within a certain boundary could cause feelings of failure and guilt, often then leading to restrictive eating or excessive exercising behaviors (eg, not allowing themselves to eat after 3 pm), simply because the app indicated that they had already reached their caloric intake for the day. This tendency to engage in purge behaviors echoes findings from previous research examining the consequences of healthy eating and fitness app use [32] and food journaling [20,63].

In addition to calorie counting, there were other aspects of self-tracking evident in the apps reviewed (eg, exercise tracking and weight tracking). The primary objective of self-tracking is around obtaining *self-knowledge through numbers* [64], and these types of self-tracking practices are used successfully by large numbers of people to achieve their goals. Indeed, participants in both our study and previous research have expressed experiencing acute benefits from short-term self-tracking [28]. However, the act of self-tracking was often seen as a *numbers game* for the young people in our study, becoming a law for them to live their lives by. In the workshops, participants describe how self-tracking led them to over focus on the calories contained in food at the expense of nutrition (eg, prioritizing ready meals as this was easier to log). Furthermore, some survey respondents described cheating the app by not accurately logging their food consumption. Although healthy eating and fitness apps offer new ways of monitoring, measuring, and representing the human body [64], they also contribute to a new way of conceptualizing one's health status and introduce a new language around digitizing one's health, making it easier to trust *the numbers* over physical observations [64]. This obsession with numbers can facilitate a shift toward unhealthy behavior in contrary to the healthy eating and fitness apps' initial goal of promoting health.

Reflecting on these points, we need to think about better ways to log food and exercise. For example, [65] have used picture logging of food plate for older people to help with monitoring of nutritional value of daily food intake in the care home environment. In the context of self-tracking, [66] also have investigated photo-based journals as a simplified self-tracking method. This style of easy logging and calorie estimation might be a more forgiving approach to take.

Furthermore, our experts highlighted a need for approaches that promote listening to one's own body, and its nutritional needs and physical limitations, rather than becoming reliant on the attainment of what are often arbitrary numerical goals. Such an approach is consistent with newly emerged evidence-based psychological approaches to improving exercise and eating behavior that are focused on the development of embodied and mindful approaches to eating and exercise [67]. Reframing the way that young people think about eating and exercise, by moving away from language embedded in self-quantification and self-objectification, could be an important first step toward understanding how we might achieve this within future technologies that support healthy eating and fitness.

Over Emphasis on Extrinsic Appearance (Including Weight Loss) Goals

A high percentage of healthy eating and fitness apps (almost two-thirds) currently available on the market focus on the achievement of appearance-related goals rather than health goals. This finding is consistent with research conducted around other health and fitness media, such as magazines [68].

Focusing on appearance-related goals in eating and exercise settings has been linked to negative body image and maladaptive eating and exercise behavior [69]. In contrast, appreciating the functionality of one's body rather than its extrinsic appearance can enhance well-being and reduce vulnerability to environmental triggers of body dissatisfaction [70,71]. Similarly, focusing on more intrinsic factors when eating and exercising is more conducive to long-term physical and mental health [72,73]. Future apps that focus more on the intrinsic value of exercise and eating behaviors (eg, how food and exercise makes us feel, how much enjoyment we get from the experience) rather than extrinsic weight loss and appearance could be a powerful tool to supporting healthier attitudes toward food and exercise.

Relatedly, and of particular concern, is the fact that around a fifth of the apps we reviewed allowed underweight goal setting. As this is a relatively underresearched area, it is unclear the extent to which the ways that an app's reaction to underlogging of calories might affect individuals (eg, if a user is being notified that they are not reaching their underweight goals or if they are trusting that a healthy eating and fitness app would not allow than to set an underweight goal in the first place). However, it is possible that these types of responses might serve as *thinspiration* (ie, digital content that is ostensibly designed to inspire thinness and is commonly found in Web-based eating disorder communities [74]), motivating weight loss among individuals who already display problematic attitudes toward eating and exercise. Future research should examine this further; however, a simple safety measure could be implemented to ensure that users are unable to set unrealistic underweight goals (ie, those that would result in a BMI less than 18.5) in the first place and that they are provided with relevant information about the dangers of setting underweight goals if they attempt to do so.

The Power of Notifications

The literature on behavior change, and the roles of notifications in supporting change, is extensive [35,75,76]. Indeed, in our study, we found that participants appreciated notifications, particularly those containing positive content, and generally considered them to be a motivational tool. However, our app review also, concerningly, saw several examples of apps that were seen to enact punishments on the users or act as a constant reminder of the failure to achieve daily goals. In the context of body-modification strategies for weight loss or appearance-enhancement motives, guilt is generally regarded as problematic [77].

In our study, young people openly discussed the negative emotions they have experienced when interacting with healthy eating and fitness apps. They expressed feelings of guilt,

disappointment, and pressure. This echoes findings from [28,63], where feelings of guilt were associated with not achieving goals.

Although the aim of notifications is to encourage engagement, previous research suggests that these are not always effective and can instead cause additional stress on young people [63]. Rather than focusing on temporal reminders to complete a daily task, apps should encourage a process of self-reflection and habit formation. For example, users could be encouraged to think in advance about occasions that could interrupt the attainment of daily goals, facilitating a personalized approach for allowing users to plan strategies for how to cope when this happens. Following a preplanned strategy would avoid *extreme measure* behaviors (eg, not eating anything for the rest of the day once calorie allowance has been reached), instead promoting better coping mechanisms (eg, if a workout is missed in the evening because a user is trying to meet a deadline, they can plan to conduct the exercise the next day). Consequently, engaging users in the creation of preplanned statements that mitigate negative feelings such as guilt triggered by not meeting goals might be a useful way of supporting users.

Social Impact

Healthy eating and fitness apps offer the opportunity to enhance the social life of the user by giving access to group memberships and opportunities to find like-minded people or friends on the internet (sharing and comparing data through the app) and offline (engaging in physical activities with groups of preexisting friends or new ones through the app) [78]. Many apps provide the opportunity for users to rank themselves among other users, thus encouraging competition. For example, FitBit has a community section where users can add each other as friends and *compete* toward weekly step counts. Research suggests that perceived social norms can be a powerful motivator of behavior change in relation to eating and exercise, in that individuals are more likely to engage in behavior if they perceive others to be doing it [42]. However, comparing the self with others can also be demotivating and detrimental to self-perceptions, especially if users perceive themselves as comparing unfavorably with others [33].

Our app review indicates that the majority of apps facilitated social features, that is, data sharing with other app users (n=32) or through social media (n=59). However, although healthy eating and fitness apps are seen as socially mediated experiences to connect people with common interests and goals for healthy eating and physical activity [78], our qualitative data suggested that individual app engagement can result in antisocial behavior (eg, not attending social functions) that can lead into loneliness and impact the psycho-social development of a young person, for whom the development of social relationships is an important and psychologically salient developmental task [79]. Reflecting on these points, it is important that future healthy eating and fitness apps not only facilitate creation of a digital connection and communities but also provide tools to encourage social connectivity. This might be through the provision of a space to reflect on their social interactions or by providing them with practical tips and strategies (eg, advance planning for social occasions or integrating healthy eating and fitness activities in people's social lives).

Consequences of Sustained Use

Most healthy eating and fitness apps are designed to reinforce continuous use. Typically, this is achieved through gamification features that render daily routines into games [80]. Gamification is defined as the use of game design elements (eg, game-like rewards, leaderboards [81], and storytelling [82]) in nongame context (eg, healthy eating behavior) [83]. Gamification is increasingly being used in healthy eating and fitness apps, yet the effectiveness of this technique has only been assessed against revenue generation and not long-term behavior change [84].

Although the aim of gamifications is to promote sustained use, studies have shown that long-term engagement is not always achievable, for example, only 11.6% of app users in Great Britain are still engaging with an app a week after installing it on their phones (findings from other countries are similar) [85]. Our survey findings indicate that most app users do not report long-term engagement with apps, typically withdrawing from apps because of finding them too demanding, lacking motivation, boredom, annoyance of notifications, or feeling negative emotions when failing to meet the set targets. This was further supported in the workshops where participants described how the reinforcement of continuous use can have unintended negative consequences, such as experiencing guilt and experiencing the desire to stop using the app entirely.

One possible option for the future app developer would be to consider encouraging more long-term behavior change, rather than focusing on engagement [86]. Furthermore, given that obsessive engagement with apps may reflect problematic use, apps could provide informative feedback based on usage patterns (eg, frequency of app checking or time spent engaging with app) that might be indicative of problem use. For example, to overcome both the obsession and indeed lack of motivation and boredom, the apps could encourage a *period of detox* when unusual patterns of use are detected, as part of the gamification model. In addition, providing a framework within the app that encourages users to self-reflect on their use patterns and to re-evaluate their goals regularly may help users who might be struggling to re-gain a sense of much needed control.

Limitations

This research adopted a mixed-methods approach to develop a holistic understanding of the potential negative impact of healthy eating and fitness app use among young people. Though we aimed to recruit both male and female participants to engage in our qualitative research, we struggled to recruit individuals identifying as male to engage in workshop activities. That said, male perspectives are still represented in the survey data and in fact made up the majority of respondents. One reason for this could be that we recruited largely from computer science, which is a predominantly male field; however, for future work, a gender balance would be preferable. We also did recruit a male eating disorder campaigner, activist, and writer with a personal experience of eating disorder who advocates for greater recognition of eating disorder service delivery and the male experience to participate in the interviews. However, future research should aim to more deeply explore the male experience of healthy eating and fitness app use further. Similarly, it may be important for future research in this field to consider how

other individual difference variables, such as race, class, body size, and disability, affect participants' engagement with healthy eating and fitness apps and their potential for misuse.

Conclusions

Through our study, we have offered a deepened understanding of young people's experiences of healthy eating and fitness apps and the potential harm that their use might have. We have offered a set of guidelines for future apps that can be responsibly developed to prevent the formation of maladaptive eating and exercise behaviors. Although we understand that an app developer would never set out to meaningfully bring about negative emotional responses in their users, we must also be mindful of the fact that these responses are happening. Through this study, we hope to open a dialogue around how use of these apps could have the potential to become the seed that develops into a more serious issue. As the target user demographic for these types of apps, young people are most vulnerable to the development of poor body image and maladaptive eating and exercise patterns. As such, we need to take care in the type of language we are using and the type of sustained behaviors we are promoting.

Acknowledgments

The authors would like to thank all study participants for sharing their perspectives in the course of this study. In particular, Dr Leo Kroll for his continuous support throughout the project. This study presents independent research funded by the Engineering and Physical Sciences Research Council-National Institute for Health Research NewMind Network Plus. The views expressed are those of the authors and not necessarily those of the NewMind Network. The funders played no role in the study design, conduct, or interpretation of the data. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

References

1. Polivy J, Herman CP. Causes of eating disorders. Annu Rev Psychol 2002;53:187-213. [doi: 10.1146/annurev.psych.53.100901.135103] [Medline: 11752484]
2. UNESCO. What do we mean by "youth"? URL: http://www.unesco.org/new/en/social-and-human-sciences/themes/youth/youth-definition/ [accessed 2019-06-07] [WebCite Cache ID 78wqoJ4fR]
3. Rodin J, Silberstein L, Striegel-Moore R. Women and weight: a normative discontent. 1984 Presented at: Nebraska Symposium on Motivation; February 1984; Nebraska.
4. Tantleff-Dunn S, Barnes R, Larose J. It's not just a "woman thing:" the current state of normative discontent. Eat Disord 2011;19(5):392-402 [FREE Full text] [doi: 10.1080/10640266.2011.609088] [Medline: 21932970]
5. Hausenblas H, Fallon E. Relationship among body image, exercise behavior, and exercise dependence symptoms. Int J Eat Disord 2002 Sep;32(2):179-185. [doi: 10.1002/eat.10071] [Medline: 12210660]
6. Neumark-Sztainer D, Paxton S, Hannan P, Haines J, Story M. Does body satisfaction matter? Five-year longitudinal associations between body satisfaction and health behaviors in adolescent females and males. J Adolesc Health 2006 Aug;39(2):244-251. [doi: 10.1016/j.jadohealth.2005.12.001] [Medline: 16857537]
7. Smolak L, Levine M. Body image, disordered eating, and eating disorders: Connections and disconnects. In: The Wiley Handbook of Eating Disorders. United States: Wiley Online Library; 2015:1-10.
8. Levine M, Murnen S. "Everybody knows that mass media are/are not [pick one] a cause of eating disorders": A critical review of evidence for a causal link between media, negative body image, and disordered eating in females. J Soc Clin Psychol 2009 Jan;28(1):9-42. [doi: 10.1521/jscp.2009.28.1.9]
9. Keel PK, Eddy KT, Thomas JJ, Schwartz MB. Vulnerability to eating disorders across the lifespan. In: Vulnerability to Psychopathology, Second Edition: Risk Across the Lifespan. New York: The Guilford Press; 2010:489-494.
10. Culbert K, Racine S, Klump K. Research review: what we have learned about the causes of eating disorders - a synthesis of sociocultural, psychological, and biological research. J Child Psychol Psychiatry 2015 Nov;56(11):1141-1164. [doi: 10.1111/jcpp.12441] [Medline: 26095891]
11. Keel P, Forney K. Psychosocial risk factors for eating disorders. Int J Eat Disord 2013 Jul;46(5):433-439. [doi: 10.1002/eat.22094] [Medline: 23658086]
12. Keery H, van den Berg P, Thompson J. An evaluation of the Tripartite Influence Model of body dissatisfaction and eating disturbance with adolescent girls. Body Image 2004 Sep;1(3):237-251. [doi: 10.1016/j.bodyim.2004.03.001] [Medline: 18089156]
13. Deighton-Smith N, Bell B. Objectifying fitness: a content and thematic analysis of #fitspiration images on social media. Psychol Popular Media Cult 2018 Oct;7(4):467-483. [doi: 10.1037/ppm0000143]
14. Willis L, Knobloch-Westerwick S. Weighing women down: messages on weight loss and body shaping in editorial content in popular women's health and fitness magazines. Health Commun 2014;29(4):323-331. [doi: 10.1080/10410236.2012.755602] [Medline: 23844558]

15. Damiano S, Hart L, Paxton S, Cornell C, Sutherland F. Are body image and eating attitudes, behaviours, and knowledge of parents of pre-schoolers associated with parent feeding practices? J Eat Disord 2014 Nov 24;2(S1). [doi: 10.1186/2050-2974-2-S1-O34]
16. Huon G, Lim J. The emergence of dieting among female adolescents: age, body mass index, and seasonal effects. Int J Eat Disord 2000 Sep;28(2):221-225. [Medline: 10897085]
17. Neumark-Sztainer D, Wall M, Guo J, Story M, Haines J, Eisenberg M. Obesity, disordered eating, and eating disorders in a longitudinal study of adolescents: how do dieters fare 5 years later? J Am Diet Assoc 2006 Apr;106(4):559-568. [doi: 10.1016/j.jada.2006.01.003] [Medline: 16567152]
18. McLean S, Paxton S, Wertheim E, Masters J. Photoshopping the selfie: self photo editing and photo investment are associated with body dissatisfaction in adolescent girls. Int J Eat Disord 2015 Dec;48(8):1132-1140. [doi: 10.1002/eat.22449] [Medline: 26311205]
19. Tiggemann M, Churches O, Mitchell L, Brown Z. Tweeting weight loss: a comparison of #thinspiration and #fitspiration communities on Twitter. Body Image 2018 Jun;25:133-138. [doi: 10.1016/j.bodyim.2018.03.002] [Medline: 29567619]
20. Eikey E, Reddy M. It's Definitely Been a Journey: A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI'17; May 6-11, 2017; Denver, Colorado, USA.
21. Statista. Most popular Apple App Store categories in September 2018, by share of available apps URL: https://www.statista.com/statistics/270291/popular-categories-in-the-app-store/ [accessed 2019-06-06]
22. AppBrain. Most popular Google Play categories URL: https://www.appbrain.com/stats/android-market-app-categories [accessed 2019-06-06] [WebCite Cache ID 78vmcaKzb]
23. MyFitnessPal. URL: https://www.myfitnesspal.com/ [accessed 2019-06-06] [WebCite Cache ID 78vmSmvup]
24. Fitbit. URL: https://www.fitbit.com/uk/home [accessed 2019-06-06]
25. Strava. URL: https://www.strava.com/ [accessed 2019-06-06] [WebCite Cache ID 78vmEtSGV]
26. sweatcoin. URL: https://sweatco.in/ [accessed 2019-06-06] [WebCite Cache ID 78vmNgbma]
27. Sullivan A, Lachman M. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. Front Public Health 2016;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
28. Goodyear V, 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]
29. Eikey E, Reddy M, Booth K, Kvasny L, Blair J, Li V, et al. Desire to be underweight: exploratory study on a weight loss app community and user perceptions of the impact on disordered eating behaviors. JMIR Mhealth Uhealth 2017 Oct 12;5(10):e150 [FREE Full text] [doi: 10.2196/mhealth.6683] [Medline: 29025694]
30. Jakicic J, Rogers R, Donnelly J. The health risks of obesity have not been exaggerated. Med Sci Sports Exerc 2019 Dec;51(1):222-225. [doi: 10.1249/MSS.0000000000001747] [Medline: 30550507]
31. Gaesser G, Blair S. The health risks of obesity have been exaggerated. Med Sci Sports Exerc 2019 Dec;51(1):218-221. [doi: 10.1249/MSS.0000000000001746] [Medline: 30550506]
32. Bacon L, Stern J, van Loan MD, Keim N. Size acceptance and intuitive eating improve health for obese, female chronic dieters. J Am Diet Assoc 2005 Jun;105(6):929-936. [doi: 10.1016/j.jada.2005.03.011] [Medline: 15942543]
33. Avalos L, Tylka T. Exploring a model of intuitive eating with college women. J Counsel Psychol 2006 Oct;53(4):486-497. [doi: 10.1037/0022-0167.53.4.486]
34. Augustus-Horvath C, Tylka T. The acceptance model of intuitive eating: a comparison of women in emerging adulthood, early adulthood, and middle adulthood. J Couns Psychol 2011 Jan;58(1):110-125. [doi: 10.1037/a0022129] [Medline: 21244144]
35. Stawarz K, Cox A, Blandford A. Beyond Self-Tracking and Reminders: Designing Smartphone Apps That Support Habit Formation. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea.
36. West J, Hall P, Hanson C, Barnes M, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. J Med Internet Res 2012 May 14;14(3):e72 [FREE Full text] [doi: 10.2196/jmir.1977] [Medline: 22584372]
37. Lally P, van Jaarsveld CH, Potts H, Wardle J. How are habits formed: modelling habit formation in the real world. Eur J Soc Psychol 2009 Jul 16;40(6):998-1009. [doi: 10.1002/ejsp.674]
38. Anderson I, Maitland J, Sherwood S, Barkhuus L, Chalmers M, Hall M, et al. Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. Mobile Netw Appl 2007 Aug 3;12(2-3):185-199. [doi: 10.1007/s11036-007-0011-7]
39. Wong C, Kwok RCW. The effect of gamified mHealth app on exercise motivation and physical activity. 2016 Jun 27 Presented at: PACIS 2016; June 27-28, 2016; Taiwan.
40. Lupton D. The diverse domains of quantified selves: self-tracking modes and dataveillance. Econ Soc 2016 Apr 15;45(1):101-122. [doi: 10.1080/03085147.2016.1143726]
41. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. 2011 May 7 Presented at: CHI'11; May 7-12, 2006; Vancouver, BC, Canada.
42. Nelson R. Assessment and therapeutic functions of self-monitoring. Progr Behav Modific 1977;5:263-308. [doi: 10.1016/B978-0-12-535605-3.50012-1]

43. Toner J. Exploring the dark-side of fitness trackers: normalization, objectification and the anaesthetisation of human experience. Perform Enhance Health 2018 Aug;6(2):75-81. [doi: 10.1016/j.peh.2018.06.001]
44. Statista. Percentage of US adults who use an app to track their fitness as of 2017, by age URL: https://www.statista.com/statistics/699197/us-adults-that-use-an-app-to-track-their-fitness-by-age/
45. 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]
46. Spiel K, Kayali F, Horvath L, Penkler M, Harrer S, Sicart M. Fitter, Happier, More Productive?: The Normative Ontology of Fitness Trackers. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI EA'18; April 21-26, 2018; Montreal QC, Canada.
47. Pater J, Reining L, Miller A, Toscos T, Mynatt E. "Notjustgirls": Exploring male-related eating disordered content across social media platforms. In: Proceedings of CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI'19; May 4-9, 2019; Glasgow, UK.
48. GirlGuiding. URL: https://www.girlguiding.org.uk/
49. Lopez-Lorca A, Miller T, Pedell S, Mendoza A, Keirnan A, Sterling L. One size doesn't fit all: diversifying. In: Proceedings of the 6th International Workshop on Social Software Engineering. 2014 Presented at: SSE 2014; November 17, 2014; Hong Kong, China.
50. Nunes F, Silva P, Abrantes F. Human-computer interaction and the older adult: an example using user research and personas. In: Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments. 2010 Presented at: PETRA'10; June 23-25, 2010; Samos, Greece.
51. Schoeppe S, Alley S, Rebar A, Hayman M, Bray N, 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 Dec 24;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
52. Chen J, Cade J, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. JMIR Mhealth Uhealth 2015 Dec 16;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]
53. Franco R, Fallaize R, Lovegrove J, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 1;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
54. Bondaronek P, Alkhaldi G, Slee A, Hamilton F, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
55. Frie K, Hartmann-Boyce J, Jebb S, Albury C, Nourse R, Aveyard P. Insights from Google Play Store user reviews for the development of weight loss apps: mixed-method analysis. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e203 [FREE Full text] [doi: 10.2196/mhealth.8791] [Medline: 29273575]
56. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 4;128(1421):73-76. [Medline: 26370762]
57. Stoyanov S, Hides L, Kavanagh D, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
58. Venter W, Coleman J, Chan V, Shubber Z, Phatsoane M, Gorgens M, et al. Improving linkage to HIV care through mobile phone apps: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 17;6(7):e155 [FREE Full text] [doi: 10.2196/mhealth.8376] [Medline: 30021706]
59. Comscore. 5 Interesting Facts About Millennials' Mobile App Usage from URL: https://www.comscore.com/Insights/Blog/5-Interesting-Facts-About-Millennials-Mobile-App-Usage-from-The-2017-US-Mobile-App-Report [accessed 2019-06-06] [WebCite Cache ID 78vhvezX9]
60. Hsieh H, Shannon S. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
61. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
62. Boyatzis R. Transforming Qualitative Information: Thematic Analysis And Code Development. United States: Sage Publications, Inc; 1998.
63. Cordeiro F, Epstein D, Thomaz E, Bales E, Jagannathan A, Abowd G, et al. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2015 Apr Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea p. 1159-1162 URL: http://europepmc.org/abstract/MED/26894233 [doi: 10.1145/2702123.2702155]
64. Lupton D. Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Critic Pub Health 2013 Dec;23(4):393-403. [doi: 10.1080/09581596.2013.794931]
65. Comber R, Weeden J, Hoare J, Lindsay S, Teal G, Macdonald A. Supporting visual assessment of food and nutrient intake in a clinical care setting. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 May Presented at: CHI'06; April 22-27, 2012; Austin, Texas, USA.

66. Cordeiro F, Bales E, Cherry E, Fogarty J. Rethinking the mobile food journalxploring opportunities for lightweight photo-based capture. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Apr 18 Presented at: CHI'15; April 18-23, 2015; Seoul, Republic of Korea.
67. Tylka T, Wood-Barcalow N. What is and what is not positive body image? Conceptual foundations and construct definition. Body Image 2015 Jun;14:118-129. [doi: 10.1016/j.bodyim.2015.04.001] [Medline: 25921657]
68. Aubrey J. Looking good versus feeling good: An investigation of media frames of health advice and their effects on women's body-related self-perception. Sex Roles 2010 Jun 5;63(1-2):50-63. [doi: 10.1007/s11199-010-9768-4]
69. Hurst M, Dittmar H, Banerjee R, Bond R. "I just feel so guilty": The role of introjected regulation in linking appearance goals for exercise with women's body image. Body Image 2017 Mar;20:120-129 [FREE Full text] [doi: 10.1016/j.bodyim.2016.12.002] [Medline: 28193554]
70. Alleva J, Martijn C, van Breukelen GJ, Jansen A, Karos K. Expand Your Horizon: a programme that improves body image and reduces self-objectification by training women to focus on body functionality. Body Image 2015 Sep;15:81-89. [doi: 10.1016/j.bodyim.2015.07.001] [Medline: 26280376]
71. Alleva J, Veldhuis J, Martijn C. A pilot study investigating whether focusing on body functionality can protect women from the potential negative effects of viewing thin-ideal media images. Body Image 2016 Jun;17:10-13. [doi: 10.1016/j.bodyim.2016.01.007] [Medline: 26878220]
72. Grabe S, Ward L, Hyde J. The role of the media in body image concerns among women: a meta-analysis of experimental and correlational studies. Psychol Bull 2008 May;134(3):460-476. [doi: 10.1037/0033-2909.134.3.460] [Medline: 18444705]
73. Stice E, Presnell K, Spangler D. Risk factors for binge eating onset in adolescent girls: a 2-year prospective investigation. Health Psychol 2002 Mar;21(2):131-138. [Medline: 11950103]
74. Crossley J. Academia.edu. Chewing the Fat: Ideology, Identity and Group Membership in Pro-Eating Disorder Communities URL: https://www.academia.edu/36050404/Chewing_the_Fat_Ideology_Identity_and_Group_Membership_in_Pro-Eating_Disorder_Communities [accessed 2019-06-06] [WebCite Cache ID 78vgdYm8C]
75. Morrison L, Hargood C, Lin S, Dennison L, Joseph J, Hughes S, et al. Understanding usage of a hybrid website and smartphone app for weight management: a mixed-methods study. J Med Internet Res 2014 Oct 22;16(10):e201 [FREE Full text] [doi: 10.2196/jmir.3579] [Medline: 25355131]
76. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
77. Verstuyf J, Patrick H, Vansteenkiste M, Teixeira P. Motivational dynamics of eating regulation: a self-determination theory perspective. Int J Behav Nutr Phys Act 2012 Mar 2;9(1):21 [FREE Full text] [doi: 10.1186/1479-5868-9-21] [Medline: 22385782]
78. Trnka S. Digital care: agency and temporality in young people's use of health app. Engaging STS 2016 Dec 04;2:248. [doi: 10.17351/ests2016.119]
79. Fitton D, Bell B. Working with teenagers within HCI research: understanding teen-computer interaction. In: Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI. 2014 Presented at: BCS-HCI'14; September 9-12, 2014; Southport, UK.
80. Lupton D. The thing-power of the human-app health assemblage: thinking with vital materialism. Soc Theory Health 2019 Feb 22;17(2):125-139. [doi: 10.1057/s41285-019-00096-y]
81. Ferrara J. Games for persuasion: argumentation, procedurality, and the lie of gamification. Games Culture 2013 Aug 27;8(4):289-304. [doi: 10.1177/1555412013496891]
82. Honary M, McNaney R, Lobban A. Designing video stories around the lived experience of severe mental illness. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction. 2018 Sep 29 Presented at: NordiCHI'18; September 29-October 3, 2018; Oslo, Norway.
83. Deterding S, Björk S, Nacke L, Dixon D, Lawley E. Designing gamification: creating gameful and playful experiences. In: Extended Abstracts on Human Factors in Computing Systems. Designing gamification: creating gamefulplayful experiences. CHI'13 Extended Abstracts on Human Factors in Computing Systems; 2013 Presented at: CHI'13; April 27-May 2, 2013; Paris, France.
84. King D, Greaves F, Exeter C, Darzi A. 'Gamification': influencing health behaviours with games. J R Soc Med 2013 Mar;106(3):76-78 [FREE Full text] [doi: 10.1177/0141076813480996] [Medline: 23481424]
85. Johnston R. Gizmodo. Australians Only Use Most Apps For Two Weeks, Tops URL: https://www.gizmodo.com.au/2016/05/australians-only-use-most-apps-for-two-weeks-tops/ [accessed 2019-06-06] [WebCite Cache ID 78veXnmBM]
86. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 02.04.19; peer-reviewed by G Signorelli; comments to author 23.04.19; revised version received 13.05.19; accepted 14.05.19; published 18.06.19.

Please cite as:
Honary M, Bell BT, Clinch S, Wild SE, McNaney R
Understanding the Role of Healthy Eating and Fitness Mobile Apps in the Formation of Maladaptive Eating and Exercise Behaviors in Young People
JMIR Mhealth Uhealth 2019;7(6):e14239
URL: http://mhealth.jmir.org/2019/6/e14239/
doi:10.2196/14239
PMID:31215514

Original Paper

A Community-Based Short Message Service Intervention to Improve Mothers' Feeding Practices for Obesity Prevention: Quasi-Experimental Study

Hong Jiang[1,2*], PhD; Mu Li[3,4*], PhD; Li Ming Wen[3,5], PhD; Louise Baur[3,6], PhD; Gengsheng He[1,2], PhD; Xiaoying Ma[1,2], MS; Xu Qian[1,2], PhD

[1]School of Public Health, Global Health Institute, Fudan University, Shanghai, China

[2]Key Lab of Health Technology Assessment, National Health Commission of the People's Republic of China, Fudan University, Shanghai, China

[3]School of Public Health, University of Sydney, Sydney, Australia

[4]China Studies Centre, University of Sydney, Sydney, Australia

[5]Health Promotion Unit, Sydney Local Health District, Sydney, Australia

[6]Discipline of Child & Adolescent Health, University of Sydney, Sydney, Australia

[*]these authors contributed equally

Corresponding Author:
Hong Jiang, PhD
School of Public Health
Global Health Institute
Fudan University
175 Mailbox
138 Yixueyuan Road
Shanghai, 200032
China
Phone: 86 2164179976
Fax: 86 2164179976
Email: h_jiang@fudan.edu.cn

Related Article:

This is a corrected version. See correction statement: http://mhealth.jmir.org/2019/7/e15046/

Abstract

Background: The prevalence of childhood obesity is increasing in China, and the effect of mobile phone short message service (SMS) interventions to prevent early childhood obesity needs to be evaluated.

Objective: The objective of this study was to assess the effect of an SMS intervention on the prevention of obesity in young children.

Methods: A quasi-experimental design SMS intervention was carried out in 4 community health centers (CHCs) in Shanghai, China. A total of 2 CHCs were assigned to the intervention group, and 2 CHCs were assigned to the control group. Mothers in the intervention group received weekly SMS messages on breastfeeding and infant feeding from the third trimester to 12 months postpartum. The primary outcomes were children's body mass index (BMI), BMI z-score, and weight-for-length z-score at 12 and 24 months. Factors associated with higher BMI and weight-for-length z-score at 24 months were also assessed.

Results: A total of 582 expectant mothers were recruited at the beginning of the third gestational trimester. 477 (82.0%) and 467 (80.2%) mothers and their children were followed up to 12 and 24 months postpartum, respectively. There were no significant differences in children's BMI, BMI z-score, and weight-for-length z-score at 12 and 24 months between the 2 groups. Factors associated with higher BMI, BMI z-score, and weight-for-length z-score at 24 months included higher birth weight, introduction of solid foods before 4 months, and taking a bottle to bed at 12 months.

Conclusions: The SMS intervention did not show a significant effect on children's BMI, BMI z-score, or weight-for-length z-score at 12 and 24 months. Introduction of solid foods before 4 months and taking a bottle to bed at 12 months were significantly

and positively correlated with a higher BMI, BMI z-score, and weight-for-length z-score at 24 months. Further studies with more rigorous design are needed to evaluate the effect of SMS interventions on preventing early childhood obesity.

(JMIR Mhealth Uhealth 2019;7(6):e13828) doi:10.2196/13828

KEYWORDS

short message service; child development; body mass index; BMI; childhood obesity

Introduction

Background

Childhood obesity is a major public health concern globally. In China, the prevalence of obesity in children and adolescents aged 0 to 18 years increased from 0.4% in 1985 to 7.5% in 2010, a more than 18-fold increase in 25 years [1]. Early childhood, particularly the prenatal period, is a critical period to prevent obesity in later life [2,3]. There has been an increased interest in research on early childhood obesity interventions in recent years, including home- or caregiver-based interventions [4,5], interactive training modules for parents, and environmental change to promote healthy eating and active playing [6,7]. Such interventions have reportedly led to improvements in infant feeding practices (eg, increased duration of breastfeeding and a decrease in soft drink consumption) and a reduction in TV viewing while eating [8,9]. One study has also shown a reduction in body mass index (BMI) at age 24 months [8].

Mobile health (mHealth) refers to delivering health care and health promotion through mobile devices [10]. Mobile phone short message service (SMS) can deliver health care inexpensively through text messages wherever the person is located [11]. SMS text messaging has attracted global attention for its ability to enhance health care services [12,13]. Reported SMS text messaging in maternal and child health care are extensive, such as promoting prenatal care utilization, promoting exclusive breastfeeding, and improving the training of maternal and child health care providers [10,11,14]. The effects of these interventions, however, are underevaluated.

Aims

We implemented an SMS intervention to first-time mothers in Shanghai, China. The overall aims of the study were to promote healthy infant feeding practices to new mothers and to examine whether the intervention is effective in preventing early onset of childhood obesity of their children. We have previously reported significant improvements in median exclusive breastfeeding (EBF) duration and rate of EBF at 6 months among the mothers receiving the intervention [15]. In this paper, we report the effect of this SMS intervention on the BMI, BMI z-score, and weight-for-length z-score of the children at 12 and 24 months. We also present factors associated with higher BMI, BMI z-score, and weight-for-length z-score at 24 months.

Methods

Study Design

This SMS text messaging intervention, using a quasi-experimental design, was conducted in Shanghai, China. A total of 4 community health centers (CHCs) were purposively selected from 2 administrative districts; 2 CHCs were assigned as the intervention group, and the other 2 were assigned to the control group. The intervention group received a weekly mobile phone SMS on appropriate infant feeding practices, plus routine child health care provided by CHCs. Participating mothers in the control group only received the routine child health care services. The intervention was carried out between December 2010 and October 2012. The 24-month anthropometric data collection was completed in October 2013. Written informed consent was obtained from each participant. Ethics approval was obtained from the Institutional Review Board of the School of Public Health, Fudan University, Shanghai, China, and the Human Research Ethics Committee of the University of Sydney, Sydney, Australia [15].

Primary and Secondary Outcomes

The primary outcomes of this intervention were child's BMI, BMI z-score, and weight-for-length z-score at 12 and 24 months. Length and weight data at 12 and 24 months were analyzed. Secondary outcomes were the proportion of infants introduced to solid foods before 4 months and taking a bottle to bed at 12 months; their associations with BMI at 24 months were also investigated.

Participants and Recruitment

The inclusion criteria and recruitment processes have been reported previously [15]. Essentially, eligible expectant mothers, owning a mobile phone, from the 4 communities were recruited to the study when they attended their first antenatal checkup in the CHCs after giving informed consent. They were asked to complete a baseline self-administered questionnaire, including questions on their initial awareness of the World Health Organization (WHO) breastfeeding guidelines [16]. CHC staff reviewed the participants' eligibility again at the beginning of the third trimester before the commencement of the SMS intervention.

Sample Size

To detect a difference in BMI of 0.4 units between the intervention and control groups at 24 months at the significance level of $P<.05$ with 80% power, a sample size of 446 (223 per arm) was needed. The estimation was made on the assumption of SDs of 1.4 and 1.6 for BMI at 24 months based on a published survey. Given an estimated loss to follow-up rate of 20% at the end of 24 months, 558 expectant mothers were required for the study.

Short Message Service Intervention

One weekly text message was sent to the participants in the intervention group from a computer-based platform from 28 weeks gestation to 12 months postpartum [15]. The total duration of the intervention was thus 66 weeks. Each text

message contained approximately 180 to 210 characters. The messages were developed from the WHO breastfeeding guidelines and infant and young child feeding recommendations, in consultation with child health care experts and informed by a formative study [17]. The messages covered different stages of infant growth and development, providing anticipatory knowledge and guidance for appropriate infant feeding practices [15]. In addition, messages were sent periodically inquiring about the breastfeeding status, the time to return to work, or the timing of introducing semisolids or solids, so that appropriate messages could be sent in response to each woman's feeding situation. The control group received routine maternal and child health care in the CHCs. The control group received maternal and child health care routinely provided by the CHCs, which included anthropometric measurements of length and weight periodically, responses to child feeding enquires, and assessment of the child development.

Data Collection and Main Measurements

Child health care doctors from the 4 CHCs were trained for standard measurement and data collection procedures at the beginning of the study. The data on EBF, breastfeeding, and timing of introducing solids were collected and recorded by doctors via face-to-face interviews [15]. Child weight (to the nearest 0.1 kg) and length (to the nearest 0.1 cm) at 12 and 24 months were measured by the doctors using established methods and recorded using a standard protocol as part of a routine child health checkup [18]. Information on breastfeeding, timing of introducing semisolids and solids, and anthropometric measurement were extracted from the children's health records. Children's BMI at 12 and 24 months was calculated as weight in kg/(length in m)^2, and BMI z-score and weight-for-length z-score were calculated using the lambda-mu-sigma method based on the WHO Child Growth Standards [19]. At 12 months postpartum, a face-to-face interview was held with mothers by CHC staff in each CHC using a questionnaire adapted and translated from the Healthy Beginnings Trial [20]. Information on care givers' feeding practices, such as drinking from a cup, giving a bottle at bedtime, and using food for reward, was collected [15].

Statistical Analysis

Data quality control was performed before data analysis. The data were excluded if children's weight gain or length was less than 0.5 kg or 3 cm between birth weight and 12 months and between 12 and 24 months without any clinical indications.

Statistical analyses were made using the SPSS for Windows, version 17.0. Continuous variables were compared using independent samples *t* test, proportions using the Pearson chi-squared test, and trends in proportions using Mantel-Haenszel chi-squared tests. Statistical significance level was set at $P<.05$.

Awareness of WHO breastfeeding guidelines was assessed with 6 questions in the baseline questionnaire. Each correct answer was scored 1 and an incorrect scored 0. The total scores ranged from 0 to 6, categorized into *high* or *low* score groups according to the median score.

BMI, BMI z-score, length, weight, and weight-for-length z-score at 12 and 24 months were analyzed as continuous variables. The effect of intervention on BMI, BMI z-score and weight-for-length z-scores was determined by analysis of variance (ANOVA) after controlling for baby's birth weight; mothers' age, education level, preconception maternal BMI; household registration status, whether living in rental accommodation; and baseline awareness of WHO breastfeeding guidelines. The effect of the intervention on the child's BMI and BMI z-score was analyzed in the same way, with additional adjustment for baby's sex. As children might not be measured exactly when turning 12 or 24 months, further adjustment was made for the anthropometric data collection by exact child age (months to 2 decimal point). Similarly, the analysis of length and weight were further controlled for mother's preconception height and weight, respectively.

Multivariate linear regression was applied to analyze the factors associated with BMI, BMI z-score, and weight-for-length z-score at 24 months. Factors in the model included baby's birth weight; mothers' age and education level, awareness of the WHO breastfeeding guidelines at baseline, preconception BMI, EBF status at 6 months, breastfeeding duration (weeks), and introduction of solid foods before 4 months; care givers' feeding practices at 12 months, such as drinking from a cup, taking a bottle to bed, and using food for reward; and whether the family had Shanghai household registration and whether the family was living in a rental accommodation.

Results

Recruitment and Follow-Up

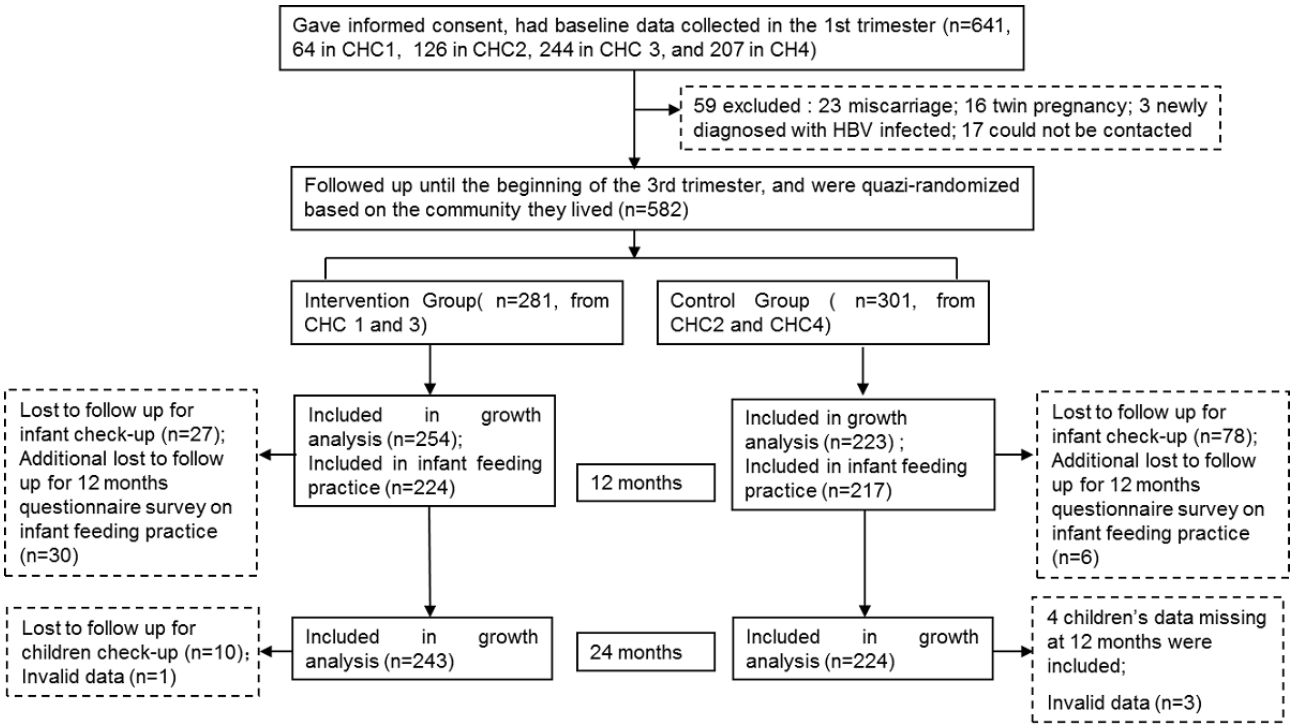
Of the 641 expectant mothers recruited from the 4 CHCs during their first visit, 582 were eligible at the beginning of the third trimester, 281 in the intervention group and 301 in the control group. There was no significant difference on main characteristics between women recruited in the first trimester and eligible in the third trimester. At 12 months, 82.0% (478/582) of the children's weight and length measurements were obtained, and 75.7% (441/582) of the mothers completed the questionnaire survey. At 24 months, 80.2% (467/582) of the children's anthropometric data were collected (Figure 1).

Participants' Main Characteristics

Multimedia Appendix 1 shows the main characteristics of participants at baseline and at 12 and 24 months. At baseline, the mean age of the mothers was 28 years (range: 20-39 years). 75.6% (440/582) of the mothers did not hold Shanghai household registration, and 86.4% (503/582) had tertiary education. Nearly one-half of mothers (44.7%, 260/582) had a low awareness of WHO breastfeeding guidelines. Mothers who were lost to follow-up at 12 months were more likely to be living in a rental accommodation ($P<.001$) and had a higher awareness score of WHO breastfeeding guidelines ($P=.02$). Similarly, mothers who were lost to follow-up at 24 months were more likely to be living in a rental accommodation ($P<.001$) and had a non-Shanghai household registration status ($P=.03$). At 12 months, the average BMI of the children was 17.04 kg/m^2 (SE 0.06) ranging from 13.65 to 21.04, and BMI

z-score was 0.46 (SE 0.04), with the range of −1.96 to 2.85. At 24 months, the average BMI of the children was 15.92 kg/m^2 (SE 0.05) ranging from 12.62 to 20.22, and the mean BMI z-score 0.05 (SE 0.04) ranging from −2.69 to 2.83.

Figure 1. Participant recruitment and retention. CHC: community health center; BMI: body mass index; HBV:hepatitis b virus.



Effect of the Intervention on Body Mass Index, Body Mass Index Z-Score, and Weight-for-Length Z-Score

By ANOVA, there was no statistically significant difference between the intervention and control groups in BMI (mean 17.04 kg/m^2, SE 0.08 vs mean 17.04 kg/m^2, SE 0.08), BMI z-score (mean 0.47, SE 0.05 vs mean 0.46, SE 0.06), or weight-for-length z-score (mean 0.44, SE 0.05 vs mean 0.44, SE 0.05) at 12 months. Similar results were found at 24 months for BMI (mean 15.93 kg/m^2, SE 0.07 vs mean 15.90 kg/m^2, SE 0.08), BMI z-score (mean 0.06, SE 0.06 vs mean 0.04, SE 0.06), or weight-for-length z-score (mean 0.44, SE 0.05 vs mean 0.23, SE 0.06), respectively (Multimedia Appendix 2). No difference was found in length and weight between children in intervention and control groups at 12 months and 24 months.

Factors Associated With Body Mass Index and Body Mass Index Z-Score at 24 Months

Table 1 shows the results of multiple linear regression analysis of factors associated with BMI, BMI z-score, and weight-for-length z-score at 24 months. The factors significantly and positively correlated with BMI, BMI z-score, and weight-for-length z-score at 24 months included higher birthweight (all P<.001), introduction of solid foods before 4 months (P=.003, P=.004, and P=.003, respectively), and taking a bottle to bed at 12 months (P=.048, P=.044, and P=.036, respectively).

Table 1. Factors associated with body mass index (BMI) and BMI z-score at 24 months (n=359).

Variables	Beta[a] (95% CI)	*P* value[a]
BMI[b] at 24 months		
Household registration	−.144 (−.416 to .129)	.30
Rental accommodation	−.162 (−.495 to .172)	.34
Awareness of WHO[c] breastfeeding guidelines at baseline	.008 (−.222 to .248)	.95
Maternal pre-pregnancy BMI	.020 (−.027 to .067)	.40
Maternal education	.013 (−.265 to .291)	.93
Birthweight	.612 (.343 to .881)	<.001
Baby's sex	.113 (−.121 to .347)	.34
Exact child's age at follow-up (month)	−.139 (−.378 to .099)	.25
Maternal age at recruitment	−.045 (−.252 to .163)	.67
EBF[d] at 6 months	−.276 (−.641 to .090)	.14
Introduction of solid foods before 4 month	1.225 (.429 to 2.021)	.003
Food used as a reward at 12 month	.063 (−.180 to .307)	.61
Drinking from a cup at 12 month	−.101 (−.341 to .138)	.41
Taking a bottle to bed at 12 month	.239 (.002 to .477)	.048
BF[e] duration (month)	.001 (−.031 to .034)	.94
BMI z-score at 24 months		
Household registration	−.110 (−.319 to .098)	.30
Rental accommodation	−.114 (−.370 to .142)	.38
Awareness of WHO breastfeeding guidelines at baseline	.010 (−.173 to .194)	.91
Maternal pre-pregnancy BMI	.010 (−.026 to .046)	.58
Maternal education	−.009 (−.221 to .203)	.93
Birthweight	.458 (.252 to .664)	<.001
Maternal age at recruitment	−.038 (−.197 to .121)	.64
EBF at 6 months	−.230 (−.510 to .050)	.11
Introduction of solid foods before 4 month	.888 (.278 to 1.498)	.004
Food used as a reward at 12 month	.033 (−.153 to .219)	.73
Drinking from a cup at 12 month	−.064 (−.247 to .120)	.495
Taking a bottle to bed at 12 month	.187 (.005 to .369)	.04
BF duration (month)	.003 (−.022 to .028)	.80
Weight-for-length z-score at 24 months		
Household registration	−.113 (−.312 to .085)	.26
Rental accommodation	−.119 (−.363 to .124)	.34
Awareness of WHO breastfeeding guidelines at baseline	.012 (−.163 to .186)	.90
Maternal pre-pregnancy BMI	.010 (−.024 to .044)	.56
Maternal education	.004 (−.198 to .206)	.97
Birthweight	.518 (.322 to .714)	<.001
Maternal age at recruitment	−.029 (−.181 to .122)	.70
EBF at 6 months	−.241 (−.507 to .025)	.08
Introduction of solid foods before 4 month	.894 (.313 to 1.474)	.003
Food used as a reward at 12 month	.036 (−.141 to .213)	.69

Variables	Beta[a] (95% CI)	*P* value[a]
Drinking from a cup at 12 month	−.057 (−.231 to .117)	.52
Taking a bottle to bed at 12 month	.186 (.013 to .359)	.04
BF duration (month)	−.002 (−.026 to .022)	.86

[a]Multiple linear regression: The number of participants with both infant feeding practice data at 12 months and BMI data at 24 months was 359 in total, 197 from the intervention group, and 162 from the control group.

[b]BMI: body mass index.

[c]WHO: World Health Organization.

[d]EBF: exclusive breastfeeding.

[e]BF: breastfeeding.

Discussion

Principal Findings

Our study found that the SMS intervention, delivered to mothers on a weekly basis from the third trimester until 12 months postpartum, had no significant effect on their children's BMI, BMI z-score, or weight-for-length z-score at 12 and 24 months. Higher birth weight, introduction of solid foods before 4 months, and taking a bottle to bed at 12 months were found significantly and positively associated with higher BMI, BMI z-score, and weight-for-length z-score at 24 months.

Strength and Limitations

The study was a quasi-experimental design, community-based intervention, with a high follow-up rate at 12 and 24 months. The SMS intervention was proven to be feasible and practical in improving breastfeeding practices, including significantly higher EBF rate at 6 months, longer median duration of EBF, and lower rate of instruction of solid foods before 4 months [15,17]. Although the intervention did not demonstrate significant effect on children's BMI or BMI z-score at 24 months, several infant feeding practices were identified as potentially important influencing factors. There are several limitations of the study. Participants were not randomized to the intervention and control groups, and some characteristics, for example, maternal age, education, and birth weight, were significantly different between the 2 groups. Therefore, multiple regression was used to control confounding factors. Moreover, although SMS intervention is seen to be low cost, and potentially suitable for low resource settings, the cost-effectiveness of SMS intervention was not explored in our study. Information on other possible risk factors of obesity at 24 months, such as physical activity levels and TV viewing time, were not collected.

Interpretation of the Research Findings

Among the published early childhood obesity prevention studies, the Healthy Beginnings Trial is the only study that showed a significant reduction in BMI (0.38 kg/m^2 mean decrease, *P*=.01) in children of the intervention group at 24 months [8,21]. This trial was a multicomponent intervention delivered through 8 home visits by trained community nurses [8]. The NOURISH trial, comprising 2 interactive group education modules, each delivering over 3 months to new mothers to promote healthy eating patterns, demonstrated a small but not statistically significant reduction in mean BMI z-score of 0.14 (*P*=.12) [22]. Compared with these 2 studies, our intervention, with 1 weekly text message with no more than 210 characters, had much lower intensity. This could be one explanation for not showing effect on BMI z-score or the BMI. The results could also be due to a type 2 error-inadequate statistical power to detect the difference because there was a limited reference for postulating a meaningful expected difference for estimating sample size.

Whether early introduction of solid foods affects BMI in later childhood is still a contested topic. Some studies suggest that early introduction of solids is associated with increased childhood BMI [23], whereas others did not support such a conclusion [24,25]. We found that the introduction of solid foods before 4 months was positively associated with a higher BMI, BMI z-score, and weight-for-length z-score at 24 months. This finding is in keeping with previous reported studies [26,27]. No association between EBF or duration of breastfeeding with BMI, BMI z-score, and weight-for-length z-score at 24 months was found, which was consistent with 1 prospective study [28], but different to other studies' findings [26,29]. Further research is needed to confirm breastfeeding promotion as a robust childhood obesity prevention strategy [30].

Our study showed that giving an infant a bottle to take to bed at 12 months, which is a less reported practice, was positively associated with a higher BMI, BMI z-score, and weight-for-length z-score at 24 months. However, we found no significant association between drinking from a cup at 12 months and anthropometric status at the age of 24 months. This is consistent with a randomized controlled trial that also failed to show an association between reduced bottle usage from 12 months and adiposity at 24 months [31]. However, links were found between bottle use at 24 months and higher obesity prevalence at the age of 5.5 years in an observational study [32].

Future Research

mHealth intervention as an effective strategy to promote breastfeeding and infant feeding and childhood obesity prevention still requires further research using stronger study designs and larger sample sizes. Research on motivations and barriers for mothers to adopt healthy child feeding behaviors is needed to inform the development of effective interventions. The cost-effectiveness of SMS interventions compared with facility-based interventions also needs to be determined.

Conclusions

Overweight and obesity in children are affected by many factors. Although no significant effect of this SMS intervention on children's BMI, BMI z-score, or weight-for-length z-score at 12 and 24 months was found, we have identified several influencing factors. The introduction of solid foods before 4 months, taking a bottle to bed at 12 months, and high birth weight were significantly and positively correlated with a higher BMI, BMI z-score, and weight-for-length z-score at 24 months. Furthermore, large-scale randomized controlled trials are needed to evaluate the effectiveness and cost- effectiveness of SMS text messaging interventions on the prevention of early childhood obesity, including the usability, intensity, frequency, and duration of the interventions.

Acknowledgments

The authors thank all participants and staff in CHCs involved in the study. They also thank Qiaozhen Hu, Dongling Yang, and Yang Li for data collection during the research. This study was funded by the Nestle Foundation (Small Research Grant) and Shanghai Municipal Health Bureau (no.12GWZX0301 and no. 15GWZK0402).

Authors' Contributions

ML, XQ, and HJ conceptualized and designed the study. LMW, ML, XQ, GH, and HJ designed the data collection instruments. XQ and HJ coordinated and supervised data collection at 4 study sites. HJ carried out the initial analyses. HJ and ML drafted the initial manuscript, and LMW and LB provided critical comments on intellectual content. All authors reviewed and revised the manuscript and approved the final manuscript as submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of mothers and babies in the study, Shanghai China 2010-13.

[DOCX File, 21KB - mhealth_v7i6e13828_app1.docx]

Multimedia Appendix 2

Body mass index (BMI), BMI z-score and weight for length z-score (kg/m2) between children in the intervention and control groups at 12 and 24 months.

[DOCX File, 21KB - mhealth_v7i6e13828_app2.docx]

References

1. Yu Z, Han S, Chu J, Xu Z, Zhu C, Guo X. Trends in overweight and obesity among children and adolescents in China from 1981 to 2010: a meta-analysis. PLoS One 2012;7(12):e51949 [FREE Full text] [doi: 10.1371/journal.pone.0051949] [Medline: 23284829]
2. Baird J, Fisher D, Lucas P, Kleijnen J, Roberts H, Law C. Being big or growing fast: systematic review of size and growth in infancy and later obesity. Br Med J 2005 Oct 22;331(7522):929 [FREE Full text] [doi: 10.1136/bmj.38586.411273.E0] [Medline: 16227306]
3. Glavin K, Roelants M, Strand BH, Júlíusson PB, Lie KK, Helseth S, et al. Important periods of weight development in childhood: a population-based longitudinal study. BMC Public Health 2014 Feb 13;14:160 [FREE Full text] [doi: 10.1186/1471-2458-14-160] [Medline: 24524269]
4. Salvy S, de la Haye K, Galama T, Goran MI. Home visitation programs: an untapped opportunity for the delivery of early childhood obesity prevention. Obes Rev 2017 Dec;18(2):149-163 [FREE Full text] [doi: 10.1111/obr.12482] [Medline: 27911984]
5. Wen LM, Rissel C, Baur LA, Hayes AJ, Xu H, Whelan A, et al. A 3-arm randomised controlled trial of Communicating Healthy Beginnings Advice by Telephone (CHAT) to mothers with infants to prevent childhood obesity. BMC Public Health 2017 Dec 14;17(1):79 [FREE Full text] [doi: 10.1186/s12889-016-4005-x] [Medline: 28088203]
6. Campbell KJ, Hesketh KD, McNaughton SA, Ball K, McCallum Z, Lynch J, et al. The extended Infant Feeding, Activity and Nutrition Trial (InFANT Extend) Program: a cluster-randomized controlled trial of an early intervention to prevent childhood obesity. BMC Public Health 2016 Feb 18;16:166 [FREE Full text] [doi: 10.1186/s12889-016-2836-0] [Medline: 26888759]
7. Bean MK, Thornton LM, Jeffers AJ, Gow RW, Mazzeo SE. Impact of motivational interviewing on engagement in a parent-exclusive paediatric obesity intervention: randomized controlled trial of NOURISH+MI. Pediatr Obes 2018 Dec 4:e12484. [doi: 10.1111/ijpo.12484] [Medline: 30515995]

8. Wen LM, Baur LA, Simpson JM, Rissel C, Wardle K, Flood VM. Effectiveness of home based early intervention on children's BMI at age 2: randomised controlled trial. Br Med J 2012;344:e3732 [FREE Full text] [doi: 10.1136/bmj.e3732] [Medline: 22735103]
9. Daniels LA, Mallan KM, Battistutta D, Nicholson JM, Meedeniya JE, Bayer JK, et al. Child eating behavior outcomes of an early feeding intervention to reduce risk indicators for child obesity: the NOURISH RCT. Obesity (Silver Spring) 2014 May;22(5):E104-E111 [FREE Full text] [doi: 10.1002/oby.20693] [Medline: 24415390]
10. Premji S. Mobile health in maternal and newborn care: fuzzy logic. Int J Environ Res Public Health 2014 Jun;11(6):6494-6503 [FREE Full text] [Medline: 25003177]
11. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
12. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
13. McCartney P. Global maternal-child mobile health. MCN Am J Matern Child Nurs 2012 Sep;37(5):347. [doi: 10.1097/NMC.0b013e3182619089] [Medline: 22895209]
14. 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] [Medline: 22635061]
15. Jiang H, Li M, Wen LM, Hu Q, Yang D, He G, et al. Effect of short message service on infant feeding practice: findings from a community-based study in Shanghai, China. JAMA Pediatr 2014 May;168(5):471-478 [FREE Full text] [doi: 10.1001/jamapediatrics.2014.58] [Medline: 24639004]
16. Jiang H, Li M, Yang D, Wen LM, Hunter C, He G, et al. Awareness, intention, and needs regarding breastfeeding: findings from first-time mothers in Shanghai, China. Breastfeed Med 2012 Dec;7(6):526-534 [FREE Full text] [doi: 10.1089/bfm.2011.0124] [Medline: 22424469]
17. Jiang H, Li M, Wen LM, Baur LA, He G, Ma X, et al. A short message service intervention for improving infant feeding practices in Shanghai, China: planning, implementation, and process evaluation. JMIR Mhealth Uhealth 2018 Oct 29;6(10):e11039 [FREE Full text] [doi: 10.2196/11039] [Medline: 30373728]
18. Cogill B. Food and Nutrition Technical Assistance. Anthropometric Indicators Measurement Guide URL: https://www.fantaproject.org/sites/default/files/resources/anthropometry-2003-ENG.pdf [accessed 2019-05-14] [WebCite Cache ID 78MvIDHOF]
19. World Health Organization. Child growth standards URL: https://www.who.int/childgrowth/standards/bmi_for_age/en/ [accessed 2019-05-14] [WebCite Cache ID 78MvK3hdV]
20. Wen LM, Baur LA, Simpson JM, Rissel C, Flood VM. Effectiveness of an early intervention on infant feeding practices and. Arch Pediatr Adolesc Med 2011 Aug;165(8):701-707. [doi: 10.1001/archpediatrics.2011.115] [Medline: 21810633]
21. Wen LM, Baur LA, Simpson JM, Xu H, Hayes AJ, Hardy LL, et al. Sustainability of effects of an early childhood obesity prevention trial over time: a further 3-year follow-up of the Healthy Beginnings Trial. JAMA Pediatr 2015 Jun;169(6):543-551. [doi: 10.1001/jamapediatrics.2015.0258] [Medline: 25893283]
22. Daniels LA, Mallan KM, Nicholson JM, Battistutta D, Magarey A. Outcomes of an early feeding practices intervention to prevent childhood obesity. Pediatrics 2013 Jul;132(1):e109-e118. [doi: 10.1542/peds.2012-2882] [Medline: 23753098]
23. Huh SY, Rifas-Shiman SL, Taveras EM, Oken E, Gillman MW. Timing of solid food introduction and risk of obesity in preschool-aged children. Pediatrics 2011 Mar;127(3):e544-e551 [FREE Full text] [doi: 10.1542/peds.2010-0740] [Medline: 21300681]
24. Moorcroft KE, Marshall JL, McCormick FM. Association between timing of introducing solid foods and obesity in infancy and childhood: a systematic review. Matern Child Nutr 2011 Jan;7(1):3-26. [doi: 10.1111/j.1740-8709.2010.00284.x] [Medline: 21143583]
25. Symon B, Crichton GE, Muhlhausler B. Does the early introduction of solids promote obesity? Singapore Med J 2017 Nov;58(11):626-631 [FREE Full text] [doi: 10.11622/smedj.2017024] [Medline: 28367583]
26. Wen LM, Baur LA, Rissel C, Xu H, Simpson JM. Correlates of body mass index and overweight and obesity of children aged 2 years: findings from the healthy beginnings trial. Obesity (Silver Spring) 2014 Jul;22(7):1723-1730 [FREE Full text] [doi: 10.1002/oby.20700] [Medline: 24415528]
27. Kramer MS. Breastfeeding, complementary (solid) foods, and long-term risk of obesity. Am J Clin Nutr 2010 Mar;91(3):500-501. [doi: 10.3945/ajcn.2010.29199] [Medline: 20107200]
28. Neutzling MB, Hallal PR, Araújo CL, Horta BL, Vieira MD, Menezes AM, et al. Infant feeding and obesity at 11 years: prospective birth cohort study. Int J Pediatr Obes 2009;4(3):143-149. [doi: 10.1080/17477160802453530] [Medline: 19353369]
29. Harder T, Bergmann R, Kallischnigg G, Plagemann A. Duration of breastfeeding and risk of overweight: a meta-analysis. Am J Epidemiol 2005 Sep 1;162(5):397-403. [doi: 10.1093/aje/kwi222] [Medline: 16076830]
30. Lumeng JC, Taveras EM, Birch L, Yanovski SZ. Prevention of obesity in infancy and early childhood: a National Institutes of Health workshop. JAMA Pediatr 2015 May;169(5):484-490. [doi: 10.1001/jamapediatrics.2014.3554] [Medline: 25775180]
31. Bonuck K, Avraham SB, Lo Y, Kahn R, Hyden C. Bottle-weaning intervention and toddler overweight. J Pediatr 2014 Feb;164(2):306-12.e1. [doi: 10.1016/j.jpeds.2013.09.029] [Medline: 24183206]

32. Gooze RA, Anderson SE, Whitaker RC. Prolonged bottle use and obesity at 5.5 years of age in US children. J Pediatr 2011 Sep;159(3):431-436. [doi: 10.1016/j.jpeds.2011.02.037] [Medline: 21543085]

Abbreviations

ANOVA: analysis of variance
BMI: body mass index
CHC: community health center
EBF: exclusive breastfeeding
mHealth: mobile health
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 26.02.19; peer-reviewed by J Chen, G Humphrey; comments to author 10.04.19; revised version received 19.04.19; accepted 20.04.19; published 03.06.19.

Please cite as:
Jiang H, Li M, Wen LM, Baur L, He G, Ma X, Qian X
A Community-Based Short Message Service Intervention to Improve Mothers' Feeding Practices for Obesity Prevention: Quasi-Experimental Study
JMIR Mhealth Uhealth 2019;7(6):e13828
URL: http://mhealth.jmir.org/2019/6/e13828/
doi:10.2196/13828
PMID:

Original Paper

SmokefreeTXT for Homeless Smokers: Pilot Randomized Controlled Trial

Travis P Baggett[1,2,3], MPH, MD; Claire McGlave[4], BA; Gina R Kruse[1,3], MPH, MD; Awesta Yaqubi[5], BA; Yuchiao Chang[1,3], PhD; Nancy A Rigotti[1,3], MD

[1]Tobacco Research and Treatment Center, Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA, United States

[2]Institute for Research, Quality, and Policy in Homeless Health Care, Boston Health Care for the Homeless Program, Boston, MA, United States

[3]Harvard Medical School, Boston, MA, United States

[4]Harvard TH Chan School of Public Health, Boston, MA, United States

[5]Boston University School of Medicine, Boston, MA, United States

Corresponding Author:
Travis P Baggett, MPH, MD
Tobacco Research and Treatment Center
Division of General Internal Medicine
Massachusetts General Hospital
100 Cambridge St
16th Floor
Boston, MA, 02114
United States
Phone: 1 617 643 9314
Email: tbaggett@mgh.harvard.edu

Abstract

Background: Homeless smokers want to quit smoking but face numerous barriers to doing so, including pervasive smoking among peers and a lack of social support for quitting. An SMS (short message service) text messaging intervention could address these challenges by providing virtual daily support for homeless smokers who are trying to quit but coping with multiple triggers to smoke.

Objective: This study aimed to assess whether a free SMS text messaging program, added to evidence-based pharmacotherapy and counseling, improved smoking abstinence among homeless adult smokers.

Methods: From October 2015 to June 2016, we conducted an 8-week pilot randomized controlled trial (RCT) of nicotine patch therapy and weekly in-person counseling with (n=25) or without (n=25) SmokefreeTXT, a free SMS text messaging service administered by the National Cancer Institute (NCI) at Boston Health Care for the Homeless Program. All participants were provided with a mobile phone and a 2-month prepaid voice and text plan at no cost. SmokefreeTXT enrollees were sent 1 to 5 automated SMS text messages daily for up to 8 weeks and could receive on-demand tips for managing cravings, mood symptoms, and smoking lapses. The primary outcome was smoking abstinence, defined as an exhaled carbon monoxide count of <8 parts per million, assessed 14 times over 8 weeks of follow-up, and analyzed using repeated-measures logistic regression with generalized estimating equations. Other outcomes were use of SmokefreeTXT, assessed by data obtained from NCI; perceptions of SmokefreeTXT, assessed by surveys and qualitative interviews; and mobile phone retention, assessed by self-report.

Results: Of the SmokefreeTXT arm participants (n=25), 88% (22) enrolled in the program, but only 56% (14) had confirmed enrollment for ≥2 weeks. Among 2-week enrollees, the median response rate to interactive messages from SmokefreeTXT was 2.1% (interquartile range 0-10.5%). Across all time points, smoking abstinence did not differ significantly between SmokefreeTXT and control arm participants (odds ratio 0.92, 95% CI 0.30-2.84). Of SmokefreeTXT enrollees who completed exit surveys (n=15), two-thirds were very or extremely satisfied with the program. However, qualitative interviews (n=14) revealed that many participants preferred in-person intervention formats over phone-based, found the SMS text messages impersonal and robotic, and felt that the messages were too frequent and repetitive. Only 40% (10/25) of SmokefreeTXT arm participants retained their study-supplied mobile phone for the 8-week duration of the trial, with phone theft being common. Storing and charging phones were cited as challenges.

Conclusions: SmokefreeTXT, added to nicotine patch therapy and in-person counseling, did not significantly improve smoking abstinence in this 8-week pilot RCT for homeless smokers. SMS text messaging interventions for this population should be better tuned to the unique circumstances of homelessness and coupled with efforts to promote mobile phone retention over time.

Trial Registration: ClinicalTrials.gov NCT02565381; https://clinicaltrials.gov/ct2/show/NCT02565381 (Archived by WebCite at http://www.webcitation.org/78PLpDptZ)

(JMIR Mhealth Uhealth 2019;7(6):e13162) doi:10.2196/13162

KEYWORDS

homeless persons; cigarette smoking; smoking cessation; text messaging

Introduction

Background

An estimated 2.3 to 3.5 million people experience homelessness each year in the United States [1], with over 553,000 homeless on any given night [2]. The prevalence of smoking among homeless adults is about 4 times higher than in the general population [3-8], contributing to over 2-fold higher rates of lung cancer [9] and 3- to 5-fold higher rates of tobacco-attributable death [10].

Studies have consistently demonstrated that homeless smokers want to quit smoking [4,11-14], but prior randomized controlled trials (RCTs) have not yet revealed the optimal approach to promoting smoking cessation in this vulnerable population. The largest such study to date examined the effect of motivational interviewing to promote nicotine patch adherence, which was not significantly better than control in producing smoking abstinence at 6 months [15].

The pervasiveness and social acceptability of smoking in the setting of homelessness [14,16] may contribute to low social support for quitting and frequent cues to continue smoking. In a survey of homeless smokers in Dallas, participants reported contact with a mean of 43 other smokers each day [17]. Treatment strategies must contend with this and other prosmoking influences that homeless smokers encounter daily outside the domain of a traditional health care setting.

Smoking cessation interventions delivered via a mobile device might enhance traditional interventions by providing virtual daily support to homeless smokers who are trying to quit but coping with multiple social and environmental triggers to relapse. Mobile phone possession is common among homeless people [18], and mobile technologies have been used to deliver appointment reminders to homeless veterans [19] and to collect ecological momentary assessment data from homeless smokers making a quit attempt [20,21]. Although SMS (short message service) text messaging interventions for smoking cessation have demonstrated efficacy across a range of settings and populations [22-33], no studies to our knowledge have examined SMS text messaging interventions for homeless smokers.

Objective

To address these gaps in the literature, our objective was to assess whether a free SMS text messaging program, SmokefreeTXT, improved smoking abstinence among homeless adult smokers when added to evidence-based pharmacotherapy and in-person counseling. Developed by the National Cancer Institute (NCI), SmokefreeTXT provides around-the-clock and on-demand support for adults who want to quit smoking [34] through SMS text messaging content that incorporates a variety of behavior change techniques [35]. Although SmokefreeTXT is not targeted specifically to homeless or low-income smokers, its accessibility and potential for immediate dissemination, if effective, made it appealing for testing in a highly impoverished population.

Methods

Study Design

We conducted a 3-arm, parallel group, nonblinded, 8-week pilot RCT that tested 2 separate smoking cessation interventions: (1) SMS text messaging delivered via SmokefreeTXT to support smoking abstinence and (2) financial incentives for smoking abstinence against (3) a shared control condition consisting of counseling and nicotine replacement therapy. We originally planned a 2-arm trial examining financial incentives against control treatment. However, additional funding enabled us to add a third study arm examining SMS text messaging relative to the same control condition. We prespecified a plan to analyze the SMS text messaging and financial incentives interventions separately because of the pilot nature of the trial and the differing rationales for each approach. This paper compares the effect of the SMS text messaging intervention with the control condition. The financial incentives findings are published elsewhere, as is a detailed description of the participants and setting, enrollment and randomization procedures, and baseline and follow-up assessments [36]. The following sections summarize those elements while emphasizing the details of the SMS text messaging intervention. The study protocol was approved by the Partners Human Research Committee and registered with ClinicalTrials.gov (NCT02565381) before commencement. All study procedures occurred from October 2015 to June 2016 at Boston Health Care for the Homeless Program (BHCHP) headquarters.

Enrollment and Randomization

Participants were recruited through in-person advertisement in the BHCHP lobby, flyers posted in BHCHP clinics, and referrals from BHCHP clinicians. Eligibility criteria were aged 18 years or older, lifetime smoking of 100 or more cigarettes [37] with current smoking of 5 or more cigarettes/day, verified by an exhaled carbon monoxide (CO) level of ≥8 parts per million (ppm) [38], readiness to quit smoking within the next month, current homelessness, and self-reported English proficiency. We defined current homelessness as usually staying in an

emergency shelter, transitional shelter, abandoned building, place of business, car or other vehicle, church or mission, hotel or motel, or anywhere outside during the past 7 days, or if currently in a residential treatment program, in the 7 days before program entry. In addition, individuals were considered currently homeless if they usually stayed in somebody else's place in the past 7 days because of not having their own place to stay. This definition is generally concordant with the US federal definition of homelessness [39] and identical to the definition that we and others have used in prior studies [40-44].

Exclusion criteria were current pregnancy, past-month use of any smoking cessation medication, prior serious adverse reaction to the nicotine patch, myocardial infarction or undiagnosed chest pain in the past 2 weeks, and inability to read a sentence written at a Flesch-Kincaid 4th-grade level. Participants were not excluded because of active substance use or mental illness.

We used a multistep enrollment process described elsewhere [36] to ensure that participants sufficiently understood the study and were committed to participating. All participants provided written informed consent to participate. Enrolled participants were randomized 1:1:1 to the SmokefreeTXT arm, financial incentives arm, or control arm. The allocation sequence was computer-generated in random permuted blocks and concealed from study staff.

Baseline Measures

Participants completed a baseline assessment of self-reported sociodemographic, health, and smoking characteristics at the time of enrollment. Sociodemographic measures included age, sex, and race and ethnicity. We asked participants to rate their general health status. We used the Addiction Severity Index-5th edition [45], which has been validated in homeless populations [46-48], to assess past-month alcohol use, drug use, and psychiatric symptom severity, with a problem in each of these domains based on score cut-offs described elsewhere [49]. Smoking characteristics included current cigarette dependence (assessed with the Fagerstrom test of nicotine dependence, FTND; range 0-10) [50], confidence to quit and perceived importance of quitting (each assessed using 10-point scales), and previous smoking cessation attempts.

Mobile Device

Participants in all study arms received a mobile phone with a QWERTY keyboard and a 2.4-inch display (AT&T ZTE Z432; retail US $29.99 each). The phones were activated by study staff at the time of randomization and loaded with a prepaid 2-month voice and text plan. We provided the same mobile device and plan to all participants in all arms, regardless of whether they already had a cell phone of their own, to deliver brief reminder phone calls and/or SMS text messages about upcoming study appointments, ensure uninterrupted mobile phone service for everyone, and standardize the end-user interface. Mobile devices and plans were provided at no cost to participants. Participants were informed that they would be given only 1 device to last the duration of the study. To incentivize phone retention, participants were allowed to keep their study phones if they still had them at the end of the trial but were told that they would need to cover the cost of any additional usage beyond the 2-month study period.

Control Arm

Participants assigned to the control arm were offered 8 weeks of nicotine patch therapy and weekly in-person counseling.

Nicotine Patch Therapy

Nicotine patches were distributed in 1-week allotments at no cost to participants. Participants who smoked 10 or more cigarettes per day were started on 21 mg/day patches and tapered to 14 mg/day at week 6. Participants who smoked less than 10 cigarettes per day were given 14 mg/day patches throughout the 8-week study.

In-Person Counseling

Participants were offered 8 weekly in-person counseling sessions lasting up to 15 min each. The counseling curriculum was developed by our study team in collaboration with a certified Master Tobacco Treatment Specialist (TTS) and was structured around the American Lung Association "Freedom from Smoking" program theme of addressing the physical, mental, and social aspects of tobacco addiction [51]. Counseling sessions incorporated elements of motivational interviewing and cognitive behavioral therapy and were tailored to the unique needs and circumstances of people experiencing homelessness. Before the trial, the study counselor completed a 9-module Web-based training course on basic skills for working with smokers, 6 hours of case-based didactics conducted by a certified Master TTS, 6 hours of observing a certified TTS provide counseling to smokers, and 3 hours of observing a clinician interact with homeless patients.

SmokefreeTXT Arm

Participants assigned to the SmokefreeTXT arm received nicotine patch therapy and in-person counseling in a fashion identical to the control arm. In addition, participants in this arm were offered assistance with enrolling in SmokefreeTXT. At enrollment, SmokefreeTXT prompted participants to set a quit date within the next 2 weeks, although the date could be adjusted later if desired. The program then sent 1 to 5 automated SMS text messages daily that provided encouragement, advice, and tips for quitting smoking. Message frequency and content varied according to where participants were in relation to their specified quit date as well as whether and how long they abstained from smoking following their quit date. SmokefreeTXT message content was updated periodically by NCI and was not tailored to homeless or low-income individuals. A behavior change technique analysis of the SmokefreeTXT library conducted contemporaneously with our study found that 14 of 16 behavioral technique groups outlined in the Behaviour Change Technique Taxonomy version 1 were present in SmokefreeTXT, with the most prevalent being feedback and monitoring, natural consequences, social support, and shaping knowledge [35]. Most SmokefreeTXT messages were unidirectional, but beginning on the quit date, a subset (approximately 23%) of messages were interactive in nature and solicited brief participant responses (eg, "Feelings can be a smoking trigger. If you feel cranky or grouchy, it is only temporary, so stay strong. How is your mood? Reply: GOOD, OK, or BAD"). In

addition, participants could spontaneously text keywords ("crave," "mood," or "slip") at any time to signal the need for automated supportive messaging around the management of cravings to smoke, distressing mood symptoms, or lapses in smoking abstinence. Participants could unsubscribe from SmokefreeTXT at any time. If participants lost their study-issued mobile phone but wished to continue using SmokefreeTXT and had another mobile device of their own, we offered to assist them in re-enrolling in SmokefreeTXT under their other phone number.

Assessment Procedure

Following randomization, all participants were asked to make 14 assessment visits over 8 weeks: thrice weekly during weeks 1 and 2, twice weekly during weeks 3 and 4, and once weekly during weeks 5 and 8. At each visit, study staff measured participants' exhaled CO levels using a Micro+ Smokerlyzer CO monitor (Bedfont Scientific Ltd; Maidstone, Kent, United Kingdom). The intensive nature of the follow-up assessment scheme ensured equal attention to all study arms, as the financial incentives intervention required frequent abstinence monitoring for maximum effect [52]. Control and SmokefreeTXT arm participants received US $10 payments onto study-supplied debit cards for each assessment visit they attended, regardless of whether they were abstinent. Participants in both arms received brief SMS text message reminders from study staff about their assessment visit appointments, and they also received public transportation tickets to facilitate attendance.

Outcomes

Smoking Outcomes

The prespecified primary study outcome was a repeated measure of brief smoking abstinence, defined as an exhaled CO <8 ppm [38] and assessed 14 times over 8 weeks. We used exhaled CO rather than nicotine metabolites to define smoking abstinence because the latter can be affected by nicotine replacement therapy [38], which was provided to all participants. We did not incorporate self-report into the primary outcome definition because the financial incentives arm provided abstinence-contingent financial rewards that created the potential for differential misreporting of smoking status. In a sensitivity analysis, we combined exhaled CO data with self-reported past-week smoking behavior to create 2 alternative abstinence outcomes, each assessed 8 times over 8 weeks: (1) Past 7-day cigarette abstinence: CO <8 ppm and last smoked all or part of a cigarette ≥7 days ago and (2) Past 7-day puff abstinence: CO <8 ppm and last smoked even a puff of a cigarette ≥7 days ago. Other smoking-related outcomes included past-month 24-hour quit attempts, assessed at 4 and 8 weeks.

Attendance

We assessed the number of study visits and the number of in-person counseling sessions attended by participants in each study arm.

SmokefreeTXT Engagement and Perceptions

We used a mixed methods approach to assess SmokefreeTXT arm participants' engagement with and perceptions of the texting program.

First, we obtained data from NCI to confirm whether participants enrolled in SmokefreeTXT, whether and when they unenrolled from the program, and whether they responded to interactive SMS text messages. Data on participants' spontaneous use of SmokefreeTXT keywords ("crave," "mood," or "slip") were not available.

Second, SmokefreeTXT arm participants who attended the final study visit were asked to complete a survey containing Likert-type items assessing their satisfaction with the SMS text messaging program, the perceived importance of the program in helping them to quit, the applicability of the messages to their everyday lives, and the likelihood that they would recommend the program to others.

Finally, 14 weeks after the study opened for enrollment, we added an in-person qualitative exit interview to assist in contextualizing and explaining the RCT results. Participants who had already completed the study were retroactively contacted about participating in the interview, whereas the remainder were prospectively offered the opportunity to complete the exit interview on or shortly after their final study visit. Study staff followed a semistructured interview guide that prompted all participants to reflect on their experience with each of the intervention components as well as with the mechanics of the study itself to inform future work.

SmokefreeTXT arm participants were additionally asked open-ended questions about whether SmokefreeTXT was useful in helping them to quit or cut back smoking, what they liked most about the SMS text messages they received, the frequency of the messages, and what they would change about the program. Participants received US $10 for completing the exit interview.

Mobile Phone Retention and Perceptions

The potential benefits of an SMS text messaging program can be realized only if participants retain a mobile device capable of receiving SMS text messages for the intended duration of the program and feel comfortable using that device to achieve health-related goals. To assess mobile phone retention, we asked participants at each assessment visit whether they still had their study-issued mobile phone. At the final study visit, we asked participants to rate their level of satisfaction with using a mobile phone to participate in a research study. In addition, the qualitative exit interviews described above asked participants about their perceptions of the features and benefits of their study-issued mobile phone, particularly in relation to their own phone if they had one, to better understand end-user mobile device preferences.

Analysis

Smoking Outcomes

We used repeated-measures logistic regression with generalized estimating equations (GEEs) to estimate the overall effect of SmokefreeTXT on the primary outcome of brief smoking abstinence (exhaled CO <8 ppm) across 14 measurements and on alternative definitions of smoking abstinence that combined exhaled CO data with self-reported smoking behavior (past 7-day cigarette and puff abstinence) across 8 weekly measurements. Our main analysis (1) included only treatment

effect in the GEE model, (2) assumed that those with missing abstinence data at any given time point were nonabstinent, and (3) was based on the intention to treat principle. In sensitivity analyses, we examined the impact of (1) multivariable adjustment for age, sex, race, nicotine dependence score, and baseline alcohol use, drug use, and psychiatric symptom severity scores; (2) multiple imputation of missing abstinence outcomes [53,54] based on nonmissing abstinence values in addition to age, sex, race/ethnicity, nicotine dependence score, and baseline alcohol use, drug use, and psychiatric symptom severity scores; and (3) reassessing the effect of SmokefreeTXT among those who enrolled in the program for at least 2 consecutive weeks (the point at which a quit date had to be set and interactive messaging began) to estimate the effect of treatment on the treated. We used repeated-measures linear regression with GEE to estimate the effect of SmokefreeTXT on past-month 24-hour quit attempts across 2 measurements (4 and 8 weeks).

Power Analysis

Before the start of the study, we used simulated data to estimate statistical power based on assumptions derived from the largest smoking cessation RCT for homeless individuals [15] and the most recent Cochrane systematic review of mobile phone–based interventions for smoking cessation [55]. The power analysis was based on the primary outcome of brief smoking abstinence measured 14 times over 8 weeks. The sample size (N=25 per arm) was dictated by the financial resources available for this pilot study. We assumed a similar pattern of missing data across both study arms, with 90% attendance at the first assessment, decreasing to 50% at the study midpoint, and staying flat thereafter until increasing to 80% at the last assessment. For the control arm, we assumed that CO-defined abstinence would be 10% at the first assessment and 8% thereafter. For the SmokefreeTXT arm, we assumed 25% abstinence at the first assessment, decreasing to 15 to 16% abstinence over subsequent visits. On the basis of these assumptions, we had 82% power to detect the specified differences in brief smoking abstinence.

Other Quantitative Analyses

We compared study visit attendance and counseling session attendance between the 2 arms using Wilcoxon tests.

We combined enrollment and unenrollment data from SmokefreeTXT with participant-confirmed mobile phone possession dates to estimate participants' actual exposure time to the program (ie, dates of SmokefreeTXT exposure during which cell phone possession could be confirmed). Among those with confirmed enrollment in SmokefreeTXT for at least 2 consecutive weeks, we calculated individual-level response rates to interactive SMS text messages by dividing the number of times a participant replied to an interactive message by the total number of interactive messages sent by the program during time periods of confirmed phone possession.

We used descriptive statistics to present the percentages of SmokefreeTXT arm participants who were "very" or "extremely" satisfied with the program, who rated the program as "very" or "extremely" important for quitting, who were "very" or "extremely" likely to recommend the program to others, and who "agreed" or "strongly agreed" that the messages were applicable to their lives. We also descriptively examined mobile phone retention duration as well as the proportion of SmokefreeTXT arm participants who were "very" or "extremely" satisfied with using a mobile phone for a research study.

Qualitative Analyses

Exit interviews were audio-recorded and transcribed verbatim by a member of the study staff, and the portions related to SmokefreeTXT and mobile phone usage were extracted for content analysis [56]. Using an inductive approach [57], 2 study staff members independently coded the transcripts. Coding was performed at the sentence level, and multiple coding was permitted. During iterative team meetings, major and minor themes were identified, refined, and organized into a hierarchical thematic framework that guided subsequent iterations of coding until all text had been categorized with an overall inter-rater reliability of kappa=.80. The overall kappa was calculated as the average kappa for all themes, with sources and themes weighted equally.

We used SAS 9.4 (SAS Institute) to conduct quantitative analyses and NVivo 10 (QSR International) to conduct qualitative analyses. Inferential analyses of smoking outcomes used a 2-sided significance level of 0.05.

Results

Screening, Enrollment, and Randomization

A total of 83 (67%) of 123 eligible individuals enrolled in the study and completed the baseline assessment (Figure 1). A total of 8 enrollees (10%) did not return for randomization; these participants reported higher baseline confidence to quit (P=.01) than the 75 participants who returned for randomization, but they did not differ significantly in other ways. The remainder of the results focus on the 50 participants randomized to the control arm (n=25) or SmokefreeTXT arm (n=25).

Figure 1. Consolidated Standards of Reporting Trials flow diagram. Sum of exclusion reasons totals greater than 117 because individuals could be ineligible for more than one reason. Smoking inclusion criteria were: (1) lifetime smoking of ≥100 cigarettes, (2) current daily smoking of ≥5 cigarettes per day, and (3) exhaled carbon monoxide level of ≥8 parts per million on 2 separate occasions.

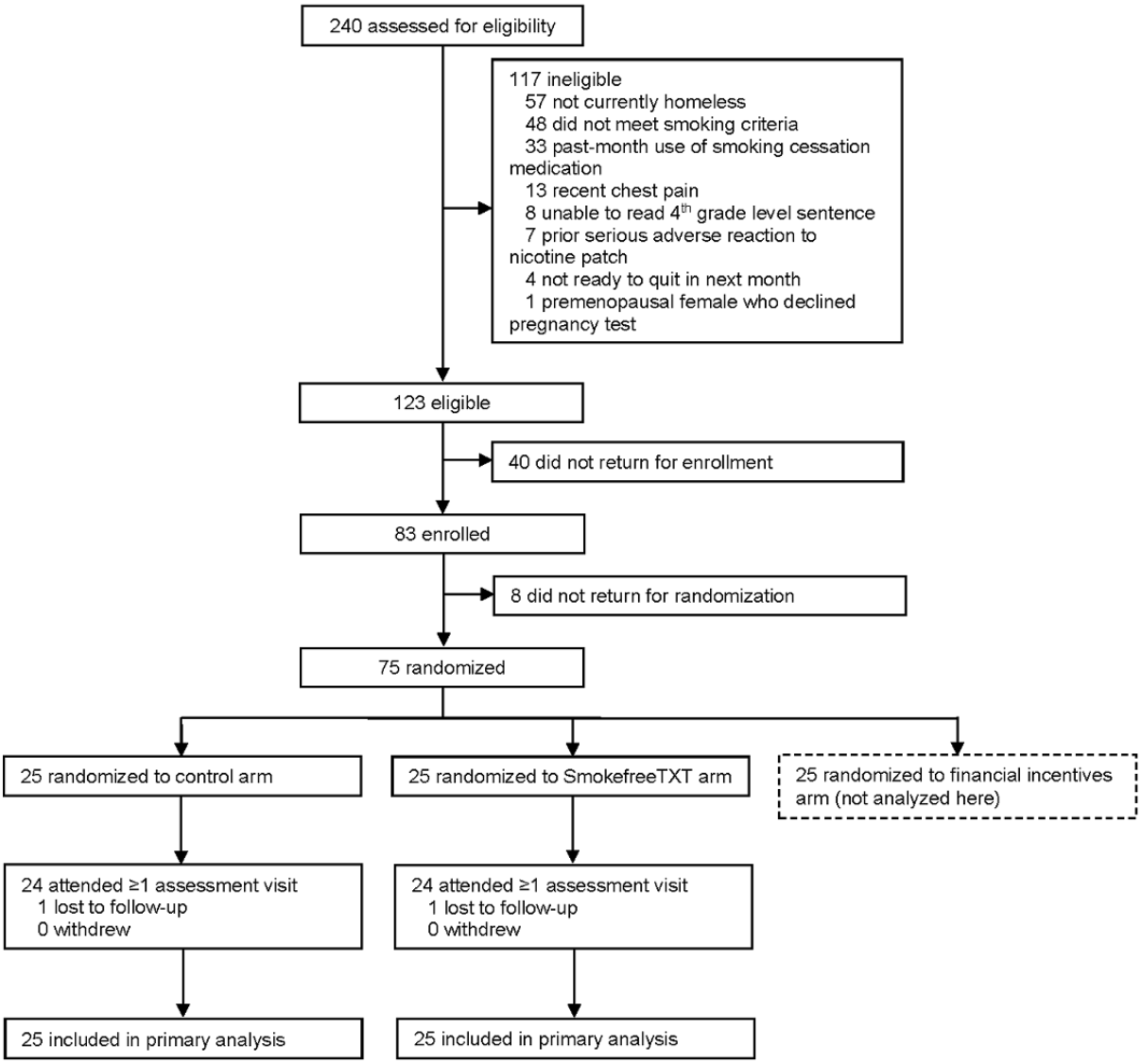


Table 1. Characteristics of study participants (n=50).

Characteristics	All	Control (n=25)	SmokefreeTXT (n=25)
Sociodemographic			
Age, years, mean (SD)	45.6 (9.3)	45.1 (9.6)	46.1 (9.2)
Female, n (%)	29 (58)	14 (56)	15 (60)
Race/ethnicity, n (%)			
White, non-Hispanic	20 (40)	10 (40)	10 (40)
Black, non-Hispanic	19 (38)	9 (36)	10 (40)
Hispanic	7 (14)	5 (20)	2 (8)
Other	4 (8)	1 (4)	3 (12)
Health			
Fair or poor health, n (%)	24 (48)	13 (52)	11 (44)
Alcohol problem, n (%)	10 (20)	3 (12)	7 (28)
Drug problem, n (%)	29 (59)	14 (58)	15 (60)
Psychiatric problem, n (%)	30 (60)	15 (60)	15 (60)
Smoking			
Cigarettes per day, mean (SD)	15.8 (6.4)	16.2 (6.3)	15.4 (6.7)
Nicotine dependence (0-10), mean (SD)	5.0 (1.9)	5.1 (1.8)	4.9 (2.0)
Past quit attempts, median (IQR[a])	3 (1-5)	2 (1-4)	3 (2-6)
Quitting importance (1-10), mean (SD)	8.9 (1.5)	8.8 (1.5)	8.9 (1.6)
Quitting confidence (1-10), mean (SD)	6.4 (2.3)	6.8 (2.2)	5.9 (2.4)

[a]IQR: interquartile range.

Baseline Characteristics

The mean age was 45.6 years, 58% of participants (n=29) were female, and 60% (n=30) were non-white (Table 1). Almost half (48%) rated their health as fair or poor. Considerable proportions had a current alcohol use problem (20%), current drug use problem (59%), or current psychiatric problem (60%). Participants smoked an average of 15.8 cigarettes per day and had a mean FTND score of 5.0.

Attendance

Participants attended a median of 10 (interquartile range, IQR 7-12) of 14 possible assessment visits, with no significant difference between arms (*P*=.72). A total of 96% of participants attended at least one assessment visit, and 76% attended at least half of the assessment visits. Participants attended a median of 1 (IQR 0-3) of 8 possible in-person counseling sessions, with no significant difference between arms (*P*=.77).

SmokefreeTXT Uptake and Engagement

A total of 22 (88%) of the 25 SmokefreeTXT arm participants enrolled in SmokefreeTXT; the remaining 3 declined to enroll but still participated in the study. Median confirmed exposure to SmokefreeTXT was estimated at 25 days (IQR 5-50 days), with 56% (n=14) being enrolled for at least 2 consecutive weeks. Among those 14 participants, the median response rate to interactive messages from SmokefreeTXT was 2.1% (IQR 0-10.5%).

Smoking Outcomes

Across the 14 follow-up visits, brief smoking abstinence (exhaled CO <8 ppm) did not differ significantly between the SmokefreeTXT and control arms (ranges 4-20% vs 0-20%; overall odds ratio [OR] 0.92, 95% CI 0.30-2.84; Figure 2). The treatment effect estimate was not substantively altered by either multivariable adjustment (OR 0.96, 95% CI 0.32-2.95) or multiple imputation of missing smoking status data (OR 1.06, 95% CI 0.72-1.58). Brief smoking abstinence among the 14 participants who were enrolled in SmokefreeTXT for at least 2 consecutive weeks did not differ significantly from control arm abstinence (OR 1.06, 95% CI 0.48-2.37).

When combining exhaled CO data with self-reported past-week smoking behavior over 8 weekly measurements (Table 2), there were no statistically significant differences between the SmokefreeTXT and control arms in either 7-day cigarette abstinence (range 0-12% vs 0-12%; OR 0.65, 95% CI 0.12-3.45) or 7-day puff abstinence (range 0-4% vs 0-4%; OR 2.01, 95% CI 0.19-21.1).

Across 2 monthly measurements, SmokfreeTXT and control arm participants did not differ significantly in past-month 24-hour quit attempts (mean difference –0.06, 95% CI –0.96 to 0.83).

Figure 2. Brief carbon monoxide–defined smoking abstinence by study arm. CO: carbon monoxide; OR: odds ratio; ppm: parts per million.

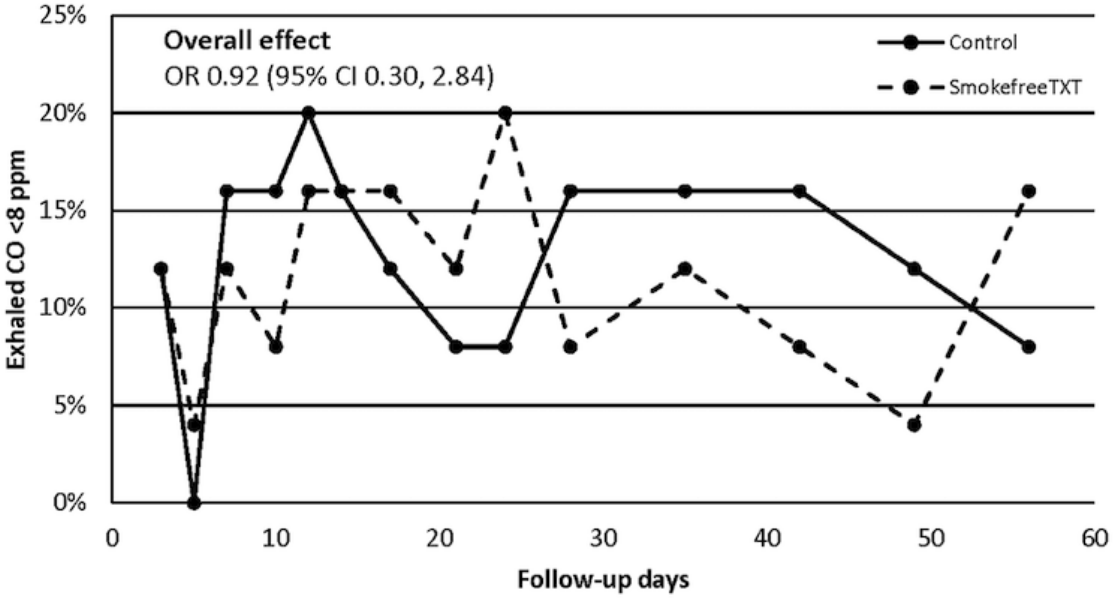


Table 2. Seven-day smoking abstinence by study arm. Individuals with missing carbon monoxide or self-reported smoking data are assumed not to meet the criteria for abstinence. Repeated-measures logistic regression with generalized estimating equations demonstrated no statistically significant overall differences between the SmokefreeTXT and control arms for either abstinence measure (see text).

Study week	Exhaled CO[a] <8 ppm[b] *and* last cigarette ≥7 days ago		Exhaled CO <8 ppm *and* last puff ≥7 days ago	
	Control, n (%)	SmokefreeTXT, n (%)	Control, n (%)	SmokefreeTXT, n (%)
1	1 (4)	1 (4)	0 (0)	1 (4)
2	3 (12)	2 (8)	1 (4)	0 (0)
3	0 (0)	0 (0)	0 (0)	0 (0)
4	2 (8)	1 (4)	0 (0)	0 (0)
5	2 (8)	3 (12)	0 (0)	1 (4)
6	2 (8)	1 (4)	0 (0)	0 (0)
7	1 (4)	0 (0)	0 (0)	0 (0)
8	1 (4)	0 (0)	0 (0)	0 (0)

[a]CO: carbon monoxide.

[b]ppm: parts per million.

SmokefreeTXT Perceptions

Of SmokefreeTXT enrollees who attended the final study visit (n=15), 67% were very/extremely satisfied with the program (0% not at all satisfied), 53% rated the program as very/extremely important in helping them to quit (0% not at all important), 93% were very/extremely likely to recommend the program to others (0% not at all likely), and 60% agreed or strongly agreed that the messages were applicable to their lives (0% strongly disagree).

Qualitative interviews (n=14) revealed more nuanced feedback on SmokefreeTXT. Major themes and representative quotes are displayed in Table 3. With respect to the program format, almost all interviewees expressed a preference for in-person rather than phone-based support (“I’d rather come in and talk.”). Indeed, 1 participant cited this as the reason they did not want to sign up for SmokefreeTXT (“Well I just wanted to do it in person instead of on the phone.”). Some felt that the SMS text messages were too impersonal (“Like a robot, everyone gets the same text.”). Views on the interactive aspects were mixed; some admitted to making little use of this feature, whereas others liked it and wanted more. The timing of messages was viewed as important, but many felt that there were too many SMS text messages overall (“They always came poppin’ in all the time.”). With respect to message content, some users found the SMS text messages informative and useful, although many reported that the content was too repetitive (“Different wordings, same thing.”). Occasional interviewees commented on memorable SMS text messages containing recommendations that were at odds with their living circumstances (“There’s one that said ‘Take a bath’...and I’m like ‘Really?... This is a homeless study folks.’”). Participants generally found the message tone to be encouraging, although 1 participant described having a strongly negative reaction to some messages (“I got mad at them texts.”).

Table 3. Qualitative findings related to SmokefreeTXT.

Thematic domains	Findings and sample quotes
Program format	
Comparison with in-person format	Universal preference for in-person over phone-based counseling: "I'd rather come in and talk. … I like to do things in person." "I'd rather just be, up in person. …So that way you can see my expressions, x, y, z." "I'd love it in person so we can see eye to eye."
Automation	Format felt robotic and impersonal to some: "They're kind of, like, impersonal, you know? Like a robot, everyone gets the same text."
Interactiveness	Some made little use of available interactive features: "Um, I never, I don't think I've ever texted them back." "I responded a couple of times. A lot of times I started to um, kind of honestly disregard it."
	Others liked interactive features but wanted more: "The interactive stuff is good. …If they had more of that I think people would get involved more."
Frequency of messages	Many felt there were too many: "…they always came poppin' in all the time. Like 'ring ring ring.'" "No more than once a day. That's just crazy. What you got to talk about?"
	Some found the frequency appropriate or wanted more: "I think I needed more."
Timing of messages	Viewed as important; desire for customizing timing: "Earlier in the day, morning, I would change them to the morning instead of evening/afternoon."
Program content	
Usefulness for quitting	Many found them useful: "They kept me going… when I was getting ready to smoke it 'binged' then and saying, 'don't pick up' and I would put the cigarette down."
	Some found them unhelpful: "It was kind of like 'ok, thank you.' I'm moving on with my day." "I mean I read them, but they didn't really change my… they didn't really change anything for me."
Informativeness	Some learned new information: "I loved it… lot of information of things that I wasn't aware of, related to smoking."
	Others reported the information or tips were not new: "Some of them were stuff I already knew." "It's like, some of them, I already do that."
Applicability	Generally deemed applicable, but non-applicable texts were memorable: "Um, do some laundry type of questions... I was like 'yeah ok.'" "Like there's one that said 'take a bath' … And I'm like 'Really? This is a homeless study folks.'"
Repetitiveness	Many found content repetitive: "A few of them started getting repetitive. I'd think 'oh dammit I saw that one two days ago.'" "Majority of what they said – different wordings, same thing."
Tone	Generally found messages encouraging: "I like the message that they're trying to send and trying to convey."
	One found texts critical at times: "I got mad at them texts. …I feel like they were being hard on me in some of the texts. …It was – this one sticks with me – 'nobody told you it was going to be easy.' …Sticks with me."

Of note, SmokefreeTXT arm participants who completed exit interviews (n=14) were less nicotine dependent at baseline (mean FTND score 4.2 vs 5.8; *P*=.04), had better study attendance (median visits 11.5 vs 9.0; *P*=.04), and were more likely to attain the primary outcome of brief CO-defined smoking abstinence (OR 4.41, 95% CI 1.20-16.2) than SmokefreeTXT arm participants who did not complete exit interviews (n=11).

Mobile Phone Retention and Perceptions

Over three-quarters (77%) of SmokefreeTXT arm participants who attended the final study visit reported being very/extremely satisfied with using a mobile phone to participate in a study. However, median study phone retention was 41 days (IQR 12-57), and only 40% of SmokefreeTXT arm participants retained their phone for the entire 8-week study. Theft was the most commonly reported reason for phone loss. This was reiterated in qualitative interviews, where participants also cited logistical challenges with charging and storing their phones ("Honest to goodness that phone was mostly in my bag...in that bag in my locker at the shelter."). A total of 76% of SmokefreeTXT arm participants already had another mobile phone before the study, and some of these individuals reported that they preferred their own phone to the study phone ("Most of the time I use my regular cell phone...It's just what I use and check more often."). Others preferred the study phone, especially if they did not have their own phone or if their own phone had limited features. Keyboard layout and screen size were cited as important factors in phone usability ("Maybe bigger for the punches...'cause sometimes when you press one button it presses another one.").

Discussion

Principal Findings and Implications

This is one of very few controlled studies to assess the effect of SmokefreeTXT on smoking outcomes, the first to assess biochemically verified smoking abstinence in SmokefreeTXT users, and the first study of an SMS text messaging intervention for homeless smokers. We found that enrollment in SmokefreeTXT was high, but sustained use of the program was modest, interaction with the program was minimal, and perceptions of the program were mixed. When added to nicotine replacement therapy and in-person counseling, SmokefreeTXT had no significant effect on any measure of smoking abstinence or self-reported quit attempts.

Our findings add to the mixed results seen in prior trials conducted in health care settings of SMS text messaging interventions for smoking cessation [32,58-61]. Our results also extend those of prior observational studies of SmokefreeTXT [62,63] and a related NCI SMS text messaging program designed specifically for veterans [64] demonstrating the challenges of sustaining end-user engagement in this program. End of treatment CO-defined brief smoking abstinence rates in our study were comparable with self-reported abstinence rates in these large-scale, noncontrolled, observational studies [62-64] but were notably lower than self-reported abstinence rates in an emergency department-based pilot study of a multicomponent smoking cessation program that included SmokefreeTXT [65].

Although quantitative scales suggested that the majority of participants were satisfied with SmokefreeTXT, qualitative exit interviews revealed more nuanced perceptions and provided insights about potential reasons for the absence of a treatment effect. Most participants expressed a strong preference for in-person treatment modalities and did not care for the automated and impersonal aspects of an SMS text messaging program, underscoring the importance of personal relationships when working with a highly marginalized population. Although most participants considered the messages to be a good source of support for quitting, several found them to be too frequent and too repetitive. In addition, selected messages struck some participants as being memorably incompatible with their daily lives. As SmokefreeTXT arm participants who completed qualitative interviews were more likely to achieve brief smoking abstinence than those who were not interviewed, these mixed views may represent a best-case scenario of end-user experience with the program. Altogether, these findings suggest that SMS text messaging interventions for homeless individuals might be more successful if the delivery format were more customizable and the content better targeted to the unique circumstances of homelessness. Such changes might enhance SMS text messaging program engagement, which appears to be an important determinant of treatment response in other settings [22,62-64].

The itinerancy of homelessness and the high prevalence of mobile phone possession among homeless people has suggested a potentially important role for mobile health interventions targeting this vulnerable population [66]. However, our study underscores the difficulty that homeless individuals may have in retaining a single mobile device over a period of time sufficient to receive an intervention through it. Half of all 75 trial participants and only 40% of SmokefreeTXT arm participants still had their study-issued phone after 8 weeks. Qualitative interviews highlighted several challenges to successfully using these phones because of loss, theft, and problems with storage and recharging. As a result, efforts to develop mobile health interventions for homeless people will need to be coupled with innovations to help promote mobile device retention and use.

The negligible effect of both study conditions on complete 7-day smoking abstinence at any time point is disappointing and highlights the challenges to achieving long-term smoking cessation in this vulnerable group of smokers. This finding underscores that the optimal approach to promoting smoking cessation in this population remains uncertain but may ultimately require a combination of traditional and nontraditional intervention modalities targeting multiple levels of influence [16].

Limitations

An important limitation of our study is that SmokefreeTXT was added to a relatively robust evidence-based tobacco treatment regimen of nicotine patch therapy and in-person counseling as well as frequent in-person abstinence monitoring. The high-contact nature of the study could have overwhelmed a small but meaningful treatment effect for SmokefreeTXT, which might be more optimally suited for smokers not already engaged (or able to be engaged) in more intensive treatment modalities. Another limitation is that participants in the control arm could have gained access to SmokefreeTXT without our assistance as it is free and publicly accessible, introducing the possibility of contamination. However, control arm participants were not told about this program, and surveys of control arm participants at 4 and 8 weeks found that none reported using an SMS text messaging program to support smoking cessation in the prior month. Other limitations of the trial include the small sample size and short duration of the study. Although our repeated-measures approach to outcome measurement offered adequate power to detect what we deemed to be a clinically important effect on CO-defined abstinence, we may have been underpowered to detect smaller treatment effects. Our qualitative interviews were wide-ranging in scope and not focused exclusively on SmokefreeTXT, and sampling was dictated by the number of trial participants in each arm who were willing to be interviewed rather than by attainment of thematic saturation. However, each of the identified themes emerged from more than 1 source, and later interviews generally did not uncover new thematic areas. Finally, our study was conducted at a large homeless health care program in Boston, so the findings may not be generalizable to other settings.

Conclusions

SmokefreeTXT, added to nicotine patch therapy and in-person counseling, did not improve smoking abstinence in this 8-week pilot RCT for homeless smokers. Although program uptake was high, interaction with the program was minimal, and mobile device loss was common. Our qualitative findings suggest that future SMS text messaging interventions for this population should be better tuned to the unique circumstances of

homelessness and coupled with innovative efforts to promote mobile phone retention over time.

Acknowledgments

This study was supported by awards K23DA034008 (TPB) and K23DA038717 (GRK) from the National Institute on Drug Abuse at the National Institutes of Health, and by the Massachusetts General Hospital Department of Medicine Transformative Scholars Program (TPB). These entities had no role in any aspect of the study. The study content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or Massachusetts General Hospital. The authors would like to thank the study participants in addition to the staff and patients of BHCHP.

Conflicts of Interest

TPB receives royalties from UpToDate for authorship of a topic review on the health care of homeless people in the United States. GRK has a family financial interest in Dimagi Inc and is a paid consultant for Click Therapeutics. NAR has consulted without pay for Pfizer and for pay for Achieve Life Sciences regarding smoking cessation medications. She receives royalties from UpToDate for authorship of topic reviews on smoking cessation. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1

CONSORT-eHealth checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i6e13162_app1.pdf]

References

1. Burt MR, Aron LY, Lee E, Valente J. How Many Homeless People Are There? Helping America's Homeless: Emergency Shelter or Affordable Housing?. Washington, DC: Urban Institute; 2001:23-54.
2. Henry M, Watt R, Rosenthal L, Shivji A, Abt Associates. HUD Exchange. 2017 Dec. The 2017 Annual Homeless Assessment Report (AHAR) to Congress; Part 1: Point-in-Time Estimates of Homelessness URL: https://www.hudexchange.info/resource/5639/2017-ahar-part-1-pit-estimates-of-homelessness-in-the-us/ [accessed 2019-05-14] [WebCite Cache ID 78Mz5OedS]
3. Baggett TP, Rigotti NA. Cigarette smoking and advice to quit in a national sample of homeless adults. Am J Prev Med 2010 Aug;39(2):164-172. [doi: 10.1016/j.amepre.2010.03.024] [Medline: 20621264]
4. Connor SE, Cook RL, Herbert MI, Neal SM, Williams JT. Smoking cessation in a homeless population: there is a will, but is there a way? J Gen Intern Med 2002 May;17(5):369-372 [FREE Full text] [Medline: 12047734]
5. Snyder LD, Eisner MD. Obstructive lung disease among the urban homeless. Chest 2004 May;125(5):1719-1725. [Medline: 15136382]
6. Szerlip MI, Szerlip HM. Identification of cardiovascular risk factors in homeless adults. Am J Med Sci 2002 Nov;324(5):243-246. [Medline: 12449444]
7. Tsai J, Rosenheck RA. Smoking among chronically homeless adults: prevalence and correlates. Psychiatr Serv 2012 Jun;63(6):569-576. [doi: 10.1176/appi.ps.201100398] [Medline: 22476200]
8. Chau S, Chin M, Chang J, Luecha A, Cheng E, Schlesinger J, et al. Cancer risk behaviors and screening rates among homeless adults in Los Angeles County. Cancer Epidemiol Biomarkers Prev 2002 May;11(5):431-438 [FREE Full text] [Medline: 12010856]
9. Baggett TP, Chang Y, Porneala BC, Bharel M, Singer DE, Rigotti NA. Disparities in cancer incidence, stage, and mortality at Boston Health Care for the Homeless Program. Am J Prev Med 2015 Nov;49(5):694-702 [FREE Full text] [doi: 10.1016/j.amepre.2015.03.038] [Medline: 26143955]
10. Baggett TP, Chang Y, Singer DE, Porneala BC, Gaeta JM, O'Connell JJ, et al. Tobacco-, alcohol-, and drug-attributable deaths and their contribution to mortality disparities in a cohort of homeless adults in Boston. Am J Public Health 2015 Jun;105(6):1189-1197. [doi: 10.2105/AJPH.2014.302248] [Medline: 25521869]
11. Baggett TP, Lebrun-Harris LA, Rigotti NA. Homelessness, cigarette smoking and desire to quit: results from a US national study. Addiction 2013 Nov;108(11):2009-2018 [FREE Full text] [doi: 10.1111/add.12292] [Medline: 23834157]
12. Arnsten JH, Reid K, Bierer M, Rigotti N. Smoking behavior and interest in quitting among homeless smokers. Addict Behav 2004 Aug;29(6):1155-1161. [doi: 10.1016/j.addbeh.2004.03.010] [Medline: 15236817]
13. Butler J, Okuyemi KS, Jean S, Nazir N, Ahluwalia JS, Resnicow K. Smoking characteristics of a homeless population. Subst Abus 2002 Dec;23(4):223-231. [doi: 10.1080/08897070209511495] [Medline: 12438835]
14. Okuyemi KS, Caldwell AR, Thomas JL, Born W, Richter KP, Nollen N, et al. Homelessness and smoking cessation: insights from focus groups. Nicotine Tob Res 2006 Apr;8(2):287-296. [doi: 10.1080/14622200500494971] [Medline: 16766421]

15. Okuyemi KS, Goldade K, Whembolua G, Thomas JL, Eischen S, Sewali B, et al. Motivational interviewing to enhance nicotine patch treatment for smoking cessation among homeless smokers: a randomized controlled trial. Addiction 2013 Jun;108(6):1136-1144 [FREE Full text] [doi: 10.1111/add.12140] [Medline: 23510102]
16. Baggett TP, Tobey ML, Rigotti NA. Tobacco use among homeless people--addressing the neglected addiction. N Engl J Med 2013 Jul 18;369(3):201-204. [doi: 10.1056/NEJMp1301935] [Medline: 23863048]
17. Businelle MS, Cuate EL, Kesh A, Poonawalla IB, Kendzor DE. Comparing homeless smokers to economically disadvantaged domiciled smokers. Am J Public Health 2013 Dec;103(Suppl 2):S218-S220. [doi: 10.2105/AJPH.2013.301336] [Medline: 24148069]
18. McInnes DK, Li AE, Hogan TP. Opportunities for engaging low-income, vulnerable populations in health care: a systematic review of homeless persons' access to and use of information technologies. Am J Public Health 2013 Dec;103(Suppl 2):e11-e24. [doi: 10.2105/AJPH.2013.301623] [Medline: 24148036]
19. McInnes DK, Petrakis BA, Gifford AL, Rao SR, Houston TK, Asch SM, et al. Retaining homeless veterans in outpatient care: a pilot study of mobile phone text message appointment reminders. Am J Public Health 2014 Sep;104(Suppl 4):S588-S594. [doi: 10.2105/AJPH.2014.302061] [Medline: 25100425]
20. Reitzel LR, Kendzor DE, Nguyen N, Regan SD, Okuyemi KS, Castro Y, et al. Shelter proximity and affect among homeless smokers making a quit attempt. Am J Health Behav 2014 Mar;38(2):161-169 [FREE Full text] [doi: 10.5993/AJHB.38.2.1] [Medline: 24629545]
21. Businelle MS, Ma P, Kendzor DE, Reitzel LR, Chen M, Lam CY, et al. Predicting quit attempts among homeless smokers seeking cessation treatment: an ecological momentary assessment study. Nicotine Tob Res 2014 Oct;16(10):1371-1378 [FREE Full text] [doi: 10.1093/ntr/ntu088] [Medline: 24893602]
22. Abroms LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of Text2Quit: a text messaging program for smoking cessation. Am J Prev Med 2014 Sep;47(3):242-250. [doi: 10.1016/j.amepre.2014.04.010] [Medline: 24913220]
23. Bock B, Heron K, Jennings E, Morrow K, Cobb V, Magee J, et al. A text message delivered smoking cessation intervention: the initial trial of TXT-2-Quit: randomized controlled trial. JMIR Mhealth Uhealth 2013;1(2):e17 [FREE Full text] [doi: 10.2196/mhealth.2522] [Medline: 25098502]
24. Brendryen H, Drozd F, Kraft P. A digital smoking cessation program delivered through internet and cell phone without nicotine replacement (happy ending): randomized controlled trial. J Med Internet Res 2008;10(5):e51 [FREE Full text] [doi: 10.2196/jmir.1005] [Medline: 19087949]
25. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]
26. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
27. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin R, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
28. Scott-Sheldon LAJ, 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;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
29. Ybarra ML, Jiang Y, Free C, Abroms LC, Whittaker R. Participant-level meta-analysis of mobile phone-based interventions for smoking cessation across different countries. Prev Med 2016 Aug;89:90-97 [FREE Full text] [doi: 10.1016/j.ypmed.2016.05.002] [Medline: 27154349]
30. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
31. Pollak KI, Lyna P, Bilheimer A, Farrell D, Gao X, Swamy GK, et al. A pilot study testing SMS text delivered scheduled gradual reduction to pregnant smokers. Nicotine Tob Res 2013 Oct;15(10):1773-1776 [FREE Full text] [doi: 10.1093/ntr/ntt045] [Medline: 23569007]
32. Cobos-Campos R, Apiñaniz FD, Sáez DL, Parraza DN, Aizpuru BF. Effectiveness of text messaging as an adjuvant to health advice in smoking cessation programs in primary care. A randomized clinical trial. Nicotine Tob Res 2017 Aug 1;19(8):901-907. [doi: 10.1093/ntr/ntw300] [Medline: 27838659]
33. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]
34. Smokefree.gov. SmokefreeTXT - About URL: https://smokefree.gov/smokefreetxt-about [accessed 2019-05-14] [WebCite Cache ID 78Mzpx2n4]
35. Stoyneva I, Coa K, Pugatch J, Sanders A, Schwarz M, Cole-Lewis H. SmokefreeTXT behaviour change technique analysis. J Smok Cessat 2016 Dec 20;12(4):231-243. [doi: 10.1017/jsc.2016.24]

36. Baggett TP, Chang Y, Yaqubi A, McGlave C, Higgins ST, Rigotti NA. Financial incentives for smoking abstinence in homeless smokers: a pilot randomized controlled trial. Nicotine Tob Res 2018 Nov 15;20(12):1442-1450. [doi: 10.1093/ntr/ntx178] [Medline: 29059442]
37. Jamal A, Agaku IT, O'Connor E, King BA, Kenemer JB, Neff L. Current cigarette smoking among adults--United States, 2005-2013. MMWR Morb Mortal Wkly Rep 2014 Nov 28;63(47):1108-1112 [FREE Full text] [Medline: 25426653]
38. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
39. HUD Exchange. Sec. 1003. Definition of Homelessness. Homeless Emergency Assistance and Rapid Transition to Housing Act of 2009. S. 896, 34-35 URL: https://www.hudexchange.info/resource/1717/s-896-hearth-act/ [accessed 2019-05-14] [WebCite Cache ID 78N0jmZ4U]
40. Hwang SW, Colantonio A, Chiu S, Tolomiczenko G, Kiss A, Cowan L, et al. The effect of traumatic brain injury on the health of homeless people. CMAJ 2008 Oct 7;179(8):779-784 [FREE Full text] [doi: 10.1503/cmaj.080341] [Medline: 18838453]
41. Grinman MN, Chiu S, Redelmeier DA, Levinson W, Kiss A, Tolomiczenko G, et al. Drug problems among homeless individuals in Toronto, Canada: prevalence, drugs of choice, and relation to health status. BMC Public Health 2010 Feb 24;10:94 [FREE Full text] [doi: 10.1186/1471-2458-10-94] [Medline: 20181248]
42. Baggett TP, Campbell EG, Chang Y, Magid LM, Rigotti NA. Posttraumatic stress symptoms and their association with smoking outcome expectancies among homeless smokers in Boston. Nicotine Tob Res 2016 Dec;18(6):1526-1532 [FREE Full text] [doi: 10.1093/ntr/ntv238] [Medline: 26508393]
43. Baggett TP, Campbell EG, Chang Y, Rigotti NA. Other tobacco product and electronic cigarette use among homeless cigarette smokers. Addict Behav 2016 Dec 19;60:124-130 [FREE Full text] [doi: 10.1016/j.addbeh.2016.04.006] [Medline: 27128808]
44. Baggett TP, Rigotti NA, Campbell EG. Cost of smoking among homeless adults. N Engl J Med 2016 Feb 18;374(7):697-698 [FREE Full text] [doi: 10.1056/NEJMc1508556] [Medline: 26886544]
45. McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, et al. The Fifth Edition of the Addiction Severity Index. J Subst Abuse Treat 1992;9(3):199-213. [Medline: 1334156]
46. Zanis DA, McLellan AT, Cnaan RA, Randall M. Reliability and validity of the Addiction Severity Index with a homeless sample. J Subst Abuse Treat 1994;11(6):541-548. [Medline: 7884837]
47. Argeriou M, McCarty D, Mulvey K, Daley M. Use of the Addiction Severity Index with homeless substance abusers. J Subst Abuse Treat 1994;11(4):359-365. [Medline: 7966506]
48. Drake RE, McHugo GJ, Biesanz JC. The test-retest reliability of standardized instruments among homeless persons with substance use disorders. J Stud Alcohol 1995 Mar;56(2):161-167. [Medline: 7760561]
49. Burt MR, Aron LY, Douglas T, Valente J, Lee E, Iwen B. Homelessness: Programs and the People They Serve. Findings of the National Survey of Homeless Assistance Providers and Clients: Technical Report. Washington, DC: The Urban Institute. US Department of Housing and Urban Development, Office of Policy Development and Research; 1999. URL: https://tinyurl.com/yyeynwt5 [WebCite Cache ID 78YujqMAe]
50. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
51. Freedom from Smoking. URL: http://www.freedomfromsmoking.org/ [accessed 2019-05-14] [WebCite Cache ID 78NMEWmkb]
52. Higgins ST, Redner R, White TJ. Contingency management and the community reinforcement approach. In: Ries RK, Fiellin DA, Miller SC, Saitz R, editors. The ASAM Principles of Addiction Medicine. Fifth Edition. Philadelphia: Wolters Kluwer; 2014:877-893.
53. Nelson DB, Partin MR, Fu SS, Joseph AM, An LC. Why assigning ongoing tobacco use is not necessarily a conservative approach to handling missing tobacco cessation outcomes. Nicotine Tob Res 2009 Jan;11(1):77-83. [doi: 10.1093/ntr/ntn013] [Medline: 19246444]
54. Hedeker D, Mermelstein RJ, Demirtas H. Analysis of binary outcomes with missing data: missing = smoking, last observation carried forward, and a little multiple imputation. Addiction 2007 Oct;102(10):1564-1573. [doi: 10.1111/j.1360-0443.2007.01946.x] [Medline: 17854333]
55. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
56. Trochim WM, Donnelly JP. Qualitative and Mixed Methods Analysis. The Research Methods Knowledge Base. Third Edition. Mason, OH: Thomson; 2007.
57. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
58. Naughton F, Cooper S, Foster K, Emery J, Leonardi-Bee J, Sutton S, et al. Large multi-centre pilot randomized controlled trial testing a low-cost, tailored, self-help smoking cessation text message intervention for pregnant smokers (MiQuit). Addiction 2017 Jul;112(7):1238-1249 [FREE Full text] [doi: 10.1111/add.13802] [Medline: 28239919]

59. Naughton F, Jamison J, Boase S, Sloan M, Gilbert H, Prevost AT, et al. Randomized controlled trial to assess the short-term effectiveness of tailored web- and text-based facilitation of smoking cessation in primary care (iQuit in practice). Addiction 2014 Jul;109(7):1184-1193 [FREE Full text] [doi: 10.1111/add.12556] [Medline: 24661312]
60. Yingst JM, Veldheer S, Hrabovsky S, Hammett E, Nicholson J, Berg A, et al. Pilot randomized trial of an automated smoking cessation intervention via mobile phone text messages as an adjunct to varenicline in primary care. J Health Commun 2018;23(4):370-378. [doi: 10.1080/10810730.2018.1453890] [Medline: 29578832]
61. Tseng T, Krebs P, Schoenthaler A, Wong S, Sherman S, Gonzalez M, et al. Combining text messaging and telephone counseling to increase varenicline adherence and smoking abstinence among cigarette smokers living with HIV: a randomized controlled study. AIDS Behav 2017 Jul;21(7):1964-1974 [FREE Full text] [doi: 10.1007/s10461-016-1538-z] [Medline: 27605365]
62. Augustson E, Cole-Lewis H, Sanders A, Schwarz M, Geng Y, Coa K, et al. Analysing user-reported data for enhancement of SmokefreeTXT: a national text message smoking cessation intervention. Tob Control 2017 Nov;26(6):683-689. [doi: 10.1136/tobaccocontrol-2016-052945] [Medline: 27852892]
63. Coa KI, Wiseman KP, Higgins B, Augustson E. Associations between engagement and outcomes in the SmokefreeTXT program: a growth mixture modeling analysis. Nicotine Tob Res 2018 Apr 16:16. [doi: 10.1093/ntr/nty073] [Medline: 29668984]
64. Christofferson DE, Hertzberg JS, Beckham JC, Dennis PA, Hamlett-Berry K. Engagement and abstinence among users of a smoking cessation text message program for veterans. Addict Behav 2016 Nov;62:47-53. [doi: 10.1016/j.addbeh.2016.06.016] [Medline: 27318948]
65. Bernstein SL, Rosner J, Toll B. A multicomponent intervention including texting to promote tobacco abstinence in emergency department smokers: a pilot study. Acad Emerg Med 2016 Dec;23(7):803-808 [FREE Full text] [doi: 10.1111/acem.12990] [Medline: 27146116]
66. Post LA, Vaca FE, Doran KM, Luco C, Naftilan M, Dziura J, et al. New media use by patients who are homeless: the potential of mHealth to build connectivity. J Med Internet Res 2013 Sep 3;15(9):e195 [FREE Full text] [doi: 10.2196/jmir.2724] [Medline: 24001876]

Abbreviations

BHCHP: Boston Health Care for the Homeless Program
CO: carbon monoxide
GEE: generalized estimating equation
IQR: interquartile range
NCI: National Cancer Institute
OR: odds ratio
ppm: parts per million
RCT: randomized controlled trial
SMS: short message service
TTS: Tobacco Treatment Specialist

Edited by K Eddens, G Eysenbach; submitted 16.12.18; peer-reviewed by N Machado, P Krebs, S Linke, F Fries; comments to author 04.02.19; revised version received 27.02.19; accepted 23.04.19; published 04.06.19.

Please cite as:
Baggett TP, McGlave C, Kruse GR, Yaqubi A, Chang Y, Rigotti NA
SmokefreeTXT for Homeless Smokers: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(6):e13162
URL: https://mhealth.jmir.org/2019/6/e13162/
doi:10.2196/13162
PMID:31165717

Original Paper

Mindfulness-Based Smoking Cessation Enhanced With Mobile Technology (iQuit Mindfully): Pilot Randomized Controlled Trial

Claire Adams Spears[1], PhD; Lorien C Abroms[2], ScD; Carol R Glass[3], PhD; Donald Hedeker[4], PhD; Michael P Eriksen[1], ScD; Cherell Cottrell-Daniels[1], MPH; Binh Q Tran[5], PhD; David W Wetter[6], PhD

[1]Department of Health Policy and Behavioral Sciences, Georgia State University School of Public Health, Atlanta, GA, United States

[2]Prevention and Community Health, Milken Institute School of Public Health, George Washington University, Washington, DC, United States

[3]Department of Psychology, The Catholic University of America, Washington, DC, United States

[4]Department of Public Health Sciences, The University of Chicago, Chicago, IL, United States

[5]Department of Biomedical Engineering, The Catholic University of America, Washington, DC, United States

[6]Center for Health Outcomes and Population Equity, University of Utah and Huntsman Cancer Institute, Salt Lake City, UT, United States

Corresponding Author:
Claire Adams Spears, PhD
Department of Health Policy and Behavioral Sciences
Georgia State University School of Public Health
140 Decatur St SE
Atlanta, GA, 30303
United States
Phone: 1 404 413 9335
Email: cspears@gsu.edu

Abstract

Background: Mindfulness training shows promise for improving smoking cessation and lapse recovery, and between-session mobile health messages could enhance treatment engagement and effectiveness. Personalized, in-the-moment text messaging support could be particularly useful for low-income smokers with fewer smoking cessation resources.

Objective: This pilot study examined the feasibility of a text messaging program (*iQuit Mindfully*) as an adjunct to in-person Mindfulness-Based Addiction Treatment (MBAT) for smoking cessation.

Methods: A total of 71 participants were randomly assigned to MBAT (n=33) or iQuit Mindfully (n=38; MBAT + between-session text messages); of these, 70% (50/71) were African American, and 61% (43/71) had an annual household income of US $30,000 or less. All participants received 8 weekly therapist-led group counseling sessions, nicotine patches, and self-help materials. Outcomes were feasibility (attrition, engagement, and participants' ratings), participants' feedback regarding the text messaging intervention, and smoking cessation (assessed in person).

Results: Strong retention was achieved (76% [54/71] at the end of treatment, and 89% [63/71] at 1-month follow-up). In the iQuit Mindfully group, engagement was high (88% [29/33] indicated reading all or most texts, and 89% [34/38] engaged in interactive texting), and participants provided positive ratings (on a 1-10 scale, average rating for recommending the program to others was 8.4 [SD 2.5]). Participants indicated benefiting from the texts (eg, appreciating encouraging reminders, coping strategies, and social support) and suggested improvements (eg, more personalization). Overall, biochemically confirmed smoking cessation rates were 22% (12/55) at the end of treatment and 19% (12/62) at 1-month follow-up, with no differences between conditions. Living below the poverty level predicted worse cessation outcomes at 1-month follow-up among participants receiving in-person only treatment (P=.03) but not among those receiving iQuit Mindfully.

Conclusions: Text messaging appears to be a feasible and acceptable modality for supporting mindfulness-based smoking cessation treatment. The availability of 24/7 text messaging might be particularly helpful for low-income smokers who have access to fewer cessation resources and experience significant day-to-day barriers to quitting.

Trial Registration: ClinicalTrials.gov NCT03029819; https://clinicaltrials.gov/ct2/show/NCT03029819

(JMIR Mhealth Uhealth 2019;7(6):e13059) doi:10.2196/13059

KEYWORDS

text messaging; smoking cessation; low-income populations

Introduction

Background

Smoking is the leading cause of premature death in the United States [1]. Only 7% of smokers quit each year despite most indicating an interest in quitting [2], and profound tobacco-related health disparities exist [3]. Smoking prevalence remains disproportionately high among adults with low socioeconomic status (SES) despite substantial decline in the smoking rate in the general US population [3-5]. One-fourth (25.3%) of adults living below the federal poverty line smoke, compared with only 14.3% of those at or above the poverty level [5]. Low-SES smokers and members of certain racial and ethnic minority groups, including African Americans, often have greater difficulty quitting and have higher incidence and mortality rates for tobacco-related cancers [2,3,6,7]. Smokers with poorer financial, structural, and social resources face formidable day-to-day barriers, including societal (eg, low health care access), community (eg, tobacco advertising and neighborhood stress), interpersonal (eg, social norms for smoking and low social support), and intrapersonal factors (eg, high stress and low self-efficacy) [3,6,8], which promote addiction and impede efforts to quit. Improving on evidence-based smoking cessation interventions for low-SES and racial and ethnic minority populations will be critical for targeting tobacco-related health disparities.

Mindfulness-Based Programs for Smoking Cessation

Training in mindfulness (ie, purposeful, nonjudgmental, present-focused attention [9,10]) shows promise for increasing rates of smoking cessation and lapse recovery [11,12]. Mindfulness refers to one's relationship with his or her thoughts and emotions (ie, observing these experiences nonjudgmentally, without reacting or trying to change them) rather than to their content. A meta-analysis of randomized controlled trials (RCTs) found that participants receiving mindfulness interventions were almost twice as likely to achieve smoking abstinence for more than 4 months compared with those receiving usual care (25.2% vs 13.6%) [13]. There is a dearth of research on mindfulness in low-SES and racial and ethnic minority groups, but mindfulness does show promise for smoking cessation in these populations [14,15]. Programs that teach nonjudgmental, self-compassionate awareness could be particularly useful for racial and ethnic minority populations [16] and have been perceived as empowering among low-SES and racial minority adults [17]. However, additional between-session support may be needed for low-SES and racial and ethnic minority smokers, who experience significant day-to-day barriers to quitting and have lower access to smoking cessation resources.

Mobile Health and Smoking Cessation

Mobile health (mHealth) interventions have promise for encouraging skills on a real-time, real-life basis, thus increasing skill level, self-efficacy, and the likelihood that the skill will become a part of daily routines [18]. mHealth messages could encourage participants to use mindfulness and other smoking cessation strategies in moments of high stress or craving. This type of in-the-moment support could be especially beneficial for populations (eg, low-SES smokers) with fewer cessation resources. Recent research supports the promise of technology (eg, Web-based training and mobile apps) for teaching mindfulness [19,20], including for smoking cessation [21-23]. Although mindfulness apps have been proliferating, most mHealth mindfulness programs have not been rigorously tested [24].

There is strong empirical support for text messaging programs for smoking cessation [25-30], although none to our knowledge has focused on mindfulness. In a systematic review, mobile phone interventions (most using text messaging) increased smoking abstinence at 6 months (risk ratio [RR]=1.67), with even more positive findings for biochemically verified abstinence (RR=1.83) [26]. Text messaging does not require a smartphone, internet access, or high technical literacy, thus meeting the needs of many adults with lower SES. For example, the vast majority (91%) of college graduates own a smartphone compared with only 57% of adults with less than high school education [31]. However, 90% of Americans with less than high school education own a mobile phone [31]. Furthermore, low-SES and certain racial and ethnic minority adults use text messaging particularly often. In a Pew Research Center study, mean number of texts sent/received per day for Caucasians, African Americans, and Latinos were 31.2 (median 10), 70.1 (median 20), and 48.9 (median 20), respectively. Whereas mean texts per day among adults with college education or greater was 23.8 (median 10), those with less than high school education sent/received 69.4 texts per day (median 20) [32]. Text messaging can provide strategies and encouragement in the context of everyday life and in real time (eg, in moments of high stress or craving), and the content of messages can be personalized. As an adjunct to in-person mindfulness treatment, between-session text messaging could increase treatment engagement and provide vital 24/7 support for smokers from disadvantaged backgrounds.

iQuit Mindfully

Recognizing that most mHealth programs have been developed without adequate feedback from the target population [33], we took a user-centered design approach [34,35] to develop a text messaging smoking cessation program for predominantly low-SES, racially/ethnically diverse smokers. As described in detail elsewhere [36], we conducted 2 phases of formative research (initial focus groups before developing text message content, and then an abbreviated 1-week trial of text messages) to gather qualitative data to inform and improve the text messaging program. User feedback was elicited throughout the process of developing and refining the messages. The text messages were designed to be sent between weekly in-person mindfulness treatment sessions for smoking cessation.

This study is a pilot investigation of mindfulness-based smoking cessation that incorporates this between-session text messaging (*iQuit Mindfully*). To our knowledge, this is the first study to

use text messaging to enhance mindfulness-based smoking cessation treatment. Participants were randomly assigned to 1 of 2 groups: Mindfulness-Based Addiction Treatment (MBAT) or iQuit Mindfully (MBAT with the addition of between-session text messages). Feasibility and acceptability outcomes critical to this pilot study were attrition, participant engagement with text messages, and participant ratings and feedback regarding the text messaging program. The primary smoking cessation outcome was 7 consecutive days of abstinence from smoking at the end of treatment. In addition, secondary analyses examined associations among engagement with text messaging, mindfulness practice, and smoking cessation, as well as cessation outcomes by poverty status.

Methods

Participants

Recruitment targeted a racially/ethnically diverse sample of smokers with relatively low income levels in the Atlanta, GA, area. Inclusion criteria were age 18 to 65 years; current smoker with history of ≥5 cigarettes per day for the past year (and expired carbon monoxide [CO] ≥6 ppm); motivated to quit within the next 30 days; valid home address in the greater Atlanta area; functioning telephone number; owning a mobile phone with text messaging capacity; ability to speak, read, and write in English; and marginal/adequate health literacy (at least a sixth grade level) as determined by the Rapid Estimate of Adult Literacy in Medicine [37]. Exclusion criteria were contraindication for the nicotine patch; past 30-day use of recreational drugs, alcohol-related problems (positive response on 2 or more of the 5 Patient Health Questionnaire [PHQ] Alcohol Abuse/Dependence Scale items [38]), self-reported current diagnosis of schizophrenia or bipolar disorder or use of antipsychotic medications, score of ≥3 on the PHQ-2 [39] depression screening instrument, regular use of tobacco products other than cigarettes (electronic cigarette users were not excluded), current use of tobacco cessation medications, pregnancy or lactation, or another household member enrolled in the study. The study was approved by the university's institutional review board, and written informed consent was obtained from all participants.

Procedures

Participants were recruited through flyers (posted at venues including the university's downtown campus, local hospitals/community health centers, and near bus and train stops), Web-based sources (eg, Craigslist, listservs), and word of mouth. After screening and informed consent procedures, participants were randomly assigned to iQuit Mindfully (MBAT with text messaging) or MBAT (without text messaging). Randomization took place at the end of the baseline session, after baseline assessments had been administered. Permuted block randomization, with stratification based on age (ages 18-49 vs 50-65 years), was used to allocate participants to treatment condition. Co-author DH used SAS software (SAS Institute Inc) to generate the random number sequence. A graduate research assistant (unaware of the size of the blocks) allocated interventions through opaque sealed envelopes marked according to the allocation schedule. The majority of study personnel were masked to treatment condition. Limited staff were unmasked to handle randomization codes (ie, the graduate research assistant) and delivery of interventions (ie, the study therapist). Participants were recruited between January and June 2017; interventions were delivered between February and September 2017; and follow-up assessments were conducted between May and October 2017.

Study Interventions

All participants received in-person group treatment based on the 8-week MBAT protocol [12], in addition to the 6 weeks of nicotine patch therapy and self-help materials based on the Treating Tobacco Use and Dependence Clinical Practice Guideline [40]. Patch therapy (beginning on the quit day) for participants who smoked more than 10 cigarettes per day consisted of 4 weeks of 21 mg patches, 1 week of 14 mg patches, and 1 week of 7 mg patches. Patch therapy for those who smoked 5 to 10 cigarettes per day consisted of 4 weeks of 14 mg patches and 2 weeks of 7 mg patches.

MBAT was provided by a master's level licensed professional counselor with formal training in mindfulness and addictions. MBAT closely follows Mindfulness-Based Cognitive Therapy [41] procedures but replaces the depression-related material with nicotine dependence–related material. The program consists of 8 weekly 2-hour sessions. Aims are to help participants increase moment-to-moment awareness of thoughts, feelings, and sensations; observe these sensations nonjudgmentally; and learn to disengage their attention and choose more skillful responses (rather than automatic reactions) to uncomfortable sensations (including cravings) and high-risk situations. To provide additional support on the quit date and encourage further mindfulness practice, session 5 (quit date) was an extended 4-hour session. MBAT emphasizes daily practice in several forms: formal sitting meditation, body scan meditation, walking meditation, eating meditation, and gentle yoga. There were between 8 and 15 participants in the MBAT group sessions.

Participants in the iQuit Mindfully condition also received approximately 2 to 6 text messages per day on each day between treatment sessions. Texts were sent using the Mobile Commons/Upland Mobile Messaging platform. The content and frequency of messages were revised based on focus groups and pilot testing with low-income, racially/ethnically diverse smokers [36]. Text messages reminded participants to practice mindfulness (eg, reminders for informal practice, such as awareness of the breath throughout the day, and reminders for formal practice, such as the body scan and sitting meditation). Text messages were personalized (eg, reminding participants of their personal reasons to quit and amount of money to be saved as a result of quitting; incorporating first names) and interactive (eg, participants were asked questions such as "Good morning, John! Would you like to try a mindfulness exercise?" and "There are people, places, and things that make you want to smoke. What are your top 3 triggers to smoke?"), and automated text responses were sent based on their replies. Texts also provided specific strategies to aid in cessation (eg, reminders to get rid of cues to smoke, reach out for social support, and coping strategies taught in MBAT).

Participants received approximately 2 messages per day during week 1, 3 per day during week 2, 4 per day during week 3, 5 per day during week 4, 6 per day during week 5, 4 per day during week 6, and 3 per day during week 7. This message schedule was based on our earlier qualitative work [36]. Participants could also text specified words (CRAVE, STRESS, or SLIP) at any point to receive additional text message support for coping with cravings, stress, or smoking lapses, respectively. Participants received a relatively small number of texts (1-3 per week) during the 1-month follow-up period and had the opportunity to text the CRAVE/STRESS/SLIP keywords during this time.

Measures

Demographics and Baseline Smoking Behavior

At baseline, participants indicated their gender, age, education, income, and employment status. Poverty status (below vs at or above the federal poverty line) was calculated according to US Census Bureau Guidelines based on family size and number of children [42]. The Heaviness of Smoking Index, a strong indicator of nicotine dependence [43], assessed self-reported average number of cigarettes smoked per day and the time to first cigarette on waking.

Participants' Ratings and Feedback Regarding iQuit Mindfully

Participants in the iQuit Mindfully condition completed program evaluation forms in person at the end of treatment. Participants were asked, "Of all of the text messages that you received as part of this program, how many did you read?" (response options: *none*, *some*, *most*, or *all*). They were also asked, "Overall, how helpful were the text messages in getting you to try to quit smoking?" (rated from 1=not at all helpful to 10=extremely helpful), and then specifically asked about the helpfulness of the CRAVE, STRESS, and SLIP keywords using the same 10-point scales. They were prompted to "Please circle the number that best represents whether you would recommend that other people receive the text messages that you received in this program (or similar texts) as a way to help them quit smoking" (rated from 1=would not recommend to 10=would definitely recommend). It was determined a priori that scores of 6 or more on these 10-point scales would indicate that the texts were acceptable. Participants were also asked, "Please rate the overall *number* of text messages that you received as part of this program": (response options: "not enough texts," "prefer more texts," "about the right number of texts," "prefer fewer texts," and "way too many texts").

Finally, participants were asked the following open-ended questions:

- What did you like the most about the text messages?
- How, if at all, did you find the text messages to be helpful? (these 2 items were combined for analysis given considerable overlap)
- What, if anything, did you dislike about the text messages?
- What recommendations do you have to improve the text messages?

Smoking Abstinence

Smoking abstinence at the end of treatment (3 weeks after quit date) and 1-month follow-up (7 weeks after quit date) was defined as self-reported complete abstinence for 7 days that was biochemically confirmed in person with CO <6 ppm. Participants who denied smoking for the past 7 days but had CO levels ≥6 ppm were coded as not abstinent (n=4 at the end of treatment and n=2 at follow-up). Missing data were not coded as smoking because of the potential for severe bias that has been demonstrated in prior studies [44,45].

Weekly Mindfulness Practice

Each week from treatment session 2 to session 8, participants completed a weekly mindfulness practice log [12] to indicate the number of days that they practiced each of 5 core mindfulness techniques taught in treatment (sitting meditation, body scan, walking meditation, yoga/mindful stretching, and mindful awareness of breath during the day) during the past week. Responses were averaged for each mindfulness practice over the course of treatment.

Data Analysis

Descriptive statistics provided information on feasibility and acceptability (ie, attrition, engagement, and participants' ratings). Open-ended responses to program evaluations were coded using QSR International's NVivo 11 software. The first author (CAS) and coauthor CCD each reviewed the responses to develop an initial set of themes and then collaborated to define specific codes and refine the coding manual. CAS and CCD then each separately coded all the responses, with an overall kappa of .95, indicating high interrater reliability. Discrepancies were resolved through discussion, with final decisions made by the first author. Chi-square tests examined group differences in smoking cessation outcomes at the end of treatment and 1-month follow-up. Independent samples *t* tests examined group differences in mindfulness practice over the course of treatment.

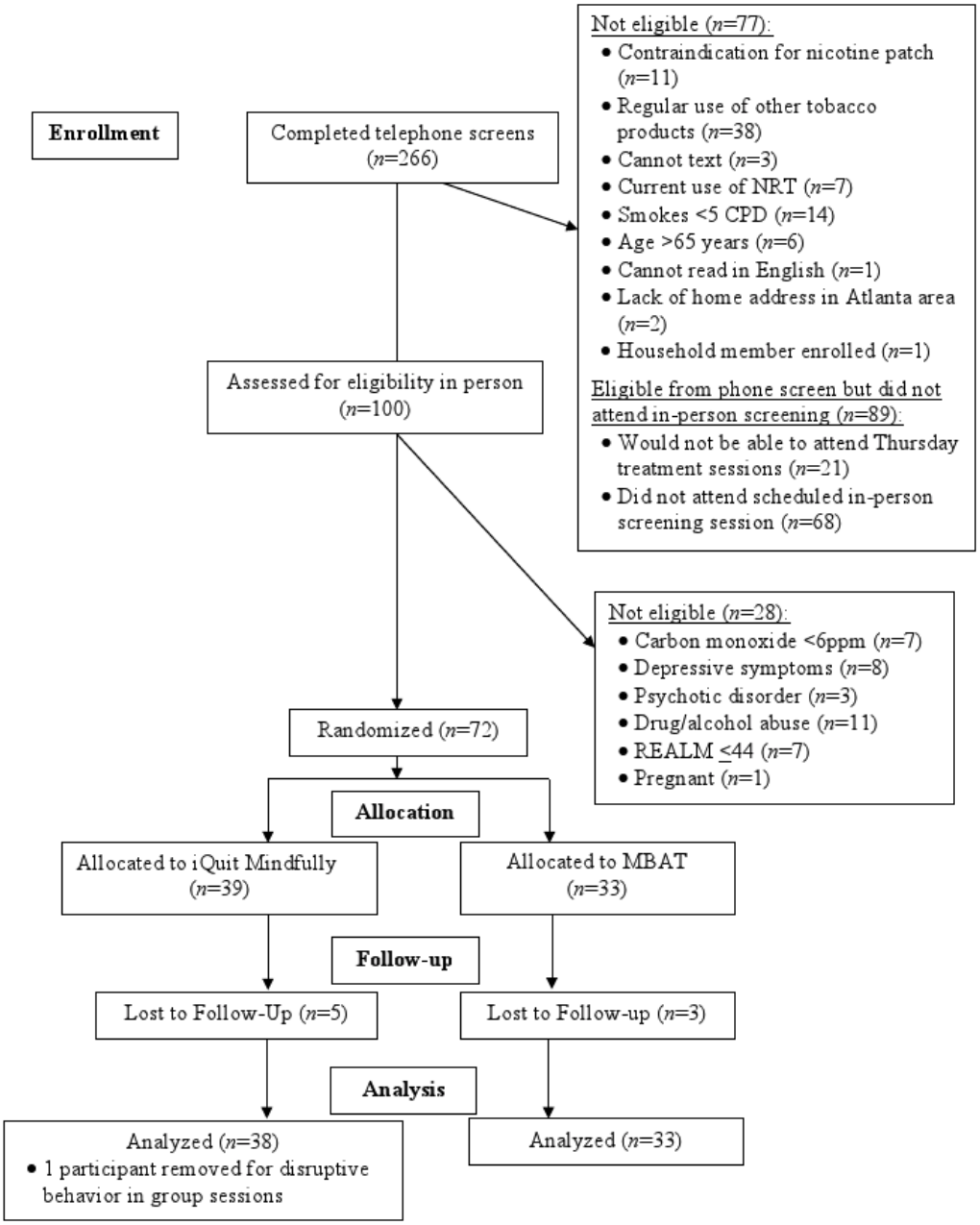
In addition, 2 ancillary analyses were conducted. First, to examine whether participants who were more engaged with the text messages and/or practiced mindfulness more frequently between sessions had better outcomes, associations among text message engagement (based on the number of times participants texted the system), weekly mindfulness practice variables, and abstinence were examined using chi-square, *t* tests, and logistic regression. Second, because the text messages were specifically designed to target low-SES smokers, analyses examined results separately by poverty status (below vs at or above the federal poverty line).

Results

Screening and Enrollment

A total of 266 individuals completed telephone screening, 100 completed in-person screening, and 72 participants were enrolled in the study (Figure 1). Overall, 1 participant in the iQuit Mindfully condition was removed because of disruptive behavior in the in-person group treatment for a final analytic sample of 71.

Figure 1. Consolidated Standards of Reporting Trials flow diagram. CPD: cigarettes per day; MBAT: Mindfulness-Based Addiction Treatment; NRT: nicotine replacement therapy; REALM: Rapid Estimate of Adult Literacy in Medicine.



Participants' Characteristics

Demographic and smoking-related characteristics at baseline are shown in Table 1. Mean age of the participants was 45.6 years (SD 12.1), and about half (37/71, 52%) were female. Most participants were African American (50/71, 70%), 15 (21%) were Caucasian, and 4 (6%) reported more than 1 race. Most (43/71, 61%) indicated having an annual household income of US $30,000 or less, and 41% (27/66) were living below the federal poverty line. On average, participants smoked 16.5 (SD 9.6) cigarettes per day at baseline and had been smoking daily for 23.6 (SD 14.1) years. The majority (57/71, 80%) reported smoking their first cigarette within 30 min of waking, and most (57/71, 80%) smoked primarily menthol cigarettes. Although statistical tests were not conducted to examine baseline differences between treatment groups [46,47], examination of the descriptive statistics suggests that participants in the iQuit Mindfully condition were more likely to be African American and have lower SES than those in MBAT.

Table 1. Participants' characteristics.

Demographic characteristics	Full sample (N=71)	iQuit Mindfully (n=38)	Mindfulness-Based Addiction Treatment (n=33)
Age (years), mean (SD)	45.6 (12.1)	45.6 (12.4)	45.6 (12.0)
Gender, female, n (%)	37 (52)	17 (45)	20 (61)
Race/ethnicity, n (%)			
Black/African American	50 (70)	34 (89)	16 (49)
Caucasian	15 (21)	4 (11)	11 (33)
Asian	1 (1)	0 (0)	1 (3)
More than 1 race	4 (6)	0 (0)	4 (12)
Other	1 (1)	0 (0)	1 (3)
Hispanic/Latino	3 (4)	1 (3)	2 (6)
Employment, n (%)			
Regular full-time work (40+ hours/week)	17 (24)	6 (16)	11 (33)
Regular part-time work	13 (18)	3 (8)	10 (30)
Temporary part-time work	1 (1)	1 (3)	0 (0)
Self-employed	1 (1)	1 (3)	0 (0)
Student	4 (6)	2 (5)	2 (6)
Unemployed	19 (27)	15 (39)	4 (12)
Retired	10 (14)	5 (13)	5 (15)
Unable to work or disabled	6 (9)	5 (13)	1 (3)
Education, n (%)			
Less than high school degree	12 (17)	7 (18)	5 (15)
High school degree or General Education Development	14 (20)	7 (18)	7 (21)
Some college/technical school	19 (27)	12 (32)	7 (21)
Associates degree	10 (14)	4 (11)	6 (18)
Bachelor's degree	10 (14)	5 (13)	5 (15)
Some postbac school	3 (4)	3 (8)	0 (0)
Graduate degree	3 (4)	0 (0)	3 (9)
Annual household income in US dollars (n=66), n (%)			
≤$12,000	20 (30)	12 (36)	8 (24)
$12,001-$18,000	10 (15)	7 (21)	3 (9)
$18,001-$30,000	13 (20)	5 (15)	8 (24)
$30,001-$42,000	7 (11)	2 (6)	5 (15)
$42,001-$54,000	2 (3)	0 (0)	2 (6)
$60,001-$84,000	4 (6)	0 (0)	4 (12)
>$84,000	10 (15)	7 (21)	3 (9)
Poverty status (n=66), n (%)			
Below poverty threshold	27 (41)	16 (48)	11 (33)
At or above poverty threshold	39 (59)	17 (52)	22 (67)
Cigarettes per day, mean (SD)	16.5 (9.6)	14.4 (9.4)	18.8 (9.3)
Years smoking daily, mean (SD)	23.6 (14.1)	20.7 (13.1)	27.0 (14.6)
Time to first cigarette, n (%)			
Within 5 min	27 (38)	17 (45)	10 (30)

Demographic characteristics	Full sample (N=71)	iQuit Mindfully (n=38)	Mindfulness-Based Addiction Treatment (n=33)
6-30 min	30 (42)	15 (39)	15 (45)
31-60 min	7 (10)	3 (8)	4 (12)
After 60 min	7 (10)	3 (8)	4 (12)
Menthol cigarettes as regular brand, n (%)	57 (80)	33 (87)	24 (73)

Treatment Attendance

Participants attended an average of 5.7 of 8 treatment sessions (SD 2.7), with no difference between conditions (iQuit Mindfully: 5.9 sessions, SD 2.7; MBAT: 5.4 sessions, SD 2.8), *P*=.50. Greater treatment attendance predicted higher likelihood of smoking cessation at the end of treatment (β=4.05, *P*=.049) and approached significance in predicting cessation at 1-month follow-up (β=1.76, *P*=.058).

Feasibility and Acceptability

Attrition

A priori, it was deemed that 35% attrition would be acceptable (based on 34% attrition in trial of MBAT by Vidrine et al [12]). Overall attrition rates were 24% (17/71) at the end-of-treatment assessment and 11% (8/71) at 1-month follow-up, with no differences between groups (attrition at the end of treatment: iQuit Mindfully 21% [8/38], MBAT 27% [9/33], *P*=.38; attrition at 1-month follow-up: iQuit Mindfully 11% [4/38], MBAT 12% [4/33], *P*=.83).

iQuit Mindfully Engagement

Level of engagement was determined based on (1) the proportion of texts that participants indicated reading (we expected that at least 75% would read most or all texts, based on the study by Abroms et al [48]) and (2) responses to interactive text messages (we expected that at least 85% of participants would respond to at least one of the interactive text messages or use the CRAVE, STRESS, or SLIP keywords at least once, based on the study by Heminger et al [49]). These benchmarks were achieved. The majority (88%, 29/33) indicated reading all or most text messages, and 89% (34/38) responded to at least one of the interactive text messages or used the CRAVE, STRESS, or SLIP keywords.

Participants' Ratings and Feedback

On a scale of 1 to 10, participants' average rating of text helpfulness was 8.0 (SD 2.4). Ratings regarding the helpfulness of the keywords CRAVE, STRESS, and SLIP were 7.8 (SD 0.29), 7.9 (SD 2.7), and 7.8 (SD 2.8), respectively. On a scale of 1 to 10, participants' average rating of the extent to which they would recommend the text messaging program to others was 8.4 (SD 2.5). Most (58%, 19/33) indicated receiving about the right number of texts, 30% (10/33) preferred fewer, and 12% (4/33) preferred more.

Table 2 shows themes and illustrative quotations from participants' open-ended responses on iQuit Mindfully program evaluations. Overall, participants reported positive experiences with the text messages (eg, "I loved them; sometimes I read them going to bed instead of smoking"). Almost all (97%, 32/33) provided positive responses when asked what they liked and what was most helpful about the texts. Themes were appreciating the positive tone (n=13), receiving reminders (n=11), benefiting from mindfulness (n=9), perceiving a sense of social support (n=9), perceiving that the messages had good timing (n=6), noting that the messages encouraged self-compassion in the face of smoking lapses (n=6), and receiving specific strategies to cope with cravings and stress (n=5). When asked what they disliked, 61% (20/33) responded *nothing*, *N/A*, or left the question blank. Whereas 9 participants indicated that there were too many text messages or that they were repetitive, 7 indicated wanting more text messages. When asked their suggestions for improvements, 67% (22/33) responded *none*, *N/A*, or left the question blank. Of those who provided feedback, suggestions included connecting participants to additional outside resources (n=5; eg, connect to phone call, emergency resources, one-on-one support), incorporating more personalization (n=3), and including more religion/spirituality in the messages (n=3).

Smoking Abstinence

There were no significant differences between groups either at the end of treatment (iQuit Mindfully: 26% [8/31]; MBAT: 17% [4/24], *P*=.42) or follow-up (iQuit Mindfully: 16% [5/32]; MBAT: 23% [7/30], *P*=.44).

Weekly Mindfulness Practice

Mean number of days engaging in each type of mindfulness practice per week ranged from 2.02 (SD 1.72) for yoga to 3.25 (SD 2.11) for mindful awareness of breathing. There were no significant between-group differences, *P*>.59.

Associations Among Text Engagement, Weekly Mindfulness Practice, and Abstinence

Among iQuit Mindfully participants, the mean number of times participants texted the system was 55.2 (SD 63.1; median 37) and was highly skewed (7 participants texted over 100 times, with 2 of these texting over 200 times). Thus, a dichotomized text engagement variable was created based on the median. Participants categorized as having high engagement had significantly higher abstinence rates at the end of treatment than those with low engagement, χ^2_1=7.8, *P*=.005. Whereas 44% (8/18) of participants with high engagement were abstinent at the end of treatment, none of those with low engagement was abstinent. Engagement was not significantly associated with abstinence at 1-month follow-up (7% [1/14] of those with low engagement vs 22% [4/18] of those with high engagement were abstinent, *P*=.24).

Table 2. Example quotations from open-ended iQuit Mindfully program evaluation responses.

Themes	Example quotations
Most helpful aspects	
Positive tone	• "Positive response was really cool for my confidence" • "Positive and uplifting"
Reminders	• "They became an integral part of your day and served as gentle reminders and encouragement" • "They reminded me of my goals and told me why I was choosing to quit smoking"
Mindfulness	• "Helped me to stay mindful" • "Kept me aware" • "Stop breathe think"
Social support	• "I felt that someone cared how I was feeling" • "I was able to reach out for support and it was very helpful" • "It let me know somebody out there to help me"
Good timing	• "Sometimes they came right on time. I would start thinking about smoking and here comes that text." • "Every time I thought about smoking I get that text of encouragement to not smoke."
Self-compassion in the context of smoking lapses	• "They encouraged me to continue with my journey and don't worry about the slip up and just start over." • "Made you not beat yourself up about a slip"
Strategies for coping with cravings and stress	• "If you text CRAVE and actually do what the text message says you will successfully overcome that current craving" • "You were given techniques to help overcome the stress"
Dislikes	
Too many text messages/repetitive	• "Came a little too quick sometimes" • "Repetitive" • "Less texts would be better"
Not enough text messages	• "[I disliked] when they became less frequent" • "Even more would be helpful"
Suggestions	
Connect to outside resources	• "One on one support. Additional resources" • "Maybe a call"
More personalization	• "Really try to find out what best suits each individual" • "More intuitive and spontaneous and less generically programmed"
Religion/spirituality	• "Send Bible verses/scriptures" • "More spiritual texts"

Associations between engagement with text messages and mindfulness practice were also examined. Participants who showed high text engagement reported practicing informal mindfulness more frequently than those with low engagement (high engagement: mean 3.9 [SD 2.0] days vs low engagement: mean 2.0 [SD 1.6] days), t_{32}=2.91, P=.006. The association between text engagement and frequency of sitting meditation practice approached significance (high engagement: mean 3.2 [SD 1.7] days vs low engagement: mean 2.2 [SD 1.5] days), t_{32}=1.85, P=.07. There were no differences in frequency of other mindfulness practices by the level of text engagement. When mindfulness practice variables were entered simultaneously into a logistic regression analysis predicting abstinence outcomes, mindful awareness of the breath uniquely predicted greater likelihood of abstinence at the end of treatment (β=1.60, P=.04) and 1-month follow-up (β=1.99, P=.008).

Associations Between Poverty Status and Abstinence Outcomes by Condition

Among participants living in poverty, 23% (3/13) of those in iQuit Mindfully were abstinent at the end of treatment and 1-month follow-up, whereas none of the participants in the control group receiving MBAT quit smoking at either time point. Fisher's exact tests examined associations between poverty status and abstinence separately by condition. Poverty status was not significantly associated with abstinence at the

end of treatment (MBAT: *P*=.26; iQuit Mindfully: *P*=.67). At 1-month follow-up, living below the poverty level was associated with worse cessation outcomes among MBAT participants (*P*=.03) but not among those receiving iQuit Mindfully (*P*=.65).

Discussion

Principal Findings

This pilot study examined the feasibility of a mindfulness-based smoking cessation program incorporating between-session text messaging (*iQuit Mindfully*). To our knowledge, this is the first study to use text messaging to enhance mindfulness-based smoking cessation treatment, and preliminary results support the feasibility and acceptability of text messaging for providing day-to-day smoking cessation support to low-SES, racially/ethnically diverse adults. Strong retention was achieved (76% [54/71] at the end of treatment, and 89% [63/71] at 1-month follow-up); engagement in iQuit Mindfully was high (88% [29/33] indicated reading all or most text messages, and 89% [34/38] texted the system); and participants provided positive ratings and feedback about the text messages. Between-session text messaging could be particularly beneficial for promoting smoking cessation among low-SES adults. Participants provided suggestions for further improving the text messaging program, and the results of this study warrant additional investigation in a larger RCT.

The overall biochemically confirmed smoking cessation rates were 22% (12/55) at the end of treatment and 19% (12/62) at 1-month follow-up, with no differences between conditions. Living below the poverty level predicted worse cessation rates at 1-month follow-up in participants receiving in-person MBAT only, but not among those receiving iQuit Mindfully text messages. Smokers living in poverty not only have lower health care access but also are continually confronted with more tobacco advertising, higher social norms for smoking, and lower social support for quitting and experience higher stress and lower self-efficacy for quitting [3,6,8]. The availability of 24/7 text messaging support could be vital for helping low-SES smokers to overcome these chronic, day-to-day barriers. This study is limited by small sample size, and further investigation in a larger, appropriately powered trial is needed. Extant studies do support the use of mobile phone–delivered interventions for smoking cessation specifically among low-income smokers [50].

Comparison With Prior Work

Participants who were more engaged with the iQuit Mindfully text messaging program practiced informal mindfulness more frequently and were more likely to quit smoking at the end of treatment. This is consistent with past research, suggesting that higher engagement with mHealth programs predicts better smoking cessation outcomes [51-53]. However, low user engagement is a pervasive problem with mHealth programs [24,54], and efforts are needed to increase engagement with the ultimate goal of improving outcomes. One strategy suggested by our participants is to further personalize text messages, and extant research suggests that tailoring interventions to users' needs and preferences can indeed increase engagement and efficacy [55-57]. We also examined associations between treatment condition and in-person session attendance. There were no differences in attendance between MBAT and iQuit Mindfully participants, but those who attended more sessions were more likely to quit smoking. Future research might consider how technology could increase in-person session attendance as well as engagement during these sessions.

Greater informal mindfulness practice (ie, mindful attention to breathing throughout the day) predicted higher likelihood of smoking abstinence at both end of treatment and 1-month follow-up. Although more frequent personal mindfulness practice is hypothesized to confer psychosocial benefits, the literature on associations between mindfulness practice and clinical outcomes has been somewhat mixed [58-60]. Past research has shown positive associations between mindfulness practice and better smoking cessation outcomes [11]. The findings of this study suggest that apart from formal meditation practice, mindful attention to breathing in the context of daily activities is uniquely associated with better smoking cessation outcomes. This informal practice (eg, taught through the STOP [*Stop*, *Take a breath*, *Observe*, *Proceed*] acronym in MBAT and other mindfulness programs) could be especially useful for coping with cravings and other stressors during the cessation process.

Overall, iQuit Mindfully participants noted positive experiences with the text messages. Themes included appreciating the positive tone, reminders, mindfulness techniques, social support, timing of the messages, self-compassion in the face of smoking lapses, and coping strategies. This is consistent with our past qualitative work developing iQuit Mindfully [36] as well as other qualitative studies of adults' experiences with text messaging for smoking cessation [61,62]. A common theme across studies is that although participants understand that the texts are automated, they often describe a sense of social support (eg, "It's like having a friend who texts you when you are feeling stressed or having a feeling like you want to smoke" [36]). Our participants suggested that the program be even more personalized, and future iterations of the program might provide more flexibility and personalization in terms of frequency, timing, and content of text messages (eg, varying the number and timing of texts based on individual preferences and triggers). As suggested by participants, text messages could also include more religious/spiritual content (this could also be personalized based on individual preferences) and connect them to outside resources (eg, direct connection to quitlines or other support as needed). In addition, the research team noted some logistical issues with participants using their own mobile phones during the study (eg, service interruptions and changing phone numbers), and future studies might offer participants mobile phones with wireless plans to use for the duration of the study. It is possible that these issues may become less common over time, as mobile phone access continues to increase in low-SES populations [63,64].

Limitations

This pilot study is limited by a small sample size without statistical power to detect group differences in smoking cessation or other outcomes. Although modeling-based approaches can

help to address pitfalls of as-treated analyses [65], our sample was too small to fit such models, and results should be viewed as preliminary evidence of feasibility that will need to be tested in larger trials. In addition, iQuit Mindfully text messages were designed to supplement (rather than replace) in-person MBAT sessions, and thus, this program also involves the substantial time and resources associated with in-person treatment. Our decision to include both in-person treatment and text messaging was based on our formative work with low-SES smokers, who noted that text messaging alone (in the absence of other resources such as in-person treatment) would not be sufficient [36]. However, future iterations might consider fully implementing the program through mHealth to increase scalability and reduce costs. In addition, results may or may not generalize to those not included based on the eligibility criteria (eg, people using multiple forms of tobacco, those with psychotic disorders or drug/alcohol abuse). Finally, frequency of between-session mindfulness practice was relatively low in both MBAT and iQuit Mindfully conditions. This is consistent with other research finding that participants often do not practice mindfulness as frequently as directed [12]. Research is needed to examine strategies to promote mindfulness practice among smokers, and as discussed above, more tailored messaging could be one such strategy. Despite limitations, this pilot study is strengthened by the use of biochemical confirmation of smoking behavior; recruitment of predominantly low-SES, racial minority adults; and RCT design, all of which support feasibility for conducting an efficacy RCT.

Conclusions

Overall, this proof-of-concept study provides strong evidence for feasibility and acceptability of iQuit Mindfully text messages to enhance MBAT by providing between-session support. Offering tailored 24/7 text messaging support could be helpful for low-SES smokers, who have lower access to cessation support and face formidable day-to-day barriers to quitting. Preliminary findings warrant further investigation in an appropriately powered RCT to determine efficacy.

Acknowledgments

This work was supported by the National Center for Complementary and Integrative Health (K23AT008442). DWW is supported by the Huntsman Cancer Foundation, the National Cancer Institute (P30CA042014), and the National Center for Advancing Translational Sciences (UL1TR001067). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Mark Dannenfelser for leading MBAT groups and Brittani Carter, Sharrill Bell, Charlayne Scarlett, Maitreyi Bandlamudi, Amanda Grant, and Jasmine Guo for their help with data collection. They also thank Drs Sarah Bowen, Bradley N Collins, and Danielle E McCarthy for serving on the independent monitoring committee for this study.

Conflicts of Interest

LCA has stock in Welltok Inc and receives royalties from the licensing of Text2Quit to Welltok, Inc. MPE receives research funding support from Pfizer, Inc (“Diffusion of Tobacco Control Fundamentals to Other Large Chinese Cities,” MPE, Principal Investigator). The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i6e13059_app1.pdf]

References

1. The Health Consequences of Smoking-50 Years of Progress: A Report of the Surgeon General. Washington, DC: US Department of Health and Human Services; 2014.
2. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting smoking among adults - United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017 Jan 6;65(52):1457-1464 [FREE Full text] [doi: 10.15585/mmwr.mm6552a1] [Medline: 28056007]
3. US National Cancer Institute. Division of Cancer Control and Population Sciences. 2017. Monograph 22. A sociological approach to addressing tobacco-related health disparities URL: https://cancercontrol.cancer.gov/brp/tcrb/monographs/22/
4. Drope J, Liber AC, Cahn Z, Stoklosa M, Kennedy R, Douglas CE, et al. Who's still smoking? Disparities in adult cigarette smoking prevalence in the United States. CA Cancer J Clin 2018;68(2):106-115 [FREE Full text] [doi: 10.3322/caac.21444] [Medline: 29384589]
5. Jamal A, Phillips E, Gentzke AS, Homa DM, Babb SD, King BA, et al. Current cigarette smoking among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Jan 19;67(2):53-59 [FREE Full text] [doi: 10.15585/mmwr.mm6702a1] [Medline: 29346338]
6. Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. Ann N Y Acad Sci 2012;1248:107-123. [doi: 10.1111/j.1749-6632.2011.06202.x] [Medline: 22092035]

7. Centers for Disease Control and Prevention. 2013. CDC Health Disparities and Inequalities Report - United States, 2013 URL: https://www.cdc.gov/mmwr/pdf/other/su6203.pdf
8. Businelle MS, Kendzor DE, Reitzel LR, Costello TJ, Cofta-Woerpel L, Li Y, et al. Mechanisms linking socioeconomic status to smoking cessation: a structural equation modeling approach. Health Psychol 2010;29(3):262-273 [FREE Full text] [doi: 10.1037/a0019285] [Medline: 20496980]
9. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
10. Kabat-Zinn J. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. New York: Delacorte Press; 1990.
11. Brewer JA, Mallik S, Babuscio TA, Nich C, Johnson HE, Deleone CM, et al. Mindfulness training for smoking cessation: results from a randomized controlled trial. Drug Alcohol Depend 2011;119(1-2):72-80 [FREE Full text] [doi: 10.1016/j.drugalcdep.2011.05.027] [Medline: 21723049]
12. Vidrine JI, Spears CA, Heppner WL, Reitzel LR, Marcus MT, Cinciripini PM, et al. Efficacy of mindfulness-based addiction treatment (MBAT) for smoking cessation and lapse recovery: a randomized clinical trial. J Consult Clin Psychol 2016;84(9):824-838 [FREE Full text] [doi: 10.1037/ccp0000117] [Medline: 27213492]
13. Oikonomou MT, Arvanitis M, Sokolove RL. Mindfulness training for smoking cessation: a meta-analysis of randomized-controlled trials. J Health Psychol 2017;22(14):1841-1850. [doi: 10.1177/1359105316637667] [Medline: 27044630]
14. Davis JM, Goldberg SB, Anderson MC, Manley AR, Smith SS, Baker TB. Randomized trial on mindfulness training for smokers targeted to a disadvantaged population. Subst Use Misuse 2014 Apr;49(5):571-585 [FREE Full text] [doi: 10.3109/10826084.2013.770025] [Medline: 24611852]
15. Davis JM, Manley AR, Goldberg SB, Smith SS, Jorenby DE. Randomized trial comparing mindfulness training for smokers to a matched control. J Subst Abuse Treat 2014;47(3):213-221 [FREE Full text] [doi: 10.1016/j.jsat.2014.04.005] [Medline: 24957302]
16. Witkiewitz K, Greenfield BL, Bowen S. Mindfulness-based relapse prevention with racial and ethnic minority women. Addict Behav 2013 Dec;38(12):2821-2824 [FREE Full text] [doi: 10.1016/j.addbeh.2013.08.018] [Medline: 24018224]
17. Spears CA, Houchins SC, Bamatter WP, Barrueco S, Hoover DS, Perskaudas R. Perceptions of mindfulness in a low-income, primarily African American treatment-seeking sample. Mindfulness (N Y) 2017 Dec;8(6):1532-1543 [FREE Full text] [Medline: 29333200]
18. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
19. Sliwinski J, Katsikitis M, Jones CM. A review of interactive technologies as support tools for the cultivation of mindfulness. Mindfulness 2017;8(5):1150-1159. [doi: 10.1007/s12671-017-0698-x]
20. Fish J, Brimson J, Lynch S. Mindfulness interventions delivered by technology without facilitator involvement: what research exists and what are the clinical outcomes? Mindfulness (N Y) 2016;7(5):1011-1023 [FREE Full text] [doi: 10.1007/s12671-016-0548-2] [Medline: 27642370]
21. Davis JM, Manley AR, Goldberg SB, Stankevitz KA, Smith SS. Mindfulness training for smokers via web-based video instruction with phone support: a prospective observational study. BMC Complement Altern Med 2015 Mar 29;15:95 [FREE Full text] [doi: 10.1186/s12906-015-0618-3] [Medline: 25886752]
22. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 01;143:87-94 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
23. Garrison KA, Pal P, O'Malley SS, Pittman BP, Gueorguieva R, Rojiani R, et al. Craving to quit: a randomized controlled trial of smartphone app-based mindfulness training for smoking cessation. Nicotine Tob Res 2018 Jun 18 (forthcoming). [doi: 10.1093/ntr/nty126] [Medline: 29917096]
24. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
25. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
26. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
27. Armanasco AA, Miller YD, Fjeldsoe BS, Marshall AL. Preventive health behavior change text message interventions: a meta-analysis. Am J Prev Med 2017 Mar;52(3):391-402. [doi: 10.1016/j.amepre.2016.10.042] [Medline: 28073656]
28. Abroms LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of Text2Quit: a text messaging program for smoking cessation. Am J Prev Med 2014 Sep;47(3):242-250 [FREE Full text] [doi: 10.1016/j.amepre.2014.04.010] [Medline: 24913220]

29. 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. [doi: 10.2196/mhealth.5436] [Medline: 27207211]
30. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]
31. Pew Research Center. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 6zycrGIuw]
32. Smith A. Pew Research Center. 2011. Americans and Text Messaging URL: http://pewinternet.org/Reports/2011/Cell-Phone-Texting-2011.aspx [WebCite Cache ID 6zycTUE04]
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] [Medline: 23039777]
34. Abroms LC, Whittaker R, Free C, Mendel van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
35. Bock BC, Rosen RK, Barnett NP, Thind H, Walaska K, Foster R, et al. Translating behavioral interventions onto mHealth platforms: developing text message interventions for smoking and alcohol. JMIR Mhealth Uhealth 2015 Feb 24;3(1):e22 [FREE Full text] [doi: 10.2196/mhealth.3779] [Medline: 25714907]
36. Spears CA, Bell SA, Scarlett CA, Anderson NK, Cottrell-Daniels C, Lotfalian S, et al. Text messaging to enhance mindfulness-based smoking cessation treatment: program development through qualitative research. JMIR Mhealth Uhealth 2019 Jan 7;7(1):e11246 [FREE Full text] [doi: 10.2196/11246] [Medline: 30617043]
37. Davis TC, Crouch MA, Long SW, Jackson RH, Bates P, George RB, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med 1991 Aug;23(6):433-435. [Medline: 1936717]
38. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
39. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
40. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons, and Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A US public health service report. Am J Prev Med 2008 Aug;35(2):158-176 [FREE Full text] [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
41. Segal ZV, Williams JM, Teasdale JD. Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse. New York: The Guilford Press; 2001.
42. United States Census Bureau. 2018. Poverty Thresholds URL: https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html
43. Kozlowski LT, Porter CQ, Orleans CT, Pope MA, Heatherton T. Predicting smoking cessation with self-reported measures of nicotine dependence: FTQ, FTND, and HSI. Drug Alcohol Depend 1994 Feb;34(3):211-216. [Medline: 8033758]
44. Hedeker D, Mermelstein RJ, Demirtas H. Analysis of binary outcomes with missing data: missing = smoking, last observation carried forward, and a little multiple imputation. Addiction 2007 Oct;102(10):1564-1573. [doi: 10.1111/j.1360-0443.2007.01946.x] [Medline: 17854333]
45. Blankers M, Smit ES, van der Pol P, de Vries H, Hoving C, van Laar M. The missing=smoking assumption: a fallacy in internet-based smoking cessation trials? Nicotine Tob Res 2016 Jan;18(1):25-33. [doi: 10.1093/ntr/ntv055] [Medline: 25744969]
46. Senn S. Testing for baseline balance in clinical trials. Stat Med 1994 Sep 15;13(17):1715-1726. [Medline: 7997705]
47. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2002 Oct 15;21(19):2917-2930. [doi: 10.1002/sim.1296] [Medline: 12325108]
48. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17(Suppl 1):44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]
49. Heminger CL, Schindler-Ruwisch JM, Abroms LC. Smoking cessation support for pregnant women: role of mobile technology. Subst Abuse Rehabil 2016;7:15-26 [FREE Full text] [doi: 10.2147/SAR.S84239] [Medline: 27110146]
50. Daly AT, Deshmukh AA, Vidrine DJ, Prokhorov AV, Frank SG, Tahay PD, et al. Cost-effectiveness analysis of smoking cessation interventions using cell phones in a low-income population. Tob Control 2019;28(1):88-94. [doi: 10.1136/tobaccocontrol-2017-054229] [Medline: 29886411]
51. Heminger CL, Boal AL, Zumer M, Abroms LC. Text2Quit: an analysis of participant engagement in the mobile smoking cessation program. Am J Drug Alcohol Abuse 2016;42(4):450-458. [doi: 10.3109/00952990.2016.1149591] [Medline: 27120396]

52. Zeng EY, Heffner JL, Copeland WK, Mull KE, Bricker JB. Get with the program: adherence to a smartphone app for smoking cessation. Addict Behav 2016 Dec;63:120-124 [FREE Full text] [doi: 10.1016/j.addbeh.2016.07.007] [Medline: 27454354]
53. Bricker JB, Copeland W, Mull KE, Zeng EY, Watson NL, Akioka KJ, et al. Single-arm trial of the second version of an acceptance & commitment therapy smartphone application for smoking cessation. Drug Alcohol Depend 2017 Jan 1;170:37-42 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.10.029] [Medline: 27870987]
54. 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] [Medline: 26537656]
55. Thrul J, Klein AB, Ramo DE. Smoking cessation intervention on Facebook: which content generates the best engagement? J Med Internet Res 2015 Nov 11;17(11):e244 [FREE Full text] [doi: 10.2196/jmir.4575] [Medline: 26561529]
56. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008 Nov 4;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
57. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
58. Lloyd A, White R, Eames C, Crane R. The utility of home-practice in mindfulness-based group interventions: a systematic review. Mindfulness (N Y) 2018;9(3):673-692 [FREE Full text] [doi: 10.1007/s12671-017-0813-z] [Medline: 29875880]
59. Vettese LC, Toneatto T, Stea JN, Nguyen L, Wang JJ. Do mindfulness meditation participants do their homework? And does it make a difference? A review of the empirical evidence. J Cogn Psychother 2009;23(3):198-225 [FREE Full text] [doi: 10.1891/0889-8391.23.3.198]
60. Parsons CE, Crane C, Parsons LJ, Fjorback LO, Kuyken W. Home practice in mindfulness-based cognitive therapy and mindfulness-based stress reduction: a systematic review and meta-analysis of participants' mindfulness practice and its association with outcomes. Behav Res Ther 2017 Aug;95:29-41 [FREE Full text] [doi: 10.1016/j.brat.2017.05.004] [Medline: 28527330]
61. Grau LE, Pham T, O'Leary T, Weiss J, Toll B, Bernstein SL. Smokers' perspectives on texting for tobacco dependence treatment: a qualitative analysis. Nicotine Tob Res 2017 Mar 1;19(3):307-313. [doi: 10.1093/ntr/ntw184] [Medline: 27613935]
62. Douglas N, Free C. 'Someone batting in my corner': experiences of smoking-cessation support via text message. Br J Gen Pract 2013 Nov;63(616):e768-e776 [FREE Full text] [doi: 10.3399/bjgp13X674459] [Medline: 24267860]
63. Smith A. Pew Research Center. 2017. Record Shares of Americans Now Own Smartphones, Have Home Broadband URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [WebCite Cache ID 73lXLweZn]
64. Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. J Med Internet Res 2016 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
65. Little RJ, Yau LH. Statistical techniques for analyzing data from prevention trials: treatment of no-shows using Rubin's causal model. Psychol Methods 1998;3(2):147-159 [FREE Full text]

Abbreviations

CO: carbon monoxide
MBAT: Mindfulness-Based Addiction Treatment
mHealth: mobile health
PHQ: Patient Health Questionnaire
RCT: randomized controlled trial
RR: risk ratio
SES: socioeconomic status

Edited by G Eysenbach; submitted 11.12.18; peer-reviewed by S Comulada, J McClure; comments to author 23.01.19; accepted 01.05.19; published 24.06.19.

Please cite as:
Spears CA, Abroms LC, Glass CR, Hedeker D, Eriksen MP, Cottrell-Daniels C, Tran BQ, Wetter DW
Mindfulness-Based Smoking Cessation Enhanced With Mobile Technology (iQuit Mindfully): Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(6):e13059
URL: http://mhealth.jmir.org/2019/6/e13059/
doi:10.2196/13059
PMID:31237242

Original Paper

Factors Influencing User Engagement of Health Information Disseminated by Chinese Provincial Centers for Disease Control and Prevention on WeChat: Observational Study

Yan Zhang[1*], MS; Tingsong Xia[2*], PhD; Lingfeng Huang[1], MS; Mingjuan Yin[1], MS; Mingwei Sun[1], MS; Jingxiao Huang[1], MS; Yu Ni[3], BS; Jindong Ni[1], PhD

[1]Guangdong Medical University, Dongguan, China

[2]Baoan District Shajing Health Supervision Office, Shenzhen, China

[3]Beijing Jiaotong University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Jindong Ni, PhD
Guangdong Medical University
No. 1 Xincheng Road, Songshan Lake
Dongguan,
China
Phone: 86 769 2289 6570
Fax: 86 769 2289 6570
Email: nijd-gw@gdmu.edu.cn

Abstract

Background: Social media is currently becoming a new channel for information acquisition and exchange. In China, with the growing popularity of WeChat and WeChat official accounts (WOAs), health promotion agencies have an opportunity to use them for successful information distribution and diffusion online.

Objective: We aimed to identify features of articles pushed by WOAs of Chinese provincial Centers for Disease Control and Prevention (CDC) that are associated with user engagement.

Methods: We searched and subscribed to 28 WOAs of provincial CDCs. Data for this study consisted of WeChat articles on these WOAs between January 1, 2017 and December 31, 2017. We developed a features frame containing title type, article content, article type, communication skills, number of marketing elements, and article length for each article and coded the data quantitatively using a coding scheme that assigned numeric values to article features. We examined the descriptive characteristics of articles for every WOA and generated descriptive statistics for six article features. The amount of reading and liking was converted into the level of reading and liking by the 75% position. Two-category univariate logistic regression and multivariable logistic regression were conducted to explore associations between the features of the articles and user engagement, operationalized as reading level and liking level.

Results: All provincial CDC WOAs provided a total of 5976 articles in 2017. Shanghai CDC articles attracted the most user engagement, and Ningxia CDC articles attracted the least. For all articles, the median reading was 551.5 and the median liking was 10. Multivariable logistic regression analysis revealed that article content, article type, communication skills, number of marketing elements, and article length were associated with reading level and liking level. However, title type was only associated with liking level.

Conclusions: How social media can be used to best achieve health information dissemination and public health outcomes is a topic of much discussion and study in the public health community. Given the lack of related studies based on WeChat or official accounts, we conducted this study and found that article content, article type, communication skills, number of marketing elements, article length, and title type were associated with user engagement. Our study may provide public health and community leaders with insight into the diffusion of important health topics of concern.

(JMIR Mhealth Uhealth 2019;7(6):e12245) doi:10.2196/12245

KEYWORDS

WeChat; WeChat official accounts; user engagement; CDC; health education

Introduction

Acquisition and dissemination of health information play a significant role in promoting positive health behavior change [1]. Social media is currently becoming a new channel for information acquisition and exchange [2,3]. Nearly one-third of the world's population uses social media for entertainment, study, work, and socializing [4]. Compared with traditional types of print and broadcast media, social media has a unique advantage in facilitating two-way communication, allowing organizations to personalize content and interact with the public [5]. With the popularity of social media in the public, the use of these tools by health education organizations will have many opportunities to influence and change health behaviors [6].

WeChat, a free mobile app released in 2011, has become the most widely and frequently used social media platform in China [2]. WeChat has many functions, including instant messaging, free phone calls, mobile payments, etc [7]. In addition, a new functional module of WeChat called WeChat official accounts (WOAs) can be freely used by governments, companies, and organizations to provide information called pushed articles to the public; individuals can freely read these articles and communicate with others via these official accounts [3,8]. According to the latest data, the number of monthly active users of WeChat has reached 650 million and the number of WOAs exceeds 10 million. Nearly 80% of WeChat users have subscribed to the WOA [9]. Studies have found that WeChat can successfully encourage health improvement and behavior change [10]. For example, Wei et al [3] used WOAs to improve malaria health literacy. Cao et al [8] found that giving additional education and instruction via WOAs can improve therapy outcomes of patients. A weight loss intervention campaign based on a WOA was found to be effective for males [10]. With the growing popularity of WeChat and WOA in health knowledge acquisition, their roles in health education and health intervention are gradually receiving more attention [11].

The widespread public engagement with WeChat creates a ready platform for health promotion agencies for successful information distribution and diffusion online [12]. The China Centers for Disease Control and Prevention (CDC) are professional institutions that conduct health education and health promotion work for the public. In order to broaden health communication and make it easier for users to obtain information, the China CDC opened the Chinese disease control dynamics WOA in April 2014 [13]. Most of the provincial CDCs have now opened WOAs [14]. Generating user engagement is vital for effective information diffusion and health promotion. Identifying predictors of social media engagement can guide the development of content and use of features that have high appeal for the public [15]. User engagement was defined as users reacting to (ie, reading, liking) any content [16,17]. Past research has identified strategies for successful user engagement on Facebook and Twitter, including using multimedia content, highlighting celebrity involvement, using humor or shock appeals, etc [17,18]. However, there is little evidence establishing the best ways to engage with the public using WeChat, highlighting the importance of further exploration of this area. The literature base exploring user engagement through the WOAs of CDCs is even more limited. CDCs have the potential to enable broad dissemination of health information and messaging online and promote healthy behaviors, contributing to the development of social health [5].

In this study, we reviewed the use of WOAs by Chinese provincial CDCs. The study aim was to identify features of their articles that are associated with user engagement, which we defined as the level of reading and liking. Ultimately, we sought to formulate predictors of user engagement that would inform health education of public health organizations, so they can make use of WOAs to engage their target market and increase the effectiveness of changing health-related behaviors.

Methods

Data Source

We used the mobile WeChat app to search the official accounts by the name of the province and the key words "Centers for Disease Control and Prevention" and "CDC." We found 28 WOAs of provincial CDCs on January 15, 2018, and subscribed to them. Data for this study consisted of WeChat articles found on these WOAs between January 1, 2017, and December 31, 2017.

Main Indicators and Article Features Frame

For each article, we recorded the code of official accounts, push time, and amounts of article reading and liking. We developed a features frame for each article, referring to the research of Kite et al [17]. Specifically, we added article content, title type, and article length, which are features that may affect the effectiveness of the WOAs. Next, we conducted a presurvey on 100 articles. Based on the results of the survey, we have merged some of the less frequently used categories into other and made some modifications during iterative testing to make the features frame more relevant to public health communication. The final features framework with definitions is shown in Table 1. Then we invited six experts in relevant fields to evaluate the content validity of the features frame and calculated the content validity index (CVI). Experts used a 4-point Likert scale to assess the degree of consistency between the content of each item and the corresponding article features. Very unrelated was counted as 1 point, comparatively unrelated was counted as 2 points, comparatively related was counted as 3 points, and very relevant was counted as 4 points. The CVI of the item level was 0.83 to 1.00, and the CVI of the entire framework was 0.98.

Table 1. Final features frame with definitions.

Item	Definition
Title type	
Declarative sentence	The sentence contains subject, predicate, and object, usually with a period, aiming to state a fact. Example: The Centers for Disease Control and Prevention organized a series of women's day activities.
Exclamation or emphatic sentence	The sentence contains the subject, predicate, and object with an exclamation point. Example: These things will affect all doctors!
Question sentence	The sentence contains the subject, predicate, and object with a question mark. Example: Can the leftovers be eaten?
Imperative sentence	The sentence contains predicate and object, without subject, with an exclamation point. Example: Beware of these invisible drugs!
Phrase	Not a complete sentence, no predicate. Example: Common cold and flu.
Other	The title contains two or more sentences that are belong to different types. Example: Forget your flu shot? It is never too late!
Article content	
Infectious disease	Topic is related to infectious diseases. Example: Deciphering new knowledge about AIDS.
Chronic diseases	Topic is related to chronic diseases. Example: Women are more likely to develop diabetes. How can we prevent this?
Food safety and nutrition	Topic is related to food safety or nutrition. For example, the topic introduces food effects and gives recommendations.
Vaccination	Topic is related to vaccination, such as the introduction and guidance of some vaccines. Example: Forget your flu shot? It is never too late!
Environmental and occupational health	Topic is related to air hygiene, drinking water hygiene, soil hygiene, housing hygiene, occupational diseases, etc. Example: How air pollution affects the view.
Health education activities	Topic is related to events organized to conduct health education. Example: AIDS Day campaign.
Healthy lifestyle	Topic is related to popular knowledge of life, such as how to lose weight, the benefits of drinking water, the dangers of staying up late, etc.
Research progress	Topic is related to scientific research projects, including the laboratory construction of institutions.
National health policy, conferences, etc	Topic is related to national meetings and policies related to the health industry, such as the National Health and Wellness Conference.
Other	Content of the article is related to the unit meetings, work arrangements, introduction of advanced deeds, etc.
Article type	
Text only	Article contains only text.
Text and pictures	Article contains pictures and text.
Text, pictures, and links	Article contains pictures, links, and text.
Text, pictures, and videos	Article contains pictures, videos, and text.
Other	Other types or other combinations.
Communication skills	
Informative/instructive	Provides information on a health issue or instruction on how to do a behavior, such as health-related behavioral guidance.
Questioning	Remains skeptical and professionally corrects knowledge about widely circulated opinions, remarks or certain articles. Example: No rumors! True facts about dietary supplements.
Positive emotional appeal	Aims to elicit positive emotions, such as using positive examples to convey hope and excitement.
Fear appeal	Uses negative cases to cause user fear or other negative emotions, such as discussing serious consequences of long-term smoking, etc.
Humor	Uses any humorous technique (such as sarcastic, jokes, etc) to convey health messages, including funny pictures.
Other	Uses two and more of the skills mentioned above.
Marketing elements	
Persons of authority	Any person used for the purpose of lending their personal or positional authority to the health issue (eg, doctor, academic, scientist, politician).
Celebrities and sports people	Connecting people related to entertainment media or sports profiles with health events.

Item	Definition
Information sources	Citing sources of information such as books, guides, references, or instructions from other platforms.
Competitions, prizes, or giveaways	Any contest involving a participant entry, including minimal requirements such as liking or commenting on a article.
Hot topics	Using internet or real-life hot topics such as news coverage.
Tests	Collecting or testing participants' health knowledge.
Votes	Involving various types of voting for events or selections.
Article length	
0-1000 words	Amount of words in the article text.
1000-1499 words	Amount of words in the article text.
1500-1999 words	Amount of words in the article text.
2000 or more words	Amount of words in the article text.

Coding Method

We coded the data quantitatively using a coding scheme that assigned numeric values to article features. Title type, article content, article type, communication skills, and article length contained mutually exclusive categories, and they were coded with the corresponding number. However, multiple marketing elements are possible in a single article, so not all categories were mutually exclusive. Each marketing element was coded as 0 or 1 (0 represents not using, 1 represents using). Next we tested the interrater reliability between the two coders. After receiving training and getting familiar with the content, LFH and MJY independently coded the same subset of articles (n=300) from 28 WOAs. Any disagreement between the coders was resolved by discussion. Once interrater reliability reached 80%, LFH and MJY then individually coded half of the WOAs each.

Statistical Analyses

For every WOA, we examined the descriptive characteristics of articles, including the number and percentage of articles pushed in a year, median reading, and median liking per article. Next we generated descriptive statistics for each title type, article content, article type, communication skill, marketing element, and article length. We then investigated associations between the features of the articles and user engagement, operationalized as reading and liking. Since the data distribution is skewed, we categorized the amount of reading and liking by the 75% position, defining an amount of reading and liking less than the 75% position as low-level reading and liking and above the 75% position as high-level reading and liking. Given that the marketing elements were multichoice, we converted the features of marketing elements into the number of marketing elements. Next, we applied logistic regression analyses to assess associations between the features of the articles and the user engagement with the level of reading and the level of liking as the outcome variables and six features as categorical independent variables and used the *P* value to represent the result of the hypothesis test, which decided whether to reject the null hypothesis (the regression coefficient is equal to 0). We conducted a series of two-category univariate logistic regression analyses to perform the initial screening of variables. Variables that gave $P<.10$ in the univariate analyses were evaluated further using multivariable logistic regression. For the multivariable regression, $P<.05$ was considered the statistically significant level. Risks were expressed as adjusted odd ratios with 95% confidence intervals. EpiData 3.1 software (EpiData Association) was used to establish the database; SPSS Statistics version 15.0 (IBM Corp) was used for the statistical analysis.

Results

Characteristics of WeChat Official Accounts and Article Features

Overall, 28 provincial CDC WOAs pushed a total of 5976 articles in 2017 (Table 2). Guangdong CDC pushed the most articles (565/5976, 9.45%), and Hainan CDC pushed the fewest articles (4/5976, 0.07%). Shanghai CDC articles attracted the most user engagement (median reading: 3777, median liking: 37), and Ningxia CDC articles attracted the least (median reading: 21, median liking: 0). For all articles, the median reading was 551.5 and the median liking was 10.

As described in Table 3, article titles were usually declarative sentences (2358/5976, 39.45%). Most articles were related to healthy lifestyle (1355/5976, 22.67%), and text with pictures (3686/5976, 61.68%) was the most common article type. The most common communication skill was informative or instructive (3096/5976, 51.81%). Only 60.66% (3625/5976) of the articles contained any marketing elements, and the most articles used only one kind of marketing element (2543/5976, 42.55%). The article length was usually 1000 words or fewer (3615/5976, 60.49%).

Table 2. Characteristics of the 28 included WeChat official accounts (N=5976).

Province of the WeChat official channel	Articles, n (%)	Median reading	Median liking
Zhejiang	31 (0.5)	409	9
Fujian	205 (3.4)	281	9
Guangdong	565 (9.5)	1715	20
Hunan	210 (3.5)	2239	19.5
Hubei	529 (8.9)	956	16
Hainan	4 (0.1)	49.5	1.5
Yunnan	475 (7.9)	413	11
Guizhou	328 (5.5)	205.5	4
Sichuan	78 (1.3)	418.5	13
Qinghai	160 (2.7)	34	0
Gansu	185 (3.1)	270	4
Shanxi	231 (3.9)	1568	13
Jilin	223 (3.7)	163	3
Liaoning	358 (6.0)	211.5	3
Hebei	321 (5.4)	197	2
Jiangxi	154 (2.6)	200.5	8
Jiangsu	445 (7.4)	2498	31
Anhui	66 (1.1)	791	17
Henan	82 (1.4)	226.5	4
Shandong	211 (3.5)	393	9
Shanxi	55 (0.9)	127	4
Beijing	355 (5.9)	750	12
Shanghai	323 (5.4)	3777	37
Chongqing	59 (1.0)	525	23
Xinjiang	20 (0.3)	429	15
Guangxi	19 (0.3)	80	2
Ningxia	165 (2.8)	21	0
Xizang	119 (2.0)	180	2

Table 3. Frequencies by category of six article features (n=5976).

Article features	Value, n (%)
Title type	
Declarative sentence	2358 (39.45)
Exclamation or emphatic sentence	1011 (16.92)
Question sentences	858 (14.36)
Imperative sentence	517 (8.65)
Phrase	704 (11.78)
Other	528 (8.84)
Article content	
Infectious disease	910 (15.23)
Chronic diseases	385 (6.44)
Food safety and nutrition	578 (9.67)
Vaccination	350 (5.86)
Environmental and occupational health	221 (3.70)
Health education activities	625 (10.46)
Healthy lifestyle	1355 (22.67)
Research progress	89 (1.49)
National health policy, conferences, etc	421 (7.04)
Other	1042 (17.44)
Article type	
Text only	264 (4.42)
Text and pictures	3686 (61.68)
Text, pictures, and links	1576 (26.37)
Text, pictures, and videos	268 (4.48)
Other	182 (3.05)
Communicative skills	
Informative/instructive	3096 (51.81)
Questioning	314 (5.25)
Positive emotional appeal	1019 (17.05)
Fear appeal	704 (11.78)
Humor	258 (4.32)
Other	585 (9.79)
Number of marketing elements	
None	2351 (39.34)
One	2543 (42.55)
Two	910 (15.23)
Three or more	172 (2.88)
Article length	
0-1000 words	3615 (60.49)
1000-1499 words	1573 (26.32)
1500-1999 words	449 (7.51)
2000 or more words	339 (5.67)

Association Between Article Features and Level of Reading and Liking

As shown in Table 4, univariate logistic regression analysis revealed that title type, article content, article type, communication skills, number of marketing elements, and article length were significantly related to reading level (P<.001). Liking level displayed a similar pattern. All article features were evaluated further using multivariable logistic regression.

Articles for which the title type was an exclamation or emphatic sentence were less likely to obtain a high level of liking than those whose title types were declarative sentences. With regard to article content, articles about infectious diseases, vaccination, food safety and nutrition, and healthy lifestyle were more likely to receive high-level reading and liking than those about other topics. Articles related to health education activities or national health policy conferences were less likely to obtain high-level reading and liking. Environmental and occupational health–related articles were more likely to obtain high-level reading, unlike articles describing research progress, but both showed no effect for level of liking. For article type, articles containing text, pictures, and links were 5.21 times and 5.64 times more likely to receive high-level reading and liking, respectively, than text-only articles. Those containing text and pictures, similar to those containing text, pictures, and videos, were also more likely to receive high-level reading. With regard to communication skills, when compared with informative or instructive articles, those using traits of humor, questioning, fear appeal, and other were more likely to receive high-level reading and liking. Articles with positive emotional appeal were more likely to obtain a high-level liking, but this had no effect observed for level of reading. Articles using three or more kinds of marketing elements were more likely to obtain high-level reading and liking than those using none, while articles using one and two kinds were both less likely to obtain high levels of reading and liking. Compared with articles containing 1000 words or fewer, those containing 1000 to 1500 words and 1500 to 2000 words were more likely to obtain high-level reading and liking.

Table 4. Univariate and multivariable logistic regression analysis.

Article features	Univariate logistic regression		Multivariable logistic regression			
	Reading level, *P* value	Liking level, *P* value	Reading level		Liking level	
			Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Title type	<.001	<.001	—	—	—	—
Declarative sentence (reference)	—	—	1	—	1	—
Exclamation or emphatic sentence	.18	.60	—	—	0.80 (0.65-0.97)	.02
Question sentence	<.001	<.001	—	—	—	—
Imperative sentence	.04	.14	—	—	—	—
Phrase	.26	.22	—	—	0.69 (0.55-0.87)	.001
Other	<.001	<.001	—	—	—	—
Article content	<.001	<.001	—	—	—	—
Infectious disease	<.001	<.001	3.33 (2.60-4.25)	<.001	2.15 (1.69-2.73)	<.001
Chronic diseases	.57	.05	—	—	—	—
Food safety and nutrition	<.001	<.001	1.76 (1.34-2.31)	<.001	1.54 (1.18-2.01)	.001
Vaccination	<.001	.14	2.93 (2.15-4.00)	<.001	1.76 (1.29-2.41)	<.001
Environmental and occupational health	.48	.85	1.54 (1.05-2.26)	.03	—	—
Health education activities	<.001	<.001	0.65 (0.48-0.88)	.01	0.64 (0.48-0.86)	.002
Healthy lifestyle	<.001	<.001	1.89 (1.50-2.38)	<.001	1.39 (1.11-1.74)	.004
Research progress	<.001	.01	0.20 (0.05-0.82)	.03	—	—
National health policy, conferences, etc	<.001	<.001	0.30 (0.18-0.50)	<.001	0.38 (0.25-0.59)	<.001
Other (reference)	—	—	1	—	1	—
Article type	<.001	<.001	—	—	—	—
Text only (reference)	—	—	1	—	1	—
Text and pictures	.02	<.001	1.62 (1.09-2.40)	.02	2.31 (1.48-3.62)	<.001
Text, pictures, and links	<.001	<.001	6.21 (4.12-9.35)	<.001	6.64 (4.20-10.48)	<.001
Text, pictures, and videos	<.001	<.001	2.24 (1.36-3.69)	.002	2.60 (1.51-4.47)	.001
Other	.32	.42	—	—	—	—
Communication skills	<.001	<.001	—	—	—	—
Informative/instructive (reference)	—	—	1	—	1	—
Questioning	<.001	<.001	1.83 (1.40-2.39)	<.001	2.05 (1.57-2.67)	<.001
Positive emotional appeal	.01	.03	—	—	1.22 (1.00-1.49)	.048
Fear appeal	<.001	<.001	1.34 (1.10-1.64)	.004	1.44 (1.18-1.77)	<.001
Humor	<.001	<.001	2.93 (2.20-3.90)	<.001	4.13 (3.11-5.49)	<.001
Other	<.001	<.001	1.59 (1.25-2.02)	<.001	1.55 (1.22-1.97)	<.001
Number of marketing elements	<.001	<.001	—	—	—	—
None (reference)	—	—	1	—	1	—
One	.03	<.001	0.83 (0.72-0.96)	.01	0.76 (0.66-0.88)	<.001
Two	.01	<.001	0.62 (0.50-0.76)	<.001	0.60 (0.49-0.74)	<.001
Three or more	<.001	<.001	1.63 (1.13-2.36)	.01	1.86 (1.30-2.67)	.001
Article length	<.001	<.001	—	—	—	—
0-1000 words	—	—	1	—	1	—
1000-1499 words	<.001	<.001	1.45 (1.25-1.68)	<.001	1.42 (1.23-1.65)	<.001
1500-1999 words	<.001	<.001	1.67 (1.32-2.11)	<.001	1.55 (1.22-1.96)	<.001

Article features	Univariate logistic regression		Multivariable logistic regression			
	Reading level, *P* value	Liking level, *P* value	Reading level		Liking level	
			Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
2000 or more words	.97	.78	—	—	—	—

Discussion

Principal Findings

To our knowledge, this is the first study to investigate the current status of WOAs of CDCs for health information dissemination. Our analysis identified several features of articles that were associated with better information dissemination, which we defined as high-level user engagement. In the field of public health research and practice, WeChat represents a convenient and accessible tool for health education in China, which historically has been difficult using traditional methods [2,19]. Health promotion organizations should be aware of strategies to engage their target audience. The results presented in this paper may provide these organizations with some guidance on how to improve engagement with WeChat users.

According to our results, the content of articles was associated with user engagement. Indeed, article content was identified as an essential factor in determining whether WeChat users forward or share articles with friends [20]. Research on other social media platforms has also revealed that the content of posts seems to have a significant effect on user engagement [16,21]. Our findings further showed that what the public liked to read and praise were articles about infectious diseases; vaccination; food safety and nutrition; and healthy lifestyles, which reflects that they seem to be interested in popular science articles about daily health knowledge. In contrast, the public was relatively less interested in articles related to research progress and health education activities of institutions or national health policy, which may seem to be far away from daily life.

As the results showed, article type was an important covariate of user engagement. This is consistent with previous research. Research found that article type is an important indicator that alters the effectiveness of article dissemination on WOAs [22]. Among studies based on Facebook, the effect of the post type is also displayed. One found higher engagement to be associated with pictures, videos, and links [16]. Another showed that video posts were more attractive post types than picture posts [17]. Our study seems to give another possibility related to the combination of article types. A combination of text, pictures, and links was the most engaging article type, 5 times more popular than text only. Although the combination of text, pictures, and videos and the combination of text and pictures were also more likely to get a high-level reading and liking, the effect was far lower. This seems to indicate that links play a more important role in the combination of article types for increasing the user engagement. The effect of the links can be easily found in the advertising field. Ads may work better if the users can directly access the relevant pages through the links in the social media website pages instead of being forced to visit the external website [23]. And in the field of health promotion using networking platforms, relevant links are also identified as a key strategy for successful user engagement [24]. However, only 26.4% of all articles we coded included the combination of text, pictures, and links and the combination of text and pictures accounted for 61.7%, which suggested that public health organizations were trailing behind marketers in advertising field. The administrators of the CDC official accounts should probably add links to the traditional combination of text and pictures.

The communication skills used were also related to user engagement. Compared with the use of a peaceful tone to convey health knowledge, articles using humor were more likely to receive high levels of reading and liking. This result is similar to the findings of previous research: Klassen et al [25] found that using an optimistic tone was associated with more interactions on Facebook, and posts from health promotion organizations that were more serious in tone had minimal engagement from fans. This may reflect the effect of emotions [26]. Some researchers think that emotion is an important motivator of social sharing, and when participants experience positive emotions viewing a post on social media, they are more likely to engage with that post than are those who do not experience positive emotions [27,28]. In addition, articles using questioning skills were also nearly two times more likely to achieve high-level reading and liking than those using informative or instructive skills. Some surveys have found that many respondents believe that internet health information is not reliable [2,29]. Indeed, there are potential dangers associated with using social media for health communication, such as sharing of misleading or inaccurate information [5]. As professional health organizations, CDCs use questioning expressions to correct this information, which may attract readers' attention. According to our results, articles with fear appeal were likely to show similar but slightly lower results than those using the above two skills. In the past, fear has been employed by many health promotion organizations to induce behavior change [25]; however, Soames Job [30] and Keller [31] think that fear is only likely to work under particular circumstances. Therefore, the effectiveness of fear appeal may require further research.

The number of marketing elements was associated with the level of user engagement. Articles that used one or two marketing elements were less likely to obtain high-level reading and liking than those that used none. However, articles using three or more kinds of element obtained the opposite result. This showed that some marketing elements were associated with lower levels of engagement and some with higher engagement. Previous research has discussed the impact of some marketing elements on user engagement in other social media. The effect of celebrities is usually positive. Chapman [32] believed that celebrity involvement in public health campaigns delivers long-term benefit. Similarly, Kite [17] found that the use of celebrities and athletes led to higher average engagement.

However, the effect of authoritative people is still controversial. One study found that, compared with Facebook posts that have no marketing elements at all, using authoritative people reduced likes, shares, and comments [17]. But Preece et al [21] thought that using charismatic leaders or respected authorities increased the chances of user engagement. Therefore, which marketing methods can play positive roles in increasing the user engagement on articles of WOAs remains to be further discussed.

The results of this study showed that article length was also a covariate of user engagement: articles containing 1000 to 1500 words and 1500 to 2000 words were more likely to obtain high-level reading and liking. A study of WeChat articles considered that text above 1500 words can be considered as a long text [33]. Our study showed that users seem to have a preference for articles of moderate length.

In our study, title type was negatively associated with user engagement. When the title was an exclamation or an emphatic sentence or phrase, the articles were less likely to receive high-level liking compared with titles involving declarative sentences. Zhang [34] analyzed the influence of WeChat article titles on people's willingness to open the full text and found that the content of the title could affect the dissemination of the article. However, we do not have sufficient evidence to demonstrate what type of title was more popular with readers. Therefore, the link between title type and user engagement seems to present an interesting direction for future research.

Limitations

Our findings should be interpreted with consideration of some study limitations. First, the study involved a short time frame of data collection. Evaluating a longer period may identify time differences and improve strategies. Second, this study used reading and liking as indicators of user engagement, and we further transformed them into two-category variables: high-level and low-level. We may thus be obscuring the effects of the original variables. Future analyses should expand our findings by quantitatively evaluating the indicators associated with engagement. Last, other factors that we have not focused on in this study may also shape user engagement such as individual-level factors. These individual factors and article features should be analyzed comprehensively in future research.

Conclusions

Social media involves technologies that facilitate opportunities for engaging with the audience [35]. How social media can best be used to achieve health information dissemination and public health outcomes is a topic of much discussion and study in the public health community. Given the lack of related studies based on WeChat or official accounts, we conducted this study and found that the article content, article type, communication skills, the number of marketing elements, and article length were associated with reading level and liking level. However, title type was only associated with liking level. Our results may provide public health and community leaders with insight into the diffusion of important health topics of concern. Because this study was focused primarily on user engagement and article features, future studies might improve our understanding of other factors that contribute to the dissemination of specific key themes. For example, the agency must identify what audience they are trying to reach, how that audience uses WeChat, and what goals and objectives are most appropriate.

Acknowledgments

We thank International Science Editing (www.internationalscienceediting.com) for editing this manuscript. Funding was provided by grants from the social science and technology development project of Dongguan, Guangdong, China (grant number 201650715000528), Guangdong Higher Education Teaching Reform Project (grant number 4G17154), and Guangdong Medical University Teaching Reform Research Project (grant number 4SG17143).

Conflicts of Interest

None declared.

References

1. Adam M, McMahon SA, 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-e12163 [FREE Full text] [doi: 10.2196/12128] [Medline: 30698531]
2. Zhang X, Wen D, Liang J, Lei J. How the public uses social media wechat to obtain health information in china: a survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
3. Li W, Han LQ, Guo YJ, Sun J. Using WeChat official accounts to improve malaria health literacy among Chinese expatriates in Niger: an intervention study. Malar J 2016 Nov 24;15(1):567 [FREE Full text] [doi: 10.1186/s12936-016-1621-y] [Medline: 27881122]
4. Mejova Y, Weber I, Fernandez-Luque L. Online health monitoring using Facebook advertisement audience estimates in the United States: evaluation study. JMIR Public Health Surveill 2018 Mar 28;4(1):e30 [FREE Full text] [doi: 10.2196/publichealth.7217] [Medline: 29592849]

5. Alber JM, Bernhardt JM, Stellefson M, Weiler RM, Anderson-Lewis C, Miller MD, et al. Designing and testing an inventory for measuring social media competency of certified health education specialists. J Med Internet Res 2015 Sep 23;17(9):e221 [FREE Full text] [doi: 10.2196/jmir.4943] [Medline: 26399428]
6. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
7. Man L. Characteristics and skills of text editing under WeChat public platform. Editorial Transactions (Chinese) 2017:116-120.
8. Cao Y, Lin S, Zhu D, Xu F, Chen Z, Shen H, et al. WeChat public account use improves clinical control of cough-variant asthma: a randomized controlled trial. Med Sci Monit 2018 Mar 14;24:1524-1532 [FREE Full text] [doi: 10.12659/msm.907284] [Medline: 29536984]
9. Liang X, Yang Y. An experimental study of Chinese tourists using a company-hosted WeChat official account. Electronic Commerce Research and Applications 2018 Jan;27:83-89. [doi: 10.1016/j.elerap.2017.12.007]
10. He C, Wu S, Zhao Y, Li Z, Zhang Y, Le J, et al. Social media-promoted weight loss among an occupational population: cohort study using a wechat mobile phone app-based campaign. J Med Internet Res 2017 Dec 23;19(10):e357 [FREE Full text] [doi: 10.2196/jmir.7861] [Medline: 29061555]
11. Hudnut-Beumler J, Po'e E, Barkin S. The use of social media for health promotion in hispanic populations: a scoping systematic review. JMIR Public Health Surveill 2016 Jul 11;2(2):e32 [FREE Full text] [doi: 10.2196/publichealth.5579] [Medline: 27400979]
12. Lim MS, Hare JD, Carrotte ER, Dietze PM. An investigation of strategies used in alcohol brand marketing and alcohol-related health promotion on Facebook. Digit Health 2016 Apr 29;2:2055207616647305 [FREE Full text] [doi: 10.1177/2055207616647305] [Medline: 29942554]
13. Huan Z, Ming W, Wei W. Practice analysis of disease control organizations using WeChat for health communication. China Pub Health Manag (Chinese) 2017:605-607. [doi: 10.19568/j.cnki.23-1318.2017.05.006]
14. Mengyuan Z. Opportunities and challenges brought by WeChat public account for health science popularization. New Media Res (Chinese) 2017;3(5):51-52. [doi: 10.16604/j.cnki.issn2096-0360.2017.05.023]
15. Andrade EL, Evans WD, Barrett N, Edberg MC, Cleary SD. Strategies to increase latino immigrant youth engagement in health promotion using social media: mixed-methods study. JMIR Public Health Surveill 2018 Dec 19;4(4):e71 [FREE Full text] [doi: 10.2196/publichealth.9332] [Medline: 30567689]
16. Card KG, Lachowsky N, Hawkins BW, Jollimore J, Baharuddin F, Hogg RS. Predictors of Facebook user engagement with health-related content for gay, bisexual, and other men who have sex with men: content analysis. JMIR Public Health Surveill 2018 Apr 06;4(2):e38 [FREE Full text] [doi: 10.2196/publichealth.8145] [Medline: 29625953]
17. Kite J, Foley BC, Grunseit AC, Freeman B. Please like me: Facebook and public health communication. PLoS One 2016 Sep;11(9):e0162765 [FREE Full text] [doi: 10.1371/journal.pone.0162765] [Medline: 27632172]
18. Veale HJ, Sacks-Davis R, Weaver ER, Pedrana AE, Stoové MA, Hellard ME. The use of social networking platforms for sexual health promotion: identifying key strategies for successful user engagement. BMC Public Health 2015 Feb 06;15(1):85 [FREE Full text] [doi: 10.1186/s12889-015-1396-z] [Medline: 25884461]
19. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: a systematic review. J Med Internet Res 2014 Mar;16(3):e79 [FREE Full text] [doi: 10.2196/jmir.2679] [Medline: 24642014]
20. Ji H, Cai D. Analysis of influencing factors of WeChat public information information transmission. J Jimei University (Chinese) 2016;19(4):107-115. [doi: 10.3969/j.issn.1008-889X.2016.04.015]
21. Preece J, Shneiderman B. The reader-to-leader framework: motivating technology-mediated social participation. AIS Trans HCI 2009 Mar 31;1(1):13-32. [doi: 10.17705/1thci.00005]
22. Zheng B, Ding Y. Analysis of the factors affecting the spreading effect of the WeChat public account Technology Daily. AV (Chinese) 2018:112-115. [doi: 10.19395/j.cnki.1674-246x.2018.01.062]
23. Freeman B, Potente S, Rock V, McIver J. Social media campaigns that make a difference: what can public health learn from the corporate sector and other social change marketers? Pub Health Res Pract 2015 Mar 30;25(2):e2521517 [FREE Full text] [doi: 10.17061/phrp2521517] [Medline: 25848735]
24. Veale HJ, Sacks-Davis R, Weaver ER, Pedrana AE, Stoové MA, Hellard ME. The use of social networking platforms for sexual health promotion: identifying key strategies for successful user engagement. BMC Public Health 2015 Feb 06;15(1):85 [FREE Full text] [doi: 10.1186/s12889-015-1396-z] [Medline: 25884461]
25. Klassen KM, Borleis ES, Brennan L, Reid M, McCaffrey TA, Lim MS. What people “like”: analysis of social media strategies used by food industry brands, lifestyle brands, and health promotion organizations on facebook and instagram. J Med Internet Res 2018 Jun 14;20(6):e10227 [FREE Full text] [doi: 10.2196/10227] [Medline: 29903694]
26. Dunlop SM, Perez D, Cotter T. The natural history of antismoking advertising recall: the influence of broadcasting parameters, emotional intensity and executional features. Tob Control 2014 May;23(3):215-222. [doi: 10.1136/tobaccocontrol-2011-050256] [Medline: 23143868]

27. So J, Prestin A, Lee L, Wang Y, Yen J, Chou WS. What do people like to "share" about obesity? A content analysis of frequent retweets about obesity on twitter. Health Commun 2016;31(2):193-206. [doi: 10.1080/10410236.2014.940675] [Medline: 26086083]
28. Smith S. Conceptualising and evaluating experiences with brands on Facebook. Int J Market Res 2013 May;55(3):357-374. [doi: 10.2501/IJMR-2013-034]
29. Hou S, Gan Y, Bai Q. Present situation of using health information technologies in residents of Beijing. Chinese J Med Libr Inform Sci (Chinese) 2016;25(4):41-44. [doi: 10.3969/j.issn.1671-3982.2016.04.009]
30. Soames Job RF. Effective and ineffective use of fear in health promotion campaigns. Am J Public Health 1988 Feb;78(2):163-167. [doi: 10.2105/ajph.78.2.163] [Medline: 3276236]
31. Keller PA, Block LG. Increasing the persuasiveness of fear appeals: the effect of arousal and elaboration. J Consum Res 1996 Mar;22(4):448. [doi: 10.1086/209461]
32. Chapman S. Does celebrity involvement in public health campaigns deliver long term benefit? Yes. BMJ 2012 Sep 25;345(sep25 2):e6364-e6364. [doi: 10.1136/bmj.e6364] [Medline: 23015036]
33. Zhang Y, Li W. Which WeChat articles are easier to be forwarded—an exploratory research based on content analysis. AV (Chinese) 2016:117-118. [doi: 10.19395/j.cnki.1674-246x.2016.02.061]
34. Lu Z. Analysis of the influence of WeChat Tweet title on people's willingness to open. Southeast Propagation (Chinese) 2017:64-67. [doi: 10.13556/j.cnki.dncb.cn35-1274/j.2017.06.021]
35. Thackeray R, Neiger BL, Smith AK. Adoption and use of social media among public health departments. BMC Pub Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]

Abbreviations

CDCs: Centers for Disease Control and Prevention
CVI: content validity index
WOA: WeChat official account

Edited by G Eysenbach; submitted 19.09.18; peer-reviewed by A Jadhav, J Lei, C Allen; comments to author 07.01.19; revised version received 28.02.19; accepted 14.05.19; published 27.06.19.

Please cite as:
Zhang Y, Xia T, Huang L, Yin M, Sun M, Huang J, Ni Y, Ni J
Factors Influencing User Engagement of Health Information Disseminated by Chinese Provincial Centers for Disease Control and Prevention on WeChat: Observational Study
JMIR Mhealth Uhealth 2019;7(6):e12245
URL: http://mhealth.jmir.org/2019/6/e12245/
doi:10.2196/12245
PMID:31250833

Original Paper

Development and Testing of a Mobile Phone App for Risk Estimation of Gas Volume Expansion and Intraocular Pressure Elevation in Patients With Intravitreous Gas or Air Tamponade: Interobserver Assessment Study

Zhaotian Zhang[1], MD; Fei Li[1], MD, PhD; Haochuan Zhang[2], PhD; Zhipeng Miao[2], PhD; Yantao Wei[1], MD, PhD; Li Wang[1], MD; Shaochong Zhang[1], MD

[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China
[2]School of Automation, Guangdong University of Technology, Guangzhou, China

Corresponding Author:
Shaochong Zhang, MD
State Key Laboratory of Ophthalmology
Zhongshan Ophthalmic Center
Sun Yat-sen University
54S Xianlie Road
Guangzhou, 510060
China
Phone: 86 2087330415
Fax: 86 2087330415
Email: zhangshaochong@gzzoc.com

Abstract

Background: Pars plana vitrectomy (PPV) with intravitreous tamponade of gas or air has been widely used for a series of vitreoretinal diseases. It is estimated that 100,000 patients per year undergo PPV globally, and half of them were subsequently tamponaded with gas or air. According to Boyle's law ($P_1V_1=P_2V_2$), patients with an intravitreous remnant of gas or air will be under high risk of intraocular pressure (IOP) elevation and subsequent vision loss owing to the expanded intravitreous gas or air when traveling post operation to a place with a significantly higher altitude. We always explain to patients why postoperative travel is potentially risky. Emergency cases of elevated IOP caused by postoperative traveling would sometimes come to surgeons. However, there have been few disease education or reference tools for both the surgeons and patients to have better communication.

Objective: The aim of this study was to introduce and evaluate a mobile phone app developed by surgeons (the authors) for preliminary risk estimation of volume expansion and IOP elevation in patients with intravitreous gas or air when traveling to a place of higher altitude.

Methods: The app was developed on the iOS and Android operating systems. Boyle's law ($P_1V_1=P_2V_2$) was the theoretical basis of the app. Intravitreous gas or air volume and altitude values were independent factors to deduce the risk report. Consecutive patients underwent vitrectomy, and those with an intravitreous remnant of gas or air were recruited. The surgeons judged the vertical height of the fluid/gas interface through the dilated pupil; the patients were instructed to judge it according to their visual field when looking straight ahead and line it out on a chart included in the app. Finally, all the patients were required to fill a Likert scale–based questionnaire with 2 main items to evaluate the participants' user experience and attitudes toward the app.

Results: A total of 50 patients were included (30 males and 20 females). All patients could independently operate the app to complete the test. The median heights of the fluid/gas interface independently judged by the surgeon and patients were 40% (range: 10%-75%) and 41% (range: 9%-78%), respectively (P=.63). The median altitude of the participants' destinations was 150.0 m (range: 0-3490 m). The Bland-Altman analysis revealed a good agreement between the surgeons' and patients' judgments (bias of −0.3%), with 95% limits of agreement of −5.8% to 5.3%. Overall, the Likert scale revealed a positive attitude from the patients toward the app.

Conclusions: The app is reliable for patients to have preliminary risk estimation of intravitreous gas or air volume expansion and IOP elevation if travel to a place of higher altitude is planned. The surgeons could also use it as a platform for better disease communication.

(JMIR Mhealth Uhealth 2019;7(6):e14592) doi:10.2196/14592

KEYWORDS

intraocular pressure; mobile phone; vitrectomy; air

Introduction

Intravitreous gas or air is widely used by vitreoretinal surgeons to achieve anatomic resolution and thus help patients to regain vision [1,2]. Pars plana vitrectomy (PPV) with an intravitreous air/gas tamponade has significantly improved the surgical treatment outcomes for some vitreoretinal diseases, such as rhegmatogenous retinal detachment, macular holes, vitreomacular traction, and foveal retinoschisis [3-6]. Surgeons should always be aware of the effect of intraocular gas or air on the fluctuation of intraocular pressure (IOP). Patients with intravitreous gas or air have generally been advised not to travel in airplanes or travel to locations at significantly higher altitudes as discomfort and even vision loss owing to sharply elevated IOP have been reported [7-11]. The gas or air is tightly sealed within the eyewall, and the eyeball is a semirigid container. Therefore, when the operated eye filled with gas or air is exposed to lower atmospheric pressure, the volume of the intravitreous gas expands. Boyle's law ($P_1V_1=P_2V_2$) provides the theoretical basis for noticing and explaining the potential risks when patients are exposed to different atmospheric pressures [12].

Vitreoretinal surgeons sometimes need to have better communication with their postoperative patients when intravitreous gas or air is present [7-12]. The mathematical models of eyes with intravitreal gas bubbles developed by Amini et al [13-15] and Wong et al [16] were significant and inspired vitreoretinal surgeons to gain an in-depth understanding of the impact of altitude-induced IOP changes. However, there has not been any easily available reference tool for both surgeons and patients to get a preliminary estimation of the risk of volume expansion and IOP elevation when traveling to a place at a higher altitude. In addition, patients with intravitreous gas or air have few reliable tools for self-monitoring the volume after being discharged from the hospital when their surgeons' professional advice may not be readily available.

The current trend shows that the global ownership rate of mobile phones is quite high [17]. The usefulness of mobile phones has been demonstrated in some medical disciplines because of the rapid advancement of mobile phone technology. Some physicians have tried to use mobile phones and tablets to facilitate their work and have achieved positive outcomes in subspecialties of public health, psychology, oncology, medical education, and so forth [18-30]. Therefore, we hypothesized that mobile phones would be an ideal platform for vitreoretinal surgeons to perform risk estimation and patient education of intravitreous gas or air expansion. To that end, we developed a mobile app based on Boyle's law and previously reported mathematical models [8,12-16,31]. This study was designed to test the applicability of the app in a real-world setting and evaluate the interobserver agreement on gas volume estimation between surgeons and patients using the app.

Methods

Design Principle of the App

The app was developed in the iOS operating system (iOS 12.2; Apple Inc) and Android operating system (Android 8.1; Google LLC) using the programming languages Objective-C (Xcode 10.1; Apple Inc) (for iOS version) and Java (Oracle Java ME SDK 8.2; Oracle Inc) (for Android version), respectively. The app was installed and verified on an iOS mobile phone (iPhone 7+; Apple Inc) and an Android mobile phone (BlackBerry KEY^2; BlackBerry Limited) as appropriate. The app, named *Intraocular Gas*, has been released by the authors and can be downloaded for free from Apple's App Store (Apple Inc) and the Google Play app store (Google LLC).

Boyle's law ($P_1V_1=P_2V_2$) was the principle used to measure the prediction accuracy of gas expansion. The calculation depended on gas volume and altitude values. The vitreous cavity was set to be an oblate spheroid (whole volume=4.5 ml). The app calculated the intravitreous gas or air volume based on the vertical height of the fluid/gas interface (in percentages) when the patient was in a sitting position with the head in a level position. Finally, the ratio (expanded volume of intravitreous gas(air)/normal volume of the anterior chamber) was transferred by the app to report a final risk estimation of IOP elevation provided that the patient was transported directly and immediately to a place of higher altitude. The governing equations of the app are illustrated in detail (see Multimedia Appendix 1).

Interobserver Assessment

The study was approved by the Institutional Review Board of Zhongshan Ophthalmic Center of Sun Yat-sen University (reference number: 2019KYPJ059) and performed in accordance with the World Medical Association's Declaration of Helsinki. Written informed consent was obtained from each participant. Consecutive patients who had undergone PPV and had an intravitreous remnant of gas or air were recruited. Patients with significant opacification of the anterior segment and those who could not easily understand or operate the app were excluded (Figure 1).

First, the surgeons observed the horizontal fluid/gas interface with a preset lens through the dilated pupil and judged its height according to the anatomic landmarks (fovea, vascular arcades, and optic disc; Figure 2).

Second, with the fellow eye covered, the patient was instructed by the nurses to assess the vertical height of the fluid/gas interface according to his/her visual field when looking straight ahead (Figure 2). The app includes a chart for the patient to line out the interface according to his/her vision (Figure 2). A live demonstration video illustrates the entire procedure on how patients used the app, following the nurse's brief instruction.(see Multimedia Appendix 2).

Figure 1. Flowchart for study participants and analysis of the agreement between the heights of the fluid/gas interface independently judged by the surgeons and patients. PPV: pars plana vitrectomy.

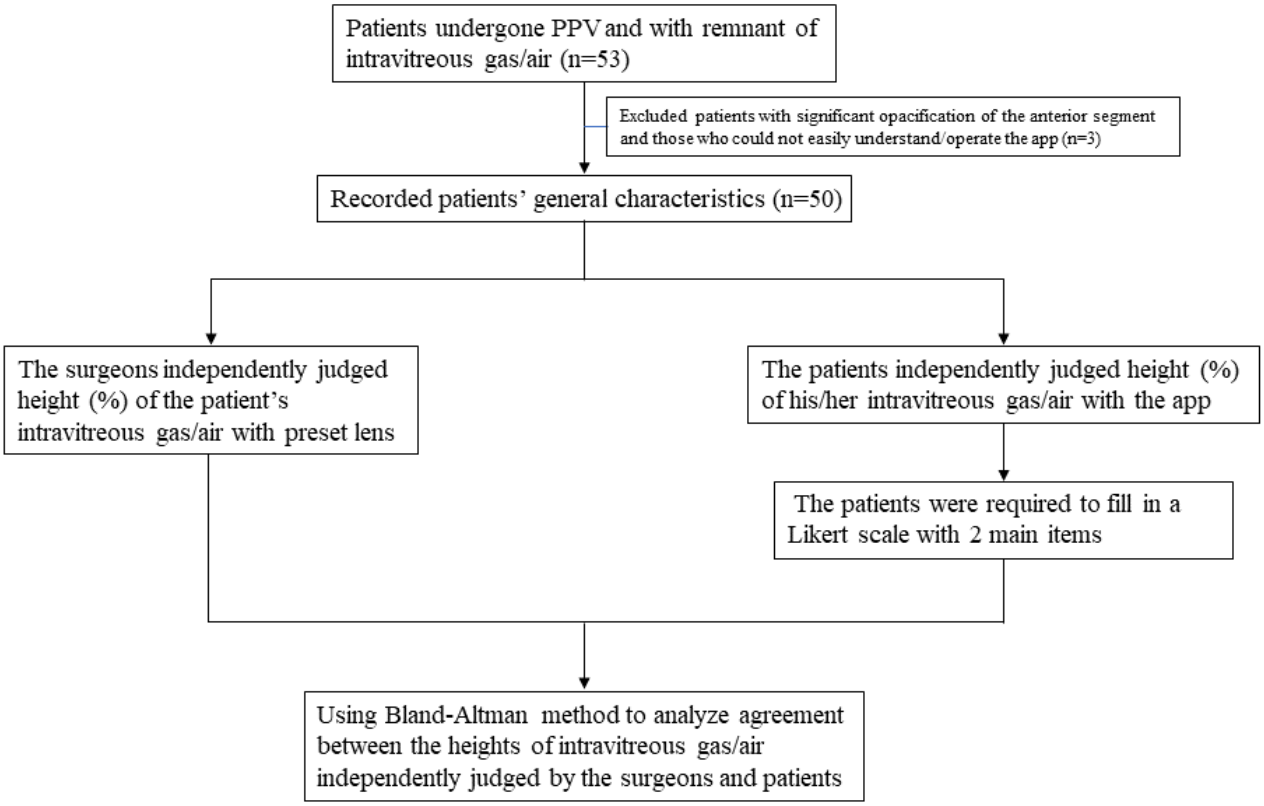


Figure 2. Brief illustration of the interobserver judgement study and usage of the app The surgeon observed the horizontal gas/fluid interface with a preset lens through the dilated pupil and judged the height of it according to the anatomic landmarks (A); With the fellow eye covered, the patient was instructed to judge the height of gas/fluid interface according to his/her visual field when looking straight forward (B); Screenshot of the app showing the brief instruction on how to line out the height of gas/fluid interface in his/her tested eye (C); The chart included in the app for the patient to line out the gas/fluid interface according to his/her judgement (D).

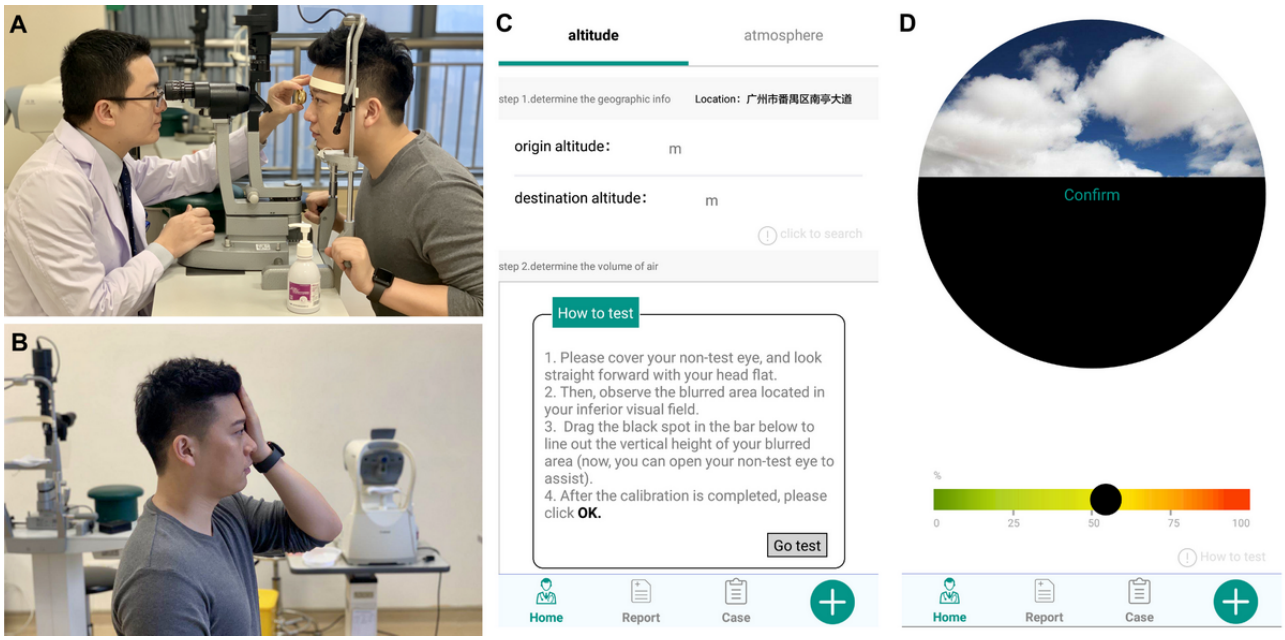
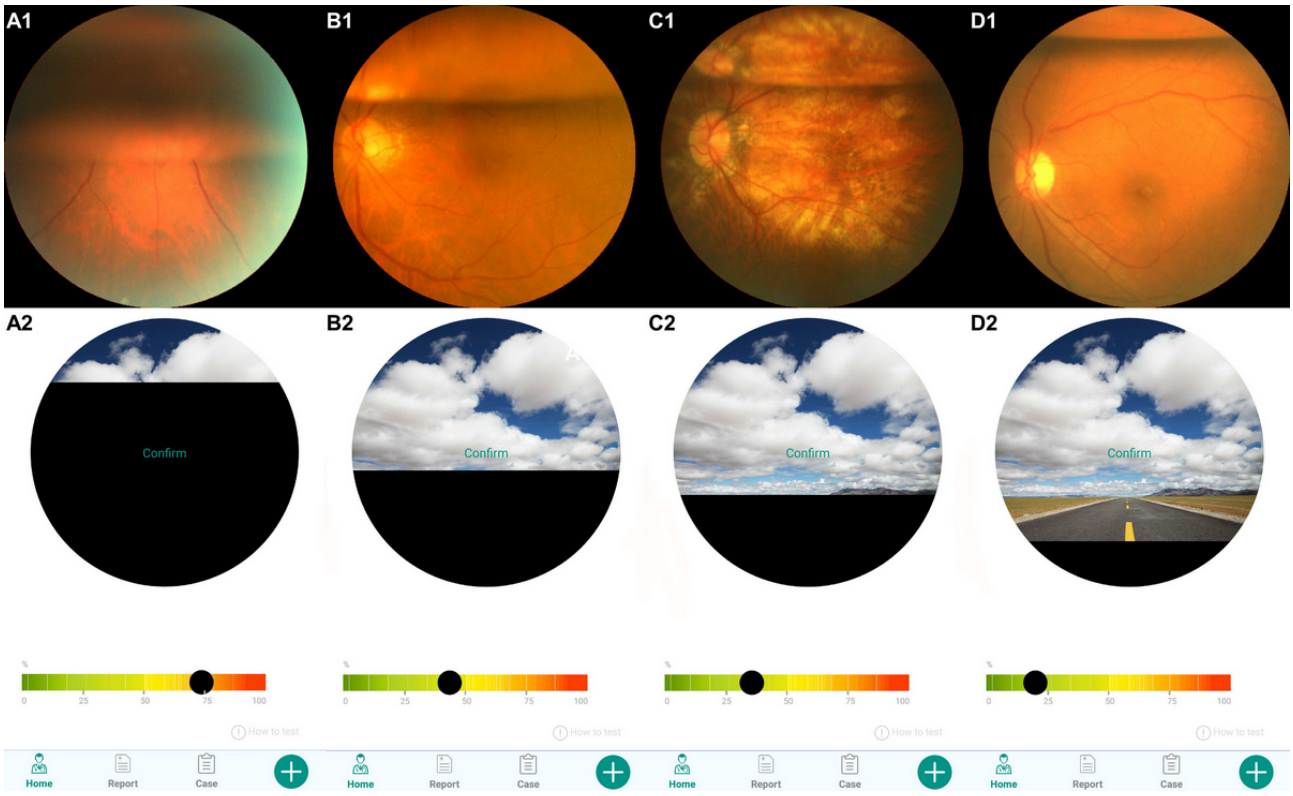


Figure 3. Representative fundoscopy photographs and the corresponding heights of gas/fluid interface independently judged by the surgeons and patients. The top row (A1, B1, C1, and D1) lists fundoscopy photographs of the patients, the height of gas/fluid interface judged by the surgeons were 75%, 45%, 35%, and 20%, respectively (from left to right). The bottom row (A2, B2, C2, and D2) lists corresponding heights of the gas/fluid interface lined out in the app's chart independently by the patients, the values were 76%, 44%, 35%, and 18%, respectively (from left to right).



Fundus photography was obtained from each patient to verify the height of the fluid/gas interface. Representative fundoscopy photographs and the corresponding heights of the horizontal fluid/gas interface independently judged by the surgeons and patients are illustrated in Figure 3.

Finally, all of the patients were required to fill a Likert scale with 2 items to assess the usability of the app (*Item 1: I think the app is reliable for me to have a postoperative travel plan to another place at a different altitude; Item 2: I will recommend the app to other patients with the same condition as me; 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree*).

Statistical Analysis

Statistical analysis was performed using MedCalc Statistical Software, version 18.9 (MedCalc Software bvba). Continuous variables are expressed as mean (SD) and median (range) according to their distributions (Shapiro-Wilk normality test). Categorical variables are described as proportions (%). Agreement between the heights of the fluid/gas interface independently judged by the surgeons and patients was assessed using the Bland-Altman method with 95% limits of agreement. A *P* value of less than .05 was considered statistically significant.

Results

Patient Characteristics

A total of 50 consecutive patients were included (30 males and 20 females; median age: 51.5 years [range: 18-70 years]; visual acuity ranged from counting fingers [CF] to 20/25). All the patients underwent surgery in the Zhongshan Ophthalmic Center of Sun Yat-sen University (Guangzhou, China; altitude=10 m). All the patients could easily and independently operate the app to complete the test following the nurse's brief instructions.

Outcomes of Interobserver Assessment

The median heights of the fluid/gas interface independently judged by the surgeon and patients were 40% (range: 10%-75%) and 41% (range: 9%-78%), respectively (with no significant statistical difference when analyzed with the Wilcoxon matched-pairs signed rank test; *P*=.63). The median altitude of the participants' places of planned destination was 150.0 m (range: 0-3490 m). The median expanded volumes of intravitreous gas or air independently judged by the surgeons and patients and calculated by the app was 0.02 ml (range: 0-2.4 ml) and 0.02 ml (range: 0-2.3 ml), respectively (with no significant statistical difference when analyzed with the Wilcoxon matched-pairs signed rank test; *P*=.55). In total, 38% (19/50) of the patients planned to travel back by airplane, 30% (15/50) by train (high-speed/ordinary=11/4), and the remaining 32% (16/50) by car/bus. The main characteristics of all the patients are summarized into a table in Multimedia Appendix 3).

The Bland-Altman analysis revealed a bias of −0.3% between the surgeons' and patients' judgments, with 95% limits of agreement, of −5.8% to 5.3% (Figure 4). The Likert scale revealed a generally positive attitude from the patients toward the app (Figure 5).

Figure 4. The Bland-Altman analysis revealed a bias of −0.3% (solid line) between the surgeons' and patients' judgements, with 95% limits of agreement, of −5.8% to 5.3% (dash lines).

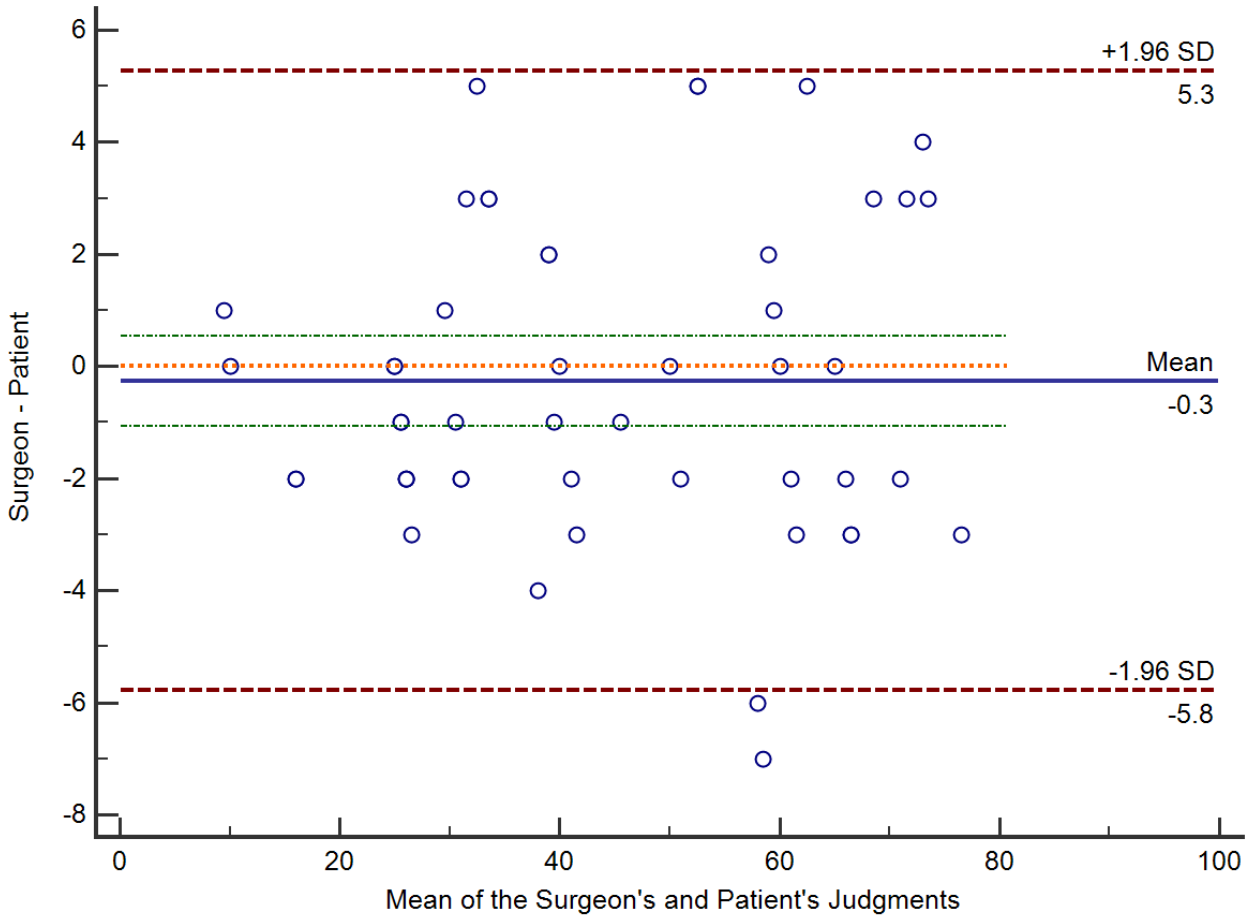
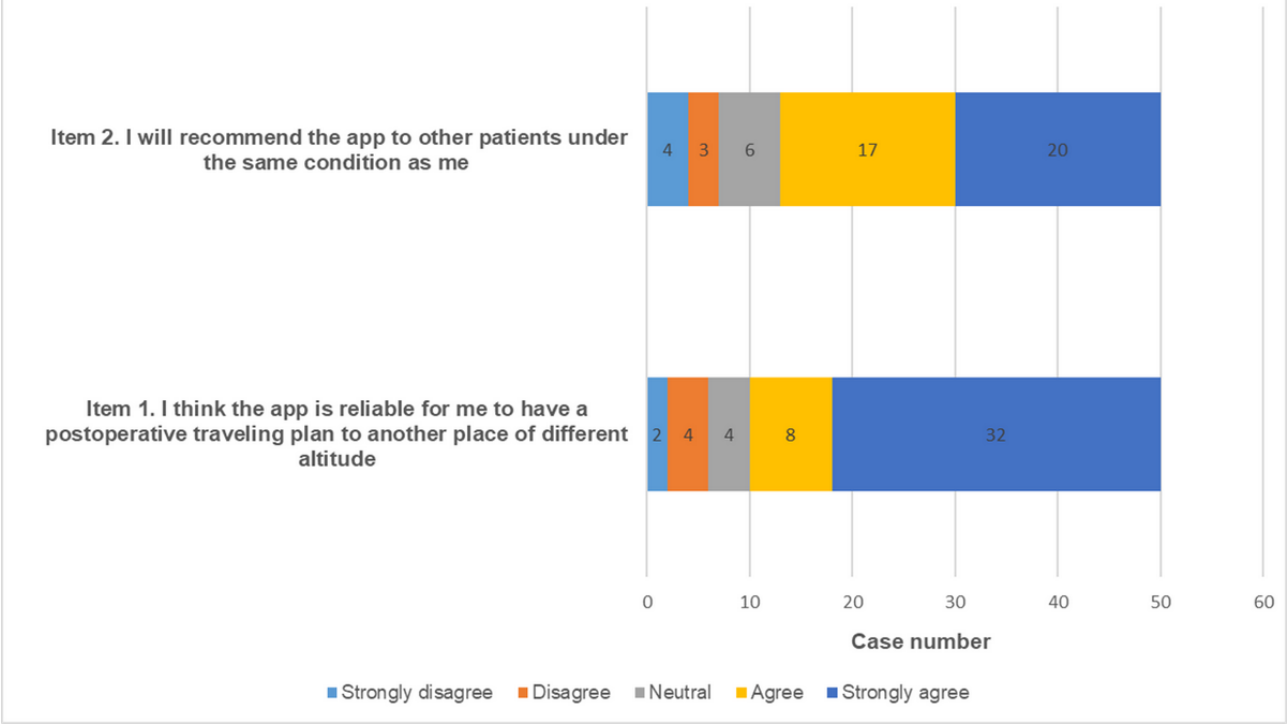


Figure 5. The Likert scale revealed a generally positive attitude from the patients toward the app.



Discussion

In this study, we demonstrated the design principle and clinical applications of a mobile app developed by the authors, and we showed the app's potential in preliminary risk estimation of volume expansion and IOP elevation in patients with an intravitreous gas or air tamponade. To the best of our knowledge, the app is the first one developed by vitreoretinal surgeons to solve the actual conundrum of intravitreous gas or air in our daily clinical work.

Theoretical Basis and Design Principle of the App

The essential factors of the app include the following: the mechanics and kinetics of intravitreous gas or air tamponading after PPV; the semirigid property of the eyeball; Boyle's law (pressure × volume=constant); the relationship between atmospheric pressure and altitude; and the ratio of the expanded volume of intravitreous gas or air to that of the anterior chamber. As proposed by Lincoff and Aronowitz, the eyeball has 3 compensation mechanisms to accommodate the expansion of intravitreous gas or air: choroidal compression, scleral expansion, and accelerated aqueous outflow [11,32]. These mechanisms are significant for us to have an in-depth understanding of elevated IOP induced by decreases in atmospheric pressure.

Unlike previously reported mathematical models [13-15,33], we developed our app based on the assumption that the patient with intravitreous gas or air would be traveling immediately from one place to another. Therefore, the abovementioned compensation mechanisms were not incorporated into the app design. Another reason for developing this app was to provide a readily available electronic tool for both surgeons and patients to calculate a preliminary estimation of the risk. Therefore, it was decided to develop the app based on a simplified and feasible principle. We contributed our best efforts to making the app suitable for both surgeons and patients to achieve more friendly interactions and higher interobserver agreement.

The App's Potential to Help Disease Communication and Preliminary Risk Estimation

There have been continuous reports on the risks of postoperative travel for vitrectomied eyes before the intravitreous gas or air is completely absorbed [7,8,34,35]. We referred to Boyle's law to explain the phenomenon, which is also the governing equation incorporated into this app. On the basis of previous studies of animal models, simulated flight experiments, mathematical models, and case reports, patients are generally advised against traveling through lower atmospheric pressure environments [7,9,10,14,33-37]. Patients are exposed to lower atmospheric pressure when they embark on postoperative travel by airplane or by car/train to areas with higher altitudes. A previous animal study by Dieckert et al [37] indicated that it is unsafe to fly with intraocular gas of any volume, unless cabin pressurization can be maintained at 2000 feet at 706 mmHg or higher. However, there is still some controversy on the seriousness of this risk because the cost of vision loss is huge and irreversible, even though the risk is theoretically low.

In our clinical work, we always explain to patients why postoperative travel is potentially risky. Emergency cases involving elevated IOP caused by postoperative travel sometimes came to our clinic, and there are anecdotal reports and authors' experience to support the existence of this risk. There are 4 possible underlying causes of the vision-threatening tragedy: absence or inadequate disease education regarding the risk of postoperative travel with an intravitreous gas or air tamponade; lack of an easily understandable education tool for surgeons and patients to engage in more effective interaction about the risk of gas or air expansion as the risk is informed according to the physical law, mathematical models, and case reports; lack of postoperative monitoring by surgeons of the volume of intravitreous gas after the patient is discharged from the hospital; and the fluke mind of some patients who lack sufficient awareness of the risk of expanded intravitreous gas or air [8]. This app can enable both surgeons and patients to better keep track of the risk of expanded intravitreous gas or air.

Where and When the App Will Serve Surgeons and Patients

We speculate that the app will be practical in many areas of the world. Indeed, previous reports indicate that significant elevation differences between 2 places should be considered a risk factor for elevated IOP in such patients, regardless of the mode of transportation they choose [7,8,10,34,35]. The app directly offers a preliminary evaluation of the risk of IOP elevation based on 2 independent factors, namely, the height of the intravitreous gas or air and the altitude. According to the data accessed from Google Earth, many large cities (especially urban areas) have significantly lower altitudes than smaller towns (especially rural and mountain areas). Medical institutions are usually located in large cities, so some patients may be discharged from the hospital to a higher altitude or to even travel by an airplane. Thus, it is likely that such problems with intravitreous gas or air expansion exist in many ophthalmic institutes around the world.

In China, as the land area is large, different cities are located at different altitudes and have different atmospheric pressure. When patients travel from one place to another at a higher altitude, they are exposed to a higher risk of gas expansion and IOP elevation. In our institute (one of the national tertiary referral centers; Zhongshan Ophthalmic Center of Sun Yat-sen University, Guangzhou, China; altitude=10 m), approximately 30% of our patients who receive vitreoretinal surgeries come from other provinces that generally are at higher altitudes. For instance, in this study, one patient was from the city of Lhasa (altitude=3490 m), 3 were from the city of Guiyang (altitude=1277 m), and 5 were from the city of Kunming (altitude=1842 m). We can imagine that there is a risk of elevated IOP if patients with intravitreous gas or air have postoperative travel back home.

High Interobserver Agreement Enhances the App's Potential for Patients to Self-Monitor

The app can provide a reliable reference tool for patients to do self-monitoring and generate a preliminary risk report. The significant advantages of the app can be summarized into the

following aspects: acceptable accuracy of the estimation as this study revealed high interobserver agreement between the surgeons and patients; the app could serve as a physician-patient communication platform as they share the same logistics for calculating the risk of expanded intravitreous gas or air; and high acceptability for the patients as the Likert scales revealed a significantly positive attitude from the patients toward the app because it is user-friendly, and it is not difficult for patients to understand the underlying principles.

To the best of our knowledge, this study is the first one to have interobserver analyses of the height of the intravitreous fluid/gas interface independently judged by surgeons and patients. Inspired by previous studies [16,31], we designed the app from 2 perspectives: the surgeons' professional judgment with dilated fundoscopy and patients' true vision perception. The surgeon directly observed the horizontal fluid/gas interface with a preset lens through the dilated pupil and judged the vertical height of the fluid/gas interface according to anatomic landmarks. It is not difficult for surgeons to make a reliable judgment of the height; however, there are few guidelines for patients to express their personal judgments, which decreases patients' awareness and limits the efficacy of disease education. The app eliminated this barrier and imbalance with its modified design.

The app is the first one designed based on the patients' true vision perception with intravitreous gas or air. Owing to the great differences in the refraction index, the intravitreous fluid and gas or air produces vastly different vision results for the patients. Intravitreous gas located in the upper part of the vitreous cavity makes vision darker and more blurred, whereas intravitreous fluid located in the lower part of the vitreous cavity makes vision almost normal. According to the law of photorefraction through the visual axis, clear vision (through the intravitreous fluid) is in the patient's upper visual field and blurred vision (through the intravitreous gas) is located in the patient's lower visual field. Therefore, a patient can easily line out the fluid/gas interface according to the 2 different visions in his/her operated eye. We are encouraged by the high interobserver agreement and believe that the app will be a reliable reference tool for patients in case professional advice from their surgeons is unavailable. In addition, the app is worth recommending to our vitreoretinal colleagues as a supplementary tool. It is important to teach patients how to evaluate their postsurgical conditions. In the app, gas volume estimation can be done easily either by retinal surgeons or by the patients themselves. As altitude information can be objectively obtained from the internet, the accuracy of gas expansion estimation will mainly be affected by gas volume estimation. According to the results, gas volume estimation by surgeons and patients themselves achieved good agreement, which means the patients are more likely to make correct estimations of gas volume through self-evaluation with the app. It is well known that patients usually have lower visual acuity after PPV with a gas or air tamponade, when the visual axis is blurred by the gas or air. Therefore, we also evaluated whether patients with lower visual acuity (finger counting or hand movement) could accurately estimate their intravitreous gas volume. The results showed that even in these patients, their self-estimations were quite close to those of the surgeons, which means the app will be applicable for different types of patients.

Limitations and Recommendation for Future Refinement of the App/Research

There are several limitations that should be acknowledged. The sample size was relatively small, so the efficacy of the app needs to be tested in a larger population. The compensation mechanisms of the eyeball to accommodate for the expansion of intravitreous gas or air under decreased atmospheric pressure were not included in the app design, which could decrease the validity of risk estimation. The volume of the vitreous cavity is set to 4.5 ml, which may not be completely consistent in some patients, such as those with hyper-myopia. In vivo experiments on humans or animals were not performed to verify the true accuracy of the risk estimation.

Although certain limitations exist in this study, the app is a reliable reference tool for both surgeons and patients to make a preliminary estimation of intravitreous gas or air expansion and IOP elevation. Furthermore, the good interobserver agreement and user-friendly interface provide a readily available platform to create better physician-patient interaction with regard to the risk of expanded intravitreous gas or air after surgery. Further refinement of the app and in vivo experiments are warranted to provide more accurate risk estimation for surgeons and patients.

Acknowledgments

This study was partially supported by the Science and Technology Program of Guangzhou (No. 201803010022). The funding organization had no role in the design or conduct of the research.

Authors' Contributions

ZZ designed and conducted this study. FL organized the main text and raised significant advice on how to improve this study. HZ and ZM developed the app and offered a lot of support for the refinement of the app. YW partially wrote and revised the whole text. LW attended all the procedures when we conducted this study. SZ designed this study and had significant guidance on the whole research team.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The governing equations of the app.

[PDF File (Adobe PDF File), 211KB - mhealth_v7i6e14592_app1.pdf]

Multimedia Appendix 2

Video demonstration of the app's usage.

[MP4 File (MP4 Video), 44MB - mhealth_v7i6e14592_app2.mp4]

Multimedia Appendix 3

The main characteristics of all the patients.

[PDF File (Adobe PDF File), 94KB - mhealth_v7i6e14592_app3.pdf]

References

1. Gupta B, Neffendorf JE, Williamson TH. Trends and emerging patterns of practice in vitreoretinal surgery. Acta Ophthalmol 2018 Nov;96(7):e889-e890. [doi: 10.1111/aos.13102] [Medline: 27213838]
2. Ramulu PY, Do DV, Corcoran KJ, Corcoran SL, Robin AL. Use of retinal procedures in medicare beneficiaries from 1997 to 2007. Arch Ophthalmol 2010 Oct;128(10):1335-1340. [doi: 10.1001/archophthalmol.2010.224] [Medline: 20938004]
3. Neffendorf JE, Gupta B, Williamson TH. The role of intraocular gas tamponade in rhegmatogenous retinal detachment: a synthesis of the literature. Retina 2018 Sep;38(Suppl 1):S65-S72. [doi: 10.1097/IAE.0000000000002015] [Medline: 29280936]
4. Lincoff H, Coleman J, Kreissig I, Richard G, Chang S, Wilcox LM. The perfluorocarbon gases in the treatment of retinal detachment. Ophthalmology 1983 May;90(5):546-551. [doi: 10.1016/S0161-6420(83)34525-5] [Medline: 6877783]
5. Chawla HB, Birchall CH. Intravitreal air in retinal detachment surgery. Br J Ophthalmol 1973 Jan;57(1):60-70 [FREE Full text] [doi: 10.1136/bjo.57.1.60] [Medline: 4705499]
6. Parver LM, Lincoff H. Mechanics of intraocular gas. Invest Ophthalmol Vis Sci 1978 Jan;17(1):77-79. [Medline: 621133]
7. Fang IM, Huang JS. Central retinal artery occlusion caused by expansion of intraocular gas at high altitude. Am J Ophthalmol 2002 Oct;134(4):603-605. [doi: 10.1016/S0002-9394(02)01631-8] [Medline: 12383821]
8. Shiramizu KM, Okada AA, Hirakata A. Transient amaurosis associated with intraocular gas during ascending high-speed train travel. Retina 2001;21(5):528-529. [Medline: 11642385]
9. Mills MD, Devenyi RG, Lam WC, Berger AR, Beijer CD, Lam SR. An assessment of intraocular pressure rise in patients with gas-filled eyes during simulated air flight. Ophthalmology 2001 Jan;108(1):40-44. [doi: 10.1016/S0161-6420(00)00515-7] [Medline: 11150262]
10. Kokame GT, Ing MR. Intraocular gas and low-altitude air flight. Retina 1994;14(4):356-358. [doi: 10.1097/00006982-199414040-00012] [Medline: 7817030]
11. Lincoff H, Weinberger D, Reppucci V, Lincoff A. Air travel with intraocular gas. I. The mechanisms for compensation. Arch Ophthalmol 1989 Jun;107(6):902-906. [doi: 10.1001/archopht.1989.01070010929043] [Medline: 2730409]
12. Williamson T. Principles of internal tamponade. In: Vitreoretinal Surgery. Berlin, Heidelberg: Springer; 2008:61-87.
13. Rashidi N, Thomas VS, Amini R. Theoretical assessment of the risk of ocular hypotony in patients with intravitreal gas bubbles who travel through subsea tunnels. Transl Vis Sci Technol 2019 Jan;8(1):4 [FREE Full text] [doi: 10.1167/tvst.8.1.4] [Medline: 30627479]
14. Gsellman L, Amini R. Patients with intravitreal gas bubbles at risk of high intraocular pressure without exceeding elevation of surgery: theoretical analysis. Invest Ophthalmol Vis Sci 2016 Dec 1;57(7):3340-3347. [doi: 10.1167/iovs.15-18010] [Medline: 27367501]
15. Amini R, Barocas VH, Kavehpour HP, Hubschman JP. Computational simulation of altitude change-induced intraocular pressure alteration in patients with intravitreal gas bubbles. Retina 2011 Sep;31(8):1656-1663. [doi: 10.1097/IAE.0b013e318217ffa5] [Medline: 21765373]
16. Wong RF, Thompson JT. Prediction of the kinetics of disappearance of sulfur hexafluoride and perfluoropropane intraocular gas bubbles. Ophthalmology 1988 May;95(5):609-613. [doi: 10.1016/S0161-6420(88)33133-7] [Medline: 3174022]
17. Pew Research Center. 2019. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 78j663tyq]
18. Can YS, Arnrich B, Ersoy C. Stress detection in daily life scenarios using smart phones and wearable sensors: a survey. J Biomed Inform 2019 Apr;92:103139. [doi: 10.1016/j.jbi.2019.103139] [Medline: 30825538]
19. Watchirs Smith L, Guy R, Degenhardt L, Yeung A, Rissel C, Richters J, et al. Meeting sexual partners through internet sites and smartphone apps in Australia: national representative study. J Med Internet Res 2018 Dec 18;20(12):e10683 [FREE Full text] [doi: 10.2196/10683] [Medline: 30563809]

20. Flickinger TE, Ingersoll K, Swoger S, Grabowski M, Dillingham R. Secure messaging through PositiveLinks: Examination of electronic communication in a clinic-affiliated smartphone app for patients living with HIV. Telemed J E Health 2019 Mar 21. [doi: 10.1089/tmj.2018.0261] [Medline: 30900961]
21. Pung A, Fletcher SL, Gunn JM. Mobile app use by primary care patients to manage their depressive symptoms: qualitative study. J Med Internet Res 2018 Sep 27;20(9):e10035 [FREE Full text] [doi: 10.2196/10035] [Medline: 30262449]
22. Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Mar 20. [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]
23. Olson D, Lamb M, Lopez MR, Colborn K, Paniagua-Avila A, Zacarias A, et al. Performance of a mobile phone app-based participatory syndromic surveillance system for acute febrile illness and acute gastroenteritis in rural Guatemala. J Med Internet Res 2017 Dec 9;19(11):e368 [FREE Full text] [doi: 10.2196/jmir.8041] [Medline: 29122738]
24. Langley A, Mar Fan G. Teaching fibreoptic bronchoscopy using smart phones. Anaesthesia 2014 Jul;69(7):793-794 [FREE Full text] [doi: 10.1111/anae.12768] [Medline: 24917345]
25. Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e200 [FREE Full text] [doi: 10.2196/mhealth.9774] [Medline: 30578215]
26. Senior K. Smart phones: new clinical tools in oncology? Lancet Oncol 2011 May;12(5):429-430. [doi: 10.1016/S1470-2045(11)70116-4] [Medline: 21536224]
27. Zhou L, Parmanto B, Alfikri Z, Bao J. A mobile app for assisting users to make informed selections in security settings for protecting personal health data: development and feasibility study. JMIR Mhealth Uhealth 2018 Dec 11;6(12):e11210 [FREE Full text] [doi: 10.2196/11210] [Medline: 30538088]
28. King D, Lee H, Darzi A. Abandoned HealthSpace. Smart phones may help. Br Med J 2010 Dec 20;341:c7216. [doi: 10.1136/bmj.c7216] [Medline: 21172982]
29. Quinn CC, Swasey KK, Crabbe JC, Shardell MD, Terrin ML, Barr EA, et al. The impact of a mobile diabetes health intervention on diabetes distress and depression among adults: secondary analysis of a cluster randomized controlled trial. JMIR Mhealth Uhealth 2017 Dec 7;5(12):e183 [FREE Full text] [doi: 10.2196/mhealth.8910] [Medline: 29217502]
30. Zhang ZT, Zhang SC, Huang XG, Liang LY. A pilot trial of the iPad tablet computer as a portable device for visual acuity testing. J Telemed Telecare 2013 Jan;19(1):55-59. [doi: 10.1177/1357633X12474964] [Medline: 23434538]
31. Lincoff H. Transient amaurosis associated with intraocular gas during ascending high-speed train travel. Retina 2002 Apr;22(2):240-241. [Medline: 11927867]
32. Aronowitz JD, Brubaker RF. Effect of intraocular gas on intraocular pressure. Arch Ophthalmol 1976 Jul;94(7):1191-1196. [doi: 10.1001/archopht.1976.03910040099017] [Medline: 938297]
33. Williamson TH, Guillemaut JY, Hall SK, Hutter JC, Goddard T. Theoretical gas concentrations achieving 100% fill of the vitreous cavity in the postoperative period: a gas eye model study. Retina 2018 Sep;38(Suppl 1):S60-S64. [doi: 10.1097/IAE.0000000000001963] [Medline: 29232331]
34. Brosh K, Strassman I, Seelenfreund M. High intraocular pressure in four vitrectomized eyes with intravitreal C3F8 without high altitude travel. Eye (Lond) 2014 Jul;28(7):892-894 [FREE Full text] [doi: 10.1038/eye.2014.83] [Medline: 24788015]
35. Muzychuk AK, Adatia FA, Ford BA, Kherani AM. Commercial air travel with a small intravitreous gas bubble. Arch Ophthalmol 2011 Jun;129(6):811-813. [doi: 10.1001/archophthalmol.2011.144] [Medline: 21670355]
36. Noble J, Kanchanaranya N, Devenyi RG, Lam WC. Evaluating the safety of air travel for patients with scleral buckles and small volumes of intraocular gas. Br J Ophthalmol 2014 Sep;98(9):1226-1229. [doi: 10.1136/bjophthalmol-2014-304923] [Medline: 24782473]
37. Dieckert JP, O'Connor PS, Schacklett DE, Tredici TJ, Lambert HM, Fanton JW, et al. Air travel and intraocular gas. Ophthalmology 1986 May;93(5):642-645. [Medline: 3725322]

Abbreviations

IOP: intraocular pressure
PPV: pars plana vitrectomy

Edited by G Eysenbach; submitted 04.05.19; peer-reviewed by B Smith, Y Zhu; comments to author 27.05.19; revised version received 30.05.19; accepted 12.06.19; published 26.06.19.

Please cite as:
Zhang Z, Li F, Zhang H, Miao Z, Wei Y, Wang L, Zhang S
Development and Testing of a Mobile Phone App for Risk Estimation of Gas Volume Expansion and Intraocular Pressure Elevation in Patients With Intravitreous Gas or Air Tamponade: Interobserver Assessment Study
JMIR Mhealth Uhealth 2019;7(6):e14592
URL: http://mhealth.jmir.org/2019/6/e14592/
doi:10.2196/14592
PMID:31244482

Original Paper

Apps to Support Self-Management for People With Hypertension: Content Analysis

Chi Yan Hui[1], BEng, MSc, PhD; Emily Creamer[2], BSc, MSc, PhD; Hilary Pinnock[1], MB, ChB, MRCGP, MD; Brian McKinstry[2], FRCPE, FRCGP, MD

[1]Asthma UK Centre for Applied Research, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, United Kingdom

[2]Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Brian McKinstry, FRCPE, FRCGP, MD
Usher Institute of Population Health Sciences and Informatics
The University of Edinburgh
Number 9 Bioquarter
9 Little France Rd
Edinburgh, EH16 4UX
United Kingdom
Phone: 44 131 650 2683
Email: Brian.mckinstry@ed.ac.uk

Abstract

Background: Home blood pressure monitoring (HBPM) is one component of effective supported self-management, which may potentially be mediated by mobile apps.

Objective: The aim of this study was to identify the self-management features (HBPM and broader support strategies) offered by currently available apps and to determine the features associated with download frequency and user ratings.

Methods: We searched Google Play store, Apple App store, National Health Services Apps Library and myhealthapps.net (first search on February 1, 2018; updated August 18, 2018). We included high blood pressure apps available in the United Kingdom and extracted their features, number of downloads, and the average users' rating from the app stores. We mapped the features to the holistic Practical Reviews In Self-Management Support (PRISMS) taxonomy of self-management support. We employed a regression analysis to determine if any features were associated with download frequency or user rating.

Results: We included 151 apps. The 3 most common features were as follows: monitoring blood pressure (BP) and charting logs; lifestyle (exercise or dietary) advice; and providing information about hypertension. The other 11 components of the PRISMS taxonomy were rarely featured. There was little evidence to support associations between specific features and the download statistics and rating scores, with only 2 uncommon features achieving borderline significant associations. The presence of social support features, such as a forum, was weakly but significantly (R^2=.04, P=.02) correlated with the number of downloads. Apps designed specifically for particular BP monitors/smart watches were weakly associated with a higher rating score (R^2=.05, P<.001). Apps with more ratings were associated with more downloads (R^2=.91, P<.001).

Conclusions: The functionality of currently available apps is limited to logging BP, offering lifestyle advice, and providing information about hypertension. Future app development should consider broadening the remit to produce a system that can respond flexibly to the diversity of support that enables people to self-manage their hypertension.

(JMIR Mhealth Uhealth 2019;7(6):e13257) doi:10.2196/13257

KEYWORDS

hypertension; self-management; telehealth; telemedicine; mobile app

Introduction

High blood pressure (HBP) currently affects 1.13 billion people worldwide [1], and this is expected to rise to 1.5 billion by 2025 [2]. Globally, hypertension contributes to 9.4 million deaths annually [3], mainly from heart disease and stroke [4]. Identifying and controlling hypertension significantly reduces cardiovascular events [5]; yet, despite effective medication, blood pressure (BP) remains to be poorly controlled [6-8], at least in part owing to poor adherence by patients and reluctance to intensify medication by clinicians [9].

Supported self-management can improve control [10] with home BP monitoring (HBPM)—a widely used component of effective interventions. This not only avoids repeated visits to health care professionals for monitoring but also improves accuracy by allowing detection of *white-coat* and *masked* hypertension, which are common in both untreated and treated patients and result in misleading clinic BP readings [11]. HBPM can increase patient engagement in the management of their condition and provide a trusted basis for shared management decisions, including medication changes [12]. Furthermore, one mechanism by which HBPM works is by bridging the gap between patients' perceptions of hypertension as a disease with multiple symptoms compared with the professional view of an asymptomatic condition, fostering confidence in the ability to self-manage [10]. There is strong evidence that when BP readings are sent electronically to and dealt with by clinicians, the outcomes are improved [13].

The development of affordable home BP monitors [14] and increasing ownership of smartphones among the general population [15] present opportunities for mobile apps to support patients' self-management of their hypertension [16]. Many hypertension apps are available in the market [17]. However, there has been considerable concern voiced over the highly variable quality and appropriateness of these apps [18].

Self-monitoring, however, is only 1 component of supported self-management. The Practical Systematic Review of Self-Management Support (PRISMS) taxonomy describes a broader approach to supporting people to proactively self-manage their condition, offering a picklist of 14 potential support strategies associated with effective self-management interventions [19]. The purpose of this study was to map the landscape of hypertension apps that are currently available and assess the extent to which these apps deliver recognized components of self-management support (PRISMS) and whether there is a correlation between the app's self-management features and their number of downloads and user ratings.

Methods

Search Strategy

We searched, identified, and screened apps on Google Play store [20], Apple iTunes preview website (using Google site search code of site:itunes.apple.com/gb/app), National Health Services (NHS) Apps Library [21], and *myhealthapps.net* [22] on February 1, 2018 (updated August 18, 2018), using the key term, *high blood pressure*. The search was restricted to the apps available in the United Kingdom.

Screening and Data Extraction

We used the Google Web Scraper (version 0.3.7) [23] to extract the app name, description, number of raters, ratings, number of downloads (not available for iOS apps), version, latest updated date, cost, and the developers' name. A reviewer (CYH) screened the extracted data using the following inclusion/exclusion criteria:

- *Inclusion criteria:* The app (1) advertised features relevant to management of hypertension, (2) was available for download through the official Android or Apple App stores/NHS Apps Library/myhealthapps.net, and (3) was written in English.
- *Exclusion criteria:* The app (1) could not be opened and/or used owing to technical problems after 3 attempts, (2) had not been updated within 1 year of the search date (ie, the last update was before February 1, 2016) to exclude those apps that were no longer being supported, (3) was designed for children, and (4) had been in the market for greater than 1 month but had no ratings and/or reviews to reduce the chance of including apps that were still under development.

A reviewer (CYH) reviewed the description, screenshots, and the developer website of the included apps and extracted their key app features using a piloted data extraction sheet under the headings of the 14 components of the PRISMS taxonomy (Textbox 1). Apps without clear information were downloaded and the features were assessed manually.

Data Analysis

App Self-Management Features

We used the PRISMS taxonomy (see Textbox 1) [19] to categorize the extracted app features. We searched for the presence of each of these 14 items within each included app, dichotomizing the outcome as *present* (=1) or *absent* (=0). We counted the total number of apps addressing the 14 components of the taxonomy.

Association Between the Practical Reviews in Self-Management Support Self-Management Components, Downloads, and Ratings

We performed a regression analysis in Microsoft Excel. We wanted to investigate the features [24] that may influence people to download and rate a BP app. Apps with apparently 500,000,000 or more recorded downloads were excluded in the regression analysis. We decided to adopt this approach to reduce the possibility of including apps that were no longer supported by developers. In addition, 1 app seemed to have such unlikely download statistics (exceeding popular apps, such as YouTube, TikTok, Pokemon GO, Snapchat [50,000,000 to 500,000,000 downloads]) that we decided to exclude it.

Textbox 1. Items included in the Practical Reviews in Self-Management Support (PRISMS) taxonomy of self-management support.

A.	Information about condition and/or its management
B.	Information about available resources
C.	Provision of/agreement on specific clinical action plans and/or rescue medication
D.	Regular clinical review
E.	Monitoring of condition with feedback
F.	Practical support with adherence (medication or behavioral)
G.	Provision of equipment
H.	Provision of easy access to advice or support when needed
I.	Training/rehearsal to communicate with health care professionals
J.	Training/rehearsal for everyday activities
K.	Training/rehearsal for practical self-management activities
L.	Training/rehearsal for psychological strategies
M.	Social support
N.	Lifestyle advice and support

The number of downloads and the mean rating score were the outcome variables that we used in the regression analysis. The app's information on the download page (rating, number of raters, cost, multiple conditions' focus, the number of features, whether the app was developed by an internationally known company, presented as part of an established clinical program or claimed to be recommended by health care professionals, supported by a patient or professional organizations, or promoted by a campaign) were used as additional predictor variables as these features are known to influence downloads more generally [25]. Both outcome and predictor variables were standardized by using an Excel standardized function [26]. A simple linear regression analysis was performed to determine the relationship between independent variables and the downloads and ratings followed by multiple regression. We considered significance at the 5% level. All variables were entered into the multiple regression. The goodness-of-fit was assessed using R-squared statistics.

Interpretation

We discussed the results of the data synthesis within a multidisciplinary team of 16 people, which included expertise in primary care (n=1), hypertension (n=2), and electronic health and digital technology (n=13). The features that were associated with the downloads and user rating score were of key interest to the multidisciplinary team and the relevance of particular aspects of the PRISMS taxonomy in relation to BP apps. There was considerable discussion about the relevance of metrics such as *number of reviewers* as a predictor of quality. The results were used to guide the design of a prototype hypertension app.

Results

Characteristics of Included Apps

A total of 151 apps were identified. Most were identified in Google Play and the Apple app store and a much smaller number from the NHS Apps Library and myhealthapp.net. The apps identified, the screening process, and the final numbers of the apps included are detailed in the flowchart (Figure 1).

Of the 151 apps, 95.4% (144/151) apps were available on both Google Play and Apple App store, 4.6% (7/151) apps were only available on the Apple App store [27-33]. Furthermore, 51.0% (77/151) apps focused solely on HBP and 47.7% (74/155) addressed multiple clinical conditions including diabetes, heart disease, respiratory disease, kidney disease, cancer, and psychological conditions. Also, 6.6% (10/151) apps were presented as part of an established clinical program, and 1 app advertised on the app store claimed that it was recommended by doctors and pharmacists. Only 5.2% (8/155) apps charged users (£0.9 to £7.90). Furthermore, 3 apps were supported by patient or professional organizations (the American Heart Association, The George Institute for Global Health, and American College of Cardiology Foundation); 1 app was designed for a campaign (Action for Happiness); 3.2% (5/155) apps were developed by BP monitor/smart watch companies (Samsung, Beurer, Braun, Omron, and Withings); 1 app was developed by an insurance company (AXA); and 28 apps were designed to be connected with a particular brand of BP monitor or smart watch.

Apps Self-Management Features

Table 1 summarizes the application features related to the 14 components of the PRISMS taxonomy [19]. Details are shown in the Multimedia Appendix 1.

Figure 1. Flowchart of app selection and screening process.

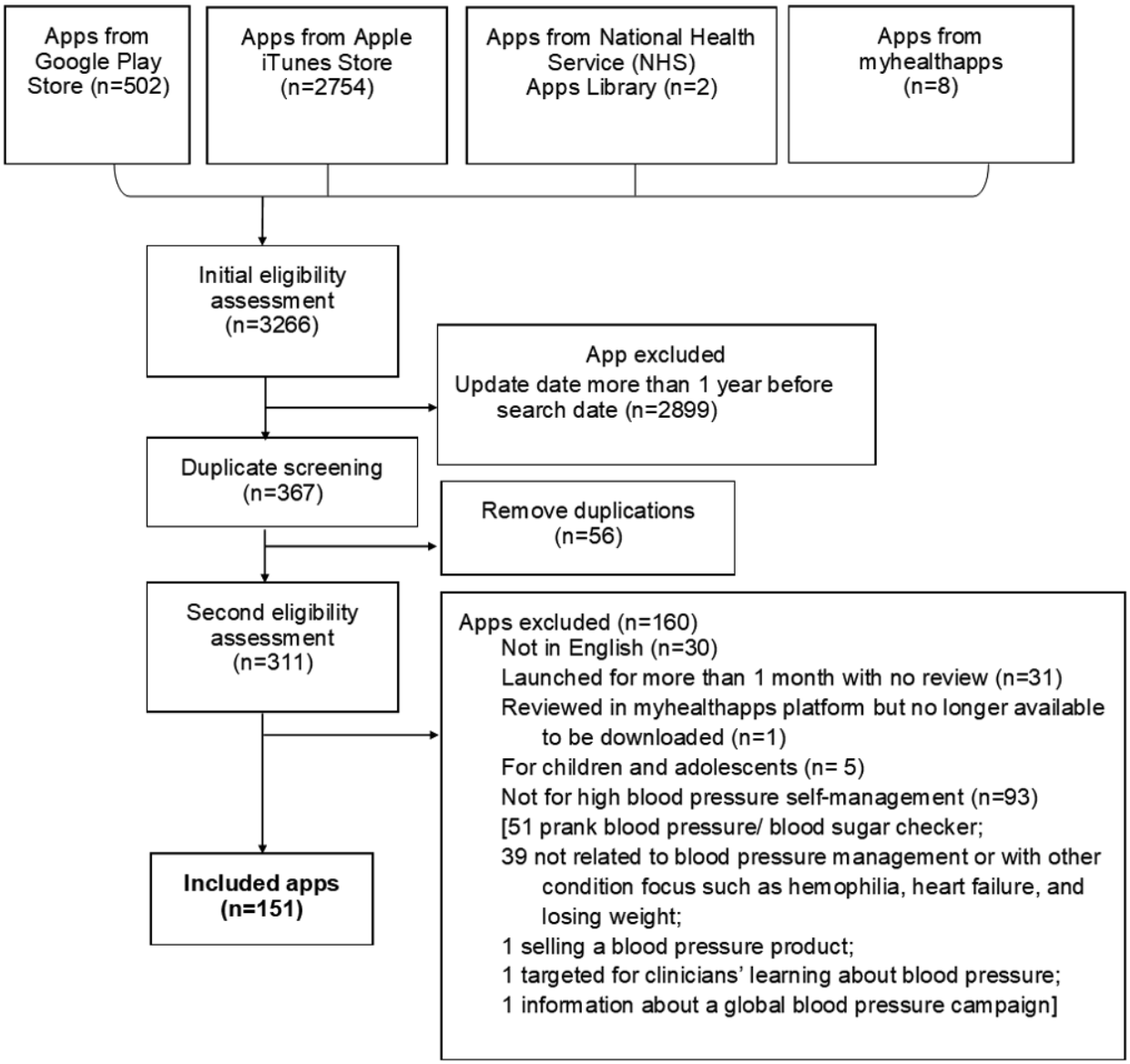


Table 1. Application features categorized by the Practical Reviews in Self-Management Support (PRISMS) taxonomy.

Taxonomy item and app features	Apps, n
Information about condition and/or its management[a]	
News about blood pressure (BP), diet advice, what qualifies as high BP, home monitoring and diagnosis of BP, specific advice for BP during pregnancy	30
Information about available resources	
Providing links to WebMD, American Heart Association, NHS[b] information pages, and other associations	2
Provision of/agreement on specific clinical action plans and/or rescue medication	
Action plan (eg, personalized agreement on what to do if BP is high/low)	1
Regular clinical review	
Manual input of appointment details (eg, date and location)	2
Booking appointments	3
Monitoring of condition with feedback[a]	
Logging systolic/diastolic BP, heart rate, medication taken, and free-text notes	70
Nonvalidated and likely misleading *measuring* of BP (via camera)	3
Nonvalidated and likely misleading *measuring* of BP (via fingerprint)	3
Voice assistant to support logging activities such as making and sending logs	3
Sending logs (xls/db/PDF) via email; saved in Dropbox, Google Drive, SD card); share logs via portal	58
Sending logs by using the app portal, SMS[c], email, or via an automated telephone call; data transferred to GP[d] systems such as Egton Medical Information Systems Web and SystemOne	2
Showing logs on charts (pie chart, line chart) only	37
Providing observations such as *your BP hasn't varied a lot in the last months*, *your BP is normal*, *you are stage 2 hypertension*.	8
Providing simple advice and suggesting actions such as *condition: stage 2 hypertension; to do: you may have high blood pressure, change your lifestyle - see your doctor, aim for a total blood pressure less than 120/80 mmHg, avoid Tobacco, etc*, *hypertensive urgency - please go to the hospital!*, *your blood pressure is dangerously high. See your GP urgently or call 111 for more advice*.	7
Showing logs on colored chart/bar/points to indicate hypertension (stage 1), hypertension, prehypertension, normal and low BP	30
Showing the person's BP at different times of the day in different places	1
Showing the number of days that meet the preset optimized BP value	1
Practical support with adherence (medication or behavioral)	
Medication reminder	11
Points/rewards to encourage logging	5
Provision of equipment	
Use the proprietary or Bluetooth smart meter to log systolic/diastolic/pulse/heart rate	19
Connecting to activity tracker gadgets or smart watches	16
Provision of easy access to advice or support when needed	
Web doctor consultation	1
Communication via texting	1
Training/rehearsal to communicate with health care professionals	
Personal health record accessible to health care professional to enhance discussions or providing patient access to health record	6
Training/rehearsal for everyday activities	
Not applicable	0
Training/rehearsal for practical self-management activities	
Tagging logs to help users to understand what causes their high BP (eg, add tag *before eating*)	3
Logging reminder to train users to make regular logs for monitoring	10

Taxonomy item and app features	Apps, n
Entering personal health profile, preparing for emergency conditions	2
Ordering repeat prescriptions	1
Emergency call to GP	1
Training/rehearsal for psychological strategies	
Using breathing exercises to relieve stress to lower BP	5
Social support	
Forum to share stories or advice about BP management	3
Lifestyle advice and support[a]	
Providing DASH[e]/low sodium diet advice/videos to lower BP, advice on exercise intensity, lifestyle factors that cause high BP, managing stress, breathing exercises, relaxing games to lower BP	54
Importing and showing weight data and sleep duration from another app (*Weight Companion*, *Sleep as Android*, *MyFitnessPal*, *Google Fit*, *Samsung Health*)/other third-party apps	7
Sending messages to users via WhatsApp and Skype	1
Log exercise intensity, BMI, and diet - make a note of the measurements, which may be associated to the HBP[f]	10
Activity tracking (bicycle/steps) via Fitbit alongside blood pressure monitoring	20

[a]Top *3* common BP self-management supports found *in* the apps.

[b]NHS: National Health Services.

[c]SMS: short message service.

[d]GP: general practitioner.

[e]DASH: Dietary Approaches to Stop Hypertension.

[f]HBP: high blood pressure.

Monitoring of Blood Pressure With Feedback

The most common self-management support provided by the apps was *monitoring of condition with feedback* (70/151, 46.4%), though only 1 UK app provided feedback integrated with current National Health Service (NHS) general practice systems such as Egton Medical Information Systems and SystemOne [34]. Monitoring involved logging systolic/diastolic BP and heart rate. Most apps (58/70, 83%) enabled sharing of logs with health care professionals, and all of them provided feedback/advice. Advice popped up immediately after people entered their BP reading as opposed to popping up when the BP breached a certain threshold. Most of the feedback took the form of simply displaying the user's BP logs back to them in a chart format (37/70, 53%). However, 9 of the 70 apps (13%) provided an additional basic evaluation of the recorded BP value, such as *your BP is normal*, and 6 apps (9%) coupled this type of evaluative feedback with advice on actions to take, such as (to a patient with a BP of 160/110) *You are hypertensive Urgency [sic]—please go to the hospital!* Furthermore, 1 app showed the patient's BP values at different times of the day and in different places [35]. In addition, 1 app showed the number of days that met the preset optimized BP value [36].

Lifestyle Advice and Support

The second most common category of self-management support was *Lifestyle advice and support* (54/151, 35.8%). Examples include DASH (Dietary Approaches to Stop Hypertension) [37] and suggestions for low-sodium cooking recipes to help lower BP. The recipes were not targeted to individuals, for example, general dietary advice was not provided in relation to the monitored weight data.

Information

The third most common self-management support was *Information about condition and/or its management* (30/151, 19.9%). The apps provided information such as news items about hypertension, diagnostic criteria for HBP, how to perform home monitoring, and specific advice for BP during pregnancy. Patients could choose what information they wanted to read and follow.

Other Less Common Features

One app that was tailored for patients with multiple conditions provided an action plan template on the app. The template was a free text template in which the patient should enter the personalized agreement with their clinician on what to do if BP was high or low and assigned a color (green, amber, and red) for the specific action. Furthermore, 1 app (incurring a charge) offered 24/7 Web consultation with a third-party health care specialist team. Another app allowed patients to text the specialist team for enquiries. None of the apps were found to provide everyday activity training support.

Features Associated With Downloads

A total of 144 apps had the number of downloads displayed on the Google Play store. Of these, the number of downloads ranged from 1 to 500,000,000. These numbers provided snapshots on the included app's adoption rate. Furthermore, 90.3% (130/144) apps had fewer than 500,000 downloads, 9.0% (13/144) apps had 500,000 to 50,000,000 downloads, 1 app was

an outlier with more than 500,000,000 downloads and was excluded in the regression analysis. Moreover, 2 apps included a service in which health care professionals gave feedback to patients. The number of downloads was available for one of the apps that was available on the Google Play store, and it did not attract a high number of downloads (number of downloads=100). However, it is not clear how restricted the recruitment was for this app. There were no missing data in the analysis. The results of the simple linear regression analysis are shown in Table 2. In addition, 1 feature (the inclusion of social support features, such as a forum) was weakly correlated but achieved statistical significance (R^2=.04, P=.02). Perhaps, unsurprisingly, the number of raters on the app stores (R^2=.91, P<.001) was associated with the number of downloads. A multiple regression analysis showed significant impact from a combination of the variables (R^2=.92, P<.001), of which *the number of raters* was the strongest contributor among the variables (coefficient=.96, P<.001). Other variables had coefficients of .01 or .001 with a P value that was more than .1. The details are shown in Multimedia Appendix 2.

Features Associated With Average Rating

Of the 148 apps, the average rating ranged from 0 to 5 (out of 5). Furthermore, 113 ratings were from the Google Play store, 35 ratings were from the Apple app store. In addition, 14 apps had an average rating of 5, 12 of them were rated 0. These apps had been available on the app markets for less than a month. Moreover, 2 apps in which health care professionals gave feedback were not associated with high overall ratings (one scored 0 and another 2.5). There were no missing data in the regression analysis. There was little or no association between the presence of PRISMS features and the number of downloads.

Table 2. R^2 and P values of a simple regression analysis at 5% significance level for downloads (Data were sorted by the P values in ascending order).

Variables	Standardized coefficient	R^2	P value
PRISMS[a] features			
Social support[b]	0.19	0.04	.02
Information about condition and/or its management	0.13	0.02	.12
Lifestyle advice and support	–0.08	0.01	.34
Training/rehearsal for practical self-management activities	0.05	<0.01	.56
Training/rehearsal for psychological strategies	–0.04	<0.01	.61
Information about available resources	–0.03	<0.01	.72
Regular clinical review	–0.03	<0.01	.76
Practical support with adherence (medication or behavioral)	–0.03	<0.01	.76
Training/rehearsal to communicate with health care professionals	–0.02	<0.01	.79
Provision of equipment	–0.02	<0.01	.81
Provision of/agreement on specific clinical action plans and/or rescue medication	–0.02	<0.01	.84
Monitoring of condition with feedback	–0.01	<0.01	.87
Provision of easy access to advice or support when needed	–0.01	<0.01	.91
Training/rehearsal for everyday activities	N/A[c]	N/A	N/A
Other possible associated factors			
Number of raters[b]	0.95	0.91	<.001
Rating	0.06	<0.01	.48
Paid	–0.05	<0.01	.57
Supported by associations or apps for a campaign	–0.04	<0.01	.65
Apps for multiple conditions	–0.03	<0.01	.72
Recommended by health care professionals/use in the NHS practices	–0.02	<0.01	.8
Numbers of features	0.01	<0.01	.89
Created by internationally BP[d] monitor/smart watch company	0.01	<0.01	.9

[a]PRISMS: Practical Reviews In Self-Management Support.

[b]The independent variables with P values <.05, which were assessed as statistically significant.

[c]N/A: not applicable.

[d]BP: blood pressure.

Table 3. R^2 and *P* values of a simple regression analysis at 5% significance level for user ratings (Data were sorted by the *P* values in ascending order).

PRISMS[a] self-management features	Standardized coefficient	R^2	*P* value
Provision of equipment[b]	–0.23	0.05	.01
Monitoring of condition with feedback	–0.16	0.02	.07
Information about condition and/or its management	0.14	0.02	.13
Training/rehearsal for psychological strategies	0.13	0.01	.15
Training/rehearsal to communicate with health care professionals	0.08	<0.01	.41
Social support	–0.07	<0.01	.42
Practical support with adherence (medication or behavioral)	–0.06	<0.01	.50
Provision of easy access to advice or support when needed	0.06	<0.01	.51
Training/rehearsal for practical self-management activities	0.04	<0.01	.64
Lifestyle advice and support	–0.03	<0.01	.72
Information about available resources	–0.03	<0.01	.73
Regular clinical review	0.02	<0.01	.83
Provision of/agreement on specific clinical action plans and/or rescue medication	0.01	<0.01	.87
Training/rehearsal for everyday activities	N/A[c]	N/A	N/A

[a]PRISMS: Practical Reviews in Self-Management Support.

[b]The independent variable with *P* values <.05, which were assessed as statistically significant.

[c]N/A: not applicable.

The linear regression analysis in Table 3 showed that apps designed specifically for particular BP monitors or smart watches were weakly associated with having higher user ratings (R^2=.05, *P*<.001). Other PRISMS features [19] were not significantly correlated with better reviews. Multiple regression showed there was no significant association between a combination of the features (Multimedia Appendix 2).

Discussion

Principal Findings

Current BP apps focus on monitoring with generic feedback, lifestyle advice, and information about the condition. Monitoring and feedback typically involve charts and summary logs. Only 7 apps provide specific advice about what to do if the BP is high or low, and of these, only 1 provided advice based on a clinical action plan agreed with a health care professional. Lifestyle advice was mainly dietary advice, advising people to reduce sodium intake to lower BP, or advice on exercise intensity, weight loss, and stress management. None of them provided any evidence for these interventions. The other 11 components of the PRISMS taxonomy were rarely featured in the apps we reviewed.

There was no evidence to support associations between specific features and the *popularity* of the app as implied by download statistics and rating scores, with only 2 features achieving borderline significant associations. Apps with more raters were associated with more downloads; however, a more frequently downloaded app will have a much larger potential pool to provide reviewers, and therefore, the direction of causation is unclear.

Strengths and Limitations

This study maps the app features onto the PRISMS taxonomy, a theoretically based list of components derived from a systematic metareview of 969 trials and 30 qualitative studies of self-management support interventions [19]. We initially performed a search in February 2018 but updated this in August 2018. Despite detecting a large number of new apps, some emerging apps may be missed in the study.

There are some methodological limitations. First, although we investigated the relationship between the app features and the number of downloads and the user rating in the app store, there are many other variables, such as the ranking in the search results in the app store or Google search engine or recommendations from personal physicians, friends, and relatives, which may also influence downloads and user ratings. Second, for resource reasons, the data extraction and review of the apps' content was completed by 1 person. However, the data were extracted by machine code to reduce the risk of human error in data extraction. The results were discussed in a multidisciplinary team to provide a broader interpretation of the results. Third, the rating in the Google Play app store is the sum of the average rating in each version of the app, whereas the rating in the Apple app store is the sum of the app rating over the total number of raters. We did not adjust the user ratings in the regression analysis as the average rating of each version was not available directly on the Google Play store. A yearly follow-up data extraction is needed to calculate this information. However, our review gives a snapshot of the current BP apps in the market. Fourth, the number of downloads and users' rating reflected user perceptions rather than clinical effectiveness or utility of the apps, nor did these measures provide information on the sustained use of those apps. Nevertheless, it provides a

consumer's view to inform app developers about possible desirable features to be developed. Fifth, we excluded those apps without continued support from their developers as a yearly analysis is needed to review those apps. We think it is important to focus the review on those that have survived the market.

Interpretation in Relation to the Published Literature

Little appears to have changed in the 2 years since Jamaladin [38] conducted the content analysis on BP apps in 2016. Logging of BP remains the most common feature in BP apps, with feedback limited to providing graphical records of logs. Research involving people with different long-term conditions shows that people want clinical and customized advice on what to do when their condition is getting worse (as published in Hui et al [39] and the study by Mendiola et al [17]). App developers' reluctance to develop apps with customized advice may relate to concerns about potential medicolegal consequences and/or the additional regulatory burden involved in meeting medical device legislation requirements [40].

More remarkable is the very narrow focus of the BP apps. Our use of the PRISMS taxonomy highlighted how rarely most of the components of self-management support were provided by the apps we reviewed. The taxonomy is described as a *tick list* and not a *check list* as different components will be relevant in different clinical and social contexts and to people with different preferences. Nevertheless, it would seem that the technology market has not yet recognized the potential breadth of self-management support and produced an app that can flexibly provide a broad range of features to support people to live with their hypertension.

We found that a greater number of raters were associated with higher downloads. This may seem obvious as a higher number of users means that there are more consumers to rate an app. However, there is evidence that online shopping customers tend to purchase products with a large number of reviews [41]. Furthermore, there is evidence that users who do not engage with the qualitative detail of app reviews tend to be greatly influenced by the *quantity* of raters on the download page [25]. Downloads of apps linked to BP monitors and smart watches may be driven by sales of the peripherals rather than the app. In addition, the device is likely to be supported by a helpline, which may be a valuable feature for people less familiar with app technology—potentially the older patients with hypertension (Hui et al [39]).

Conclusions

The number of apps for people with hypertension is increasing, but their functionality remains static, typically limited to logging BP, offering lifestyle advice, and providing information about hypertension. There is no clear evidence that these or any specific features of apps are associated with increased downloads or improved ratings.

We suggest app developers use an evidence-based framework such as PRISMS to ensure that their apps include the breadth of components that provide support for self-management. In addition, it is important to review the app's usage iteratively and collect feedback from users (both patients and professionals) to continuously improve features. Ideally, apps should be easy to use, provide immediate feedback on the BP reading possibly related to a clinician-recommended action plan, warn of dangerous BPs, provide access to information about self-management of BP and how to measure it properly, allow users to opt for reminders to check BP and to take medications, and to choose to communicate readings with professionals.

The challenge for app developers is to look outside their comfort zone of features that are technologically straightforward and to produce a flexible app that connects with health/social/community services to enable it to respond to the known breadth and diversity of support that will enable people to live with and self-manage their hypertension.

Devices that provide advice based on received readings are classed as medical devices and must pass stringent and expensive evaluations to satisfy the Medicines and Health care Products Regulatory Agency before they can be sold in the European Union. In theory, even small subsequent changes in an app (common in nonmedical apps) require recertification. In addition, some companies are concerned about the possibility of being held liable for any harm, which may befall patients using the app. These concerns have led to a reluctance in app developers to go beyond simple monitoring to provide *action plans*. Legislative and regulatory changes may be required to stimulate the development of more intelligent medical apps.

Acknowledgments

We are grateful to RP for his advice on the statistics presentation. This work is funded by European Institute of Innovation & Technology Digital.

Authors' Contributions

CYH undertook the data extraction and synthesized the data. RP was the statistical advisor. CYH with BM wrote the initial draft of the manuscript. All authors contributed to the writing of the paper and reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Application features related to the 14 components of the PRISMS taxonomy.

[XLSX File (Microsoft Excel File), 46KB - mhealth_v7i6e13257_app1.xlsx]

Multimedia Appendix 2

Multiple regression models.

[DOCX File, 17KB - mhealth_v7i6e13257_app2.docx]

References

1. World Health Organization. 2018. Global Health Observatory (GHO) data URL: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence/en/ [WebCite Cache ID 74rCA2NoC]
2. Public Health England. NHS guidance on health matters: combating high blood pressure URL: https://www.gov.uk/government/publications/health-matters-combating-high-blood-pressure/health-matters-combating-high-blood-pressure [accessed 2018-12-22] [WebCite Cache ID 74rCUs3g6]
3. Angell SY, De Cock KM, Frieden TR. A public health approach to global management of hypertension. Lancet 2015 Feb 28;385(9970):825-827 [FREE Full text] [doi: 10.1016/S0140-6736(14)62256-X] [Medline: 25752181]
4. World Health Organization. 2013. A global brief on hypertension: silent killer, global public health crisis URL: https://apps.who.int/iris/handle/10665/79059 [accessed 2018-12-22] [WebCite Cache ID 74rCgHLMJ]
5. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. Lancet 2016 Mar 05;387(10022):957-967. [doi: 10.1016/S0140-6736(15)01225-8] [Medline: 26724178]
6. Kearney PM, Whelton M, Reynolds K, Whelton PK, He J. Worldwide prevalence of hypertension: a systematic review. J Hypertens 2004 Jan;22(1):11-19. [Medline: 15106785]
7. Cottrell E, Chambers R, O'Connell P. Using simple telehealth in primary care to reduce blood pressure: a service evaluation. BMJ Open 2012;2(6):- [FREE Full text] [doi: 10.1136/bmjopen-2012-001391] [Medline: 23117563]
8. Falaschetti E, Mindell J, Knott C, Poulter N. Hypertension management in England: a serial cross-sectional study from 1994 to 2011. Lancet 2014 May 31;383(9932):1912-1919. [doi: 10.1016/S0140-6736(14)60688-7] [Medline: 24881995]
9. Elliott WJ. What factors contribute to the inadequate control of elevated blood pressure? J Clin Hypertens (Greenwich) 2008 Jan;10(1 Suppl 1):20-26 [FREE Full text] [Medline: 18174780]
10. Shahaj O, Denneny D, Schwappach A, Pearce G, Epiphaniou E, Parke HL, et al. Supporting self-management for people with hypertension: a meta-review of quantitative and qualitative systematic reviews. J Hypertens 2019 Feb;37(2):264-279. [doi: 10.1097/HJH.0000000000001867] [Medline: 30020240]
11. Stergiou GS, Asayama K, Thijs L, Kollias A, Niiranen TJ, Hozawa A, International Database on HOme blood pressure in relation to Cardiovascular Outcome (IDHOCO) Investigators. Prognosis of white-coat and masked hypertension: International Database of HOme blood pressure in relation to Cardiovascular Outcome. Hypertension 2014 Apr;63(4):675-682. [doi: 10.1161/HYPERTENSIONAHA.113.02741] [Medline: 24420553]
12. Hanley J, Ure J, Pagliari C, Sheikh A, McKinstry B. Experiences of patients and professionals participating in the HITS home blood pressure telemonitoring trial: a qualitative study. BMJ Open 2013;3(5):- [FREE Full text] [doi: 10.1136/bmjopen-2013-002671] [Medline: 23793649]
13. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: a systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
14. Sheppard JP, Schwartz CL, Tucker KL, McManus RJ. Modern management and diagnosis of hypertension in the United Kingdom: home care and self-care. Ann Glob Health 2016;82(2):274-287 [FREE Full text] [doi: 10.1016/j.aogh.2016.02.005] [Medline: 27372531]
15. Deloitte. 2017. UK public are 'glued to smartphones' as device adoption reaches new heights URL: https://www2.deloitte.com/uk/en/pages/press-releases/articles/uk-public-glued-to-smartphones.html [WebCite Cache ID 74rDBXsie]
16. Kang H, Park H. A mobile app for hypertension management based on clinical practice guidelines: development and deployment. JMIR Mhealth Uhealth 2016 Feb 02;4(1):e12 [FREE Full text] [doi: 10.2196/mhealth.4966] [Medline: 26839283]
17. 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] [Medline: 25972309]
18. Armstrong S. Which app should I use? Br Med J 2015 Sep 09;351:h4597. [doi: 10.1136/bmj.h4597] [Medline: 26353800]
19. Taylor SJC, Pinnock H, Epiphaniou E, Pearce G, Parke HL, Schwappach A, et al. NCBI. Southampton, UK: NIHR Journals Library; 2014. A rapid synthesis of the evidence on interventions supporting self-management for people with long-term conditions: PRISMS? Practical systematic Review of Self-Management Support for long-term conditions URL: https://www.ncbi.nlm.nih.gov/pubmed/25642548 [WebCite Cache ID 74rCtGGsg]
20. Google Play. URL: https://play.google.com/store?hl=en [WebCite Cache ID 74rDJQg9o]
21. National Health Services. URL: https://apps.beta.nhs.uk/ [accessed 2018-12-22] [WebCite Cache ID 74rDOpc6e]

22. myhealthapps. URL: http://myhealthapps.net/ [accessed 2018-12-22] [WebCite Cache ID 74rDTSTUW]
23. Web Scraper. URL: https://www.webscraper.io/ [accessed 2018-12-22] [WebCite Cache ID 74rDZjLXq]
24. Microsoft Office. Use the Analysis ToolPak to perform complex data analysis URL: https://support.office.com/en-us/article/use-the-analysis-toolpak-to-perform-complex-data-analysis-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6 [accessed 2018-12-22] [WebCite Cache ID 74rDf1p40]
25. Ren Z, Jiang L, Pang C. How online reviews affect consumers in mobile APP store: a conceptual framework based on elaboration likelihood model. 2017 Presented at: Information Management (ICIM), 3rd International Conference on IEEE; April 21-23, 2017; Chengdu, China p. 21-23. [doi: 10.1109/INFOMAN.2017.7950387]
26. Microsoft Office. Standardize function URL: https://support.office.com/en-us/article/standardize-function-81d66554-2d54-40ec-ba83-6437108ee775 [accessed 2018-12-22] [WebCite Cache ID 74rDj6QVs]
27. Apple. Blood Pressure Companion Pro URL: https://itunes.apple.com/gb/app/blood-pressure-companion-pro/id453210774?mt=8 [WebCite Cache ID 74rDnTD4L]
28. Apple. BMI Calculator URL: https://itunes.apple.com/gb/app/bmi-calculator-your-body-well/id1319214702?mt=8 [accessed 2019-05-18] [WebCite Cache ID 78Six1Ht6]
29. Apple. Molio URL: https://itunes.apple.com/gb/app/molio/id1220682659?mt=8 [WebCite Cache ID 74rEDIz6F]
30. Apple. What's Good: daily gratitude URL: https://itunes.apple.com/gb/app/whats-good-daily-gratitude/id1062008814?mt=8 [WebCite Cache ID 74rEQJfTG]
31. Apple. BP Watch URL: https://itunes.apple.com/gb/app/bp-watch/id1335998464?mt=8 [WebCite Cache ID 74rEZQU4L]
32. Apple. 90 Day Diet URL: https://itunes.apple.com/gb/app/90-day-diet/id871130014?mt=8 [WebCite Cache ID 74rEm8g1E]
33. Apple. Wellness Coach - Health URL: https://itunes.apple.com/gb/app/wellness-coach-health/id685831882?mt=8 [accessed 2018-12-22] [WebCite Cache ID 74rEuwZSC]
34. Google Play. My Inhealthcare URL: https://play.google.com/store/apps/details?id=uk.co.inhealthcare.ihma.inr [accessed 2018-12-22] [WebCite Cache ID 74rF5Iztd]
35. Google Play. Qardio Heart Health URL: https://play.google.com/store/apps/details?id=com.getqardio.android [accessed 2018-12-22] [WebCite Cache ID 74rF8oASA]
36. Google Play. MyBP-Blood Pressure Log URL: https://play.google.com/store/apps/details?id=com.mediitis.bptracker [accessed 2018-12-22] [WebCite Cache ID 74rFCGVZT]
37. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/NEJM199704173361601] [Medline: 9099655]
38. Jamaladin H, van de Belt TH, Luijpers LC, de Graaff FR, Bredie SJ, Roeleveld N, et al. Mobile apps for blood pressure monitoring: systematic search in app stores and content analysis. JMIR Mhealth Uhealth 2018 Nov 14;6(11):e187 [FREE Full text] [doi: 10.2196/mhealth.9888] [Medline: 30429116]
39. Hui CY, Walton R, McKinstry B, Pinnock H. Time to change the paradigm: a mixed method study of the preferred and potential features of an asthma self-management app. Health Inform J (forthcoming).
40. The Academy of the Medical Sciences and the Royal Academy of Engineering. 2014. Health apps: regulation and quality control report URL: https://acmedsci.ac.uk/policy/policy-projects/health-apps-regulation-and-quality-control [WebCite Cache ID 74rFFhbPP]
41. Cheung CM, Thadani DR. The impact of electronic word-of-mouth communication: a literature analysis and integrative model. Decis Support Syst 2012 Dec;54(1):461-470. [doi: 10.1016/j.dss.2012.06.008]

Abbreviations

BP: blood pressure
DASH: Dietary Approaches to Stop Hypertension
GP: general practitioner
HBP: high blood pressure
HBPM: home blood pressure monitoring
NHS: National Health Services
PRISMS: Practical Reviews In Self-Management Support
cSMS: short message service

Edited by C Lovis; submitted 29.12.18; peer-reviewed by K Blondon, R Padwal, T Woolf, M Schuuring, I Kedan; comments to author 26.01.19; revised version received 07.03.19; accepted 08.04.19; published 03.06.19.

Please cite as:
Hui CY, Creamer E, Pinnock H, McKinstry B
Apps to Support Self-Management for People With Hypertension: Content Analysis
JMIR Mhealth Uhealth 2019;7(6):e13257
URL: https://mhealth.jmir.org/2019/6/e13257/
doi:10.2196/13257
PMID:31162124

Original Paper

A Medication Adherence App for Children With Sickle Cell Disease: Qualitative Study

Kristina Curtis[1,2], BSc (Hons), MSc, PhD; Anastasiya Lebedev[3], MSc; Elisa Aguirre[4,5], CPsychol, PhD; Stephan Lobitz[6], MD, MSc

[1]Health Behaviour and Interventions Research, Faculty of Health and Life Sciences, Coventry University, Coventry, United Kingdom

[2]Public Health Warwickshire, Warwickshire County Council, Warwick, United Kingdom

[3]Klinik für Pädiatrie mS, Onkologie/Hämatologie, Charité-Universitätsmedizin Berlin, Berlin, Germany

[4]Division of Psychiatry, University College London, London, United Kingdom

[5]Research and Development Department, North East London Foundation Trust, Goodmayes Hospital, Essex, United Kingdom

[6]Amsterdam Street Children's Hospital, Cologne, Germany

Corresponding Author:
Kristina Curtis, BSc (Hons), MSc, PhD
Health Behaviour and Interventions Research
Faculty of Health and Life Sciences
Coventry University
Richard Crossman Building, 4th Floor
Coventry, CV1 5FB
United Kingdom
Phone: 44 07916142132
Email: kristina.curtis@coventry.ac.uk

Abstract

Background: Young people with sickle cell disease (SCD) often demonstrate low medication adherence and low motivation for effectively self-managing their condition. The growing sophistication of mobile phones and their popularity among young people render them a promising platform for increasing medication adherence. However, so far, few apps targeting SCD have been developed from research with the target population and underpinned with theory and evidence.

Objective: The aim of this study was to develop a theory-and-evidence-based medication adherence app to support children and adolescents with SCD.

Methods: The Behavior Change Wheel (BCW), a theoretically based intervention development framework, along with a review of the literature, 10 interviews with children and adolescents with SCD aged between 12 and 18 years, and consultation with experts informed app development. Thematic analysis of interviews provided relevant theoretical and evidence-based components to underpin the design and development of the app.

Results: Findings suggested that some patients had lapses in memory for taking their medication (capability); variation in beliefs toward the effectiveness of medication and confidence in self-managing their condition (motivation); a limited time to take medication; and barriers and enablers within the changing context of social support during the transition into adulthood (opportunity). Steps were taken to select the appropriate behavioral change components (involving behavior change techniques [BCTs] such as information on antecedents, prompts/cues; self-monitoring of the behavior; and social support) and translate them into app features designed to overcome these barriers to medication adherence.

Conclusions: Patients with SCD have complex barriers to medication adherence necessitating the need for comprehensive models of behavior change to analyze the problem. Children and adolescents require an app that goes beyond simple medication reminders and takes into account the patient's beliefs, emotions, and environmental barriers to medication adherence.

(JMIR Mhealth Uhealth 2019;7(6):e8130) doi:10.2196/mhealth.8130

KEYWORDS

children; adolescents; sickle cell disease; medication adherence; models, theoretical; mHealth

Introduction

Sickle Cell Disease—A Global Health Priority

Sickle cell disease (SCD) is among the most prevalent hereditary blood disorders in the world [1], leading the World Health Organization to prioritize it as a global health issue in 2006 and 2010. SCD causes red blood cells to become *sickle-shaped* which restricts the flow of blood and the transportation of oxygen [2]. Life-threatening complications include infections, acute chest syndrome, stroke, and multiorgan failure [3]. However, by far, the most frequent complications are acute vaso-occlusive events resulting in severe pain episodes. The episodes (also known as pain crises) are predominantly managed within the home environment but require hospitalization when there are complications or when the pain becomes too great [4].

In the past decade, several new medications have been developed to improve the duration and quality of patients' lives [5], resulting in medication adherence becoming fundamental to patients' self-management of their condition. However, systematic review evidence indicates that medication adherence is moderate among children and adolescents [5]. Poor adherence results in reduced effectiveness of medication, increased susceptibility to complications, and medication wastage [6,7].

It is paramount that health care professionals support young patients in developing autonomy and self-management skills [8]. The challenge globally is the serious shortage of health care services to provide support for patients with SCD. The rising migrant populations necessitate the need now more than ever to provide accessible services despite language, culture, finances, ethnicity, and geolocation. Consequently, this has led to growing interest in developing tailored electronic health technologies to support the day-to-day needs of patients. Research has demonstrated an increase in medication adherence among children using short message service (SMS) technology [2], whereas other research has achieved a medication adherence rate of 93% by using electronic directly observed therapy [9].

Sickle Cell Disease Mobile Health Apps

Mobile health (mHealth) apps *offer state-of-the-art approaches to intervention design, delivery and diffusion of treatment and prevention efforts* [10]. Key behavior change techniques (BCTs) important for self-management are optimized through this medium, such as self-monitoring techniques [11], which continue to increase in sophistication [12]. So far, there is growing evidence related to the acceptability and usability of SCD apps aimed at monitoring various symptoms such as pain and fatigue [13-15] and enabling medication reminders [9,16]. They can also enhance communication with health care providers, provide general health management [14], and provide therapeutic interventions such as cognitive behavioral therapy [17,18]. A recent systematic review [19] reported the efficacy of 1 mobile app feasibility study in improving medication adherence after a 6-month follow-up [20]. However, more research is needed to evaluate the mobile app's efficacy and effectiveness for self-managing SCD using careful methods and theoretical underpinnings [19].

A 2013 Cochrane review of asthma self-management apps [21] concluded that future app-based interventions should be underpinned with relevant theoretical frameworks to identify the impact of individual app features on patient outcomes [19]. However, a recent content analysis of 166 medication adherence apps showed that the use of evidence-based BCTs were low [22]. According to Carpenter et al, development of app features that truly implement theoretical constructs remains an undeveloped area, and most medication adherence and disease management apps fail to report any theoretical underpinnings [23], including SDC apps [21,24]. A mobile app is a cost-effective health care intervention [25,26] to support patients' medication adherence. Therefore, this study aimed to develop a theory and evidence-based medication adherence app for children and adolescents with SCD.

Intervention Development Framework

The Behavior Change Wheel (BCW), a theoretically based intervention development framework, was used to guide app development [27]. The BCW is coherent, grounded in a model of behavior (described in the Methods section), and inclusive of all possible intervention strategies. The research draws on a core component of the BCW: the Capability Opportunity, Motivation-Behaviour (COM-B) model that helps to identify important levers for change for the new behavior to occur. It then uses the next steps in intervention development to help bring about change in the new behavior through identifying intervention functions (IFs) and BCTs.

The BCW framework accepts that behavior can essentially derive from a combination of theoretical components within a behavioral system [27]. The research also draws on the next layers of the wheel, IFs, which are defined as expansive classifications through which an intervention can modify behavior [27]. The 9 IFs identified are as follows: *education* (increasing knowledge or comprehension), *persuasion* (evoking emotions to stimulate action), *incentivization* (an expectation of rewards for behavior), *coercion*, (expectation of punitive consequences and costs), *training* (transmitting skills), *restriction* (using rules and regulation to reduce behavior), *environmental restructuring* (modifying the physical or social environment), *modeling* (providing an exemplar of behavior for people to emulate), and *enablement* (increasing the means to carry out the behavior) [27]. IFs can be further broken down into strategies enabling behavior change labeled as BCTs, representing the observable, replicable, and active ingredients in an intervention that directly bring about behavior change [28].

Methods

The app development process drew on the BCW for guidance on understanding the target behavior of medication adherence and how to address this behavior through the use of relevant BCTs. The following section provides details of the steps taken during this process.

Stage 1: Understanding the Behavior

The first stage involved a number of steps to understand the behavior.

Step 1: Defining the Problem

Stakeholder meetings with pediatricians along with a review of the literature helped to define the problem.

Step 2: Selecting the Target Behavior

Stakeholder meetings helped to consider all potential self-management behaviors. However, as previously noted, the BCW recommends starting with only 1 or 2 target behaviors and gradually building on these [22,27]. Selecting the target behavior involved consultation with 2 pediatricians with expertise in SCD and e-learning and a review of the literature.

Step 3: Specifying the Target Behavior

The behavior was then specified for this target population in terms of the context in which the target behavior occurs.

Step 4: Identifying What Needs to Change

This step involved conducting empirical research using a qualitative research design guided by the COM-B model [29] and Theoretical Domains Framework (TDF [30,31]) to explore barriers and enablers to patients' capability, opportunity, and motivation toward medication adherence. The TDF is a framework that amalgamates central theoretical constructs from a wide range of behavior change theories. It classifies 14 significant domains such as skills and emotion, which influence behavior and are possible targets for change [31]. The TDF can be demarcated into 3 core elements of human behavior: capability (C), opportunity (O), and motivation (M) [32]. The COM-B model purports that behavior (B) is a consequence of the interactions between a person's physical and psychological capabilities (C) to utilize social and environmental opportunities (O) via automatic or reflective motivations (M) [33]. The qualitative research involved conducting 10 interviews with children and adolescents with SCD.

Participants

Patients were recruited through Charité University Hospital, Department of Pediatric Hematology and Oncology. All patients were considered eligible if they suffered from SCD, were treated with hydroxycarbamide, owned a smartphone, and were aged between 12 and 18 years. They were invited to take part in the development of an app at the end of their routine face-to-face consultation with the pediatrician. Upon agreement, patients were then telephoned to arrange a time to conduct the interview. Information on demographics, technology use, and smartphone ownership was collected before the interviews (see Results section). Ongoing analysis was conducted across interviews until it became clear that no new codes were emerging from the data, and therefore, recruitment for new participants ceased [34].

Procedure

Interviews took place in a private room in the hospital where parent informed consent and child informed consent were obtained where necessary, before interviews commenced. Participants were paid 30 euros for their participation in the research. A non-native German speaking female interviewer (37 years old) conducted the interviews lasting 30 min. The interviews consisted of semi-structured questions (see Table 1 for the schedule of topics) developed from a review of existing research [5,9,35-37] and structured using the COM-B and TDF to explore barriers and enablers to patients' capability, opportunity, and motivation to self-manage their condition, with a focus on medication adherence. For example, the TDF domains of memory, attention, and decision-making processes included questions such as "What are your thoughts on how well you remember to take your medication", and environmental context and resources included questions such as "What are your thoughts on the things in your environment that make it difficult to take your medication?" Upon permission of the participants, the interviews were audio-recorded, transcribed, and translated into English for analysis.

Thematic Analysis of the Interview Data

Demographic data were analyzed using descriptive statistics. Transcripts were analyzed by 2 independent researchers using recognized principles for conducting thematic analysis [38]. This involved deductively coding the data for their basic meaning before mapping to the COM-B and TDF. This analysis helped to perform a behavioral analysis of the problem wherein theoretical domains were identified as targets for change [29]. In addition, the reliability of the qualitative data was further enriched by the use of an additional trained qualitative researcher who was familiar with the BCW framework and TDF, who independently coded 10% of the data to establish interrater reliability. An agreement of 11/14 TDF domains were established on discussion, and full agreement was reached. An interrater reliability of .79 is generally considered to be an acceptable rate [39].

Stage 2: Identifying Intervention Strategies

Step 5: Identifying Intervention Functions

According to Michie et al, the *behavioral diagnosis* drawn from the COM-B and TDF tools for understanding the behavior represents the foundations for intervention design. Once the *profile* of COM-B and TDF domains has been identified as important *levers for change*, the next stage is to select from a range of IFs provided by the BCW framework [27].

The BCW framework provides a table mapping relevant IFs likely to bring about change in specific COM-B and TDF domains to help conduct this task. However, it was also necessary to review the BCTs that the BCW has mapped to IFs to see how they align with the TDF domains identified. Therefore, the mapping process underwent a cyclical process where BCTs were mapped back to IFs. The next step involves delineating these IFs into specific BCTs. The authors of the guide purposely used the term *functions* to indicate that BCTs can have more than 1 IF [27].

Table 1. Schedule of questions for interviews.

COM-B[a] model and TDF[b]	Topic question
Psychological capability	
Knowledge	• What are your thoughts on how much you know about Sickle Cell Anemia (Prompt: Can you describe what it is?) • What are the complications of the condition? • How is the condition treated? • How does the medication (specify which one) for the condition work?
Skills	• Do you know how to take your medication? • What are your thoughts on any measures that you can take to prevent the condition getting worse?
Memory, attention, and decision-making processes	• What are your thoughts on how well you remember to take your medication? • Do you normally set a reminder to take your medication?
Behavioral regulation	• Do you have a way of monitoring whether you have taken your medication every day? (If no, why not? If yes, how?)
Environmental context and resources	• What are your thoughts on the things in your environment that make it difficult to take your medication? (Prompt: lack of time, lack of privacy)
Social opportunity	
Social influences	• What are your thoughts on how much support you receive from your parents for your condition? • What kind of support do you receive from your local community? • What kind of support do you receive from your close friends? • What kind of support do you receive from your school/teachers? • Do you have any further ideas on how family and friends can support you in an app?
Reflective motivation	
Social identity	—[c]
Beliefs about capabilities	• How confident do you feel about managing your condition? (Prompt: remembering to take your medication, managing your moods, managing the pain)
Optimism	—
Beliefs about consequences	• What do you believe might happen to your body if you take your medication? (Prompts: what are your beliefs on whether medication will make your illness worse or better?) • What are your thoughts around side effects? • What are your thoughts around how serious your condition is? • What are your thoughts on whether it takes too much time and effort to take your medication during your daily routine?
Intention	• Do you intend to take your medication every day? (Prompt: If not, why not?)
Goals	• What are your thoughts on your goals for taking medication?
Automatic motivation	
Reinforcement	• What would be an incentive to take your medication? • We have had some ideas of how features in an app can help to incentivize children to take their medication
Emotion	• What are your thoughts on whether your moods make your physical symptoms worse or better? How could the app help you to manage your moods better? • Does taking your medication cause any emotional reactions and feelings?

[a]COM-B: Capability Opportunity, Motivation-Behavior.

[b]TDF: Theoretical Domains Framework.

[c]Not applicable.

Step 6: Identifying Behavior Change Techniques

Mapping BCTs to intervention functions involved 2 steps: First, the BCW table for mapping IFs to relevant BCTs provided a candidate list of BCTs to use for the intervention. As previously mentioned, selecting IFs also required looking forward to ascertain which BCTs that the BCW maps to IFs aligned with the TDF domains and context of an app. Therefore, some BCTs were already selected if they were relevant to bringing about change in the TDF domain. This step also involved reviewing a systematic review of medication adherence among pediatric patients with SCD [5].

Step 7: Translating Findings Into App Features

Consultation with the project team in the form of a workshop (app developers, pediatricians, and behavioral scientist) helped to translate the intervention mapping results into app features. The process involved the behavioral scientist (KC) presenting key findings using the intervention mapping table to the project team. Specifically, the table helped to communicate which BCTs were required to change medication adherence behavior. This then instigated discussions on how these techniques could be operationalized in the app, leading to decisions on app features. Further meetings via Skype also involved the design company (comprised user experience experts) to help with the development of engaging app features. Table 2 below presents the sample demographics and mobile phone usage.

Table 2. Demographics and mobile phone information (N=10).

Demographics	Values
Origins of parents (N=20), n (%)	
Lebanon	7 (35)
Nigeria	6 (30)
Angola	3 (15)
Sierra Leonne	2 (10)
The Congo	1 (5)
Palestine	1 (5
Gender, n (%)	
Female	6 (60)
Male	4 (40)
Age (years), mean (range)	14 (11-17)
German citizenship, n (%)	8 (80)
Living status, n (%)	
Not living alone	9 (90)
Number of people living in household, average (range)	4 (3-6)
Parents' employment status, n (%)	
Mothers unemployed (n=9)	5 (56)
Mothers employed full time (n=9)	4 (44)
Fathers employed full time (n=7)	4 (57)
Fathers employed part time (n=7)	2 (29)
Fathers self-employed (n=7)	1 (14)
Smartphone ownership and access to the internet, n (%)	10 (100)
Frequency of mobile phone use, n (%)	
Daily	9 (90)
2-3 times a week	1 (10)

Results

Step 1: Defining the Problem

The problem was defined by the stakeholder group as too many children and adolescents with SCD not successfully self-managing their condition.

Step 2: Selecting the Target Behavior

Pediatricians reported that medication nonadherence was a serious problem among their patients. This was also bolstered with systematic review evidence indicating medication adherence is low among young people, resulting in detrimental effects to their health [5-7]. Therefore, supporting children and

adolescents with their medication adherence was identified as the target behavior.

Step 3: Specifying the Target Behavior

The global term, *medication adherence*, incorporates initiating the prescription, actual dosing in relation to the prescription, and persisting with treatment. Adherence relates simply to the behavior itself—using treatment at the right time, for the right period, in the right quantity, and in the right manner [40]. Table 3 below specifies the target behavior by detailing who needs perform the behavior, when, where, how and with whom.

Step 4: Identifying What Needs to Change

The behavioral analysis (shown in Table 4) revealed barriers and enablers in all 3 COM-B domains and the following 10 TDF domains involving the following: limited knowledge related to the disease itself and how the mediation works (knowledge); Forgetting to take medication and set reminders (memory, attention, and decision processes); lack of external monitoring and reminders (behavioral regulation); limited confidence related to medication adherence (beliefs about capabilities); limited importance of regular medication adherence, priority for other influences on health such as religion and negative consequences on taste and social life (beliefs about consequences); perceived limited support required during the transition to adulthood (social identity); intrinsic goals such as excelling at sports provided motivation to medication adherence (goals); emotional responses to taking medication and managing pain (emotion); perceived limited time to take medication, being outside of the home environment and other health professionals (environmental context and resources); and an over reliance on parents to remember to take medication and importance of peer support (social influences). The subthemes taken forward for intervention development are highlighted in italics in Table 4 below.

Steps 5, 6, and 7: Intervention Mapping Table

The intervention mapping table shown below (Table 5) shows the mapping of determinants, intervention strategies, and potential app features. This is where the results from the COM-B and TDF analysis were mapped onto 15 BCTs as guided by the BCW.

Table 3. Specifying the target behavior.

Target behavior	Medication adherence
Who needs to perform the target behavior?	Children and young people with sickle cell disease
When do they need to perform the behavior?	Every day (at a time suitable for their schedules)
Where do they need to perform the behavior?	At home and outside of the home
How often do they need to perform the behavior?	Once a day
With whom do they need to perform the behavior?	By themselves (or parents)

Table 4. Behavioral analysis of the influences on patients' medication adherence. Subthemes in italics were taken forward for intervention development.

COM-B[a] model, TDF[b], and subtheme	Quote
Psychological capability	
Knowledge	
Perceived good knowledge of self-management of condition	I don't have the same stamina, as other teenagers of my age. I have to dress warmly in the winter, or else I'll have pain. I'm more susceptible when I do sports, I'm not supposed to overexert myself, I can get injured quickly. My hips are inflamed because of overexertion during sports [P10].
Perceived good knowledge of task/environment "what to do"	You have to go see Dr. Lobitz every 3 months, once a year an annual check-up, when they run all the tests. I also have to take penicillin every morning and every evening, and Siklos every evening, and Ibuprofen when I'm in pain [P10].
Limited knowledge in relation to how the medication works—scientific rationale	*Don't know [P2].* *It's a medication for not having pain. If you don't drink enough, this pill helps [P3].*
Limited disease knowledge	*Not so much. There are these white blood cells that get into your bones [P4].*
Skills	
Good knowledge and skills in relation to how to take the medication	1 ½ film-coated tablets every day. You can take them in the morning or in the evening [P6].
Good knowledge and skills in relation to prevention	I'm more susceptible when I do sports, I'm not supposed to overexert myself, I can get injured quickly. My hips are inflamed because of overexertion during sports [P10]. ...dressing warmly, if you've been swimming, put a towel on right away [P2]
Memory, attention, and decision-making processes	
Medication taking is part of a routine	It's just something I do, take a pill before I go to bed. It's become my routine. If I'm on a school trip or spending the night at a friend's house, I also bring it along and take it [P9].
Forgetting to take medication	*Sometimes I forget, pretty often, actually. My mom helps me remember, my dad and my sister too [P2].* *Very poorly. I need this app. I always forget [P6].*
Forgetting to set external reminders	*Yes, I set the alarm clock on my cell phone for nine, for example. It helps, but sometimes I forget to set the alarm clock [P2].*
Behavioral regulation	
Lack of external monitoring device for medication intake	*No – I haven't ever considered it [P5].*
Reflective motivation	
Intentions	
High intentions and adherence to morning and evening medication intake	It goes without saying. In the morning I wake up, I take my penicillin and I go out. In the evenings it's also like that, I just take them. If I'm spending the night at a friend's house, I take the medications first and then I go to my friend's house [P10].
Beliefs about capabilities	
Self-confidence: high and moderate confidence in relation to how well they can self-manage the condition	*Very confident. I'm old enough to know what I have to do [P2].* *50/50. Taking my medication is the only problem I have [P6].*
Beliefs about consequences	
High and low importance and relevance to the perceived consequences of nonadherence to treatment	*Skipping it one day is not a biggie, but if you don't take it for several months, you'll get weaker, you'll feel worse, you'll have no more strength or energy. The pain will probably also get worse [P10]* *If I don't take my medication, it doesn't make any difference [P6].*
Religion has a greater importance compared with medication	*Then my mother told me, God will not let you die at the age of 40 because you haven't taken your medication [P6].*

COM-B[a] model, TDF[b], and subtheme	Quote
Thoughts on medication intake negative consequences (physical)	*XXX is uncomfortable to swallow. It doesn't have any taste, but it starts dissolving in your mouth, that's a disgusting feeling [P10].*
Thoughts on medication intake negative consequences (social)	*It bothers me when I'm at a friend's house, and we're talking, and then I have to break up the conversation to say, I have to take my pill now. That's annoying [P8].*
Social identity	
Personal identity	*I am older and I can manage [P10].*
Goals	
Clear goals driving intrinsic motivation	*I have a goal, to be able to be the best I can be at sports. That's my motivation [P10].*
Automatic motivation	
Reinforcement	
Physical consequences of not taking medication long term	...if you skip it more often, you can get sick [P5].
Remembering to take medication is easier when it becomes a habit	I always remember, it's a question of habit [P3].
The home environment	I put the pills somewhere where I'll see them right away, someplace where I'll definitely see them this evening. That works well [P9].
Emotions	
Variation in beliefs toward the connection between emotions and pain	There's no connection. When I have pain, I just take medication; I always feel the same [P10]. Thinking of good things helps when you're in pain [P7]. When something bad happens, pain gets worse [P2].
Negative emotion toward medication taking	*It sometimes makes me sad, because I don't want to be taking medicine, but I have to [P7].*
Social opportunity	
Social influences	
Peer support during unwell periods	*My friends know about my condition. They send me messages on WhatsApp when I'm sick. When I write to my class on WhatsApp that I'm sick and not coming to school, they all write to me and tell me to get better. This makes me happy [P1].*
Practical parental support for taking medication	*If I'm watching TV, my mother comes and tells me I should take my medication [P7].* *They help me. They prepare my medication for me so that I can take it quickly and get on doing other things in my daily routine [P3].*
Emotional parental support for pain crises	*But home when I'm with my mom, I can be like, "Aaa mama I have this stomach ache!" [P6].*
(Minimal) parental support only important support explicitly perceived as needed (older children)	*My mother accompanies me sometimes to my appointments with Dr XXX. Gradually she's been helping me less and less, I'm old enough already. We've both decided that I can go to those appointments on my own [P10].*
Peer support for medication adherence and management of condition	*If I'm staying at a friend's house late, I take my medication with me and my friends tell me when to take it [P7]; If we're in PE, they tell me to sit down and tell the teacher if I'm not feeling well. When we go swimming, they come out of the water with me and go with me to fetch something to eat and drink, because I'm not supposed to stay in the water for too long. They ask me whether I've taken my medication [P5].*
Physical opportunity	
Environmental context and resources	
Limited time to take medication	*When I have to get dressed quickly and go to school, my school is far away, I have to do everything quickly. I don't have enough time and then I forget [P2].*

COM-B[a] model, TDF[b], and subtheme	Quote
Being outside of the home environment	*When I come home from the street, I don't think about it [P6]; If I'm out and about or if I have to take it at a different time [P2]; When I'm at a friend's house, every once in a while I forget to take my medication [P5].*
Other health professionals	*When [my regular] doctors aren't there, the others don't know how to treat me [P7].*

[a]COM-B: Capability Opportunity, Motivation-Behavior.

[b]TDF: Theoretical Domains Framework.

Table 5. Final intervention mapping table.

COM-B[a] model, TDF[b] and subtheme	Intervention functions	Behavior change techniques	App features
Psychological capability			
Knowledge			
Perceived disconnection between thoughts and moods and effects on their condition	Education	Information about antecedents, self-monitoring of behavior	Quiz, mood tracker
Limited knowledge in relation to how the medication works—scientific rationale	Education	Information about health consequences	Quiz related to biological causes of sickle cell disease and effects of medication
MADMP[c]			
Forgetting to take medication	Environmental restructuring	Prompts/cues	Medication reminders
Behavioral regulation			
Perceived behavioral control: Not using external monitoring device for medication intake	Environmental restructuring	Self-monitoring of the behavior, feedback on the behavior, social support (practical)	Medication diary. Record how many tablets they have taken each day and when they have missed one. Buddy messenger—A chosen close friend can also monitor medication adherence and send a reminder them to take medication
Reflective motivation			
Beliefs about consequences			
Little importance and relevance of the perceived consequences to medication adherence (religion)	Education	Information about health consequences	Quiz, persuasive messages (in-app notifications). Explain that not taking medication can result in a pain crises and serious complications
Negative consequences on medication intake (physical)	Enablement	Problem solving	Quiz, tips
Thoughts on medication intake negative consequences (social)	Enablement	Problem solving	Quiz, tips
Perceived limited time to take medication	Education enablement	Information about health consequences, problem solving	Quiz, tips
Self-confidence in remembering to take medication	Environmental restructuring enablement	Problem solving, prompts/cues	Medication reminders, tips
Social Identity			
Personal identity, transition to adulthood	Education, modelling	Information about health consequences, problem solving, social comparison, modeling	Quiz, tips, patients' stories on how they are coping with their condition. Users can view each other's avatars to see how many points they have for medication adherence and appointments
Group identity (interest in how other patients are coping)	Modeling	Social comparison, modeling	Patients stories on how they are coping with their condition
Goals			
Clear goals driving intrinsic motivation	Persuasion	Goal setting (outcome)	Encourage them to set their own goals for taking their medication (prompt: so they can take part in sports, hang out with friends)
Automatic motivation			
Emotion			
Connection between stress and pain	Enablement	Reduce negative emotions, Practice habit formation	Comical anecdotes (jokes) and videos to distract from the pain, relaxation videos, and audios

COM-B[a] model, TDF[b] and subtheme	Intervention functions	Behavior change techniques	App features
Negative emotion toward medication taking (reliance and interference with social activities)	Enablement, modeling	Reduce negative emotions, Social support, Problem solving, Action planning	Comical anecdotes (jokes), other patient stories on how they manage their medication adherence, messages to friends, encourage them to make a plan of how to take medication when they are socializing
Social Opportunity			
Social influences			
Peer support	Environmental restructuring, enablement	Social support (practical and emotional)	Tips. Advise to call or text a friend when they need to distract from the pain
Physical opportunity			
Environmental context and resources			
The home environment	Environmental restructuring	Add objects to the environment	Tips. Advise to keep their medication on show in a place in their home they will see it
Outside of the home environment	Environmental restructuring, enablement	Prompts/cues, Action planning, Behavioral practice	Medication reminders, encourage a plan to take medication when going out socially, instruct the person to practice taking medication in a social context
Other health professionals	Environmental restructuring, enablement	Adding objects to the environment	Health record

[a]COM-B: Capability Opportunity, Motivation-Behavior.

[b]TDF: Theoretical Domains Framework.

[c]Memory, attention, and decision-making processes.

Overall Concept of MyMate&Me

The following app features were chosen for the first release: *Avatar* (Figures 1 and 2)—a cartoon figure of a boy or girl accompanies the patient throughout his or her interaction with the app. The patient is encouraged to earn bonus points through being active in various app sections to dress the avatar with new clothes, accessories, or even facial hair. *Tip of the day* (Figure 3)—important information on day-to-day coping with the disease, sometimes providing gender-specific tips and dispelling myths. *Daily Quiz* (Figure 4)—points can be earned for answering quiz questions correctly. If a quiz question is answered incorrectly, the right answer is provided, increasing the patient's awareness of different aspects of his or her disease. *Mood Tracker* (Figure 5)—Users can tilt the phone to indicate which mood they are in. *Medication and Appointment Reminders* (Figure 6)—points are earned for confirming that medication has been taken and appointments have been kept. Patients can see the avatars of other patients using the app and compare their scores. *Emergency section* (Figure 7)—includes crucial information for health professionals unfamiliar with the disease and quick emergency call and text options.

Figure 1. Avatar.

Figure 2. Avatar with clothes options.

Figure 3. Tip of the Day.

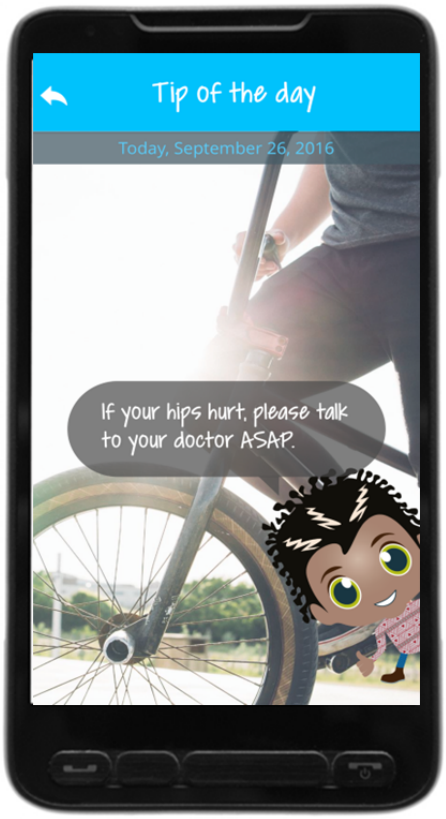


Figure 4. Daily Quiz.

Figure 5. Mood Tracker.

Figure 6. Medication Reminder.

Figure 7. Emergency card.

Discussion

Principal Findings

Previous research in relation to medication adherence supports the findings in this study, indicating that nonadherence can result from forgetting [40], children disliking the taste [41], and the transition from childhood to adulthood [35]. Chronic illnesses affect young people in a myriad of different ways as they transition into adulthood and adult care [8]. Parents have been largely responsible for ensuring their children self-manage a chronic condition such as taking their medication [8]. Furthermore, evidence suggests that parents' perceptions of whether the medication will work also predicts medication continuity which aligns with the religious influence on medication adherence reported by one of the participants [42]. Previous research suggests that parents are concerned with their child's transition into adulthood where there is a need for children to take more control and responsibility toward their condition and to self-manage [33]. Patients need to be able to self-monitor their medication adherence and self-manage their condition as they grow older, and this can supported in the MyMate&Me app through features that enable reminders, self-monitoring, feedback on their behavior, support from peers, and problem solving.

In line with previous research, some participants reported that their mood, stress, and anxiety exacerbated their pain symptoms as reported in previous research with this population [13]. This research also provides new insights among older children where they reported negative emotions around their reliance on medication and its interference with social activities. Furthermore, being outside of the home environment was also cited as a barrier to medication adherence: there were no environmental cues to remind them to take their medication and/or a lack of suitable places to take them in privacy.

Theoretical domains targeted for change were mapped onto 15 BCTs which also align with the components for self-management proposed in the Practical Reviews In Self-Management Support (PRISMS) taxonomy [43]. Although many of these BCTs have been found collectively across medication adherence apps on the market, individual apps have only been underpinned with 2 to 3 of these BCTs. In addition, most have only used Action planning and a Prompt/cue. This study also highlighted the need for BCTs: Goal setting (outcome), Problem solving, and Reducing negative emotions, which are missing in the review of medication apps [22].

Strengths of the Research

Drawing on a theoretically grounded and evidenced-based intervention development framework, along with conducting research with the target audience, has resulted in the first theory-and-evidence-based app for medication adherence for children and adolescents with SCD. The findings have highlighted barriers that go beyond simply forgetting to take their medication (as focused on in current SDC medication adherence apps), where emotions and social life are perceived to play a pivotal role in medication adherence, particularly during the transition from childhood to adulthood.

Limitations of the Research

The empirical research engaged a small purposive sample. Consequently, the identified views on the facilitators and barriers to medication adherence may be less representative of other young people with SCD. However, qualitative research may not require a high number of participants before reaching data saturation, and randomization is more suited to quantitative inquiry [44]. The qualitative approach followed, enabled a richer reflection of the barriers and facilitators to the target behavior, and assessed the contribution of the sociocognitive and external factors influencing the target behavior. In addition, the age range of participants may be considered as too wide, raising concerns about the study's ability to capture the different issues relating to medication adherence. However, the qualitative nature helps to overcome this by allowing participants to give an in-depth account of their experiences.

There are also a number of limitations with the application of the BCW approach. Selecting which BCTs to use in the intervention represented a challenging process as the BCT (v1) at the time of research did not link individual BCTs to their theoretical determinants. However, recently, a Web-based tool has been launched to support this process and specifically links BCTs with theoretical determinants known as *mechanisms of action* [45]. However, as noted by Orji and Mandryk, using a mapping process for intervention development, is always subject to interpretation [46]. The process of selecting IFs was challenging because many of the same BCTs belonged to different IFs. However, the IM table was reviewed by 2 health psychologists to help overcome these challenges.

Future Research

The app is now undergoing formal usability testing with patients. In addition, nonintrusive data collection such as usage data of app features (which correspond with BCTs) and their correlation to behavior change will help to measure engagement with the intervention as proposed in other mHealth research [47].

Conclusions

Patients with SCD have complex barriers to medication adherence which can only be identified in using comprehensive enough models and frameworks of human behavior such as the COM-B and TDF. As such, existing apps lacking a theoretical underpinning do not go far enough in supporting young people with SCD; focusing only on 1 or 2 aspects of medication adherence such as reminders and medication logs.

Acknowledgments

The authors gratefully acknowledge the help of the 10 patients who provided data for development of an SCD app. The authors confirm that all personal identifiers have been removed.

The authors acknowledge Charité for providing the funding for this study. The authors also acknowledge Condat for helping to translate the research findings into app features and conducting the technical development of this study.

Authors' Contributions

KC guided the research and app development process and wrote the first draft of the paper. AL conducted the interviews with patients and provided comments on the paper. EA performed the double coding on transcripts, reviewed the theoretical intervention mapping process, and provided comments on the paper. SL led the project and helped to recruit patients for the research, to develop the app content, and provided comments on the paper.

Conflicts of Interest

None declared.

References

1. Issom D, Zosso A, Ehrler F, Wipfli R, Lovis C, Koch S. Exploring the challenges and opportunities of eHealth tools for patients with sickle cell disease. Stud Health Technol Inform 2015;216:898. [Medline: 26262200]
2. Jacob E, Duran J, Stinson J, Lewis M, Zeltzer L. Remote monitoring of pain and symptoms using wireless technology in children and adolescents with sickle cell disease. J Am Assoc Nurse Pract 2013 Jan;25(1):42-54 [FREE Full text] [doi: 10.1111/j.1745-7599.2012.00754.x] [Medline: 23279278]
3. Estepp J, Winter B, Johnson M, Smeltzer M, Howard S, Hankins J. Improved hydroxyurea effect with the use of text messaging in children with sickle cell anemia. Pediatr Blood Cancer 2014 Nov;61(11):2031-2036 [FREE Full text] [doi: 10.1002/pbc.25177] [Medline: 25132074]
4. Silva I, Reis A, Palaré MJ, Ferrão A, Rodrigues T, Morais A. Sickle cell disease in children: chronic complications and search of predictive factors for adverse outcomes. Eur J Haematol 2015 Feb;94(2):157-161 [FREE Full text] [doi: 10.1111/ejh.12411] [Medline: 25039473]
5. Walsh K, Cutrona S, Kavanagh P, Crosby L, Malone C, Lobner K, et al. Medication adherence among pediatric patients with sickle cell disease: a systematic review. Pediatrics 2014 Dec;134(6):1175-1183 [FREE Full text] [doi: 10.1542/peds.2014-0177] [Medline: 25404717]
6. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005;353:487 [FREE Full text]

7. DiMatteo M, Giordani P, Lepper H, Croghan T. Patient adherence and medical treatment outcomes: a meta-analysis. Med Care 2002 Sep;40(9):794-811 [FREE Full text] [doi: 10.1097/01.MLR.0000024612.61915.2D] [Medline: 12218770]
8. Sattoe J, Bal M, Roelofs P, Bal R, Miedema H, van Staa AL. Self-management interventions for young people with chronic conditions: a systematic overview. Patient Educ Couns 2015 Jun;98(6):704-715 [FREE Full text] [doi: 10.1016/j.pec.2015.03.004] [Medline: 25819373]
9. Creary SE, Gladwin MT, Byrne M, Hildesheim M, Krishnamurti L. A pilot study of electronic directly observed therapy to improve hydroxyurea adherence in pediatric patients with sickle-cell disease. Pediatr Blood Cancer 2014 Jun;61(6):1068-1073. [doi: 10.1002/pbc.24931] [Medline: 24436121]
10. Tate E, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz M, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]
11. Curtis K, Karasouli E. An assessment of the potential of health promotion apps to support health behaviour change. Heal Psychol Update 2014;23(2) [FREE Full text]
12. 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(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
13. Jacob E, Pavlish C, Duran J, Stinson J, Lewis M, Zeltzer L. Facilitating pediatric patient-provider communications using wireless technology in children and adolescents with sickle cell disease. J Pediatr Health Care 2013;27(4):284-292 [FREE Full text] [doi: 10.1016/j.pedhc.2012.02.004] [Medline: 22446036]
14. Shah N, Jonassaint J, de Castro L. Patients welcome the Sickle Cell Disease Mobile Application to Record Symptoms via Technology (SMART). Hemoglobin 2014;38(2):99-103 [FREE Full text] [doi: 10.3109/03630269.2014.880716] [Medline: 24512633]
15. Jonassaint C, Shah N, Jonassaint J, de Castro L. Usability and feasibility of an mHealth intervention for monitoring and managing pain symptoms in sickle cell disease: the Sickle Cell Disease Mobile Application to Record Symptoms via Technology (SMART). Hemoglobin 2015;39(3):162-168 [FREE Full text] [doi: 10.3109/03630269.2015.1025141] [Medline: 25831427]
16. Estepp J, Winter B, Johnson M, Smeltzer M, Howard S, Hankins J. Improved hydroxyurea effect with the use of text messaging in children with sickle cell anemia. Pediatr Blood Cancer 2014 Nov;61(11):2031-2036 [FREE Full text] [doi: 10.1002/pbc.25177] [Medline: 25132074]
17. Cheng C, Brown R, Cohen L, Venugopalan J, Stokes T, Wang M. iACT--an interactive mHealth monitoring system to enhance psychotherapy for adolescents with sickle cell disease. Conf Proc IEEE Eng Med Biol Soc 2013;2013:2279-2282 [FREE Full text] [doi: 10.1109/EMBC.2013.6609992] [Medline: 24110179]
18. Schatz J, Schlenz AM, McClellan CB, Puffer ES, Hardy S, Pfeiffer M, et al. Changes in coping, pain, and activity after cognitive-behavioral training: a randomized clinical trial for pediatric sickle cell disease using smartphones. Clin J Pain 2015 Jun;31(6):536-547. [doi: 10.1097/AJP.0000000000000183] [Medline: 25503599]
19. Badawy S, Cronin R, Hankins J, Crosby L, DeBaun M, Thompson A, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018 Dec 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
20. Leonard S, Anderson L, Jonassaint J, Jonassaint C, Shah N. Utilizing a Novel Mobile Health "Selfie" Application to Improve Compliance to Iron Chelation in Pediatric Patients Receiving Chronic Transfusions. J Pediatr Hematol Oncol 2017 Dec;39(3):223-229 [FREE Full text] [doi: 10.1097/MPH.0000000000000743] [Medline: 28099398]
21. Belisario JS, Huckvale K, Greenfield G, Car J, Gunn L. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013 Nov 27;11(11):CD010013 [FREE Full text] [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
22. Morrissey E, Corbett T, Walsh J, Molloy G. Behavior change techniques in apps for medication adherence: a content analysis. Am J Prev Med 2016 May;50(5):e143-e146 [FREE Full text] [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
23. Carpenter D, Geryk L, Sage A, Arrindell C, Sleath B. Exploring the theoretical pathways through which asthma app features can promote adolescent self-management. Transl Behav Med 2016 Dec;6(4):509-518 [FREE Full text] [doi: 10.1007/s13142-016-0402-z] [Medline: 27118115]
24. Lobitz S, Curtis K, Sostmann K. Digital Health Interventions (DHIs) to support the management of children and adolescents with sickle cell disease. In: Sickle Cell Disease - Pain and Common Chronic Complications. London: InTechOpen; 2016:261.
25. Vandelanotte C, Duncan MJ, Plotnikoff RC, Mummery WK. Do participants' preferences for mode of delivery (text, video, or both) influence the effectiveness of a web-based physical activity intervention? J Med Internet Res 2012 Feb 29;14(1):e37 [FREE Full text] [doi: 10.2196/jmir.1998] [Medline: 22377834]
26. Geryk LL, Roberts CA, Sage AJ, Coyne-Beasley T, Sleath BL, Carpenter DM. Parent and clinician preferences for an asthma app to promote adolescent self-management: a formative study. JMIR Res Protoc 2016 Dec 6;5(4):e229 [FREE Full text] [doi: 10.2196/resprot.5932] [Medline: 27923777]

27. Michie S, Atkins L, West R. The Behaviour Change Wheel A Guide to Designing Interventions. London: Silverback Publishing; 2014.
28. 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 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
29. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
30. Francis J, O'Connor D, Curran J. Theories of behaviour change synthesised into a set of theoretical groupings: introducing a thematic series on the theoretical domains framework. Implement Sci 2012 Apr 24;7:35 [FREE Full text] [doi: 10.1186/1748-5908-7-35] [Medline: 22531601]
31. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
32. Steinmo S, Michie S, Fuller C, Stanley S, Stapleton C, Stone S. Bridging the gap between pragmatic intervention design and theory: using behavioural science tools to modify an existing quality improvement programme to implement "Sepsis Six". Implement Sci 2016 Feb 03;11:14 [FREE Full text] [doi: 10.1186/s13012-016-0376-8] [Medline: 26841877]
33. Alexander K, Brijnath B, Mazza D. Barriers and enablers to delivery of the Healthy Kids Check: an analysis informed by the Theoretical Domains Framework and COM-B model. Implement Sci 2014 May 23;9:60 [FREE Full text] [doi: 10.1186/1748-5908-9-60] [Medline: 24886520]
34. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
35. Kenyon C, Kavanagh P, Fiechtner L, Textor TC, Wang CJ. Setting the agenda for quality improvement in pediatric sickle cell disease. J Natl Med Assoc 2012;104(7-8):337-341. [doi: 10.1016/S0027-9684(15)30174-7] [Medline: 23092048]
36. Jenerette C, Murdaugh C. Testing the Theory of Self-care Management for sickle cell disease. Res Nurs Health 2008 Aug;31(4):355-369 [FREE Full text] [doi: 10.1002/nur.20261] [Medline: 18247376]
37. Badawy S, Thompson A, Liem R. Technology access and smartphone app preferences for medication adherence in adolescents and young adults with sickle cell disease. Pediatr Blood Cancer 2016 May;63(5):848-852 [FREE Full text] [doi: 10.1002/pbc.25905] [Medline: 26844685]
38. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
39. Salkind N. Tests & Measurement For People Who (think They) Hate Tests & Measurement. London: Sage Publications, Inc; 2017.
40. Vrijens B, Antoniou S, Burnier M, de la Sierra A, Volpe M. Current situation of medication adherence in hypertension. Front Pharmacol 2017;8:100 [FREE Full text] [doi: 10.3389/fphar.2017.00100] [Medline: 28298894]
41. Raphael J, Bernhardt M, Mahoney D, Mueller B. Oral iron chelation and the treatment of iron overload in a pediatric hematology center. Pediatr Blood Cancer 2009 May;52(5):616-620 [FREE Full text] [doi: 10.1002/pbc.21929] [Medline: 19148949]
42. Brinkman W, Sucharew H, Majcher J, Epstein J. Predictors of medication continuity in children With ADHD. Pediatrics 2018 Dec;141(6) [FREE Full text] [doi: 10.1542/peds.2017-2580] [Medline: 29794230]
43. Pearce G, Parke H, Pinnock H, Epiphaniou E, Bourne C, Sheikh A, et al. The PRISMS taxonomy of self-management support: derivation of a novel taxonomy and initial testing of its utility. J Health Serv Res Policy 2016 Apr;21(2):73-82 [FREE Full text] [doi: 10.1177/1355819615602725] [Medline: 26377727]
44. Mays N, Pope C. Rigour and qualitative research. Br Med J 1995 Jul 8;311(6997):109-112 [FREE Full text] [doi: 10.1136/bmj.311.6997.109] [Medline: 7613363]
45. Connell LE, Carey RN, de Bruin M, Rothman AJ, Johnston M, Kelly MP, et al. Links between behaviour change techniques and mechanisms of action: an expert consensus study. Ann Behav Med 2018 Nov 19:a [FREE Full text] [doi: 10.1093/abm/kay082] [Medline: 30452535]
46. Orji R, Mandryk R. Developing culturally relevant design guidelines for encouraging healthy eating behavior. Int J Hum-Comput Stud 2014 Feb;72(2):207-223 [FREE Full text] [doi: 10.1016/j.ijhcs.2013.08.012]
47. Brindal E, Hendrie G, Freyne J. Combining persuasive technology with behavioral theory to support weight maintenance through a mobile phone app: protocol for the MotiMate App. JMIR Res Protoc 2016 Jan 8;5(1):e5 [FREE Full text] [doi: 10.2196/resprot.4664] [Medline: 26747725]

Abbreviations

BCT: behavior change technique
BCW: behavior change wheel

COM-B: Capability Opportunity, Motivation-Behaviour
IF: intervention function
mHealth: mobile health
SCD: sickle cell disease
TDF: Theoretical Domains Framework

Edited by G Eysenbach; submitted 21.06.17; peer-reviewed by E Davies, F Ehrler, L Fiellin, C Lovis; comments to author 22.11.17; revised version received 19.02.19; accepted 18.04.19; published 18.06.19.

Please cite as:
Curtis K, Lebedev A, Aguirre E, Lobitz S
A Medication Adherence App for Children With Sickle Cell Disease: Qualitative Study
JMIR Mhealth Uhealth 2019;7(6):e8130
URL: http://mhealth.jmir.org/2019/6/e8130/
doi:10.2196/mhealth.8130
PMID:31215518

Original Paper

Smartphone, Social Media, and Mental Health App Use in an Acute Transdiagnostic Psychiatric Sample

Courtney Beard[1], PhD; Alexandra L Silverman[1], BA; Marie Forgeard[1,2], PhD; M Taylor Wilmer[1], PhD; John Torous[3], MD; Thro¨stur Bjo¨rgvinsson[1], PhD

[1]McLean Hospital/Harvard Medical School, Belmont, MA, United States

[2]William James College, Newton, MA, United States

[3]Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MD, United States

Corresponding Author:
Courtney Beard, PhD
McLean Hospital/Harvard Medical School
115 Mill St
Mailbox 113
Belmont, MA, 02478
United States
Phone: 1 16178553557
Email: CBEARD@MCLEAN.HARVARD.EDU

Abstract

Background: Despite high rates of smartphone ownership in psychiatric populations, there are very little data available characterizing smartphone use in individuals with mental illness. In particular, few studies have examined the interest and use of smartphones to support mental health.

Objective: This study aimed to (1) characterize general smartphone app and social media usage in an acute transdiagnostic psychiatric sample with high smartphone ownership, (2) characterize current engagement and interest in the use of smartphone apps to support mental health, and (3) test demographic and clinical predictors of smartphone use.

Methods: The survey was completed by all patients attending an adult partial hospital program, with no exclusion criteria. The primary outcomes were frequency of use of general and mental health smartphone apps (smartphone use survey) and the frequency of social media use and phone-checking behavior (mobile technology engagement scale).

Results: Overall, 322 patients (aged mean 33.49, SD 13.87 years; 57% female) reported that their most frequently used app functions were texting, email, and social media. Younger individuals reported more frequent use across most types of apps. Baseline depression and anxiety symptoms were not associated with the frequency of app use. Participants reported health care, calendar, and texting apps as most supportive of their mental health and social media apps as most negatively affecting their mental health. Most patients reported an interest in (73.9% [238/322]) and willingness to use (81.3% [262/322]) a smartphone app to monitor their mental health condition. Less than half (44%) of the patients currently had a mental health app downloaded on their smartphone, with mindfulness and meditation apps being the most common type.

Conclusions: The high interest in and willingness to use mental health apps, paired with the only moderate current reported usage, indicate a potential unmet treatment opportunity in psychiatric populations. There is potential to optimize non-mental health–specific apps to better support the needs of those with mental illness and to design a new wave of mental health apps that match the needs of these populations as well as the way they use smartphones in daily life.

(JMIR Mhealth Uhealth 2019;7(6):e13364) doi:10.2196/13364

KEYWORDS

mobile health; smartphone; social media; serious mental illness

Introduction

Estimates from 2018 suggest that 77% of the US population owns a smartphone [1]. The availability of smartphone apps designed to enhance wellness or support mental health has also grown exponentially, with over 10,000 available [2]. Studies document a strong interest in using smartphones to promote mental health in individuals with mental illness. In a

meta-analysis of individuals with psychosis, 60% endorsed interest in using their phone to monitor their mental health [3]. In addition, over half of them were interested in using their phones to obtain health care information, receive appointment or medication reminders, and facilitate contact with providers. Approximately 30% of participants recruited from community mental health centers reported using apps for enhancing physical wellness, primarily to support exercise, diet, and weight loss [4], although participants were not queried about using apps for specific mental health purposes. Individuals receiving outpatient psychiatric care report similar levels of interest in using smartphones to track mental health (56% to 84%) [5-7].

Beyond this handful of studies, most of the existing data on smartphone use in individuals with mental illness focus on social media. Reports of social media use among individuals with serious mental illness (SMI) range from 33% to 71%, with Facebook and Twitter identified as the most frequently used platforms [4,8-10]. Most users reported engaging with social media daily, with higher use for younger individuals [4,9-12]. Individuals with SMI reported using social media for reasons such as interacting with friends, family, or others outside the home as well as for searching for health-related information; in addition, interest in mental health programs delivered through social media was high [4,8,10,12].

Popular media has highlighted growing concern about the potential addictive nature of smartphone and social media use as well as the subsequent negative mental health effects. Indeed, problematic smartphone use has consistently been associated with higher levels of depression and anxiety in nonclinical samples [13,14], and greater time spent on social media was correlated with higher depression among individuals receiving inpatient psychiatric treatment [15]. In contrast, other studies indicate that individuals with mental illness disagree that smartphone and social media use necessarily negatively impact their lives. In 1 study, half of the individuals with schizophrenia agreed that social media use increased socialization and disagreed that it made their symptoms worse [10]. In a community clinic sample, social media use was not associated with the severity of psychiatric symptoms or quality of life, but instead, it was positively associated with greater community participation and civic engagement [9]; although in other studies, results have been more mixed [16]. Thus, although the potential clinical risks of smartphone and social media use warrant close monitoring and further study, individuals may also perceive benefit.

Overall, existing data suggest that individuals with mental illness increasingly own smartphones, regularly use social media, and express an interest in using smartphones and social media to support their health (though potential negative consequences also need to be considered). However, there remain several important gaps in this nascent field. First, most studies that have examined how individuals with mental illness use smartphone apps to promote their health have not focused on mental health specifically [4]. Second, although some initial studies investigated demographic and clinical correlates of smartphone *ownership*, only 1 study has characterized *use* in this manner, and it focused specifically on social media [9]. Finally, some previous studies pertaining to smartphone interest may be positively skewed. For example, in some studies, as few as 10% of patients at a clinic chose to complete the study surveys [5,6], and others surveyed individuals who were already volunteering for another study about a smartphone app [17]. Thus, studies assessing smartphone use are needed in larger naturalistic samples in which individuals complete measures as part of routine care and whose results therefore should be less affected by selection biases.

The main aim of this study was to provide a detailed account of general smartphone use (including social media) in individuals with mental illness. The secondary aim was to better understand current engagement with mental health apps specifically and to assess interest in using smartphone apps to support mental health. The results may inform efforts to harness digital technologies for delivering mental health interventions and guide research, clinical, and industry efforts [4,8,16,18]. To that end, we administered self-reported measures of smartphone use, social media use, and interest in mental health apps to all individuals attending a psychiatric partial hospital program. Given the demographic characteristics of this partial hospital program, we expected a very high rate of smartphone ownership; thus, making the sample well suited to explore patterns of use and interest in smartphone apps. On the basis of previous studies, we expected high rates of use across all 3 domains. We explored demographic and clinical correlates of smartphone use and expected that younger individuals would report higher use than older individuals [4,9,10,12]. The remaining analyses were descriptive and exploratory given the lack of previous research focused on these specific questions.

Methods

Participants and Treatment Setting

Participants included 322 patients receiving treatment at a partial hospitalization program (PHP) located in a nonprofit, insurance-based, academic psychiatric hospital in New England from September 2017 to March 2018. The PHP treats English-speaking adults (≥18 years) with a broad range of psychiatric disorders [19]. Participants were on average in young-to-middle adulthood (mean 33.49, SD 13.87 years), female (57%), and predominantly white (89%). Approximately half of the patients are referred directly from an inpatient hospitalization for further stabilization and transition to outpatient care. The other half of them are referred from outpatient providers in lieu of an inpatient hospitalization. Thus, to qualify for a partial hospital level of care, individuals are experiencing acute symptoms and significant impairment. Many patients are managing chronic mental illness, whereas others may be experiencing their first episode. The most common primary diagnosis (provided by program psychiatrists) was major depressive disorder (MDD; n=183), followed by bipolar disorder (n=72); anxiety, obsessive-compulsive disorder, or stress-related disorders (n=48); and psychotic disorder (n=19). See Table 1 for demographic characteristics.

Table 1. Demographic and clinical characteristics.

Demographic characteristic	Values
Age, years (mean, SD)	33.49 (13.87)
Gender, n (%)	
Female	183 (56.8)
Male	135 (41.9)
Gender fluid or nonbinary	4 (1.2)
Race	
White	287 (89.1)
Asian	12 (3.7)
Multiracial	11 (3.4)
Black	6 (1.9)
Did not specify	6 (1.9)
Ethnicity	
Non-Latinx	300 (93.2)
Latinx	21 (6.5)
Did not specify	1 (0.3)
Sexual orientation	
Heterosexual/straight	248 (77.0)
Bisexual	33 (10.3)
Gay/lesbian	18 (5.6)
Queer	11 (3.4)
Something else (asexual and pansexual)	12 (3.7)
Education	
High school/GED or less	17 (5.3)
Some college	120 (37.3)
4-year college graduate	91 (28.3)
Postcollege education	94 (29.2)
Current student	99 (30.7)
Employment	
Not employed	153 (47.5)
Employed part time	57 (17.7)
Employed full time	112 (34.8)
Marital status	
Never married	206 (64.0)
Separated/divorced or widowed	31 (9.6)
Married	72 (22.4)
Living with partner	13 (4.0)

Measures

Smartphone Use Survey

We developed a self-reported questionnaire to assess smartphone ownership, application usage, and interest in applications to support mental health based on a previous survey assessing smartphone usage in psychiatric outpatients [5,6]. The survey also asked participants to indicate which smartphone apps they perceived as "supportive of" versus "negatively affecting" their mental health (see Multimedia Appendix 1).

Mobile Technology Engagement Scale

The mobile technology engagement (MTE) [20] is a self-reported measure of mobile technology usage. First, the

social media usage component consists of 4 items assessing mobile use of social media platforms, including Facebook, Twitter, Instagram, and Snapchat, with each item anchored on a 7-point Likert-style scale. The social media usage score is calculated by summing the 4 items. Second, the *status updates* component consists of a single item ("How often do you post public updates—Facebook status, Tweets, Instagram uploads, etc") using a 7-point Likert-style scale, with the raw score indicating a status updates score. Third, the *phone-checking behavior* component consists of 3 items assessing relevant behaviors, with response options anchored on a 5-point Likert-style scale. The phone-checking behavior score is the mean of the 3 phone-checking behavior items. To calculate overall MTE scores, a z-score is calculated for social media usage, status updates, and phone-checking behavior, and then the mean of the z-scores is used as an individual's overall score, with higher scores indicating higher engagement. Internal consistency was acceptable (Cronbach alpha=.71).

Patient Health Questionnaire-9

The Patient Health Questionnaire-9 (PHQ-9) [21] is a self-reported measure that assesses the frequency of depressive symptoms over the past 2 weeks using a 4-point Likert-type scale anchored at 0 (*not at all*) and 3 *(nearly all the time*). A total PHQ-9 score is calculated by summing items, with higher scores indicating higher depression severity. The PHQ-9 has demonstrated strong psychometric properties in similar psychiatric populations [22]. In this sample, internal consistency was good (Cronbach alpha=.86).

Generalized Anxiety Disorder Scale—7 items

The Generalized Anxiety Disorder Scale—7 items (GAD-7) [23] is a self-reported questionnaire that assesses symptoms of generalized anxiety disorder over the past 2 weeks (0, not at all, to 3, nearly every day), with higher scores indicating greater anxiety severity. The GAD-7 has been validated as a measure of symptoms of general anxiety in a psychiatric hospital setting [23-26]. In this sample, internal consistency was good (alpha=.88).

Clinical Global Improvement Scale

The Clinical Global Improvement Scale [27] is a single item assessing patients' perceived improvement (or lack thereof) compared with their baseline status. In this study, patients rated their own improvement at discharge using the 7-point scale from 1, very much improved, to 7, very much worse. Patient ratings correlate moderately with provider ratings (Intraclass Correlation Coefficient=.65) and have comparable validity [28].

Procedure

As part of routine clinical care at the PHP, patients completed self-reported measures of symptoms and functioning. Patient data were originally used by their treatment providers for treatment planning and progress monitoring. Patient data are also used for program evaluation and quality assurance efforts. The current measures of smartphone use were included to inform ongoing treatment development and evaluation efforts at the PHP and, thus, did not require informed consent. Measures were administered using REDCap (Research Electronic Data Capture), a secure, Web-based application designed to support data collection for research studies [29].

We obtained a deidentified dataset from the PHP, and the local institutional review board deemed this research exempt. In this dataset, 10.9% of participants did not complete the discharge assessment for the following reasons: they were admitted to inpatient hospitalization (n=17), they unexpectedly did not attend the program on the day of discharge (n=14), or staff were unable to schedule assessment (n=4).

App Standardization and Categorization

We followed similar procedures as previously described [17]. We standardized the names of each mental health app. For example, Headspace, head space, and HeadSpace were all coded as the same app. We then categorized the most frequently reported apps according to their advertised purpose. To do this, 1 author (ALS) first generated potential categories based on the description in the Google Play store. Second, ALS and CB discussed the potential categories and came to a consensus about the final list as well as how each app was categorized (see Multimedia Appendix 2).

Data Analytic Strategy

We calculated descriptive statistics (frequency, %) for the smartphone use items. To facilitate comparisons with previous studies, we present the responses stratified by the age groups used by Torous et al (see Table 2) [5]. We explored whether demographic characteristics (age, gender, education, and race) and clinical characteristics (baseline depression and anxiety symptom severity, primary diagnosis [bipolar, MDD, anxiety, and psychotic], and treatment response [reported 1, very much, or 2, much improvement, vs other responses]) were associated with responses on the smartphone use survey and MTE. To simplify interpretation of results, we transformed the education variable from its 5 response options to a dichotomous variable (4-year college degree or more vs no college degree). Owing to the small number of individual racial categories endorsed, we used participants' responses (yes/no) to the *white* racial category to examine race.

Table 2. Frequency of smartphone app use in total sample and by age group.

Type of app	Never, n (%)	Rarely, n (%)	Sometimes, n (%)	Frequently, n (%)	Often, n (%)	Very often, n (%)
Texting apps (eg, WhatsApp)						
Total sample	18 (5.7)	6 (1.9)	13 (4.1)	33(10.5)	80 (25.4)	165 (52.4)
Under 30 years	5 (2.9)	3 (1.7)	7 (4)	14 (8.1)	39 (22.5)	105 (60.7)
31-45 years	6 (8)	2 (2)	3 (4)	8 (10)	24 (31)	34 (44)
46-60 years	5 (10)	0 (0)	3 (6)	7 (14)	12 (24)	24 (47)
Over 60 years	2 (14)	1 (7)	0 (0)	4 (29)	5 (36)	2 (14)
Phone/video apps (eg, Skype)						
Total sample	48 (15.3)	84 (26.8)	58 (18.5)	49 (15.7)	43 (13.7)	31 (9.9)
Under 30 years	18 (10.3)	41 (23.6)	34 (19.5)	31 (17.8)	29 (16.7)	21 (12.1)
31-45 years	15 (20)	21 (28)	14 (18)	11 (15)	9 (12)	6 (8)
46-60 years	12 (25)	16 (33)	8 (16)	7 (14)	3 (6)	3 (6)
Over 60 years	3 (21)	6 (43)	2 (14)	0 (0)	2 (14)	1 (7)
Email apps (eg, Outlook)						
Total sample	3 (.9)	9 (2.9)	22 (7)	45 (14.3)	111 (35.2)	125 (39.7)
Under 30 years	0 (0)	7 (4)	16 (9.2)	34 (19.5)	58 (33.3)	59 (33.9)
31-45 years	1 (1)	1 (1)	3 (4)	7 (9)	30 (39)	35 (46)
46-60 years	2 (4)	1 (2)	1 (2)	4 (8)	15 (30)	27 (54)
Over 60 years	0 (0)	0 (0)	2 (14)	0 (0)	8 (57)	4 (29)
Social media apps (eg, Facebook)						
Total sample	43 (13.8)	17 (5.4)	31 (9.9)	32 (10.3)	66 (21.2)	123 (39.4)
Under 30 years	11 (6.4)	4 (2.3)	20 (11.6)	16 (9.2)	40 (23.1)	82 (47.4)
31-45 years	17 (23)	4 (5)	3 (4)	9 (12)	17 (23)	25 (33)
46-60 years	11 (22)	6 (12)	5 (10)	4 (8)	8 (16)	16 (32)
Over 60 years	4 (29)	3 (21)	3 (21)	3 (21)	1 (7)	0 (0)
Calendar apps						
Total sample	46 (14.8)	30 (9.7)	41 (13.2)	54 (17.4)	83 (26.8)	56 (18.1)
Under 30 years	23 (13.4)	23 (13.4)	31 (18)	35 (20.3)	38 (22.1)	22 (12.8)
31-45 years	11 (15)	5 (7)	6 (8)	12 (16)	22 (29)	20 (26)
46-60 years	10 (21)	2 (42)	3 (6)	4 (8)	18 (38)	11 (23)
Over 60 years	2 (14)	0 (0)	1 (7)	3 (21)	5 (36)	3 (21)
Entertainment apps (eg, YouTube and radio)						
Total sample	28 (8.9)	41 (13)	56 (17.7)	62 (19.6)	72 (22.8)	57 (18)
Under 30 years	14 (8)	17 (9.8)	22 (12.6)	35 (20.1)	45 (25.9)	41 (23.6)
31-45 years	7 (9)	6 (8)	16 (21)	16 (21)	19 (25)	13 (17)
46-60 years	6 (12)	14 (28)	12 (24)	8 (16)	8 (16)	3 (6)
Over 60 years	1 (7)	4 (29)	6 (43)	3 (21)	0 (0)	0 (0)
Games						
Total sample	123 (39.7)	57 (18.4)	37 (11.9)	39 (12.6)	28 (9)	26 (8.4)
Under 30 years	60 (34.7)	32 (18.5)	21 (12.1)	24 (13.9)	19 (11.0)	17 (9.8)
31-45 years	31 (40)	17 (22)	13 (17)	7 (9)	6 (8)	3 (4)
46-60 years	23 (49)	7 (15)	2 (4)	7 (15)	2 (4)	6 (13)
Over 60 years	9 (69)	1 (8)	1 (8)	1 (8)	1 (8)	0 (0)

We conducted multiple ordinal logistic regression analyses with the 6 levels of frequency of use as the dependent variable (never, rarely, sometimes, frequently, often, and very often) and demographic and clinical variables as the predictors. All the odds ratios presented represent the effect of the predictor on the dependent variable holding all other predictor variables in the model constant. For each of these models, we tested the assumption of proportional odds (that the coefficients are equal across all levels) using the test of parallel lines in SPSS, and we report below whether this assumption was violated in any models.

We conducted multiple logistic regressions with interest and willingness to use apps to monitor mental health as the dependent variables (yes/no) and demographic and clinical variables as the predictors. Finally, we conducted multiple linear regression analyses to examine the demographic and clinical predictors of the number of mental health apps on patients' phones and MTE subscales.

Results

Aim 1: General Smartphone App and Social Media Use

Almost all the patients reported owning a smartphone (n=315/322, 98%): iPhone (n=250/322, 77.7%) and Android (n=65/322, 20.2%). The most frequently used types of apps were texting, email, and social media (see Table 2). Multimedia Appendix 3 presents the results for the MTE scale. Patients most commonly reported using Facebook, followed by Instagram, Snapchat, and Twitter.

Demographic and Clinical Associations

Overall, 3 of the models predicting frequency of the app use violated the assumption of proportional odds: calendar apps (χ^2_{40} 68.8; *P*=.003), games (χ^2_{40}=91.4; *P*<.001), and health and mental health apps (χ^2_{40}=56.9; *P*=.04). Thus, use caution when interpreting the effects for predicting use of those specific apps.

As expected, older age was associated with less frequent use of most apps: texting (OR 0.97, 95% CI 0.95-0.98; *P*<.001), phone/video (OR 0.97, 95% CI 0.95-0.98; *P*<.001), social media (OR 0.95, 95% CI 0.93-0.97; *P*<.001), entertainment (OR 0.96, 95% CI 0.94-0.98; *P*<.001), and games (OR 0.98, 95% CI 0.96-1.00; *P*=.04).

Gender also predicted the use of several types of apps, with women reporting more frequent use of social media (OR 1.92, 95% CI 1.21-3.05; *P*=.006) and physical and mental health apps (OR 1.67, 95% CI 1.05-2.64; *P*=.03) but less frequent use of entertainment apps (OR 0.49, 95% CI 0.31-0.76; *P*=.002) compared with men. Lower education was associated with less frequent use of email (OR 0.36, 95% CI 0.21-0.60; *P*<.001), calendar (OR 0.46, 95% CI 0.29-0.76; *P*=.002), and entertainment (OR 0.59, 95% CI 0.36-0.95; *P*=.03) apps. Finally, participants of color reported less frequent use of social media (OR 0.28, 95% CI 0.12-0.65; *P*=.003).

Regarding the clinical characteristics, absence of a primary diagnosis of MDD was associated with less frequent use of phone/video communication (OR 0.29, 95% CI 0.10-0.82; *P*=.02). Primary diagnosis was also associated with the use of entertainment apps, such that individuals who were not diagnosed with MDD (OR 0.29, 95% CI 0.10-0.80; *P*=.02), bipolar disorders (OR 0.29, 95% CI 0.10-0.83; *P*=.02), or anxiety disorders (OR 0.21, 95% CI 0.06-0.66; *P*=.008) used entertainment apps less frequently than individuals with these diagnoses. None of the other clinical variables (symptom severity and treatment responder) were associated with the frequency of app use.

Table 3 presents the regression models predicting the MTE subscales from demographic and clinical characteristics. More frequent social media use was predicted by younger age, female gender, white race, and not having a primary diagnosis of psychotic disorder (all *P*<.05). More frequent phone-checking behavior was predicted by younger age and greater anxiety severity. The only significant predictor of more frequent status updates was a primary diagnosis of bipolar disorder.

Table 3. Multiple regression analyses predicting mobile technology engagement subscales.

Variable	Social media usage[a]			Phone-checking[b] behavior			Status updates[c]		
	B [d]	Standard Error *B*	Beta	*B*	Standard error *B*	Beta	*B*	Standard Error *B*	Beta
Age (years)	−.032	0.004	−.444[e]	−.026	0.004	−.359[e]	−.005	0.005	−.070
Sex (male)	−.233	0.108	−.115[f]	0.127	0.114	0.063	−.179	0.12	−.091
Race (white)	0.422	0.203	.110[f]	0.394	0.215	0.103	−.077	0.225	−.021
Education	−.099	0.118	−.049	0.025	0.125	0.012	−.169	0.131	−.086
Baseline anxiety	0	0.013	−.002	0.028	0.014	.154[f]	0.008	0.015	0.045
Baseline depression	0.021	0.013	0.118	0.018	0.013	0.105	0.002	0.014	0.012
Primary bipolar disorder	−.032	0.134	−.013	−.026	0.142	−.011	0.412	0.149	.175[e]
Primary anxiety	−.015	0.17	−.005	−.111	0.18	−.036	0.045	0.189	0.015
Primary psychosis	−.676	0.25	−.150[g]	−.315	0.264	−.070	−.264	0.277	−.060
Treatment responder (yes)	0.21	0.115	0.097	0.103	0.121	0.048	0.2	0.127	0.095

[a]R^2=.293

[b]R^2=.204

[c]R^2=.075

[d]*B*: Unstandardized beta.

[e]*P*<.001.

[f]*P*<.05.

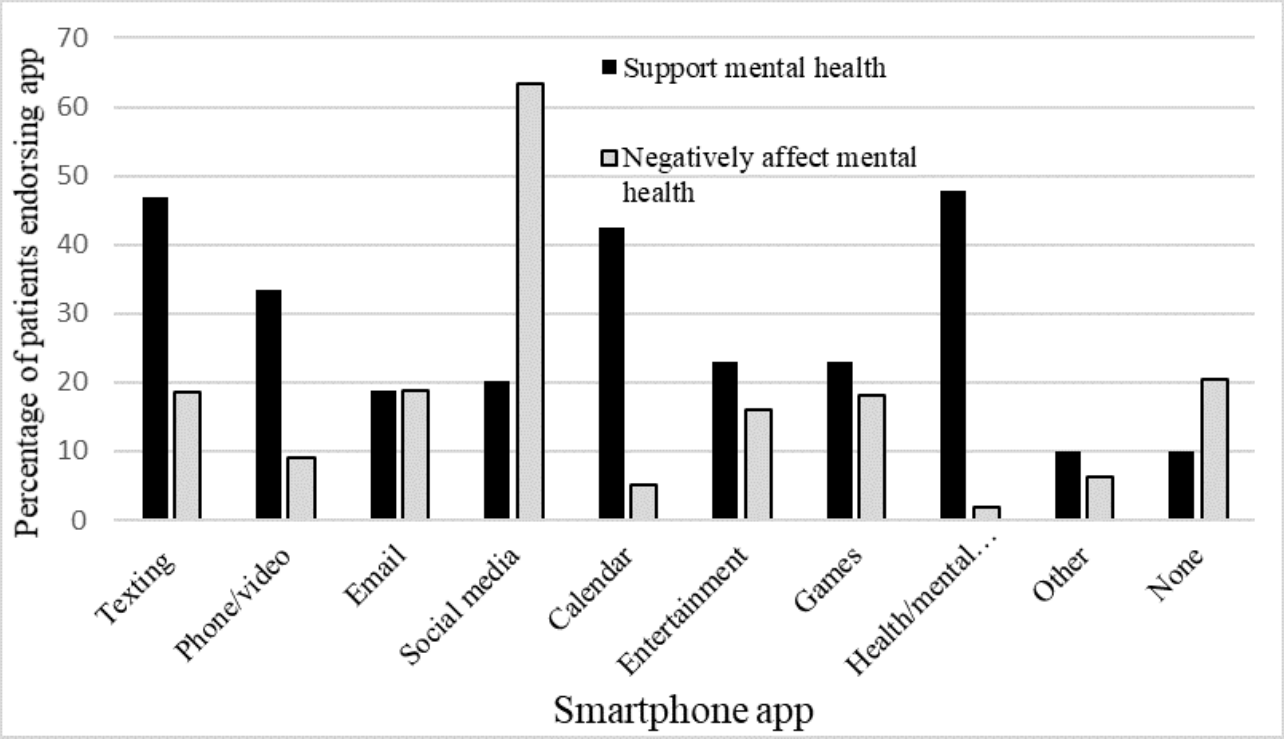
[g]*P*<.01.

Aim 2: Interest in and Use of Apps for Mental Health

Most participants reported they would want (n=238, 74%) and be willing (n=262, 81%) to use an app to monitor their mental health. Education was the only predictor of both interest (beta=−.69, SE=.31; *P*=.02; OR 0.50, 95% CI 0.28-0.91) and willingness (beta=−.73, SE .34; *P*=.03, OR 0.48, 95% CI 0.25-0.94) to use a smartphone app to monitor a mental health condition. The more educated group had higher rates of interest and willingness (79% and 85%, respectively) than the less educated group (67% and 77%, respectively).

Perceived helpfulness of general smartphone apps for mental health. Figure 1 presents the percentage of participants who endorsed the helpfulness and unhelpfulness of each app. The apps most frequently endorsed as being supportive of mental health were texting (47%) and calendars (43%). Few participants endorsed any apps as being harmful to their mental health, except for social media (63%).

Figure 1. Perceived helpfulness of smartphone apps for mental health.



Current Mental Health App Use

Approximately half of the patients (139/315, 44%) reported that they currently had at least 1 mental health app installed on their phone. None of the demographic or clinical characteristics significantly predicted the number of mental health apps downloaded (P>.08). Participants selected the primary purpose of their mental health apps as meditation/mindfulness (71%), mood tracking (10%), therapy skills (4%), and other (15%). The most commonly reported apps were Headspace (n=46), Calm (n=21), and Insight Timer (n=11; see Multimedia Appendix 2). These 3 apps were all coded as facilitating the practice of mindfulness/meditation. The 50+ other apps reported were used by 1 to 2 participants. Approximately half of the participants (154/322, 48%) endorsed health and mental health apps as being supportive of their mental health.

Discussion

Aim 1: General Smartphone App and Social Media Use

The most frequently used types of apps were texting, email, and social media. As expected and consistent with previous work, younger individuals reported more frequent smartphone use across most types of apps [4,9,10,12]. Education also predicted the frequency of use for some types of apps, specifically email, calendar, and entertainment apps. Baseline depression and anxiety severity were not associated with smartphone use nor was treatment response.

Regarding social media–specific behaviors, Facebook was the most commonly used social media app, followed by Instagram, Snapchat, and Twitter. This pattern aligns closely with previous research in similar samples [4,9,11,12]. Most patients reported checking their phone at least once per hour but posting rarely, potentially suggesting that our sample may primarily use social media passively and perhaps experience it as a platform for social comparison rather than for social connection or support [30,31]. Men and people of color reported less frequent use of social media. Younger individuals and those with greater anxiety checked their phones with greater frequency, and individuals with bipolar disorder made more frequent status posts relative to individuals with other primary diagnoses.

Aim 2: Interest in and Use of Apps for Mental Health

Most patients reported an interest (74%) and willingness (81%) to use a smartphone app to monitor their mental health. Overall, the rates of interest and willingness are slightly higher than those reported by individuals attending an outpatient psychiatric clinic [5]. In that study, 67% of patients expressed an interest in using a mobile app to track their medical condition (not mental health specifically) daily [5]. Although interest and willingness were high across education level, individuals with a higher level of education reported greater interest and willingness. This finding has implications for studies developing smartphone-delivered interventions; researchers should attend to educational background when developing and advertising such interventions. Specifically, although technology-driven interventions have the potential to reduce barriers to mental health treatment, they might also unintentionally further existing disparities in access. Future research should examine whether specific efforts are needed to encourage/assist individuals with lower levels of education to use smartphone app interventions for mental health purposes.

In contrast to the high rates of reported interest in mental health apps, only 44% of patients reported currently having a mental health app downloaded on their smartphone. This rate is higher than that obtained in a recent study of individuals with anxiety and depression enrolled in a smartphone treatment study (26%)

[17] and that in another study of psychiatric outpatients (10%) [32]. Participants who currently had mental health apps downloaded on their phones reported that they primarily used them for mindfulness or meditation (71%), with mood tracking being a distant second (10%). The uses of the self-reported app mirror the coded purposes of the most frequently reported mental health apps. Specifically, the top 3 apps reported (Headspace, Calm, and Insight Timer) are all designed to facilitate the practice of mindfulness/meditation.

Participants endorsed using general smartphone apps to support their mental health. Specifically, participants rated health care, calendar, and texting apps as the most supportive of their mental health, whereas most participants rated social media apps as negatively affecting their mental health. However, the frequency of social media use was not associated with symptom severity or treatment response in this study. Nevertheless, patients' subjective report that social media harms their mental health contrasts with results of previous research, in which individuals with SMI endorsed using social media for activities related to positive mental health, such as sharing things about themselves and feeling less lonely [9]. A recent systematic review suggested that different patterns of social media engagement may have varying impacts on mental health [33]. Experiences of social connectedness and support through social media were associated with positive mental health outcomes, whereas experiences of social comparison and rumination were associated with negative mental health outcomes. It may be that individuals in our sample are more prone to ruminative response styles and more likely to engage in social comparisons through social media, thereby perceiving a negative impact of social media on their mental health [30-31]. Individuals and societal perceptions of social media may also be quickly evolving as individuals are no longer discovering these apps but instead trying to figure out whether/how to use them in the long run. Future research in this area should carefully consider both the positive and negative clinical impacts of smartphone interventions, given the potential for addiction and negative influence on internalizing symptoms.

Together, the high interest in and perceived benefit of mental health apps, but only moderate current use, suggests a potential unmet treatment opportunity. These findings support ongoing efforts to develop and validate apps that support mental health. The naturalistic use of calendar and texting apps suggests that mental health care apps that leverage these types of functions might be especially feasible and acceptable in this population. These findings also support the development of methods for helping individuals identify existing evidence-based mental health care apps that might benefit them. Like matching an appropriate therapy or medication to the right individual is a personalized decision, there are ongoing efforts to support informed decision making around mental health apps and to help both clinicians and consumers identify safe, effective, engaging, and interoperable apps [34].

Although it is not a focus of this study, some of our findings provide support for ongoing efforts toward utilizing phone use behaviors (eg, number of social media posts, number of phone calls made, and Global Positioning System location,) in detecting mental illness phenotypes, deterioration, or relapse. Specifically, we found that a diagnosis of bipolar disorder predicted more frequent social media status updates and anxiety severity predicted more frequent phone-checking behavior. These effects mirror recent studies designed to identify personalized, objective markers of relapse via passively obtained indicators of smartphone behavior [35-36]. It might be helpful for clinicians and patients to attend to these objective indicators as part of relapse management. Individual monitoring of some phone use behaviors has become easy. For example, Apple iPhone users can monitor their phone use (eg, number of messages) through the Screen Time feature in Settings, and Android users can use apps such as Quality Time for the same purpose. Such monitoring may ultimately provide significant clinical benefit for some patients, especially if confidential methods for sharing data and alerting clinicians are developed.

Strengths and Limitations

All patients attending the psychiatric hospital program completed the measures. Thus, our results will likely generalize to other acute psychiatric populations and not only to individuals who choose to complete a survey on smartphone use. However, the results must be interpreted in the context of the hospital's demographic makeup, which is generally well educated and lacking ethnoracial diversity. In addition, although the high-smartphone ownership rate indicates that this sample was particularly well suited to examine smartphone use (ie, it would be impossible for people who do not own smartphones to answer most of the questions we asked), it also suggests that our findings may not generalize to other settings in which ownership is less common. Thus, these results regarding the frequency of use and interest in mobile health apps should be interpreted in the context of our sample of individuals who own smartphones rather than from the general population of individuals with mental illness. In addition, because of the need to minimize participant burden, we could not assess all potentially relevant variables. For example, it would have been helpful to assess the types of data plans and how these affected app use. Finally, we were unable to independently verify smartphone use, and future studies should include objective assessments.

Conclusions

People with mental illness increasingly have access to smartphones and are interested in apps to support mental health. However, the lower uptake of mental health apps today suggests the challenge of translating ongoing research on digital mental health into real-world tools that patients can easily access. This study suggests that among smartphone owners, there is high interest in using smartphone apps to improve mental health. Many nonmental health–specific apps (eg, calendars and texting) are perceived as beneficial for mental health, whereas social media apps are perceived as harmful to mental health. There is potential to optimize nonmental health–specific apps to better support the needs of those with mental illness and to design a new wave of mental health apps that match the needs of this population and the way they use these devices in daily life.

Acknowledgments

This study was supported by the Behavioral Health Partial Hospital Program at McLean Hospital. No external funding supported the study. The authors are grateful to the patients and staff of the Behavioral Health Partial Hospital Program at McLean Hospital for making this work possible and to the members of our patient advisory board for their valuable guidance in all our research. The authors thank Meghan O'Brien for her assistance with references.

Authors' Contributions

All authors contributed to the drafting and revision of the work, provided final approval of the paper, and are accountable for all aspects of the work. In addition, CB, ALS, MF, and TB contributed to the acquisition of the data. CB and MF contributed to data analysis and interpretation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Smartphone use survey.

[DOCX File, 13KB - mhealth_v7i6e13364_app1.docx]

Multimedia Appendix 2

App names and categories.

[DOCX File, 19KB - mhealth_v7i6e13364_app2.docx]

Multimedia Appendix 3

Responses to mobile technology engagement.

[DOCX File, 16KB - mhealth_v7i6e13364_app3.docx]

References

1. Pew Research Center. 2018. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-04-16] [WebCite Cache ID 77g7Y0tfs]
2. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
3. Firth J, Cotter J, Torous J, Bucci S, Firth JA, Yung AR. Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies. Schizophr Bull 2016 Mar;42(2):448-455. [doi: 10.1093/schbul/sbv132] [Medline: 26400871]
4. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
5. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]
6. Torous J, Friedman R. Smartphone use among patients age greater than 60 with mental health conditions and willingness to use smartphone applications to monitor their mental health conditions. Am J Geriatr Psychiatry 2014 Mar;22(3):S128-S129. [doi: 10.1016/j.jagp.2013.12.150]
7. Di Matteo D, Fine A, Fotinos K, Rose J, Katzman M. Patient willingness to consent to mobile phone data collection for mental health apps: structured questionnaire. JMIR Ment Health 2018 Aug 29;5(3):e56 [FREE Full text] [doi: 10.2196/mental.9539] [Medline: 30158102]
8. Lal S, Dell'Elce J, Malla AK. Technology access and use among young adults with a first episode of psychosis. Psychiatr Serv 2015 Jul;66(7):764-765. [doi: 10.1176/appi.ps.201400580] [Medline: 26130154]
9. Brusilovskiy E, Townley G, Snethen G, Salzer M. Social media use, community participation and psychological well-being among individuals with serious mental illnesses. Comput Human Behav 2016;65:232-240. [doi: 10.1016/j.chb.2016.08.036]
10. Miller BJ, Stewart A, Schrimsher J, Peeples D, Buckley PF. How connected are people with schizophrenia? Cell phone, computer, email, and social media use. Psychiatry Res 2015 Feb 28;225(3):458-463. [doi: 10.1016/j.psychres.2014.11.067] [Medline: 25563669]

11. Birnbaum ML, Rizvi AF, Correll CU, Kane JM. Role of social media and the internet in pathways to care for adolescents and young adults with psychotic disorders and non-psychotic mood disorders. Early Interv Psychiatry 2015 Mar 23;11(4):290-295. [doi: 10.1111/eip.12237] [Medline: 25808317]
12. Naslund JA, Aschbrenner KA, McHugo GJ, Unützer 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 2017 Oct 20 [Epub ahead of print]. [doi: 10.1111/eip.12496] [Medline: 29052947]
13. Elhai JD, Dvorak RD, Levine JC, Hall BJ. Problematic smartphone use: a conceptual overview and systematic review of relations with anxiety and depression psychopathology. J Affect Disord 2017 Jan 01;207:251-259. [doi: 10.1016/j.jad.2016.08.030] [Medline: 27736736]
14. Primack BA, Shensa A, Escobar-Viera CG, Barrett EL, Sidani JE, Colditz JB, et al. Use of multiple social media platforms and symptoms of depression and anxiety: a nationally-representative study among U.S. young adults. Comp Hum Behav 2017 Apr;69:1-9. [doi: 10.1016/j.chb.2016.11.013]
15. McDougall MA, Walsh M, Wattier K, Knigge R, Miller L, Stevermer M, et al. The effect of social networking sites on the relationship between perceived social support and depression. Psychiatry Res 2016 Dec 30;246:223-229. [doi: 10.1016/j.psychres.2016.09.018] [Medline: 27721061]
16. Gay K, Torous J, Joseph A, Pandya A, Duckworth K. Digital technology use among individuals with schizophrenia: results of an online survey. JMIR Ment Health 2016;3(2):e15 [FREE Full text] [doi: 10.2196/mental.5379] [Medline: 27146094]
17. Rubanovich CK, Mohr DC, Schueller SM. Health app use among individuals with symptoms of depression and anxiety: a survey study with thematic coding. JMIR Ment Health 2017 Jun 23;4(2):e22 [FREE Full text] [doi: 10.2196/mental.7603] [Medline: 28645891]
18. Torous J, Keshavan M. The role of social media in schizophrenia: evaluating risks, benefits, and potential. Curr Opin Psychiatry 2016 May;29(3):190-195. [doi: 10.1097/YCO.0000000000000246] [Medline: 26967314]
19. Forgeard M, Beard C, Kirakosian N, Björgvinsson T. Research in partial hospital settings. In: Codd R, editor. Practice-Based Research: A Guide for Clinicians. New York: Routledge; 2018:221-239.
20. Wilmer HH, Chein JM. Mobile technology habits: patterns of association among device usage, intertemporal preference, impulse control, and reward sensitivity. Psychon Bull Rev 2016 Dec;23(5):1607-1614. [doi: 10.3758/s13423-016-1011-z] [Medline: 26980462]
21. Kroenke K, Spitzer R, 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] [Medline: 11556941]
22. Beard C, Hsu KJ, Rifkin LS, Busch AB, Björgvinsson T. Validation of the PHQ-9 in a psychiatric sample. J Affect Disord 2016 Mar 15;193:267-273. [doi: 10.1016/j.jad.2015.12.075] [Medline: 26774513]
23. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
24. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
25. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
26. Beard C, Björgvinsson T. Beyond generalized anxiety disorder: psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. J Anxiety Disord 2014 Aug;28(6):547-552. [doi: 10.1016/j.janxdis.2014.06.002] [Medline: 24983795]
27. Guy W. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: US Department of Health, Education, and Welfare; 1976.
28. Forkmann T, Scherer A, Boecker M, Pawelzik M, Jostes R, Gauggel S. The Clinical Global Impression Scale and the influence of patient or staff perspective on outcome. BMC Psychiatry 2011 May 14;11:83-90 [FREE Full text] [doi: 10.1186/1471-244X-11-83] [Medline: 21569566]
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
30. Shaw A, Timpano K, Tran T, Joormann J. Correlates of Facebook usage patterns: the relationship between passive Facebook use, social anxiety symptoms, and brooding. Comput Human Behav 2015 Jul;48:575-580 [FREE Full text] [doi: 10.1016/j.chb.2015.02.003]
31. Verduyn P, Lee DS, Park J, Shablack H, Orvell A, Bayer J, et al. Passive Facebook usage undermines affective well-being: experimental and longitudinal evidence. J Exp Psychol Gen 2015 Apr;144(2):480-488. [doi: 10.1037/xge0000057] [Medline: 25706656]
32. Torous J, Wisniewski H, Liu G, Keshavan M. Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. JMIR Ment Health 2018 Nov 16;5(4):e11715 [FREE Full text] [doi: 10.2196/11715] [Medline: 30446484]

33. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
34. Torous JB, Chan SR, Gipson SYT, Kim JW, Nguyen T, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. Psychiatr Serv 2018 Feb 15;69(5):498-500. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
35. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
36. Ben-Zeev D, Brian R, Wang R, Wang W, Campbell AT, Aung MS, et al. CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatr Rehabil J 2017 Dec;40(3):266-275 [FREE Full text] [doi: 10.1037/prj0000243] [Medline: 28368138]

Abbreviations

GAD: Generalized Anxiety Disorder Scale
MDD: major depressive disorder
MTE: mobile technology engagement
PHP: partial hospitalization program
PHQ: Patient Health Questionnaire
SMI: serious mental illness

Edited by G Eysenbach; submitted 10.01.19; peer-reviewed by D Fulford, T Campellone, M Bardus; comments to author 28.03.19; revised version received 16.04.19; accepted 18.04.19; published 07.06.19.

Please cite as:
Beard C, Silverman AL, Forgeard M, Wilmer MT, Torous J, Bjo¨rgvinsson T
Smartphone, Social Media, and Mental Health App Use in an Acute Transdiagnostic Psychiatric Sample
JMIR Mhealth Uhealth 2019;7(6):e13364
URL: https://mhealth.jmir.org/2019/6/e13364/
doi:10.2196/13364
PMID:31199338

Original Paper

The Use of Smart Devices by Care Providers in Emergency Departments: Cross-Sectional Survey Design

Mohamad Alameddine[1], PhD; Hussein Soueidan[2], MBA; Maha Makki[3], MS; Hani Tamim[4], PhD; Eveline Hitti[3], MD

[1]American University of Beirut, Faculty of Health Sciences, Department of Health Management and Policy, Beirut, Lebanon

[2]American University of Beirut, Evidence Based Health Management Unit, Beirut, Lebanon

[3]American University of Beirut, Faculty of Medicine, Department of Emergency Medicine, Beirut, Lebanon

[4]American University of Beirut, Faculty of Medicine, Department of Internal Medicine, Beirut, Lebanon

Corresponding Author:
Eveline Hitti, MD
American University of Beirut
Faculty of Medicine
Department of Emergency Medicine
Bliss Street
Beirut, 11-0236
Lebanon
Phone: 961 350000
Email: eh16@aub.edu.lb

Abstract

Background: The use of smart devices (SDs) by health care providers in care settings is a common practice nowadays. Such use includes apps related to patient care and often extends to personal calls and applications with frequent prompts and interruptions. These prompts and interruptions enhance the risk of distractions caused by SDs and raise concerns about service quality and patient safety. Such concerns are exacerbated in complex care settings such as the emergency department (ED).

Objective: The objective of this study was to measure the frequency and patterns of SD use among health care providers in the ED of a large academic health center in Lebanon. The perceived consequences of care providers using SDs on provider-to-provider communication and the care quality of patients in the ED were assessed. Additionally, factors associated with the use of SDs and the approval for regulating such use were also investigated.

Methods: The study was carried out at the ED of an academic health center with the highest volume of patient visits in Lebanon. The data were collected using a cross-sectional electronic survey sent to all ED health care providers (N=236). The target population included core ED faculty members, attending physicians, residents, medical students, and the nursing care providers. The regression model developed in this study was used to find predictors of medical errors in the ED because of the use of SDs.

Results: Half of the target population responded to the questionnaire. A total of 83 of 97 respondents (86%) used one or more medical applications on their SDs. 71 out of 87 respondents (82%) believed that using SDs in the ED improved the coordination among the care team, and 71 out of 90 (79%) respondents believed that it was beneficial to patient care. In addition, 37 out of 90 respondents (41%) acknowledged that they were distracted when using their SDs for nonwork purposes. 51 out of 93 respondents (55%) witnessed a colleague committing a near miss or an error owing to the SD-caused distractions. Regression analysis revealed that age (P=.04) and missing information owing to the use of SDs (P=.02) were major predictors of committing an error in the ED. Interestingly, more than 40% of the respondents were significantly addicted to using SDs and more than one-third felt the need to cut down their use.

Conclusions: The findings of this study make it imperative to ensure the safety and wellbeing of patients, especially in high intensity, high volume departments like the ED. Irrespective of the positive role SDs play in the health care process, the negative effects of their use mandate proper regulation, in particular, an ethical mandate that takes into consideration the significant consequences that the use of SDs may have on care processes and outcomes.

(JMIR Mhealth Uhealth 2019;7(6):e13614) doi:10.2196/13614

KEYWORDS

health personnel; smart phones; emergency departments; healthcare quality; policy

Introduction

Background

A smart device (SD) is an electronic tool characterized by its mobility, connectivity, and potential to provide access to internet, geographical location, and relatively independent operation. SDs have the capability to gather information about their users and the environment to optimize users' experience [1,2]. Most commonly used SDs include smartphones, tablets, smart watches, and wearable devices, among others.

The use of SD in the health care industry is rising and gaining attention [3]. With over 70,000 health and medical SD apps available, their use for clinical practice has become common [4]. A study by Koehler shows that clinical applications include mobile result checking, drug references, clinical decision tools, case management tools, as well as staff scheduling modules, which have made SD usage integral in day to day clinical practice [5].

The use of SDs in the clinical setting, however, is associated with risks and concerns. Risks include potential breaches to patient privacy, as well as cross-contamination concerns related to cleanliness of the devices [5]. The negative impact on physician-patient relationship, as well as interprofessional relationship is another feared consequence [6,7]. Finally, there is growing evidence related to the distracting power of SD, which raises serious questions about the safety implications of use of SD in the health care settings where interruptions are common and the need to reduce error and improve safety is paramount [8].

Decision making in medicine requires significant cognitive focus integrating increasingly complex information from different sources, often under significant time constraints, in settings where interruptions are common. Previous research shows that interruptions are disruptive to care processes and may cause increased risks of error related to patient care, medication management, and communication among health care providers [9]. Emulating the aviation industry that has moved to *sterile cockpits* to reduce distractions, hospitals have started introducing separate medication rooms to reduce errors in mixing and preparing medications [10,11]. There is also increasing interest in creating similar distraction-free spaces for other high-risk activities in health care [12].

Emergency departments (EDs) are settings that are particularly prone to error, and the cases in the ED are often complex and the workflow dynamics and decisions are made under significant time pressures [13]. Furthermore, when compared with the interruptions in the primary care clinics, ED physicians are 3 times more likely to be interrupted [14]. Understanding the patterns of provider utilization of SDs in this high-risk setting and the perceptions of their impact on distraction and safety is an important first step in formulating policy or practicing recommendations for their use in health care settings.

Objectives

This study investigated self-reported health care providers' utilization of SDs in the ED and their perception regarding such usage on the patient-provider communication and on the safety of care at the ED of a large academic medical center in Beirut, Lebanon. The outcomes of the study will help in developing effective evidence-based policies to guide the use of SDs in clinical settings. The specific objectives of this study are as follows:

- Investigate the frequency and patterns of SD use among health care providers in the ED of a large academic health center in Lebanon.
- Understand the perceived consequences by care providers of SD use on patients and health care providers.
- Examine the factors associated with the use of SDs.
- Assess the perceptions toward potential SD regulatory policies in the ED of the target facility.

Methods

Ethical Approval

The study was approved by the Institutional Review Board (IRB) at the target institution under the protocol number ED.EH.06. The research team carried out the research work associated with this study after obtaining the IRB approval. ED providers were clearly instructed that this was an independent research study and that they had the full autonomy to participate without any rewards or consequences on their employment with the institution.

All participants had to read and sign an informed consent emphasizing that participating in this study was voluntary and that the participants had the full freedom to choose whether to participate or not. The participants were assured that their responses would be kept confidential and anonymous and that no one would have access to the data other than the researchers themselves. The participants were informed that the outcomes of the project would be used to guide the evidence-based use of SDs by health care providers in EDs.

Study Setting

The study was conducted in the ED of a large academic tertiary care hospital in Lebanon. The ED receives more than 57,000 visits per year and is the busiest ED in Lebanon. The ED includes 2 adult sections that separate patients according to their age and case acuity, and 1 pediatric section that receives all the pediatric cases. The ED is also considered an integral teaching setting in the education of medical students as more than 180 medical students and residents rotate through it every year.

The Survey Tool

The survey tool included, in addition to other items, questions on demographics, job position, years of experience, frequency of SD use, the types of used SDs, and applications. Using Likert scales, the health care professionals were asked to rate their perceptions toward using SDs, their perceived effect on patient care and team performance, addiction to using the SDs, and their opinion with regard to instituting a code of conduct to regulate the use of SD at the ED.

Adapted from the *Cut Down, Annoyed, Guilty and Eye Opener alcohol use disorders screening test*, the CAGE tool, the research team used a set of questions to measure the addiction score pertinent to using SDs. Similar uses of the

CAGE tool have been documented in previous research [15-18]. A respondent with a score higher than 2 positive answers out of the 5 questions was considered to be addicted to using SDs [17].

To enhance content validity, the final survey tool was reviewed by a group of experts including a health human resources management expert, a statistician, and a social scientist. The tool was also shared with international experts for opinion. All feedback was consolidated in the final version of the questionnaire, which was pilot tested on a number of care providers.

Data Collection

The email addresses of all ED providers were obtained from the ED management (as per the approved IRB protocol). ED providers were initially invited to participate in this study through an IRB-approved invitation email script. The email script included information about the study, the research team, what does it entail to participate, and the full respect for the freedom to participate, as well as the privacy and confidentiality of all answers. Should an invitee agree to participate, they were asked to click the Web-based survey link which started with an electronic consent form that needed to be signed before completing the questionnaire.

During the period January 2017 and September 2017, data were collected from ED health care providers using a 15-min electronic questionnaire on LimeSurvey (GmbH, Survey Services and Consulting, Hamburg, Germany), an electronic surveying platform.

The survey tool was made available only in English as all the target health care professionals were proficient in verbal and written English. A contact email address was provided to the participants in case they had any questions or inquires relevant to the study.

Study Population

A total of 236 health care providers working at the ED during the period of data collection were approached to participate in the study. Health care providers at the time of the study included 7 core ED faculty, 25 clinical associate faculty members, 53 ED residents, 62 nurses, 37 nursing students, and 52 rotating medical students. The invitation to participate in this study was extended to all health care providers working in the ED.

Data Analysis

The data retrieved from the LimeSurvey GmbH included 118 responses. The research team discarded 18 records as the respondents had missed more than 25% of the questions that belonged to main scales. After further examination of the dataset, the data collected from nursing students (n=3) had to be removed as such a small sample size rendered representativeness and comparability difficult. The final sample size included 97 records. IBM SPSS version 22 was used for analysis.

The continuous variables were grouped and summarized by frequency number and the same was done in the case of the categorical variables. Descriptive univariate and bivariate analyses were carried out to get insights into the different variables and the association between the dependent and the independent variables was obtained using the Pearson chi-square test. A *P* value less than .05 was considered statistically significant.

To define the predictors of making an error during an ED shift, the research team constructed a stepwise logistic regression model that included the effect of age and missing information as a result of using the SD as the predictors of making an error, while controlling for other variables (age, gender, job position, experience, whether the respondents had missed information because of using their SDs, using the SDs for personal reasons during the ED shift, and the addiction score). A *P* value less than .05 was considered to be statistically significant.

Results

Sample Demographics

The response rate in this study was 50%, with a total of 118 responses (out of 236 potential) received. However, after cleanup, 97 responses were eligible for inclusion in the final analysis.

The study sample was relatively young and equally distributed among males and females. The sample was distributed among the professions, including 32 nurses (32/97, 33%), 23 medical students (23/97, 24%), 22 residents (22/97, 23%), and 20 attending physicians (20/97, 20%). The majority of the respondents (82/97, 85%) had completed their training or residency in Lebanon. Furthermore, 19 out of 33 (58%) responded that they had been practicing professionally for more than 5 years, including almost one-third of them (10 out of 33, 30%) with experience over 10 years. Table 1 exhibits the detailed percentage distribution of the demographic and professional background of the survey respondents.

Table 1. Distribution of the demographic and professional background.

Demographic and professional background	n (%)
Age in years (N=97)	
<30	61 (63)
30-39	23 (24)
40-49	7 (7)
50-65	6 (6)
Gender (N=93)	
Male	47 (50)
Female	46 (50)
Discipline (N=97)	
Nurse	32 (33)
Medical student	23 (23)
Resident	22 (23)
Attending physician	20 (21)

Prevalence of Smart Device Use and Applications

All of the respondents reported owning a SD (Table 2). A total of 86% of the respondents (83 out of 97) used one or more medical applications on their SDs. Among those, 33 out of the 45 respondents (73%) used medical reference applications, such as Medscape, UpToDate, and others. The remaining used other applications including clinical decision applications (18%), pharmacology applications (11%), the ministry of public health application (9%), clinical practice applications (7%), and others (Table 3). The use of nonmedical applications was also very common. Among other applications, 86 out of 94 respondents (92%) used the phone for email access, 81 out of 96 respondents (84%) used their phones for Web access, 84 out of 96 used (88%) the messaging applications (including WhatsApp), 73 out of 96 (76%) used them for phone calls, 72 out of 95 respondents (76%) used the calendar, and 63 out of 95 (66%) used the social media applications (Table 4).

Table 2. Use of smart devices and medical applications.

Owning a smart device (N=97)	n (%)
Yes	97 (100)
No	0 (0)

Table 3. The most used medical applications.

Most used medical applications (N=45)	n (%)[a]
Medical reference apps	33 (73)
Clinical decision apps	8 (18)
Pharmacology apps	5 (11)
Ministry of public health app	4 (9)
Clinical practice apps	3 (7)
Other applications	4 (8)

[a]The values are not mutually exclusive.

Table 4. The most used nonmedical applications.

Most used nonmedical applications	n (%)
Email access (N=94)	86 (92)
Messaging applications (including WhatsApp; N=96)	84 (88)
Web access (N=96)	81 (84)
Phone calls (N=96)	73 (76)
Calendar (N=95)	72 (76)
Social media applications (N=95)	63 (66)

Reporting of Coworker Distraction and Error When Using the Smart Device in the Emergency Department

A total of 64% of respondents (59 out of 92) had witnessed a sizable proportion of trainees getting distracted when using their SDs for nonwork–related activities at the ED. Furthermore, 50 out 91 respondents (55%) witnessed the same for nurses and 44 out of 90 respondents (49%) witnessed attending physicians getting distracted when using their SDs for nonwork–related activities at the ED (Table 5). Consistently, 37 out of 90 respondents (41%) had acknowledged that using their SDs for nonwork–related purposes distracted them (Table 6).

In addition, 51 out of 93 respondents (55%) reported observing their colleagues having made an error or a near miss as a result of being distracted by their SDs (Table 5). Interestingly, only 15 out of 93 respondents (16%) acknowledged that they did make an error or a near miss as a result of being distracted by their SDs (Table 6).

Table 5. Performance in the emergency department when using the smart devices.

Reporting of coworker distraction and error when using the smart device in the emergency department	n (%)
I have witnessed a trainee (medical student, resident) become distracted when using his or her smart device for nonwork-related activities (N=92)	
Yes	59 (64)
No	33 (36)
I have witnessed a nurse become distracted when using his or her smart device for nonwork–related activities (N=91)	
Yes	50 (55)
No	41 (45)
Emergency department coworkers made an error or were about to make an error (near miss) as a result of being distracted by the use of the smart device during an emergency department shift? (N=93)	
Yes	51 (55)
No	42 (45)
I have witnessed an attending physician become distracted when using his or her smart device for nonwork–related activities (N=90)	
Yes	44 (49)
No	46 (51)

Table 6. Reporting of self-distraction and error when using the smart device in the emergency department.

Reporting of self-distraction and error when using the smart device (SD) in the emergency department (ED)	n (%)
Using my SD for non-work-related purposes distracts me (N=90)	
Yes	37 (41)
No	53 (59)
Have you ever made an error or were about to make an error (near miss) as a result of being distracted by the use of the smart device during an ED shift? (N=93)	
Yes	15 (16)
No	78 (84)

Impact of Using Smart Devices in the Emergency Department on the Quality of Care

A total of 71 out of 87 respondents (82%) confirmed that using SDs in the ED had improved care coordination among the health care providers and 71 out of 90 (79%) respondents believed that it was beneficial to patient care. Furthermore, 71 out of 89 respondents (80%) also agreed that using SDs at work assisted them in solving their personal issues.

Opinion and Attitudes Toward Using Smart Devices and the Need for a Code of Conduct

A total of 40 out of 91 respondents (44%; as shown in Table 7) indicated that patients feel *positive* when they see the ED providers using their SDs during their stay in the ED for *clinical purposes*, whereas 20 out of 91 (22%) responded that the patients feel *negative* and 31 out of the 91 (34%) answered *neutral*. Furthermore, 67 out of 91 respondents (74%) responded that the patients feel *negative* when they see the ED providers using their SD during their stay in the ED for *nonclinical purposes*.

Furthermore, 52 out of 91 respondents (57%) felt *neutral* toward seeing fellow colleagues using their SD during their ED shift for nonclinical purposes, whereas 34 of them (37%) felt *negative* and only 5 respondents (6%) felt *positive*.

The opinion was split with regard to medical administrators establishing a code of conduct for the use of SDs to minimize unnecessary distraction during ED shifts, with 51 out of 91 respondents (56%) disagreeing and 40 of them (44%) being in agreement (Table 8).

Smart Device Use Addiction Score

A total of 36 out of 90 respondents (40%) were found to be significantly addicted to their SDs (Table 9), with 62 out of 91 (68%) reaching for their SDs first thing in the morning and 32 out of 90 respondents (36%) feeling the need to cut down on the use of their SDs. Interestingly, 73 out of 91 respondents (80%) did not feel guilty about their overuse of the SDs at work nor felt annoyed when people criticized their use of the SDs (Table 10).

Table 7. Perceptions toward smart device use in the emergency department.

Perceptions toward smart device (SD) use in the emergency department (ED)	n (%)
How do you think patients feel when they see the ED providers using their SD during their stay in the ED for clinical purposes? (N=91)	
Positive	40 (44)
Negative	20 (22)
Neutral	31 (34)
How do you feel about fellow colleagues using their SD during their ED shift for nonclinical purposes? (N=91)	
Positive	5 (6)
Negative	34 (37)
Neutral	52 (57)
How do you think patients feel when they see the ED providers using their SD during their stay in the ED for nonclinical work–related purposes (check/send personal emails to friends or family)? (N=91)	
Positive	2 (2)
Negative	67 (74)
Neutral	22 (24)

Table 8. Opinion and attitudes toward using smart devices in the emergency department and the need for a code of conduct.

The need for a code of conduct	n (%)
Do you think that medical administrators should establish a code of conduct for the use of smart devices to minimize unnecessary distraction during emergency department shifts? (N=91)	
Yes	40 (44)
No	51 (56)

Table 9. Smart device use addiction score.

Smart device use addiction score—the modified *Cut Down, Annoyed, Guilty and Eye Opener alcohol use disorders screening test* CAGE tool (N=90)	n (%)
Respondents are clinically addicted to using smart devices	
Yes	36 (40)
No	54 (60)

Table 10. Smart device use addiction score details.

Smart device use addiction score details—the modified *Cut Down, Annoyed, Guilty and Eye Opener alcohol use disorders screening test* CAGE tool	n (%)
Do you reach for your smart device first thing in the morning?	
Yes	62 (68)
No	29 (32)
Have you ever felt you needed to cut down on use of your smart device?	
Yes	32 (36)
No	58 (64)
Have people annoyed you by criticizing your use of your smart device?	
Yes	18 (20)
No	73 (80)
Have you felt guilty about your overuse of your smart device at work?	
Yes	18 (20)
No	73 (80)

Predictors of Making an Error

The descriptive analysis of the health care providers making an error as a result of being distracted by their SDs was carried out (Multimedia Appendix 1).

Consistently, the acknowledgment of making an error was higher among the practitioners with more than 10 years of experience (3/9, 33%) compared with 1 out of 9 (11%) with 6 to 10 years of experience. The difference between the 2 groups was statistically significant (*P*=.049).

Interestingly, all the 15 respondents (100%) who acknowledged making an error used their SDs for personal matters during their ED shift (*P*=.04). In addition, out of those who acknowledged making an error, 3 out of 15 (20%) did acknowledge missing information when using their SDs during their ED shift (*P*=.03).

Furthermore, 10 out of 14 (71%) respondents who acknowledged making an error supported the need to implement a code of conduct to regulate the use of SDs in the ED (*P*=.03).

The research team developed a stepwise multivariate logistic regression model (model fitness=0.91) to help define the predictors of making an error because of using SDs in the ED (Table 11). The model shows that the respondents above 40 years of age were 5 times more likely to acknowledge making an error because of using SDs in the ED (*P*=.04), and those who missed information while using their SDs were 11 times more likely to commit an error (*P*=.02).

Table 11. Stepwise multivariate logistic regression of predictors of making error. Variables included in the stepwise regression model were with the following covariates: age (reference: <30 years); gender (reference: female); job (reference: medical student); missed information of using my smart device; used for personal purposes during your shift in the emergency department; addiction score.

Making error[a] (reference: no)	Statistical values	
	Odds ratio (95% CI)	*P* value
Age (≥40 years)	5.03 (1.09-23.15)	.04
Missed information of using my smart device	10.84 (1.49-78.91)	.02

Discussion

Principal Findings

This study revealed that the majority of the health care providers who responded to the survey did use an SD during their work shift and that a sizable proportion of respondents reported one of their colleagues or themselves getting distracted because of using SDs at the ED. Using a regression model, the analysis done by the research team shows that being above 40 years of age (*P*=.04) and missing information because of using SD (*P*=.02) are major predictors of committing an error during the work at the ED.

Findings of this study concur with those of other studies that revealed that SDs were being abundantly used in health care settings [19] and that the health care providers were aware of the medical applications [5]. In addition, the surveyed providers did hold positive views on the use of SDs with the majority stating that SDs enabled better team coordination, cohesion, and patient care. This study established that all health care providers used SDs during work shifts despite working in one of the busiest ED settings in the country. Such uses are not restricted to clinical applications that potentially support patient care, but also extend to personal uses in care settings. The propensity and the positive views on the use of SDs in patient care settings suggest that this is not a passing trend but rather one that is integrated into the health care practice environments

nowadays. Health care practitioners are very much accustomed to using their SDs and hold positive views of such use.

However, the positive attitude expressed by the ED care providers is counterbalanced by some disconcerting findings. Although our results are based on self-reported distraction, the findings are in line with studies that have shown that the use of cell phones during functions that require cognitive focus increases reaction time, lowers performance, and reduces situational awareness through *inattentional blindness* to surrounding activities [20]. This distraction potential coupled with the growing body of literature on the addictive potential of SDs with reported addiction rates ranging between 5% to 65% are concerning within a health care context. In fact, 55% of the respondents in our study reported one of their ED colleagues making an error or a near miss related to SD usage. However, respondents fell short in acknowledging making an error or a near miss themselves (15/93, 16%). Such acknowledgment of medical errors and near misses is likely to be an underestimate of the actual rate. This is because it is difficult for health care providers to acknowledge their own medical errors owing to concerns about professional discipline and job security [21]. The propensity of errors and near misses caused by SD distraction is potentially detrimental to patient care and calls for immediate and affirmative action by the ED managers and stakeholders. The study makes it evident that using SDs is negatively affecting patient care outcomes irrespective of other reported positive effects of use.

The Generation Gap

The findings of this study reveal that SD users tend to be younger and that they are more likely to accept the use of SDs in health care settings. Literature reports a generational gap with the use of SDs and technologies [22]. In our study, this gap has impacted the providers' perceptions of SD use in the ED setting and their reporting of distractions as a result of using it, as well as their acknowledgment of error.

The results of the regression have clearly showed that ED providers aged 40 years or above are 5 times more likely to acknowledge making an error owing to the use of SDs in the ED. This could either be attributed to the older providers feeling less friendly toward technology that enhances technology-induced errors, or that they are just more open and honest about themselves. This is in concordance with studies that established that older people are relatively less likely to accept SDs, which could be because of the age-related decline in cognitive abilities required for the proper use of SDs [23,24]. In addition to the age effect, older adults are less experienced with SDs than younger adults [25] and have concerns over their security and privacy, which may impact their use of SD and may restrict them from trying new functions, especially in the absence of onsite assistance [25,26].

This generation gap has also impacted the answers toward the need for a code of conduct to regulate the use of SDs in the ED. The providers who are older than 30 years supported a code of conduct compared with their younger counterparts. This difference in the way respondents viewed the regulation measures is significant (P=.03) and should be addressed by the policy makers while designing those policies and protocols. In the health care context, the different aspects of perceived usability and ease of use are highly relevant to performing the different health care tasks, and thus engaging the relatively older providers in designing and implementing SD use protocols in health care settings would probably contribute to a better acceptance and adoption of mobile health care technologies in the hospital setting [27].

The Effect of Smart Device-Driven Interruptions on Near Misses and Medical Errors

This study revisits the impact of using SDs on the quality of the health care service in a critically busy environment. Consistent with the fact that the increasing popularity of SD use in clinical settings may lead to increased distractions [6], the majority of respondents in this study have reported themselves and others getting distracted as a result of using SDs at the ED. This is highly relevant in the ED setting which demands that health care providers stay consistently focused on their jobs. What compounds the challenge is that using SDs in the clinical settings is exclusive not only to clinical applications, but also for personal purposes which could distract providers from vital patient care operations [28].

The fact that providers who acknowledged missing information because of using their SDs were 11 times more likely to acknowledge an error is quite disconcerting and requires quick action. This directly links the SD distractions to patient safety and poses serious questions on the potential negative effects of using such devices in similar busy settings. The literature is rich in examples of mobile device distractions proving detrimental to patient health and service quality, including procedural failures and decrease in care quality because of distractions by SD in clinical settings [6]. This has led the Emergency Care Research Institute (ECRI) to mention *caregiver distractions from SD and other mobile devices* as one of the top 10 hazards for 2013 [29].

The effect of distractive use of SDs in the clinical settings could be overcome by considering the human and technical factors which could be potentially compromised. In addition to the required security and disinfection measures, certain regulation procedures can be adopted to minimize the negative effects of using SDs in the clinical settings and promote the culture of using such devices for clinical purposes in a systematic way that could positively contribute to the provision and delivery of the health care service.

Regulating the Use of Smart Devices in the Health Care Settings

The use of SDs in health care settings is likely to grow in the future and thus necessitates proper regulating protocols [30]. Although SDs hold a lot of potential benefits to health care, it is critical to consider the potential harm to the process quality and thus there is need to regulate the use of SD in health care settings [30]. With around 68% of the respondents reaching out to their SDs first thing in the morning, the results of this study confirm such an addictive behavior among the health care providers.

Regulating the use of SDs in emergency and other health care settings is not an easy task. Completely banning the use of SDs

in health care institution is a difficult and highly unpopular option. In addition, it is important to recognize that imposing rules on using SDs for nonwork–related purpose would not be easy if the privately owned SDs are allowed within the clinical setting [5].

The analysis conducted by Gill and colleagues presents an interesting framework to regulate SD use in health care settings. The proposed framework aims to minimize distractions that are because of using SDs in the health care setting, keeping the patient health outcomes at the core. In addition to other measures, the framework allocates time to train the providers on security concerns and expectations, establishes clear security measures over the institutional network, and regulates the use of nonwork–related applications [6].

One of the interesting suggestions would be to set the basis for such a regulation culture during the medical education journey. Previous research suggests that hospitals should set an agenda to raise awareness on the proper use of SDs and promote professional use [31].

Regulating the use of SDs in health care is not an easy task, and hospitals should not expect it to see positive outcomes on the short run. This process is challenging as it aims at addressing addiction to SDs. This calls for a concerted multistakeholder effort involving medical organizations, health care providers, educators, and industry as well to improve the integration of SDs in the daily medical practice in an efficient, reliable, and safe way [32].

Limitations

The study has a number of shortcomings that are worth mentioning. First, despite the research team assuring respondents of the confidentiality of their responses, there remains a risk of a social desirability bias with health care providers potentially modifying their responses to reflect favorable SD use. Second, the cross-sectional nature of the study could only establish significant associations with causality requiring other types of study designs (eg, longitudinal). Third, the study questionnaire, which has been reviewed by an expert panel and pilot tested, did not undergo validation in terms of comparison with the gold standard assessment tool. Finally, the relatively low sample size may not have provided the required power to reveal significant differences. It is thus recommended that the study is replicated at a larger scale to unearth significant associations between variables.

Conclusions and Future Implications

With the wide proliferation and use of SDs in health care settings, it is imperative to ensure the safety and wellbeing of patients, particularly in high intensity, high volume department such as the ED. This study sheds light on the critical effects of using SDs in EDs, including the predictors and causes of making an error and its potential consequences on the care process from the perspective of the health care providers. Irrespective of the positive role the SDs play in the health care process, the negative effects of their use still mandate proper regulation. This is an ethical mandate taking into consideration the important consequences such use may have on care processes and outcomes.

Acknowledgments

The authors wish to acknowledge the valuable support of the administration of the Academic Health Center where the study took place. Gratitude also goes to every ED health care provider taking the time to participate in this study despite their very busy schedules. In addition, the authors wish to thank the funding of this research under the Medical Practice Plan Fund of the Faculty of Medicine of the Institution where the study took place.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Descriptive analysis of the factors associated with making an error and/or acknowledging it.

[PDF File (Adobe PDF File), 131KB - mhealth_v7i6e13614_app1.pdf]

References

1. Shin D. Cross - analysis of usability and aesthetic in smart devices: what influences users' preferences? Cross Cult Manage 2012 Oct 19;19(4):563-587. [doi: 10.1108/13527601211270020]
2. Lazar A, Koehler C, Tanenbaum J, Nguyen D. Why we use and abandon smart devices. 2015 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan.
3. Short SS, Lin AC, Merianos DJ, Burke RV, Upperman JS. Smartphones, trainees, and mobile education: implications for graduate medical education. J Grad Med Educ 2014 Jun;6(2):199-202 [FREE Full text] [doi: 10.4300/JGME-D-13-00238.1] [Medline: 24949119]
4. Statista. Number of mHealth apps available at Google Play from 2nd quarter 2015 to 3rd quarter 2018 URL: https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/ [accessed 2019-04-29] [WebCite Cache ID 780IlFjxP]

5. Koehler N. Healthcare professionals' use of mobile phones and the internet in clinical practice. JournalMTM 2013 Apr;2(1):3-13 [FREE Full text] [doi: 10.7309/jmtm.76]
6. Gill PS, Kamath A, Gill TS. Distraction: an assessment of smartphone usage in health care work settings. Risk Manag Healthc Policy 2012;5:105-114. [doi: 10.2147/RMHP.S34813] [Medline: 22969308]
7. Alsos OA, Das A, Svanæs D. Mobile health IT: the effect of user interface and form factor on doctor-patient communication. Int J Med Inform 2012 Jan;81(1):12-28. [doi: 10.1016/j.ijmedinf.2011.09.004] [Medline: 21996624]
8. Katz-Sidlow RJ, Ludwig A, Miller S, Sidlow R. Smartphone use during inpatient attending rounds: prevalence, patterns and potential for distraction. J Hosp Med 2012 Oct;7(8):595-599. [doi: 10.1002/jhm.1950] [Medline: 22744793]
9. Rivera-Rodriguez A, Karsh BT. Interruptions and distractions in healthcare: review and reappraisal. Qual Saf Health Care 2010 Aug;19(4):304-312 [FREE Full text] [doi: 10.1136/qshc.2009.033282] [Medline: 20378621]
10. Kapur N, Parand A, Soukup T, Reader T, Sevdalis N. Aviation and healthcare: a comparative review with implications for patient safety. JRSM Open 2016 Jan;7(1):2054270415616548 [FREE Full text] [doi: 10.1177/2054270415616548] [Medline: 26770817]
11. Huckels-Baumgart S, Baumgart A, Buschmann U, Schüpfer G, Manser T. Separate medication preparation rooms reduce interruptions and medication errors in the hospital setting: a prospective observational study. J Patient Saf 2016 Dec 21. [doi: 10.1097/PTS.0000000000000335] [Medline: 28009601]
12. Connor JA, Ahern JP, Cuccovia B, Porter CL, Arnold A, Dionne RE, et al. Implementing a distraction-free practice with the red zone medication safety initiative. Dimens Crit Care Nurs 2016;35(3):116-124. [doi: 10.1097/DCC.0000000000000179] [Medline: 27043397]
13. Berg LM, Källberg AS, Göransson KE, Östergren J, Florin J, Ehrenberg A. Interruptions in emergency department work: an observational and interview study. BMJ Qual Saf 2013 Aug;22(8):656-663. [doi: 10.1136/bmjqs-2013-001967] [Medline: 23584208]
14. Chisholm C, Dornfeld A, Nelson D, Cordell W. Work interrupted: a comparison of workplace interruptions in emergency departments and primary care offices. Ann Emerg Med 2001 Aug;38(2):146-151. [doi: 10.1067/mem.2001.115440] [Medline: 11468609]
15. Alosaimi F, Alyahya H, Alshahwan H, Al Mahyijari N, Shaik S. Smartphone addiction among university students in Riyadh, Saudi Arabia. Saudi Med J 2016 Jun;37(6):675-683 [FREE Full text] [doi: 10.15537/Smj.2016.6.14430] [Medline: 27279515]
16. Alrobai A, McAlaney J, Phalp K, Ali R. Online peer groups as a persuasive tool to combat digital addiction. 2016 Presented at: International Conference on Persuasive Technology; April 5-7, 2016; Salzburg, Austria.
17. Ewing J. Detecting alcoholism. The CAGE questionnaire. J Am Med Assoc 1984 Oct 12;252(14):1905-1907. [Medline: 6471323]
18. Hormes JM, Kearns B, Timko CA. Craving Facebook? Behavioral addiction to online social networking and its association with emotion regulation deficits. Addiction 2014 Dec 10;109(12):2079-2088. [doi: 10.1111/add.12713] [Medline: 25170590]
19. Tran K, Morra D, Lo V, Quan SD, Abrams H, Wu RC. Medical students and personal smartphones in the clinical environment: the impact on confidentiality of personal health information and professionalism. J Med Internet Res 2014 May 22;16(5):e132 [FREE Full text] [doi: 10.2196/jmir.3138] [Medline: 24855046]
20. De-Sola Gutiérrez J, de Fonseca FR, Rubio G. Cell-phone addiction: a review. Front Psychiatry 2016;7:175 [FREE Full text] [doi: 10.3389/fpsyt.2016.00175] [Medline: 27822187]
21. Mrayyan MT, Shishani K, Al-Faouri I. Rate, causes and reporting of medication errors in Jordan: nurses' perspectives. J Nurs Manag 2007 Sep;15(6):659-670. [doi: 10.1111/j.1365-2834.2007.00724.x] [Medline: 17688572]
22. Lane WC. The impact of personality traits on smartphone ownership and use. Int J Business Soc Sci 2011;2:22 [FREE Full text]
23. Sonderegger A, Schmutz S, Sauer J. The influence of age in usability testing. Appl Ergon 2016 Jan;52:291-300. [doi: 10.1016/j.apergo.2015.06.012] [Medline: 26360221]
24. Ma Q, Chan AH, Chen K. Personal and other factors affecting acceptance of smartphone technology by older Chinese adults. Appl Ergon 2016 May;54:62-71. [doi: 10.1016/j.apergo.2015.11.015] [Medline: 26851465]
25. Zhou J, Rau PP, Salvendy G. Older adults' use of smart phones: an investigation of the factors influencing the acceptance of new functions. Behav Inform Technol 2013 Apr 9;33(6):552-560. [doi: 10.1080/0144929x.2013.780637]
26. Gao S, Yang Y, Krogstie J. The adoption of smartphones among older adults in China. 2015 Mar 20 Presented at: International Conference on Informatics and Semiotics in Organisations; August 1-3, 2015; Toulouse, France. [doi: 10.1007/978-3-319-16274-4_12]
27. Plaza I, Martín L, Martin S, Medrano C. Mobile applications in an aging society:sStatus and trends. J Syst Soft 2011 Nov;84(11):1977-1988. [doi: 10.1016/j.jss.2011.05.035]
28. McBride DL. Distraction of clinicians by smartphones in hospitals: a concept analysis. J Adv Nurs 2015 Apr 21. [doi: 10.1111/jan.12674] [Medline: 25898861]
29. ECRI Institute. Top 10 health technology hazards for 2014. Health Devices 2013 Nov;42(11):354-380. [Medline: 24358513]
30. Thomairy NA, Mummaneni M, Alsalamah S, Moussa N, Coustasse A. Use of smartphones in hospitals. Health Care Manag (Frederick) 2015;34(4):297-307. [doi: 10.1097/HCM.0000000000000080] [Medline: 26506291]

31. Angarita F, Strickland M, Acuna S. Incorporating smartphones into clinical practice. Ann Med Surg (Lond) 2015 Jun;4(2):187-188 [FREE Full text] [doi: 10.1016/j.amsu.2015.05.005] [Medline: 26029364]
32. Sclafani J, Tirrell TF, Franko OI. Mobile tablet use among academic physicians and trainees. J Med Syst 2013 Feb;37(1):9903 [FREE Full text] [doi: 10.1007/s10916-012-9903-6] [Medline: 23321961]

Abbreviations

CAGE: Cut Down, Annoyed, Guilty and Eye Opener alcohol use disorders screening test
ED: emergency department
IRB: Institutional Review Board
SD: smart device

Edited by G Eysenbach; submitted 04.02.19; peer-reviewed by JM Wakim, B Chaudhry; comments to author 27.03.19; revised version received 04.04.19; accepted 04.04.19; published 05.05.19.

Please cite as:
Alameddine M, Soueidan H, Makki M, Tamim H, Hitti E
The Use of Smart Devices by Care Providers in Emergency Departments: Cross-Sectional Survey Design
JMIR Mhealth Uhealth 2019;7(6):e13614
URL: https://mhealth.jmir.org/2019/6/e13614/
doi:10.2196/13614
PMID:31199328

Original Paper

Factors Determining Patients' Choice Between Mobile Health and Telemedicine: Predictive Analytics Assessment

Saif Khairat[1,2], MPH, PhD; Songzi Liu[3], BA; Tanzila Zaman[1], BSc; Barbara Edson[4], RN, MBA, MHA; Robert Gianforcaro[4], DO

[1]Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]UNC Health Care, Chapel Hill, NC, United States

Corresponding Author:
Saif Khairat, MPH, PhD
Carolina Health Informatics Program
University of North Carolina at Chapel Hill
428 Carrington Hall
Chapel Hill, NC, 27514
United States
Phone: 1 9198435413
Email: saif@unc.edu

Abstract

Background: The solution to the growing problem of rural residents lacking health care access may be found in the use of telemedicine and mobile health (mHealth). Using mHealth or telemedicine allows patients from rural or remote areas to have better access to health care.

Objective: The objective of this study was to understand factors influencing the choice of communication medium for receiving care, through the analysis of mHealth versus telemedicine encounters with a virtual urgent clinic.

Methods: We conducted a postdeployment evaluation of a new virtual health care service, Virtual Urgent Clinic, which uses mHealth and telemedicine modalities to provide patient care. We used a multinomial logistic model to test the significance and predictive power of a set of features in determining patients' preferred method of telecare encounters—a nominal outcome variable of two levels (mHealth and telemedicine).

Results: Postdeployment, 1403 encounters were recorded, of which 1228 (87.53%) were completed with mHealth and 175 (12.47%) were telemedicine encounters. Patients' sex (P=.004) and setting (P<.001) were the most predictive determinants of their preferred method of telecare delivery, with significantly small P values of less than .01. Pearson chi-square test returned a strong indication of dependency between chief concern and encounter mediums, with an extremely small P<.001. Of the 169 mHealth patients who responded to the survey, 154 (91.1%) were satisfied by their encounter, compared with 31 of 35 (89%) telemedicine patients.

Conclusions: We studied factors influencing patients' choice of communication medium, either mHealth or telemedicine, for a virtual care clinic. Sex and geographic location, as well as their chief concern, were strong predictors of patients' choice of communication medium for their urgent care needs. This study suggests providing the option of mHealth or telemedicine to patients, and suggesting which medium would be a better fit for the patient based on their characteristics.

(JMIR Mhealth Uhealth 2019;7(6):e13772) doi:10.2196/13772

KEYWORDS

mHealth; telemedicine; urgent care; predictive analytics

Introduction

Background

In the United States, approximately 19.3% of the population live in rural areas. With only 9% of the nation's physicians practicing in such communities, the lack of health care providers in rural areas tends to be an intractable problem [1,2], causing rural residents to have a significantly lower health status than urban residents [3,4]. Aside from a shortage of health care staff, barriers to care due to the isolated location of residents and the lack of technology result in a poor quality of health care among rural populations [5,6]. Rural residents tend to use health care less due to the remoteness of where they live. For instance, colon cancer rates are high among rural residents, suggesting that they are less likely to receive timely cancer screening tests [7-10]. Rural populations are also at higher risk not only of cancer, but also of coronary heart disease [11]. Patients living close to a clinic tend to visit a health care provider more often than do patients living in rural areas [12]. Failure to obtain care on time may lead to a poor prognosis. These barriers to care for rural residents correlate with Hart's inverse care law, which states that underserved populations have the worst access to health care [1,13]. Therefore, it is imperative to provide care to underserved populations.

The solution to the growing problem of rural residents lacking health care access may be found in the use of telemedicine and mobile health (mHealth). In telemedicine, the doctor-patient interaction is conducted by live video consultation [14,15]. Telemedicine not only improves health care accessibility for patients living in rural areas, but it is also expected to save US $4.3 billion annually [16,17]. Another method of providing care is through mHealth. mHealth is the use of mobile devices such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices to provide medical care [18,19]. These two methods of providing care to patients remotely save significant travel costs for patients and their families, ensure that patients are seen in a timely manner, and help in-person care clinics or hospitals by reducing patient load [20].

Telemedicine is used in rural areas to educate patients, deliver teaching programs, and facilitate administrative meetings [21]. These help to reduce costs and save time. Use of teleoncology clinics in rural Kansas showed a cost reduction by almost 50%, from US $812 per consultation in 1995 to US $410 per consultation in 2000 [22]. Telemedicine can also be used to save time. The use of teleconsultation for veterans (individuals who previously served in the military) living a distance of 145 miles (233 km) from a health care facility was shown to save travel time of up to 142 minutes [23,24]. Apart from cost and time savings, telemedicine can be used to overcome barriers to health care access where conventional medical strategies do not apply [25]. Video consultation is very useful in providing consultation to patients in rural areas that lack a specialized physician. The Medical College of Georgia developed a Web-based telestroke system that enabled emergency physicians in rural areas to speak with specialists for patients with an episode of ischemic stroke. This system allowed physicians to examine patients using live video and to review medical imaging, and it recommended stroke therapies. Mean onset of stroke-to-treatment time was reduced by 20.2 minutes using the telestroke system, and only 2% of patients had a symptomatic hemorrhage [26]. Thus, patient outcomes were improved in an emergency situation. Lack of expert physicians in rural areas can create barriers for patients receiving urgent care [27]; therefore, video consultation can be effective in providing care to patients in critical conditions.

mHealth is an innovative way to deliver care. mHealth is used for remote monitoring and treating chronic diseases, to raise awareness, and for behavioral modification [28-34]. In one study, health data including blood pressure, pulse, weight, and dose of medication of patients with chronic heart failure and hypertension were transferred via a mobile phone, with an average data transfer accuracy of 83% (SD 22) [35]. This allowed physicians to remotely collect data for developing assessment and care management. Another study found that participants with controlled background displays on their mobile phones were likely to engage in a daily walk and cardiovascular exercise for 3 months, who otherwise would not have exercised [36]. Lastly, phone consultation was found to improve physical activities among women of low socioeconomic status who had high mortality rates due to high-risk behaviors [37].

Objective

These two modalities, telemedicine and mHealth, improve access to care: telemedicine enables physician intervention, and mHealth promotes patients' participation [38]. Yet less infrastructure being required for mHealth than for telemedicine, the rising popularity of mobile phones, the sophisticated third-generation network, and emerging ways to exchange information through mobile phones predict mHealth to be more promising for developing countries [39-41]. Although studies have shown the effectiveness of receiving care using mHealth and telemedicine, to the best of our knowledge, no study has compared patients' preference for phone calls versus video conferencing based on their demographics, chief concern, and time spent in consultation. The objective of this study was to understand factors influencing patients' choice of communication medium for receiving care, either through mHealth or in telemedicine encounters, when they were provided with both options in a virtual urgent clinic.

Methods

Study Design

We conducted a postdeployment evaluation of a new virtual health care service, Virtual Urgent Clinic (VUC), which uses mHealth and telemedicine modalities to provide patient care. VUC is a 24-hour-a-day, 7-day-a-week, on-demand service aimed at helping individuals with urgent medical needs to consult with a physician regarding their medical condition. The service was primarily designed to offer services regardless of the time of day or location of the patient in a more convenient form than the traditional in-person urgent care clinics. We obtained institutional review board approval from the University of North Carolina at Chapel Hill to conduct this research.

Study Setting and Participants

VUC is cloud-based platform offered through a public website. Individuals with urgent medical needs can use VUC, despite their location, as long as they have access to a phone or a computer equipped with a microphone and camera with internet connection. Inclusion criteria for this study were individuals with a medical need who were over the age of 2 years. Exclusion criteria were patients under the age of 2 years, patients with no access to a phone or a computer with microphone and camera with internet connection.

Materials

Individuals were required to create an account through the VUC website prior to scheduling a consultation. During the registration process, each individual had to fill out a short form providing basic demographic information. A secure link was sent to the individual's email address for activation of the account. Once the account was activated, the individual indicated whom the e-visit was for and the intended provider type (eg, family physician). The website provided information regarding conditions not treatable through VUC, medications that VUC physicians could not prescribe, and important information regarding children under the age of 3 years. Once the individual verified having read this information, they were asked to fill out a series of short forms on the reason for the visit, their medical history, choice of pharmacy, choice of provider, payment, and confirmation. The cost of a VUC visit was a flat fee of US $49.

After the encounter, patients were asked to voluntarily participate in a short patient satisfaction survey. The survey aimed to solicit patients' assessment of the encounter based on 4 criteria: (1) overall experience, (2) physician rating, (3) if they gave a fair or poor rating of the overall experience, their reason for the rating, and (4) open-ended patient comments.

Outcomes

The primary outcomes were two predictive models that projected the users' medium of choice given their demographics and chief concern. Secondary outcomes were encounter duration and satisfaction levels per encounter medium.

The dependent variable was encounter medium (mHealth, telemedicine). Independent variables were sex (female, male), age range (<18, 19-34, 35-49, ≥50 years), setting (urban, rural), insurance status (insured, uninsured), encounter time range (6 AM-12 PM, 12 PM-5 PM, 5 PM-12 AM, 12 AM-6 PM), day of the-week (weekday, weekend), top 20 chief concerns (Multimedia Appendix 1 shows a full list).

We included the top 20 chief concerns, which made up 68.57% (962/1403) of the total encounters, as a predictor instead of including all 148 concerns; we classified the remaining 128 encounters as others. The rationale behind this is that an excessive number of levels with a small number of data points would have added unnecessary complexity to the models.

Statistical Analysis

We used multinomial logistic regression to build and compare the two models based on the predictive power of two sets of features in determining patients' preferred method of telecare (mHealth and telemedicine) encounters. We selected the first set of independent variables to represent the demographics and socioeconomic status of the patient population. The additional feature, chief concern, captures patients' self-reported reason for the telecare visit.

For model selection purposes, we used the step function in R version 3.6.0 (R Foundation) to eliminate the least significant predictors. The process started with the full model, where all predictors were included; it ceased when the current model reached its maximum performance measured by the Akaike information criterion (AIC) [42].

To measure the features' predictive performance, we inferred the odds ratio (OR) by exponentiating the models' coefficients. However, due to the lack of a simple and intuitive explanation of OR outcomes, we decided to follow previous research by interpreting OR as the risk ratio—the relative probability of an event happening in one group compared with another group [43]. We discuss this method's limitations further below.

To evaluate the models' prediction accuracy, we performed cross-validation with 70% of the original dataset training data and using 30% as the testing set. In addition, we measured the models' efficiency and effectiveness using two common performance metrics: AIC and the simulated McFadden pseudo-R^2.

We used several R packages for advanced analysis and model building: nnet for modelling the multinomial logistic regression function; mfx for calculating the relative risk ratio; and stargazer for rendering the summary statistics. We generated visualizations using Tableau version 9.0 (Tableau Software).

Results

Demographics

Postdeployment, 1403 encounters were recorded, of which 87.53% (1228) were completed with mHealth, and 175 (12.5%) were telemedicine encounters (Table 1). We tested two predictive models: one with a set of 6 demographic features extracted from the patients' records and one with the chief concern feature as the predictor. We measured the results as the OR, indicating the magnitude of a specific feature's predictive power. In addition, we analyzed the relationship between chief concern and the two significant demographic predictors—sex and setting. Subsequently, we evaluated and compared the difference between mHealth and telemedicine encounters, specifically the duration of consultation session, chief concern, the patients' preference for alternative care-seeking options.

Table 1. Demographics of Virtual Urgent Clinic users.

Characteristics	Type of encounter		
	mHealth, n (%)	Telemedicine, n (%)	Total, n (%)
Number of encounters	1228 (87.53)	175 (12.47)	1403 (100.00)
Sex			
Male	269 (82.01)	59 (17.99)	328 (23.38)
Female	959 (89.21)	116 (10.79)	1075 (76.62)
Age range (years)			
2-18	115 (83.94)	22 (16.06)	137 (9.76)
19-34	434 (87.85)	60 (12.15)	494 (35.22)
35-49	465 (88.24)	62 (11.76)	527 (37.56)
≥50	214 (87.35)	31 (12.65)	245 (17.46)
Setting			
Rural	569 (92.22)	48 (7.78)	617 (44.04)
Urban	657 (83.80)	127 (16.20)	784 (55.96)
Insurance status			
Insured	556 (91.15)	54 (8.85)	610 (43.48)
Uninsured	672 (84.74)	121 (15.26)	793 (56.52)

Table 2. Odds ratio and significance (*P* value) of the demographic predictors[a].

Predictor	Odds ratio	*P* value
Sex: male	1.662	.004
Setting: urban	2.014	<.001
Insurance status: uninsured	1.42	.06
Constant	0.064	<.001

[a]Reference group: telemedicine.

Multinomial Logistic Regression Models

Predictive Model I: Demographics Features

Among the 6 predictors, sex and setting were the most predictive determinants of patients' preferred method of telecare delivery, with significantly small *P* values of less than .01. Insurance status was not significant (*P*<.10). With all else held constant, patients from urban areas had 1.014 times greater odds than users from rural regions of using telemedicine than of using mHealth. Similarly, male patients had 66.2% greater odds than female patients with identical features of using telemedicine than of using mHealth, as Table 2 shows.

Predictive Model II: Top 20 Chief Concerns

Among the 20 chief concerns, 6 were significant predictors of patients' preferred medium of telecare encounter (Table 3). A total of 4 predictors resulted in ORs greater than the neutral level of 1—urinary tract infection (*P*<.001), ear pain (*P*=.06), sinus infection (*P*=.04), and vaginal discharge (*P*<.001)—suggesting a lower tendency of choosing telemedicine over mHealth. Based on the model, we expected an 89% decrease in the odds of using telemedicine if a patient had a urinary tract infection. However, vaginal discharge yielded an OR of 0, indicating that no user with vaginal discharge chose telemedicine in this case. In contrast, pink eye (*P*=.05) and rash (*P*=.01) showed ORs greater than 1, suggesting a greater probability of opting for telemedicine. Based on the model, patients with pink eye were expected to have 1.39 times greater odds of choosing a telemedicine encounter.

Table 3. Odds ratio and significance (*P* value) of the chief concern predictor[a].

Predictor	Odds ratio	*P* value
Urinary tract infection	0.11	<.001
Ear pain	0.256	.06
Pink eye	2.39	.05
Rash	2.325	.01
Sinus infection	0.5	.04
Vaginal discharge	0	<.001
Constant	0.168	<.001

[a]Reference group: telemedicine.

Table 4. Evaluation metrics of multinomial logistic regression models.

Model	Akaike information criterion	McFadden R^2	Cross-validation prediction accuracy, %
Model I: demographics	1027.153	0.035	86.22
Model II: chief concerns	1030.168	0.064	86.22

Model Evaluation

The AIC of both models performed similarly, indicating that the two models were of similar complexity [44]. However, model II: chief concerns showed a slightly higher value (1030.168) than model I: demographics (1027.153), ranking model II: chief concerns lower than model I: demographics.

R^2 of model II (0.064) was almost twice that of model I (0.035; Table 4). A higher value of R^2 shows that a higher proportion of the dependent variable is explained by model II: chief concerns than by model I: demographics.

The cross-validation yielded a prediction accuracy of 86.22% (363 instances were correctly predicted out of the 421 data points in the testing set) for both models.

Chief Concerns Analysis

Pearson chi-square test returned a strong indication of dependency between chief concern and encounter mediums, with a close-to-zero *P*<.001. We further examined the relationship between chief concern and the two significant predictors—sex and setting—and found the same strong correlations between the variables.

We analyzed the top 10 chief concerns of the two encounter methods, the results of which confirmed the difference between mHealth and telemedicine users' primary reasons for seeking virtual urgent care. We observed a few extreme cases: for instance, urinary tract infection the most common concern among the mHealth users (n=147, 12.0% of a total of 1228 mHealth encounters), was absent from the telemedicine users' top 10 list (Table 5). Conversely, telemedicine users, but not mHealth users, frequently consulted about eye-related problems (pink eye and eye swollen).

Table 5. Top 10 chief concerns in mobile health (mHealth) encounters (n=1228).

Chief concerns	Encounter medium: mHealth, n (%)	Sex: female, n (%)	Setting: rural, n (%)
Urinary tract infection	147 (11.98)	147 (100.0)	62 (42.2)
Sinus infection	129 (10.51)	113 (87.6)	62 (48.1)
Sore throat	116 (9.45)	94 (81.0)	58 (50.0)
Cough	82 (6.68)	53 (65)	49 (60)
Ear pain	42 (3.42)	27 (64)	23 (55)
Rash	37 (3.02)	23 (62)	24 (65)
Fever	32 (2.61)	22 (69)	16 (50)
Nasal congestion	31 (2.53)	24 (77)	19 (61)
Cold	30 (2.44)	25 (83)	12 (40)
Animal or insect bite or scratch	28 (2.28)	18 (64)	9 (32)

Encounter Duration

The average duration of telemedicine encounters was 5.46 minutes, which is 5.4 percentage points higher than the mean duration of mHealth encounters (5.18 minutes). A Welch 2-sample *t* test refuted the null hypothesis of equal mean (P=.28, 95% CI –0.79 to 0.24) between the two samples, indicating the mean encounter durations of the two populations were significantly different.

mHealth encounter duration had a range of 1 to 15 minutes, where 70.93% (871/1228) of the total encounters fell within the 1- to 5-minute range. Telemedicine encounters had a similar range of 0 to 16 minutes. Encounters lasting longer than 10 minutes accounted for 12.6% (22/175) of all telemedicine calls, double the 6.03% (74/1228) of mHealth encounters. In addition, 14.9% (26/175) of telemedicine calls lasted less than 1 minute, in contrast to the absence of mHealth calls of this length, as shown Figure 1 shows.

Figure 1. Distribution of encounter durations by encounter methods.

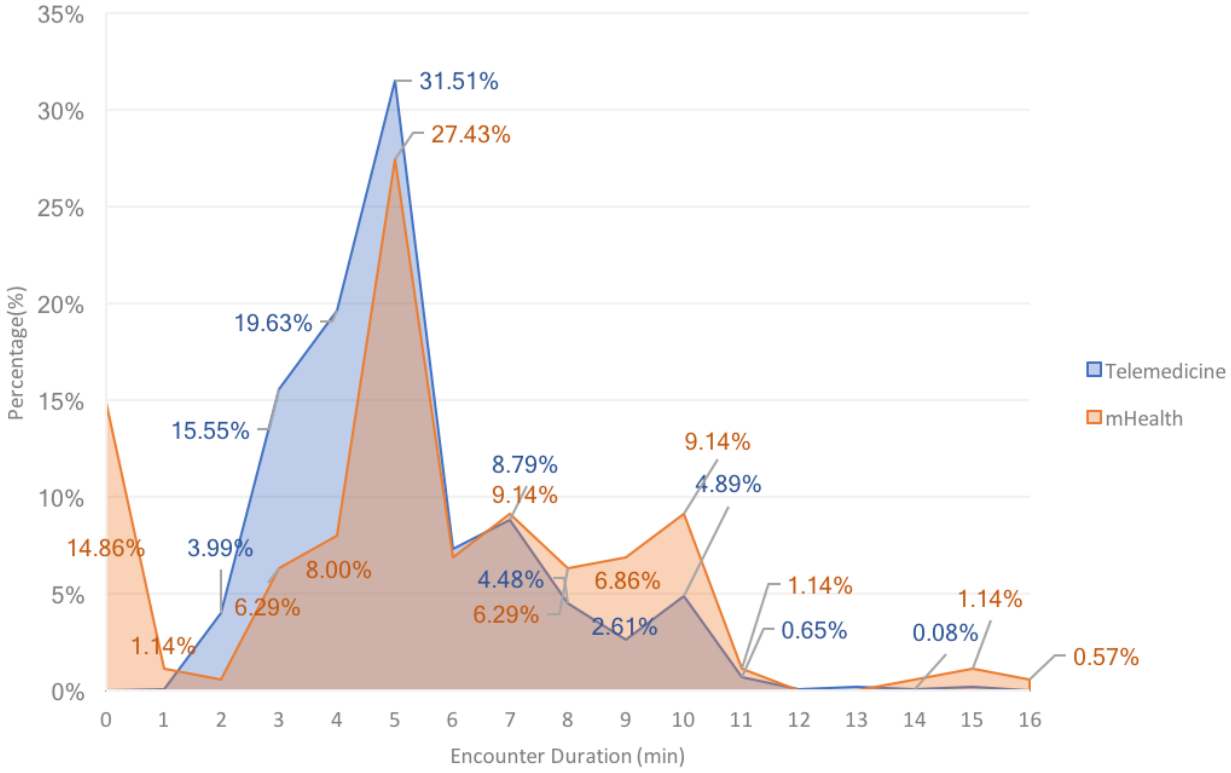
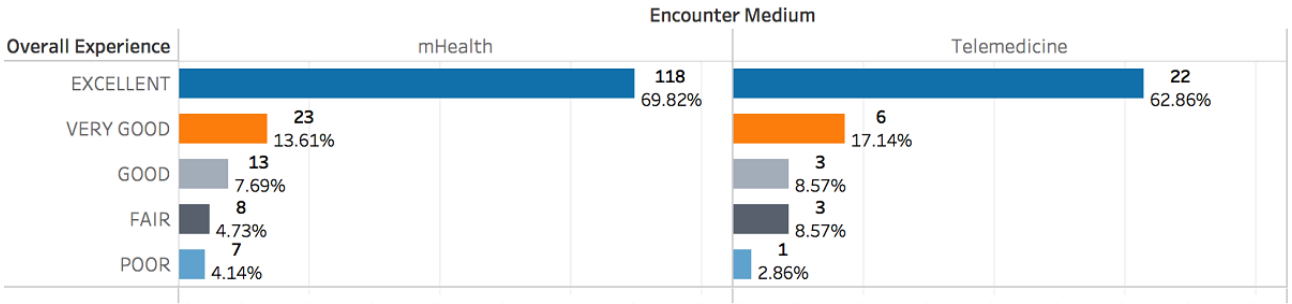


Figure 2. Self-reported overall experience satisfaction ratings. mHealth: mobile health.



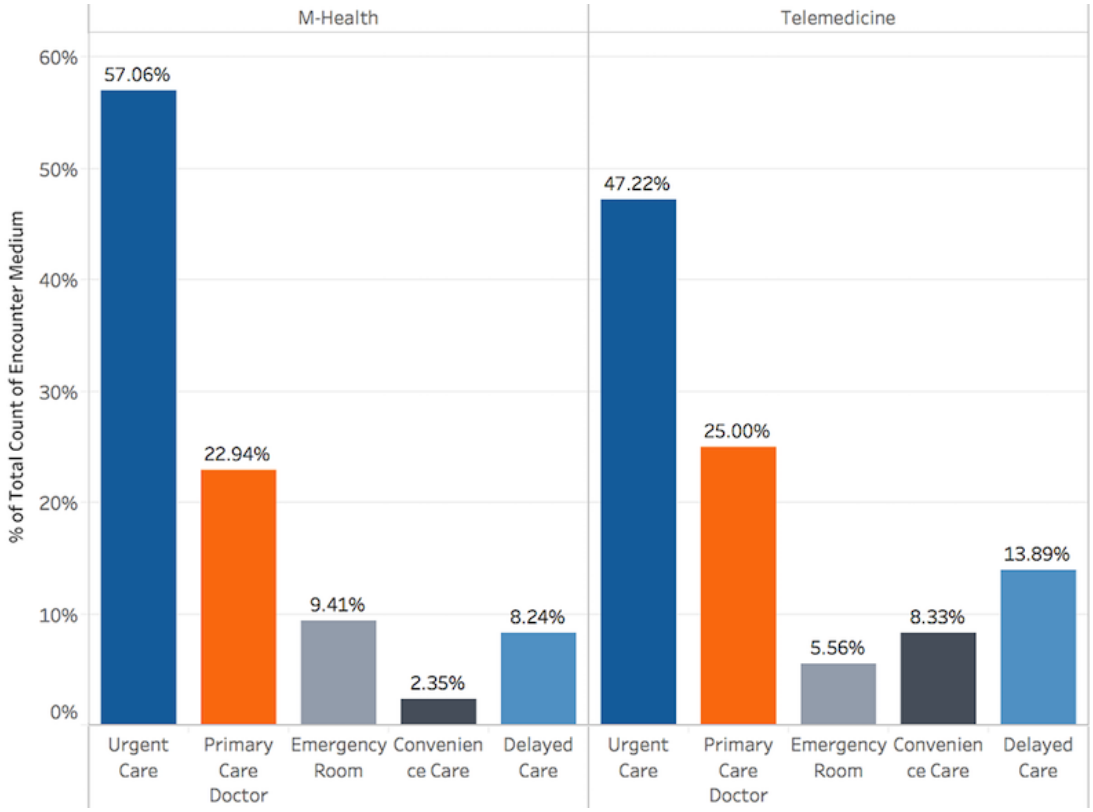
Patient Satisfaction by Encounter Medium

For participants in all 1403 encounters, 204 (14.54%) responded to the satisfaction survey. High satisfaction levels were reported among both the mHealth and telemedicine groups. Of mHealth patients, 91.1% (154/169) were satisfied by their encounter compared with 89% (31/35) of telemedicine patients. A higher proportion of telemedicine patients (4/35, 11.4%) than mHealth patients (15/169, 8.9%) rated their experience as fair or poor (Figure 2).

Alternative Care-Seeking Options

We looked further into the telemedicine and mHealth users' self-reported preferences for alternative care-seeking options. Patients were asked after their VUC consultation "if VUC was not available, which medical service would you have used?" The analysis revealed an almost identical distribution of users among the 5 options (Figure 3). In-person urgent care was the most popular alternative care option for both types of users. In addition, approximately one-fifth of the users would have delayed seeking care.

Figure 3. Alternative care-seeking choices of mobile health (mHealth) and telemedicine users.



Discussion

Principal Findings

To our knowledge, this is the first study to comprehensively assess the effectiveness of providing patients with medium choice (phone call vs video call) of either mHealth or telemedicine to consult with physicians for urgent care needs. We leveraged a data science approach, namely, data analytics, to predict what factors informed patients' choice of an mHealth or telemedicine medium. We analyzed the top 20 chief concerns in both groups to gain insight into the potential association between concern and choice of medium. Then, we analyzed the duration of encounters, self-reported alternative care-seeking options, and users' responses to satisfaction surveys between both groups.

We proposed a model to predict the preferred choice of care delivery for patients. Patients' sex and geographic location (rural or urban) significantly predicted their choice of care between mHealth and telemedicine. Patients from an urban area were twice as likely as users from rural regions to choose telemedicine over mHealth. Similarly, male patients were 1.6 times more likely than female patients with identical features to use telemedicine than mHealth. We conclude that male users from urban regions are the most likely to choose telemedicine over mHealth.

Patients' chief concern significantly correlated with their choice of medium, where chief concern strongly correlated with mHealth or telemedicine. The duration of encounters was similar between both mediums, around the 5-minute mark. Overall, telemedicine encounters had a notable difference in range, from less than 1 minute up to 16 minutes. A possible justification for telemedicine encounters to last less than 1 minute needs to be studied in the future.

We observed that patients were satisfied with their choice of medium, as well as the service provided, which suggests that providers should offer the option of mHealth or telemedicine to their patients and allow them to choose. We recommend considering patients' sex and setting as predictive factors to provide suggestions on which communication medium would best fit patients based on their characteristics. Patient satisfaction was high in both groups, with higher dissatisfaction among telemedicine users, which may be attributed to the quality of the video or audio feed. There was no significant difference between the groups in terms of their self-reported responses to alternative care-seeking options.

Strengths and Limitations

A strength of this research is the ability to alleviate the demand on in-person urgent care clinics and emergency rooms by providing a virtual clinic where patients can be seen and treated. Since VUC is an on-demand and cloud-based service, there was no purposive sampling, which allows the findings of this study to be more generalizable. The digital nature of the service may introduce bias to the sample population; however, this study focused on two digital interventions and, therefore, if any bias was introduced, it should not have influenced the study findings. Another strength is the convenience of providing both mHealth and telemedicine options to patients within the same platform without further setup. The response rate of the voluntary survey was adequate given that we provided no incentive to participate.

One limitation of this study is the lower number of telemedicine encounters relative to mHealth encounters, which can be

attributed to several factors, such as personal preference, time of the call, access to a Web camera, and internet connection speed. Another limitation is the absence of information regarding the reason for telemedicine encounters ending in less than 1 minute. This study can be further strengthened by capturing patient outcomes after the consultation visit by looking at 30-day hospitalization rates to assess the quality of care for each medium, which is a future direction of this research.

Conclusion

We studied factors influencing patients' choice of communication medium, either mHealth or telemedicine, for a virtual care clinic. Patients' preference for mHealth or telemedicine was significantly influenced by their sex and geographic location, as well as their chief concern. Despite other preferences, patients were highly satisfied by their choice of communication medium. This study showed that providing the option of mHealth or telemedicine to patients suggests which medium would be a better fit for patients based on their characteristics.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Top 20 chief concerns.

[PDF File (Adobe PDF File), 27KB - mhealth_v7i6e13772_app1.pdf]

References

1. Wilson NW, Couper ID, De Vries E, Reid S, Fish T, Marais BJ. A critical review of interventions to redress the inequitable distribution of healthcare professionals to rural and remote areas. Rural Remote Health 2009;9(2):1060 [FREE Full text] [Medline: 19530891]
2. America Counts Staff. One in five Americans live in rural areas. 2017 Aug 9. URL: https://www.census.gov/library/stories/2017/08/rural-america.html [accessed 2019-05-23] [WebCite Cache ID 78apyqzSv]
3. Laditka JN, Laditka SB, Probst JC. Health care access in rural areas: evidence that hospitalization for ambulatory care-sensitive conditions in the United States may increase with the level of rurality. Health Place 2009 Sep;15(3):731-740. [doi: 10.1016/j.healthplace.2008.12.007] [Medline: 19211295]
4. Anderson TJ, Saman DM, Lipsky MS, Lutfiyya MN. A cross-sectional study on health differences between rural and non-rural U.S. counties using the County Health Rankings. BMC Health Serv Res 2015 Oct 01;15:441 [FREE Full text] [doi: 10.1186/s12913-015-1053-3] [Medline: 26423746]
5. Brems C, Johnson ME, Warner TD, Roberts LW. Barriers to healthcare as reported by rural and urban interprofessional providers. J Interprof Care 2006 Mar;20(2):105-118. [doi: 10.1080/13561820600622208] [Medline: 16608714]
6. Doorenbos AZ, Kundu A, Eaton LH, Demiris G, Haozous EA, Towle C, et al. Enhancing access to cancer education for rural healthcare providers via telehealth. J Cancer Educ 2011 Dec;26(4):682-686 [FREE Full text] [doi: 10.1007/s13187-011-0204-4] [Medline: 21336979]
7. Davis TC, Rademaker A, Bailey SC, Platt D, Esparza J, Wolf MS, et al. Contrasts in rural and urban barriers to colorectal cancer screening. Am J Health Behav 2013 May;37(3):289-298 [FREE Full text] [doi: 10.5993/AJHB.37.3.1] [Medline: 23985175]
8. James TM, Greiner KA, Ellerbeck EF, Feng C, Ahluwalia JS. Disparities in colorectal cancer screening: a guideline-based analysis of adherence. Ethn Dis 2006;16(1):228-233. [Medline: 16599375]
9. Bennett KJ, Probst JC, Bellinger JD. Receipt of cancer screening services: surprising results for some rural minorities. J Rural Health 2012 Jan;28(1):63-72. [doi: 10.1111/j.1748-0361.2011.00365.x] [Medline: 22236316]
10. Cole AM, Jackson JE, Doescher M. Urban-rural disparities in colorectal cancer screening: cross-sectional analysis of 1998-2005 data from the Centers for Disease Control's Behavioral Risk Factor Surveillance Study. Cancer Med 2012 Dec;1(3):350-356 [FREE Full text] [doi: 10.1002/cam4.40] [Medline: 23342284]
11. Kulshreshtha A, Goyal A, Dabhadkar K, Veledar E, Vaccarino V. Urban-rural differences in coronary heart disease mortality in the United States: 1999-2009. Public Health Rep 2014;129(1):19-29 [FREE Full text] [doi: 10.1177/003335491412900105] [Medline: 24381356]
12. Seidel JE, Beck CA, Pocobelli G, Lemaire JB, Bugar JM, Quan H, et al. Location of residence associated with the likelihood of patient visit to the preoperative assessment clinic. BMC Health Serv Res 2006 Feb 22;6:13 [FREE Full text] [doi: 10.1186/1472-6963-6-13] [Medline: 16504058]
13. Marmot M. An inverse care law for our time. BMJ 2018 Jul 31;362:k3216. [doi: 10.1136/bmj.k3216] [Medline: 30065009]
14. Waller M, Stotler C. Telemedicine: a primer. Curr Allergy Asthma Rep 2018 Aug 25;18(10):54. [doi: 10.1007/s11882-018-0808-4] [Medline: 30145709]

15. Henderson P. Analysis of maternity practices in Ghana with the implementation of telemedicine. Honors Project. 712. ScholarWorks@GVSU. Allendale, MI: Grand Valley State University; 2018. URL: https://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1713&context=honorsprojects [accessed 2019-05-23]
16. Haq M, inventor; Google LLC, assignee. Virtual clinic for medical practice. US patent: 7,412,396 B1. 2013 Aug 12.
17. Cusack CM, Pan E, Hook JM, Vincent A, Kaelber DC, Middleton B. The value proposition in the widespread use of telehealth. J Telemed Telecare 2008;14(4):167-168. [doi: 10.1258/jtt.2007.007043] [Medline: 18534947]
18. van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. Bull World Health Organ 2012 May 1;90(5):393-394 [FREE Full text] [doi: 10.2471/BLT.11.099788] [Medline: 22589575]
19. Ryu S. Book review: mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). Healthc Inform Res 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231]
20. Gagnon M, Duplantie J, Fortin J, Landry R. Implementing telehealth to support medical practice in rural/remote regions: what are the conditions for success? Implement Sci 2006 Aug 24;1:18 [FREE Full text] [doi: 10.1186/1748-5908-1-18] [Medline: 16930484]
21. Eikelboom RH, Atlas MD. Attitude to telemedicine, and willingness to use it, in audiology patients. J Telemed Telecare 2005;11 Suppl 2:S22-S25. [doi: 10.1258/135763305775124920] [Medline: 16375788]
22. Doolittle GC, Williams AR, Spaulding A, Spaulding RJ, Cook DJ. A cost analysis of a tele-oncology practice in the United States. J Telemed Telecare 2004;10 Suppl 1:27-29. [doi: 10.1258/1357633042614429] [Medline: 15603601]
23. Russo JE, McCool RR, Davies L. VA telemedicine: an analysis of cost and time savings. Telemed J E Health 2016 Mar;22(3):209-215. [doi: 10.1089/tmj.2015.0055] [Medline: 26305666]
24. Bartlett BA, Smith LJ, Tran JK, Vujanovic AA. Understanding mental health among military veterans in the fire service. Psychiatry Res 2018 Dec;267:394-399. [doi: 10.1016/j.psychres.2018.06.020] [Medline: 29960261]
25. Wang H, Liu J. Mobile phone based health care technology. Recent Patents Biomed Eng 2009 Jan 01;2(1):15-21. [doi: 10.2174/1874764710902010015]
26. Switzer JA, Hall C, Gross H, Waller J, Nichols FT, Wang S, et al. A web-based telestroke system facilitates rapid treatment of acute ischemic stroke patients in rural emergency departments. J Emerg Med 2009 Jan;36(1):12-18. [Medline: 18242925]
27. Gamm L, Hutchison L. Rural Healthy People 2010—Evolving Interactive Practice. Am J Public Health 2004 Oct;94(10):1711-1712. [doi: 10.2105/ajph.94.10.1711] [Medline: 15451738]
28. Chib A, van Velthoven MH, Car J. mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries. J Health Commun 2015;20(1):4-34. [doi: 10.1080/10810730.2013.864735] [Medline: 24673171]
29. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
30. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
31. Hebden L, Cook A, van der Ploeg HP, Allman-Farinelli M. Development of smartphone applications for nutrition and physical activity behavior change. JMIR Res Protoc 2012;1(2):e9 [FREE Full text] [doi: 10.2196/resprot.2205] [Medline: 23611892]
32. Hong Y, Dahlke DV, Ory M, Hochhalter A, Reynolds J, Purcell NP, et al. Designing iCanFit: a mobile-enabled Web application to promote physical activity for older cancer survivors. JMIR Res Protoc 2013;2(1):e12 [FREE Full text] [doi: 10.2196/resprot.2440] [Medline: 23612053]
33. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. Telemed J E Health 2011 Dec;17(10):801-803. [doi: 10.1089/tmj.2011.0055] [Medline: 22010977]
34. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
35. Scherr D, Zweiker R, Kollmann A, Kastner P, Schreier G, Fruhwald FM. Mobile phone-based surveillance of cardiac patients at home. J Telemed Telecare 2006;12(5):255-261. [doi: 10.1258/135763306777889046] [Medline: 16848939]
36. Consolvo S, Klasnja P, McDonald DW, Avrahami D, Froehlich J, LeGrand L, et al. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. New York, NY: ACM; 2008 Presented at: 10th International Conference on Ubiquitous Computing; Sep 21-24, 2008; Seoul, Korea p. 54-63.
37. Albright CL, Pruitt L, Castro C, Gonzalez A, Woo S, King AC. Modifying physical activity in a multiethnic sample of low-income women: one-year results from the IMPACT (Increasing Motivation for Physical ACTivity) project. Ann Behav Med 2005 Dec;30(3):191-200. [doi: 10.1207/s15324796abm3003_3] [Medline: 16336070]
38. Omboni S, Caserini M, Coronetti C. Telemedicine and m-health in hypertension management: technologies, applications and clinical evidence. High Blood Press Cardiovasc Prev 2016 Sep;23(3):187-196. [doi: 10.1007/s40292-016-0143-6] [Medline: 27072129]
39. Curioso WH, Karras BT, Campos PE, Buendia C, Holmes KK, Kimball AM. Design and implementation of Cell-PREVEN: a real-time surveillance system for adverse events using cell phones in Peru. AMIA Annu Symp Proc 2005:176-180. [Medline: 16779025]

40. Mechael PN. The case for mHealth in developing countries. Innovations 2009 Jan;4(1):103-118. [doi: 10.1162/itgg.2009.4.1.103]
41. Curioso WH, Mechael PN. Enhancing 'M-health' with south-to-south collaborations. Health Aff 2010 Feb 01;29(2):264-267. [doi: 10.1377/hlthaff.2009.1057] [Medline: 20348071]
42. step: choose a model by AIC in a stepwise algorithm. stats v3.6.0. R-core. Vienna, Austria: R Foundation URL: https://www.rdocumentation.org/packages/stats/versions/3.6.0/topics/step [accessed 2019-05-23] [WebCite Cache ID 78aqp8Jam]
43. Davies H, Crombie I, Tavakoli M. When can odds ratios mislead? BMJ 1998 Mar 28;316(7136):989-991 [FREE Full text] [doi: 10.1136/bmj.316.7136.989] [Medline: 9550961]
44. McElreath R. Statistical Rethinking: A Bayesian Course With Examples in R and Stan. In: Statistical rethinking: a Bayesian course with examples in R and Stan. Boca Raton: Crc Press/taylor & Francis Group,; 2016.

Abbreviations

AIC: Akaike information criterion
mHealth: mobile health
OR: odds ratio
VUC: Virtual Urgent Clinic

Edited by G Eysenbach; submitted 20.02.19; peer-reviewed by C Mather, J op den Buijs; comments to author 28.03.19; revised version received 12.05.19; accepted 14.05.19; published 08.06.19.

Please cite as:
Khairat S, Liu S, Zaman T, Edson B, Gianforcaro R
Factors Determining Patients' Choice Between Mobile Health and Telemedicine: Predictive Analytics Assessment
JMIR Mhealth Uhealth 2019;7(6):e13772
URL: http://mhealth.jmir.org/2019/6/e13772/
doi:10.2196/13772
PMID:31199332

Original Paper

An Electronic Teaching Module for Improving Knowledge of Self-Management of Vaso-Occlusive Pain Crises in Patients With Sickle Cell Disease: Pilot Questionnaire Study

Tammie Tam[1], BSc; Maria R Baer[2], MD; Lewis L Hsu[3], MD, PhD; Jennie Y Law[2], MD

[1]University of Maryland School of Medicine, University of Maryland, Baltimore, MD, United States

[2]Marlene and Stewart Greenebaum Comprehensive Cancer Center, University of Maryland School of Medicine, University of Maryland, Baltimore, MD, United States

[3]Comprehensive Sickle Cell Center, Division of Pediatric Hematology/Oncology, Department of Pediatrics, University of Illinois at Chicago, Chicago, IL, United States

Corresponding Author:
Jennie Y Law, MD
Marlene and Stewart Greenebaum Comprehensive Cancer Center
University of Maryland School of Medicine
University of Maryland
22 S Greene Street S9D15C
Baltimore, MD, 21201
United States
Phone: 1 410 328 6373
Fax: 1 410 328 6896
Email: jennielaw@umm.edu

Abstract

Background: For patients with sickle cell disease (SCD), effective management of vaso-occlusive crises (VOCs) is integral to provision of care, as nearly all affected individuals will suffer from VOCs in their lifetime. A recent systematic review of technological interventions to improve self-management in the care of SCD concluded that electronic health has the potential to improve the care of individuals with SCD.

Objective: The aim of this study was to assess the value of an electronic teaching module (ETM) provided by Emmi Solutions for educating adult SCD patients on VOC self-management and treatment options for SCD.

Methods: A pretest assessed adults with SCD for baseline knowledge with regard to self-management of VOCs. Participants then watched the 35-min ETM and completed a posttest and survey on the ETM.

Results: A total of 20 adults enrolled. Their knowledge scores improved (pretest median 66.5% and posttest median 85%; $P<.001$). In total, 18 participants (18/20, 90%) agreed that they "learned a lot" or "learned something" from the ETM. The most common topic about which they reported learning was hydroxyurea. A total of 12 participants (12/20, 60%) agreed with the statement that they "would recommend the module to a friend or family member with sickle cell disease."

Conclusions: The ETM is associated with an increase in knowledge in patients with SCD. Limitations of the study include small sample size, no assessment of knowledge before premodule questionnaire completion, and no longitudinal follow-up. Identifying patients with SCD who demonstrate affinity for self-education via an ETM may further enhance utility of this tool to educate and empower patients.

(JMIR Mhealth Uhealth 2019;7(6):e13501) doi:10.2196/13501

KEYWORDS

patient education; teaching materials; portable electronic applications; sickle cell disease

Introduction

Background

For patients with sickle cell disease (SCD), effective management of vaso-occlusive crises (VOCs) is integral to provision of care, as nearly all affected individuals will suffer from VOCs in their lifetime [1]. Unfortunately, many patients with SCD lack access to specialized hematology care and, as a result, are reliant on providers who may be unaware of the myriad complications related to this chronic condition and may be uncomfortable with treating these patients [2]. Given the limited number of sickle cell day hospitals, many adult patients rely heavily on acute care services for management of their VOCs, despite often facing ineffective analgesic management and patient-perceived biases in care provision in the emergency department [3,4]. Provider desire for clinical correlates to confirm the complaint of pain in a patient with SCD leads to disbelief or skepticism of the patient and ultimately contributes to inadequate treatment of VOCs, resulting in premature discharges as well as high rates of emergency department and hospital recidivism [5,6]. In recent studies, patients with SCD report inadequate access to ambulatory care, poor communication with health care professionals, and lack of involvement in decisions regarding their own health [1,5]. Acute care encounters and rehospitalizations are frequent for patients with SCD, particularly for adolescents and young adults (AYAs) aged 18 to 30 years [7]. Given this gap in the knowledge and comfort level of health care providers managing patients with SCD, there is a clear need to improve patient health literacy and enhance patient confidence in self-care [8].

Electronic health (eHealth) may especially resonate with AYA patients, given their generational predilection for electronic media. A recent study showed that almost all AYA patients with SCD had daily internet access via either computers (84%) or mobile phones (70%) [9]. In one study examining self-management tools in SCD, the majority of participants opted to go online for their health information and preferred websites with interactive or social features (83%) [7]. Harnessing the power of Web-based modules to engage a vulnerable patient population is critical, given the positive correlation between patient confidence in self-care and improved health outcomes for individuals with SCD [8].

A recent systematic review of technological interventions to improve self-management in the care of SCD concluded that eHealth has the potential to improve the care of individuals with SCD [10]. eHealth is being used for self-management of a variety of chronic diseases, including SCD, asthma, diabetes, and hypertension [10]. Most eHealth intervention studies for SCD have been performed in children and adolescents [10]. Relatively few studies have evaluated eHealth in adults with SCD and they focus on pain or knowledge about reproductive health [10]. More studies are needed in adults with SCD in other components of care, including medication adherence, coping strategies, and clinic appointment adherence such as eHealth in pediatric SCD [10].

Objectives

Given the lack of evaluation of eHealth interventions in adults, the purpose of this pilot study was to assess the value of the electronic teaching module (ETM) produced by Emmi Solutions for self-management in individuals with SCD. The ETM is an educational video with graphics and audio written at a fifth-grade reading level, with the exception of key words such as hemoglobin and hydroxyurea. It includes information about SCD and how self-care can help mitigate VOC triggers and decrease emergency department usage; and it encourages utilization of specialized care with a hematologist and coping strategies to manage pain at home with both pharmacologic and nonpharmacologic measures. Topics covered in the ETM include an overview of basic SCD pathophysiology, SCD pain, over-the-counter medications, opioids, pain management tools, hydroxyurea, and when to call your doctor. Currently, management of SCD pain focuses on alleviation of symptoms with fluids, anti-inflammatory agents, and opioid analgesics [3,8]. As much of the initial management of early VOCs could be performed at home, the Emmi ETM is an eHealth intervention to make self-care information more accessible to patients with SCD. Specifically, this ETM may be especially helpful for individuals with SCD who lack access to a specialized sickle cell clinic, and the format should be a good fit for AYA patients with SCD who prefer to use electronic methods to receive health information. We hypothesized that there would be an improvement in scores from the premodule knowledge-based questionnaire to the postmodule questionnaire and that the postmodule survey would indicate that the ETM is well-received by patients with SCD.

Methods

Measures

The pre- and postmodule knowledge-based questionnaires included 16 identical questions about self-management of sickle cell pain. The postmodule survey added 12 questions assessing participants' satisfaction with use of the ETM and documenting age, gender, and highest level of education attained. The questionnaire (Multimedia Appendix 1) and survey (Multimedia Appendix 2) were written based on the Emmi ETM script by a professional education specialist from Emmi Solutions to assess for participant understanding of the content of the ETM. The pre- and postquestionnaire and survey responses were digitally recorded through SurveyMonkey, scored, and reviewed.

Sample

A convenience sample of adults with SCD seen at the University of Maryland Medical Center (UMMC):

- Inclusion criteria: (1) Any type of SCD (HbSS, HbSβ0-thal, HbSC, or HbSβ+-thal), (2) aged 18 years or older, and (3) willing to commit 1.5 hours to complete the study in one sitting.
- Exclusion criteria: (1) unable to read English or (2) previous use of the Emmi ETM.

Procedures

A convenience sample of patients with SCD was offered enrollment upon presentation to their scheduled Hematology outpatient visits at the UMMC between June and October 2017. After verbal consent, patients were provided with a laptop computer to view the Emmi ETM and complete the assessments in the clinic. Participants completed a pretest assessing their baseline knowledge with regard to self-management of VOCs. They then watched a 35-min ETM that described self-care techniques that can mitigate triggers for vaso-occlusion and decrease emergency department usage and also reviewed pharmacologic and nonpharmacologic measures for pain management at home. Patients were allowed to take breaks if necessary. It was noted if the participant fell asleep during the ETM. Participants then completed a posttest assessing their knowledge of self-management of VOCs and a survey regarding use of the ETM. The pre- and postmodule knowledge-based questionnaires and postmodule survey were deidentified to prevent a possible breach of privacy. Once participants completed the study, they received a US $20 Target™ gift card for the time invested in the study. They were also given an information sheet about the study with their identification number and a link to re-access the ETM. This study was approved by the University of Maryland School of Medicine Institutional Review Board.

Data Analysis

Pre- and postmodule questionnaire scores were compared using paired *t* tests. Postmodule survey responses were analyzed by frequency distribution. Data were monitored to see whether patients re-accessed the ETM.

Results

A total of 39 individuals were approached for the study: 19 declined to participate (most stating reluctance to invest the amount of time needed), none were ineligible, and 20 participated. In total, 12 patients (12/20, 60%) were female. The median age of participants was 27.5 years (range: 20 to 48 years). The highest levels of education completed were as follows: not graduated high school (3), high school diploma/Generalized Educational Development (12), Bachelor's degree (2), and beyond Bachelor's degree (3).

There was significant improvement in scores from the premodule questionnaire to the postmodule questionnaire. The median pretest knowledge score was 66.5% and the median posttest knowledge score, after completion of the ETM, was 85% (paired *t* test [n=20]=6.84; *P*<.001; Table 1). This suggests that study participants were able to learn from the ETM and improve their baseline knowledge regarding self-management of sickle cell pain. The small sample size precludes a meaningful analysis of subgroups.

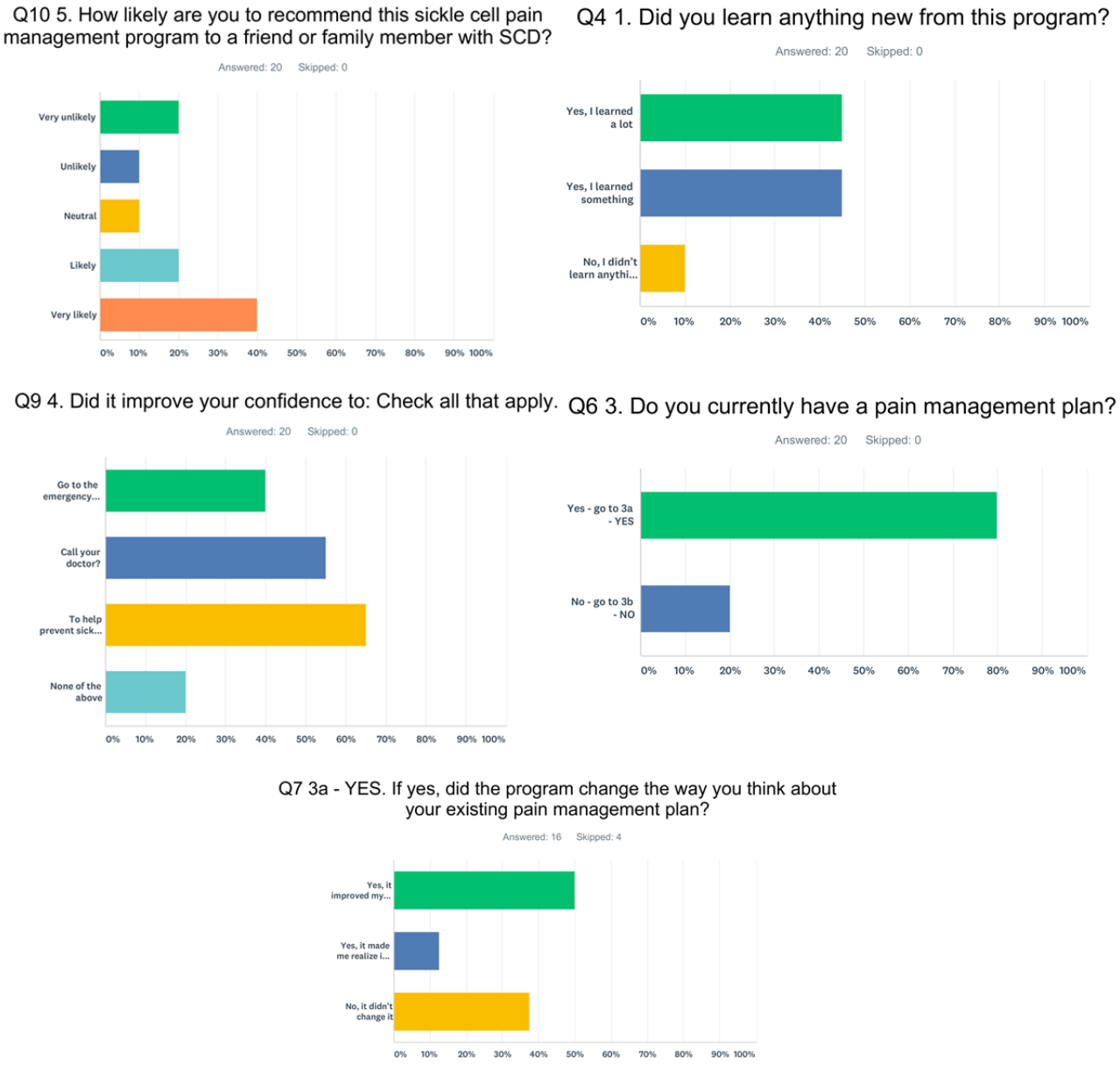
Table 1. Premodule versus postmodule questionnaire scores.

Study participant	Premodule Score (%)	Postmodule Score (%)	Change (%)
1	65	62	–4.6
2	85	97	14.1
3	68	97	42.6
4	56	85	51.8
5	76	85	11.8
6	62	85	37.1
7	74	74	0.0
8	47	59	25.5
9	65	94	44.6
10	79	88	11.4
11	65	71	9.2
12	59	82	39.0
13	38	47	23.7
14	56	82	46.4
15	74	88	18.9
16	71	88	23.9
17	51	74	45.1
18	74	79	6.8
19	68	94	38.2
20	88	97	10.2
Average	66.1	81.4	24.8

The postmodule survey (Figure 1) results indicated that the ETM was well-received. In total, 18 out of 20 participants (90%) reported that they *learned a lot* or *learned something* from the ETM. The most common topic about which they reported learning was hydroxyurea. Nine of the 20 participants commented that they gained a better understanding of how hydroxyurea works and/or of its side effects. Five participants also commented that they learned a lot about opioids. Sixteen of the 20 participants also reported that the ETM improved their confidence to help prevent sickle cell crises, call their doctor, and/or go to the emergency department when needed. Sixteen of the 20 participants had an existing pain management plan, as expected as all study participants were scheduled to see sickle cell specialists in the adult Hematology clinic at UMMC. Of the 16 participants with a pain management plan, 2 reported that the ETM made them realize that they needed to revisit or update their plan with their doctor, 8 reported that the ETM improved their confidence to manage pain, and 6 reported that the ETM had no effect on the way they thought about their existing plan. Of the 4 participants who did not have an existing pain management plan, all 4 reported that they were very likely to ask their doctor or pain specialist to create a pain management plan with them in the next 3 months. Twelve of the 20 participants (60%) were either *very likely* or *likely* to recommend the ETM to a friend or family member with SCD.

Five participants fell asleep at some point during the ETM, which at least 3 patients attributed to fatigue. One participant commented that the 25-min ETM was “way too long, especially if the target audience is users with this condition” . Patient access to the ETM was monitored for 6 months after the completion of the study and no patients re-accessed the ETM.

Figure 1. Excerpts from postmodule survey responses. SCD: sickle cell disease.



Discussion

Principal Findings

The knowledge scores showed a significant improvement from the premodule questionnaire to the postmodule questionnaire. In the postmodule survey, 90% of participants reported that they learned from the ETM, and 60% of participants were likely to recommend the ETM to others.

The small sample size precludes a meaningful analysis of subgroups defined by factors including age, sex, and level of education. The subgroup analysis might have shown differences between age groups owing to the generational predilection for electronic media and differences between the level of education subgroups owing to varying levels of comprehension of the ETM or the questionnaire. For future studies with a larger sample size, a literacy assessment might be a more accurate measure of study participants' comprehension of the ETM than the self-reported level of education. The disadvantage of adding a literacy assessment to the study is that some patients might not have the attention span to concentrate for a longer study.

Subgroups of patients who may not have benefited as much from the Emmi ETM include those who had barriers to completing the ETM, such as a low level of computer literacy, reading level below fifth grade, and short attention span, and those with a high baseline knowledge of self-care before the Emmi ETM intervention. Falling asleep during the module may be attributable to fatigue, a well-known complication of SCD that can be out of proportion to the anemia [11,12]. Although the 2 participants who reported that they did not learn anything new from the ETM scored 85% and 71% on the premodule questionnaire, both improved on the postmodule questionnaire, with postmodule scores of 97% and 88%. Other participants who scored >70% on the premodule questionnaire reported that they did learn something from the ETM.

Given that 5 participants fell asleep at some point during the ETM and no patients re-accessed the ETM within 6 months of the study, revisions may be necessary. In a recent study assessing online module development, students were more engaged when the module content was applicable to their personal life or community and provided opportunities for critical thinking [13]. The study also highlights the value of emphasizing qualitative feedback over traditionally quantitative feedback [13]. Possible changes to make the ETM more engaging include adding a playback speed control button per study participant survey comments, modifying the ETM to make it more interactive by adding questions after each section and adding a summary at the end of the ETM. In addition, we could implement an exit interview instead of an online survey to more rigorously evaluate how patients felt about the ETM. Many patients were noted to be slow at typing, so a verbal survey may encourage a more lengthy and critical assessment of the ETM.

Although the goal of health education is to demonstrate clinical benefit, this demonstration will require additional investigations beyond the scope of this pilot study. Next steps include expanding the study to include a larger sample size and multiple sites (urban vs rural and medical vs community locations). A recent systematic review concluded that evaluations of SCD educational interventions should consist of large-scale studies with many patients retained in a longitudinal follow-up, owing to the heterogeneity of SCD and the modest effect size for educational interventions [14].

Limitations

Limitations of this study include a small sample size, no assessment of literacy, no validation of the questionnaire, no assessment of patient knowledge of SCD before the initial premodule questionnaire completion, and a lack of longitudinal follow-up to track retention of knowledge gained from the ETM or further actions taken based on what participants learned from the ETM.

Conclusions

Assessment of this pilot implementation of an educational ETM for SCD in a real clinical setting revealed both benefits and drawbacks. Nearly all patients gained some knowledge from the ETM and had positive comments such as willingness to recommend the ETM to friends. The wide range of pretest scores shows that, before the ETM, there are probably subgroups of patients with greater knowledge of self-care for SCD. The patients' comments suggest that subgroups exist with different learning styles, with some requesting a more customized or streamlined ETM. Future studies should also examine whether the knowledge acquired through this ETM has an impact on disease course, although these will be difficult to conduct given the unpredictable nature of SCD. Finally, the high rate of refusal to participate (19 of 39 people approached for the study) suggests that modifications are needed to motivate people to use the ETM. Other modifications that may decrease the high rate of refusal to participate include giving patients the option to schedule the study visit at their convenience and having a trusted provider endorse the study.

Practice Implications

The educational ETM has the potential to provide an immediate impact on health literacy for an underserved patient population and potentially improve disease management and disease-specific outcomes. Identifying discrete populations of patients with SCD who demonstrate an affinity for self-education via the Emmi ETM would further enhance the utility of this tool. Innovative approaches to improving self-care for patients with SCD are sorely needed and have the potential to enhance care provision for affected individuals, as well as transform the oftentimes adversarial interaction between adult patients and their treating physicians.

As more patients with SCD are reaching adulthood and SCD has a lifelong course, providing resources to facilitate the transition from pediatric to adult SCD care is essential. Early adulthood is a high-risk period, associated with increased mortality [8,9]. The ETM may be particularly helpful in facilitating the often-fraught transition from pediatric to adult care. Barriers to successful transition include AYA-specific patient issues (lack of knowledge about SCD and about obtaining SCD care with adult health care providers, lack of financial independence and decision-making experience, and fear of leaving familiar clinics), health care provider issues

(poorly coordinated transitional care and lack of education and familiarity with SCD treatment), and systemic issues (adult care with fewer interdisciplinary services than well-established pediatric SCD programs, AYAs losing or changing insurance coverage, and a reimbursement system that may limit the willingness of specialists to see adult SCD patients) [15]. In addition, AYA patients may be expected to have a particular affinity for a Web-based educational tool.

Acknowledgments

This research was supported in part by the Proposed Research Initiated by Students and Mentors Program, University of Maryland School of Medicine Office of Student Research.

This research had no financial backing from Emmi Solutions. Emmi Solutions did provide assistance with access to the ETM and development of the questionnaire and survey.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Pre- and postmodule knowledge-based questionnaire.

[DOCX File, 18KB - mhealth_v7i6e13501_app1.docx]

Multimedia Appendix 2

Postmodule survey.

[DOCX File, 12KB - mhealth_v7i6e13501_app2.docx]

References

1. National Heart, Lung, and Blood Institute. Evidence-Based Management of Sickle Cell Diseaseexpert Panel Report, 2014 URL: https://www.nhlbi.nih.gov/sites/default/files/media/docs/sickle-cell-disease-report%20020816_0.pdf [WebCite Cache ID 75PSwBUGN]
2. Whiteman LN, Haywood C, Lanzkron S, Strouse JJ, Feldman L, Stewart RW. Primary care providers' comfort level in caring for patients with sickle cell disease. South Med J 2015 Sep;108(9):531-536. [doi: 10.14423/SMJ.0000000000000331] [Medline: 26332477]
3. Lanzkron S, Carroll CP, Hill P, David M, Paul N, Haywood C. Impact of a dedicated infusion clinic for acute management of adults with sickle cell pain crisis. Am J Hematol 2015 May;90(5):376-380 [FREE Full text] [doi: 10.1002/ajh.23961] [Medline: 25639822]
4. Colla JS, Kotini-Shah P, Brown SB, Hsu LL. Timely care for sickle cell. Jt Comm J Qual Patient Saf 2017 Dec;43(3):113-115. [doi: 10.1016/j.jcjq.2016.12.007] [Medline: 28334589]
5. Raphael JL, Oyeku SO. Sickle cell disease pain management and the medical home. Hematology Am Soc Hematol Educ Program 2013;2013:433-438. [doi: 10.1182/asheducation-2013.1.433] [Medline: 24319216]
6. Brousseau DC, Owens PL, Mosso AL, Panepinto JA, Steiner CA. Acute care utilization and rehospitalizations for sickle cell disease. J Am Med Assoc 2010 Apr 7;303(13):1288-1294. [doi: 10.1001/jama.2010.378] [Medline: 20371788]
7. Crosby LE, Ware RE, Goldstein A, Walton A, Joffe NE, Vogel C, et al. Development and evaluation of iManage: a self-management app co-designed by adolescents with sickle cell disease. Pediatr Blood Cancer 2017 Jan;64(1):139-145. [doi: 10.1002/pbc.26177] [Medline: 27574031]
8. Molter BL, Abrahamson K. Self-efficacy, transition, and patient outcomes in the sickle cell disease population. Pain Manag Nurs 2015 Jun;16(3):418-424. [doi: 10.1016/j.pmn.2014.06.001] [Medline: 25047808]
9. Quinn CT, Rogers ZR, McCavit TL, Buchanan GR. Improved survival of children and adolescents with sickle cell disease. Blood 2010 Apr 29;115(17):3447-3452 [FREE Full text] [doi: 10.1182/blood-2009-07-233700] [Medline: 20194891]
10. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
11. Ameringer S, Elswick RK, Smith W. Fatigue in adolescents and young adults with sickle cell disease: biological and behavioral correlates and health-related quality of life. J Pediatr Oncol Nurs 2014 Dec;31(1):6-17 [FREE Full text] [doi: 10.1177/1043454213514632] [Medline: 24378816]
12. Ahmadi M, Poormansouri S, Beiranvand S, Sedighie L. Predictors and correlates of fatigue in sickle cell disease patients. Int J Hematol Oncol Stem Cell Res 2018 Jan 1;12(1):69-76 [FREE Full text] [Medline: 29951181]

13. Kas-Osoka CN, Bradley LJ, Coffman R, Orpinas P. Developing online modules for a “Health and Wellness” course: adapting active learning strategies to the online environment. Pedagogy Health Promot 2017 Dec 31;4(4):254-259 [FREE Full text] [doi: 10.1177/2373379917750167]
14. Asnani MR, Quimby KR, Bennett NR, Francis DK. Interventions for patients and caregivers to improve knowledge of sickle cell disease and recognition of its related complications. Cochrane Database Syst Rev 2016 Oct 6;10:CD011175. [doi: 10.1002/14651858.CD011175.pub2] [Medline: 27711980]
15. Jordan L, Swerdlow P, Coates TD. Systematic review of transition from adolescent to adult care in patients with sickle cell disease. J Pediatr Hematol Oncol 2013 Apr;35(3):165-169. [doi: 10.1097/MPH.0b013e3182847483] [Medline: 23511487]

Abbreviations

AYA: adolescent and young adult
eHealth: electronic health
ETM: electronic teaching module
SCD: sickle cell disease
UMMC: University of Maryland Medical Center
VOCs: vaso-occlusive crises

Edited by G Eysenbach; submitted 25.01.19; peer-reviewed by I Osunkwo, C Jonassaint, MS Aslam; comments to author 21.02.19; revised version received 17.04.19; accepted 05.05.19; published 20.06.19.

Please cite as:
Tam T, Baer MR, Hsu LL, Law JY
An Electronic Teaching Module for Improving Knowledge of Self-Management of Vaso-Occlusive Pain Crises in Patients With Sickle Cell Disease: Pilot Questionnaire Study
JMIR Mhealth Uhealth 2019;7(6):e13501
URL: http://mhealth.jmir.org/2019/6/e13501/
doi:10.2196/13501
PMID:31223120

Short Paper

Utilization of Hospital Room Hospitality Features on Patient-Controlled Tablet Computers: Cohort Study

Beiqun Zhao[1], MD; Ming Tai-Seale[2], MPH, PhD; Christopher Longhurst[3,4], MS, MD; Brian Clay[4], MD

[1]Department of Surgery, University of California San Diego, La Jolla, United States

[2]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, United States

[3]Department of Pediatrics, University of California San Diego, La Jolla, CA, United States

[4]Department of Medicine, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Beiqun Zhao, MD
Department of Surgery
University of California San Diego
MC 7220
9300 Campus Point Drive
La Jolla, 92037
United States
Phone: 1 858 246 2518
Email: beiqunmzhao@gmail.com

Abstract

Background: Patient portals tethered to electronic health records can improve patient experience, activation, and outcomes. However, adoption of inpatient portals has been challenging. One way to potentially increase inpatient portal usage is to integrate it with a room control (RC) app on a common tablet computer.

Objective: The aim of this study was to perform a retrospective analysis of patient usage of an RC app provided on tablet computers in patient rooms of our new inpatient tower.

Methods: We identified all patients who were admitted for >24 hours to our new inpatient tower over a 90-day period from September 1 to November 30, 2017. After excluding newborn patients from our analysis, we then identified patients who used the RC app at least one time during their admission. We linked these data to patient demographics (including age, sex, and race) and admitting service. We then performed univariable and multivariable logistic regression to assess patterns of RC app usage.

Results: A total of 3411 patients were admitted over the course of the study period; 2242/3411 (65.73%) used the RC app during their hospitalization. Compared with white patients, other/mixed/unknown race and Asian, Hawaiian, Pacific Islander, American Indian race were significantly associated with increased use of the RC app in a multivariable analysis. Increasing age was significantly associated with increased usage of the RC app. Usage of the RC app also varied by admitting services. Compared with general medicine, bone marrow transplant and general surgery patients had increased usage of the RC app. Conversely, critical care, medical specialties, neurology, surgical subspecialties, and obstetrics/gynecology were all associated with decreased usage of the RC app.

Conclusions: Our study shows that one-third of patients are not using the RC app for critical room functions. Future initiatives to increase RC usage should take these populations into consideration. Contrary to common belief, older patients may use tablet-enabled RCs just as often, if not more often, than younger patients. Certain admitting services, such as neurology and surgical subspecialties, may have had lower usage rates owing to accessibility issues. Our study allows hospitals to tailor support for specific patient populations to increase RC app usage.

(JMIR Mhealth Uhealth 2019;7(6):e13964) doi:10.2196/13964

KEYWORDS

inpatients; electronic health records; patient satisfaction; patients' rooms

Introduction

Electronic health records with patient portals allow patients to conveniently access their health information, which can improve patient experience, patient activation, and patient care [1-3]. Although outpatient portals are becoming more common, challenges persist in the widespread adoption of inpatient portals [4]. However, it has been shown that colocation of the inpatient portal with a room control (RC) app on a tablet can increase patient utilization of the inpatient portal [5] and may increase patient engagement [6].

In addition, RC apps can centralize frequently used and critical room functions (eg, lighting, curtains, and television controls), thereby enabling patients to better control their environment while being confined to a hospital bed. This has led some to consider RC apps as a hospitality feature for patients, which may be a point of contention in today's health care environment owing to increased concerns over hospital spending on nonclinical amenities [7]. However, hospitality features have been associated with improved patient satisfaction [8,9], which can potentially affect hospital reimbursement [10,11]. In addition, hospitality features and improved patient satisfaction can lead to better patient outcomes [12-15]. For example, empowering patients with RC apps may reduce calls to nursing staff for room comfort needs, which can decrease patient call-light burden, leading to improved patient care [16].

However, it is not known which patients are more or less likely to use RCs. Answering this question will allow hospitals to tailor support for specific patient populations to increase RC usage. In this study, we have explored our initial experience with implementing an RC app on tablets in patient rooms at our institution.

Methods

In designing the patient rooms for our new inpatient tower, our institution made the conscious decision to integrate multiple RCs (eg, curtains, lighting, and television controls) into a central patient-facing app (Crestron Electronics) installed on tablet computers (Apple Inc) in every patient room (Figures 1 and 2).

We sought to examine patient factors associated with using RC features on the tablet. After obtaining institutional review board exemption, we analyzed data over a 90-day period from September 1 to November 30, 2017, for all hospital admissions of >24 hours. We examined the proportion of patients who accessed the RC app during their hospitalization by linking RC usage data with patient admission data. After patient linkage, we were able to collect patient demographic information, including sex, age, and race. We also identified the length of stay and admitting service for each patient. This was then categorized into general medicine (including family medicine, internal medicine, and hospitalist services), bone marrow transplant, critical care (including surgical and medical critical care services), general surgery (including transplant, surgical oncology, colorectal, vascular, minimally invasive, and plastic surgery services), medical specialty (including cardiology, pulmonary, medical oncology, and gastroenterology services), neurology (including stroke and neurology services), obstetrics and gynecology, and surgical subspecialty (including urology, head and neck, neurosurgery, and orthopedic surgery services). We excluded newborn patients from our analysis. We performed the Mann-Whitney U test and univariable and multivariable logistic regression analysis using IBM SPSS Statistics (IBM Corp, version 25.0). The level of significance was set at .05 for all analyses.

Figure 1. The main menu of the room control app.

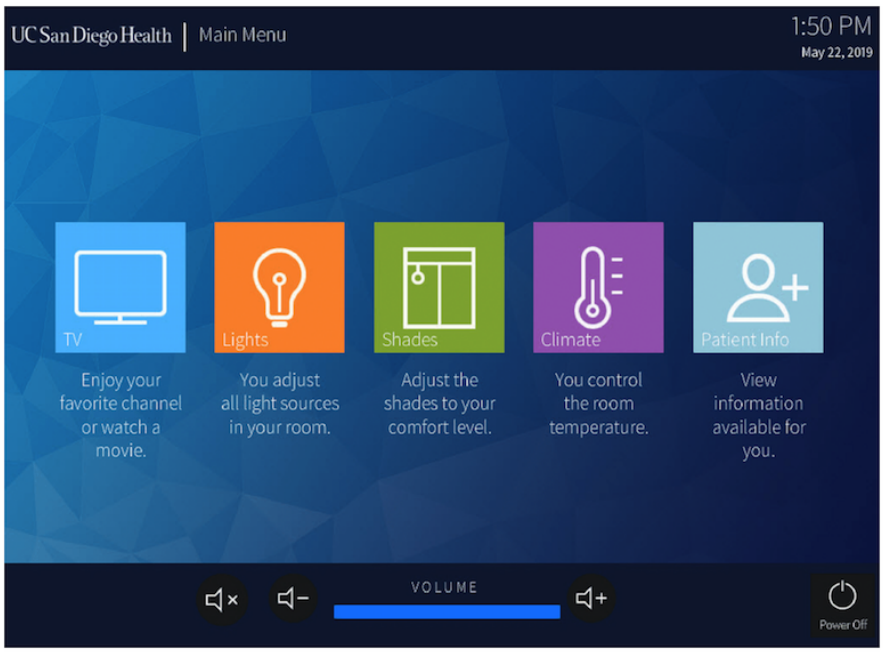
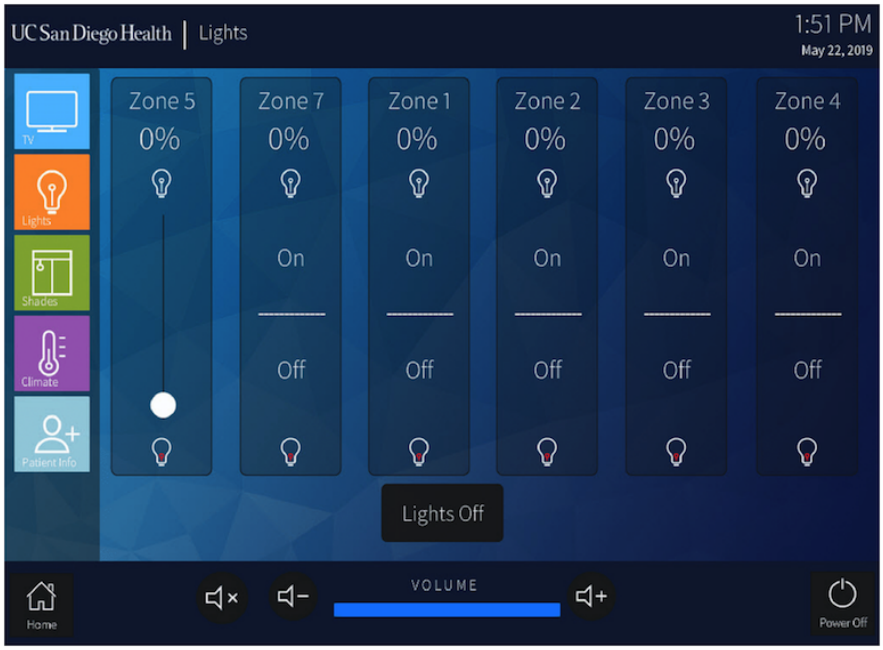


Figure 2. The lighting control submenu of the room control app.



Results

During our study period, 3411 patients were identified. Of these patients, 2242/3411 (65.73%) patients used the RC app at some point during their hospitalization. A comparison of sex, age, race, and admitting service is shown in Table 1. Patients who used the RC app had significantly longer hospital length of stay than patients who did not use the RC app (P<.001). Univariable and multivariable logistic regression is shown in Table 2. In a multivariable analysis, other/mixed/unknown race and Asian, Hawaiian, Pacific Islander, American Indian (AHPIAI) race were associated with increased use of the RC app (odds ratio [OR]=1.27; P=.008 and OR=1.51; P=.006, respectively) compared with white patients. In addition, increasing age was associated with increased usage of the RC app. Compared with general medicine, bone marrow transplant and general surgery patients had increased usage of the RC app. Conversely, critical care, medical specialty, neurology, surgical subspecialty, and obstetrics/gynecology were all associated with decreased usage.

Table 1. Patient demographics stratified by room control app usage (N=3411).

Patient demographic	Used RC[a] app	Did not use RC app
Total number of patients, N (%)	2242 (65.73)	1169 (34.27)
Sex, n (%)		
Female	1542 (68.78)	949 (81.18)
Male	700 (31.22)	220 (18.82)
Race, n (%)		
White	1286 (57.36)	723 (61.85)
Black	97 (4.33)	49 (4.19)
Asian, Hawaiian, Pacific Islander, American Indian	216 (9.63)	82 (7.01)
Other/Mixed/Unknown	643 (28.68)	315 (26.95)
Age (years), n (%)		
16-30	365 (16.28)	337 (28.83)
31-45	740 (33.00)	449 (38.41)
45-65	651 (29.04)	208 (17.79)
>65	486 (21.68)	175 (14.97)
Admission Service, n (%)		
General medicine	378 (16.86)	97 (8.30)
Medicine specialty	14 (0.62)	65 (5.56)
General surgery	419 (18.69)	37 (3.17)
Surgical subspecialty	246 (10.97)	124 (10.61)
Bone marrow transplant	228 (10.17)	13 (1.11)
Critical care	27 (1.20)	51 (4.36)
Neurology	70 (3.12)	43 (3.68)
Obstetrics/Gynecology	860 (38.36)	739 (63.22)
Hospital length of stay, median (range)	5 (1-492)	2 (1-205)

[a]RC: room control.

Table 2. Univariable and multivariable logistic regression.

Variables	Univariable analysis		Multivariable analysis	
	P value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)
Female	<.001	0.511 (0.430-0.606)	.28	—[a]
Race				
White	—	Reference	—	Reference
Black	.555	—	.31	—
Asian, Hawaiian, Pacific Islander, American Indian	.004	1.481 (1.130-1.940)	.006	1.510 (1.129-2.020)
Other/Mixed/Unknown	.097	—	.008	1.274 (1.064-1.526)
Age (years)				
16-30	—	Reference	—	Reference
31-45	<.001	1.522 (1.260-1.838)	<.001	1.494 (1.225-1.822)
45-65	<.001	2.890 (2.330-3.583)	.002	1.591 (1.179-2.149)
>65	<.001	2.564 (2.042-3.219)	.006	1.586 (1.140-2.204)
Admission service				
General medicine	—	Reference	—	Reference
Medicine specialty	<.001	0.055 (0.030-0.103)	<.001	0.036 (0.018-0.071)
General surgery	<.001	2.906 (1.9414.350)	<.001	3.288 (2.186-4.944)
Surgical subspecialty	<.001	0.509 (0.3730.694)	.001	0.588 (0.427-0.808)
Bone marrow transplant	<.001	4.501 (2.4668.215)	<.001	3.288 (2.186-4.944)
Critical care	<.001	0.136 (0.081-0.228)	<.001	0.104 (0.060-0.180)
Neurology	<.001	0.418 (0.269-0.649)	.001	0.472 (0.302-0.739)
Obstetrics/Gynecology	<.001	0.299 (0.234-0.381)	<.001	0.490 (0.352-0.681)
Hospital length of stay	<0.001	1.056 (1.043-1.069)	<.001	1.040 (1.025-1.054)

[a]Not applicable.

Discussion

Principal Findings

We found that 65.7% of patients admitted to our new inpatient tower used an integrated RC app installed on tablet computers inside patient rooms. After controlling for other predictors, we found that the AHPIAI race and other/mixed/unknown race was associated with increased RC app use. Surprisingly, we also found that increasing age was associated with *increased* RC app use. In addition, longer hospital length of stays were predictive of RC app usage. Finally, compared with general medicine, patients admitted to general surgery and bone marrow transplant had more RC app use, whereas patients admitted to medicine specialty, surgical subspecialty, critical care, neurology, and obstetrics/gynecology all had less RC app use than general medicine.

There has been growing concerns that elderly patients may be disadvantaged by the influx of technology in health care today [17]. However, in our study, we found that as age increased, use of the RC app also increased. This may be partly explained by the intuitive nature of the tablet interface, and previous literature has shown success in using tablet-based apps in elderly patients [18]. Furthermore, this finding showcases that RC apps can potentially function as a gateway for elderly patients to access other patient-facing technology.

Secondary Findings

Another interesting finding is the large variation in RC use among admitting services. Patients admitted to the bone marrow transplant service may have had the highest rate of RC app usage because these patients are often admitted to the hospital for prolonged periods of time. However, even after controlling for length of stay, these patients were still more likely to use the RC app. This is most likely because their movement in and out of their unit is limited owing to their disease process and are thereby more likely to explore hospitality features. Increased use of the RC app in general surgery patients may be because these patients are often confined to their hospital beds after surgery and are unable to control the blinds or lights through the usual physical switches on the wall. In this case, the RC app gives patients more control over their environment, which may improve the patient experience during a highly vulnerable time. This pattern is not seen in surgical subspecialty patients, potentially because these patients, especially neurosurgery, head and neck surgery, and orthopedic surgery patients, may not be physically able to operate a tablet computer after surgery. This may also explain the decreased usage of RCs in neurology patients. Therefore, further work will be needed to increase

accessibility in these patients (perhaps through voice-enabled features). Patients admitted to the critical care service are clinically very sick, which may limit their usage of ancillary technology such as tablet computers. Similarly, medicine subspecialty patients may also represent a group of patients with high clinical acuity. Obstetrics and gynecology represent a very diverse group of patients, making interpretations difficult with regard to their usage of the RC app. Future studies will focus on barriers to RC app usage unique to each admitting service and exploring potential ways to increase RC app usage for each service.

Limitations

Given its retrospective nature, we are unable to ensure that RCs were used by the patient and not by a family member. However, RC app use by family members may still encourage use of other apps found on the tablet. It is also possible that the RC app was used by the nursing staff to demonstrate the RCs to the patient, though we suspect this occurrence is rare. Interpretation of use patterns of obstetrics and gynecology patients was difficult because they represent a diverse group of patients, with different indications and acuity. This service includes patients admitted to labor and delivery for observation and patients undergoing postoperative care after a large cancer operation. The former group may be more similar to general medicine patients, whereas the latter group may be more similar to general surgery patients. Unfortunately, we are unable to separate the different obstetrics/gynecology patients in the current analysis. Although RC apps may increase inpatient portal usage [5], we are unable to track the usage of other apps on the tablet at this time. In addition, we do not have granular data of RC app usage, such as average daily use length and specific features accessed. Obtaining this type of data may require a prospective analysis (ie, installing a tracking software) and will be a source of future studies. In addition, patient-level data, such as income and education, were not available. These factors may play a role in the adoption of the RC apps.

Conclusions

Our study shows that approximately one-third of patients are not using the RC app to control critical room functions, and future initiatives to increase RC app usage should take these populations into consideration. Despite its intuitive interface, there may still be accessibility limitations to the current RC app, especially for patients admitted to certain services. A more thorough exploration of why the RC app usage is low in these patient populations is needed in the future.

Acknowledgments

The authors would like to acknowledge the following for their uncompensated assistance in obtaining data for this study: Marc Sylwestrzak BS, Shuxiang Liu MS, and Calvin Fong BS, Information Technology Services Department, University of California San Diego, La Jolla, CA.

BZ is funded by the National Library of Medicine Training Grant: National Institutes of Health grant T15LM011271.

Conflicts of Interest

None declared.

References

1. Vawdrey DK, Wilcox LG, Collins SA, Bakken S, Feiner S, Boyer A, et al. A tablet computer application for patients to participate in their hospital care. AMIA Annu Symp Proc 2011;2011:1428-1435 [FREE Full text] [Medline: 22195206]
2. Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. How outcomes are achieved through patient portals: a realist review. J Am Med Inform Assoc 2014;21(4):751-757. [doi: 10.1136/amiajnl-2013-002501] [Medline: 24503882]
3. Collins SA, Rozenblum R, Leung WY, Morrison CR, Stade DL, McNally K, et al. Acute care patient portals: a qualitative study of stakeholder perspectives on current practices. J Am Med Inform Assoc 2017 Apr 1;24(e1):e9-17. [doi: 10.1093/jamia/ocw081] [Medline: 27357830]
4. Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. What do we know about developing patient portals? A systematic literature review. J Am Med Inform Assoc 2016 Apr;23(e1):e162-e168 [FREE Full text] [doi: 10.1093/jamia/ocv114] [Medline: 26335985]
5. Tai-Seale M, Downing NL, Jones VG, Milani RV, Zhao B, Clay B, et al. Technology-enabled consumer engagement: promising practices at four health care delivery organizations. Health Aff (Millwood) 2019 Mar;38(3):383-390. [doi: 10.1377/hlthaff.2018.05027] [Medline: 30830826]
6. McAlearney AS, Sieck CJ, Hefner JL, Aldrich AM, Walker DM, Rizer MK, et al. High touch and high tech (HT2) proposal: transforming patient engagement throughout the continuum of care by engaging patients with portal technology at the bedside. JMIR Res Protoc 2016 Nov 29;5(4):e221 [FREE Full text] [doi: 10.2196/resprot.6355] [Medline: 27899338]
7. Rosenthal E. New York Times. 2013 Sep 21. Is This a Hospital or a Hotel? URL: https://www.nytimes.com/2013/09/22/sunday-review/is-this-a-hospital-or-a-hotel.html
8. Wu Z, Robson S, Hollis B. The application of hospitality elements in hospitals. J Healthc Manag 2013;58(1):47-62; discussion 62-3. [Medline: 23424818]

9. Wensley C, Botti M, McKillop A, Merry AF. A framework of comfort for practice: an integrative review identifying the multiple influences on patients' experience of comfort in healthcare settings. Int J Qual Health Care 2017 Apr 1;29(2):151-162. [doi: 10.1093/intqhc/mzw158] [Medline: 28096279]
10. Mehta SJ. Patient satisfaction reporting and its implications for patient care. AMA J Ethics 2015 Jul 1;17(7):616-621 [FREE Full text] [doi: 10.1001/journalofethics.2015.17.7.ecas3-1507] [Medline: 26158808]
11. Medicare. 2019. Survey of Patients' Experience (HCAHPS) URL: https://www.medicare.gov/hospitalcompare/Data/Overview.html [accessed 2019-03-07]
12. Sacks GD, Lawson EH, Dawes AJ, Russell MM, Maggard-Gibbons M, Zingmond DS, et al. Relationship between hospital performance on a patient satisfaction survey and surgical quality. JAMA Surg 2015 Sep;150(9):858-864. [doi: 10.1001/jamasurg.2015.1108] [Medline: 26108091]
13. Glickman SW, Boulding W, Manary M, Staelin R, Roe MT, Wolosin RJ, et al. Patient satisfaction and its relationship with clinical quality and inpatient mortality in acute myocardial infarction. Circ Cardiovasc Qual Outcomes 2010 Mar;3(2):188-195 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.109.900597] [Medline: 20179265]
14. Boulding W, Glickman SW, Manary MP, Schulman KA, Staelin R. Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. Am J Manag Care 2011 Jan;17(1):41-48 [FREE Full text] [Medline: 21348567]
15. Trzeciak S, Gaughan JP, Bosire J, Mazzarelli AJ. Association between medicare summary star ratings for patient experience and clinical outcomes in US hospitals. J Patient Exp 2016 Apr 7;3(1):6-9. [doi: 10.1177/2374373516636681]
16. Tzeng HM. Perspectives of staff nurses of the reasons for and the nature of patient-initiated call lights: an exploratory survey study in four USA hospitals. BMC Health Serv Res 2010 Feb 26;10:52. [doi: 10.1186/1472-6963-10-52] [Medline: 20184775]
17. Anthony DL, Campos-Castillo C, Lim PS. Who isn't using patient portals and why? Evidence and implications from a national sample Of US adults. Health Aff (Millwood) 2018 Dec;37(12):1948-1954. [doi: 10.1377/hlthaff.2018.05117] [Medline: 30633673]
18. Mehra S, Visser B, Cila N, van den Helder J, Engelbert RH, Weijs PJ, et al. Supporting older adults in exercising with a tablet: a usability study. JMIR Hum Factors 2019 Feb 1;6(1):e11598. [doi: 10.2196/11598] [Medline: 30707106]

Abbreviations

AHPIAI: Asian, Hawaiian, Pacific Islander, American Indian
OR: odds ratio
RC: room control

Edited by G Eysenbach; submitted 08.03.19; peer-reviewed by A Chan, M Whelan; comments to author 06.05.19; revised version received 23.05.19; accepted 23.05.19; published 20.06.19.

Please cite as:
Zhao B, Tai-Seale M, Longhurst C, Clay B
Utilization of Hospital Room Hospitality Features on Patient-Controlled Tablet Computers: Cohort Study
JMIR Mhealth Uhealth 2019;7(6):e13964
URL: http://mhealth.jmir.org/2019/6/e13964/
doi:10.2196/13964
PMID:31223118

Original Paper

Mobile Health Management Platform–Based Pulmonary Rehabilitation for Patients With Non–Small Cell Lung Cancer: Prospective Clinical Trial

Wonjun Ji[1*], MD; Hee Kwon[2*], BS; Sungin Lee[2], PhD; Seulgi Kim[1], MA; Jeong Sook Hong[1], BA; Yu Rang Park[3], PhD; Hyeong Ryul Kim[3], MD, PhD; Jae Cheol Lee[4], MD, PhD; Eun Ji Jung[2], PhD; Donghyun Kim[2], MS; Chang-Min Choi[1,4], MD, PhD

[1]Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[2]LifeSemantics, Seoul, Republic of Korea
[3]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea
[4]Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Chang-Min Choi, MD, PhD
Department of Pulmonary and Critical Care Medicine
Asan Medical Center
University of Ulsan College of Medicine
88 Olympic-ro 43-gil, Songpa-gu
Seoul, 05505
Republic of Korea
Phone: 82 2 3010 6968
Email: ccm@amc.seoul.kr

Abstract

Background: Lung cancer patients experience various symptoms during treatment. Although pulmonary rehabilitation is an effective way to improve these symptoms, a medical environment of limited availability makes it difficult to provide seamless and adequate rehabilitation for lung cancer patients.

Objective: This study aimed to investigate the effects of a personalized pulmonary rehabilitation program using real-time mobile patient health data for patients with non–small cell lung cancer.

Methods: We conducted a prospective clinical trial in 64 patients with non–small cell lung cancer aged between 20 and 80 years at a large tertiary hospital in Seoul, South Korea. A 12-week personalized pulmonary rehabilitation program, called efil breath, was administered to determine the effectiveness of the newly developed rehabilitation app. Participants were randomly allocated to the fixed exercise or fixed-interactive exercise group (which received the personalized program). We measured changes in 6-minute walk distance (6MWD) and dyspnea (modified Medical Research Council [mMRC] score) at 6 weeks; and quality of life and service satisfaction at 12 weeks. We used the paired *t* test to analyze the variables.

Results: Patients used the newly developed mobile health pulmonary rehabilitation app and a real-time patient monitoring website. In all participants, significant changes were observed in 6MWD at 12 weeks from a mean of 433.43m (SD 65.60) to 471.25m (SD 75.69; *P*=.001), and mMRC from a mean score of 0.94 (0.66) to 0.61 (SD 0.82; *P*=.02). The intervention significantly improved their quality of life (EuroQol-visual analog scale [EQ-VAS]) compared with baseline (mean score 76.05, SD 12.37 vs 82.09, SD 13.67, respectively; *P*=.002).

Conclusions: A personalized mobile health–based pulmonary rehabilitation app for recording and monitoring real-time health data of patients with non–small cell lung cancer can supplement traditional health care center–based rehabilitation programs. This technology can encourage improvement of physical activity, dyspnea, and quality of life.

(JMIR Mhealth Uhealth 2019;7(6):e12645) doi:10.2196/12645

KEYWORDS

mHealth; pulmonary rehabilitation; lung cancer; telemedicine; telerehabilitation; carcinoma, non-small-cell lung

Introduction

Background

Lung cancer is a leading cause of death from cancer worldwide [1], with high rates of morbidity and mortality, as well as having a high burden of symptoms, such as dyspnea, fatigue, anxiety, depression, and pain [2,3]. These symptoms persist after diagnosis [4] but occur even in early stages of the disease [5]. A previous study showed that the survival of patients with lung cancer is improved when these symptoms are effectively managed [6]. Therefore, it is very important to provide appropriate palliative care for patients with lung cancer.

Pulmonary rehabilitation (PR) is a multidisciplinary, comprehensive treatment intervention for patients with chronic respiratory diseases [7]. PR has been shown to improve the lives of patients with chronic obstructive pulmonary disease [8], interstitial lung disease [9], and lung cancer. Although many studies have addressed the use of PR to reduce postoperative complications or improve postoperative outcomes among patients with lung cancer undergoing surgery [10-17], some studies have reported that PR can also improve exercise capacity, alleviate symptoms, and improve the quality of life (QoL) of patients with lung cancer who have received chemotherapy or radiation therapy [18-21]. Despite the reported effectiveness of PR, the application of PR to patients with lung cancer in real clinical practice environments is limited due to long waiting lists for rehabilitation services [22]. To overcome these limitations, studies on the effects of home-based PR on lung cancer patients have been reported [23-27], giving rise to the development of mobile technologies in health care delivery.

Objective

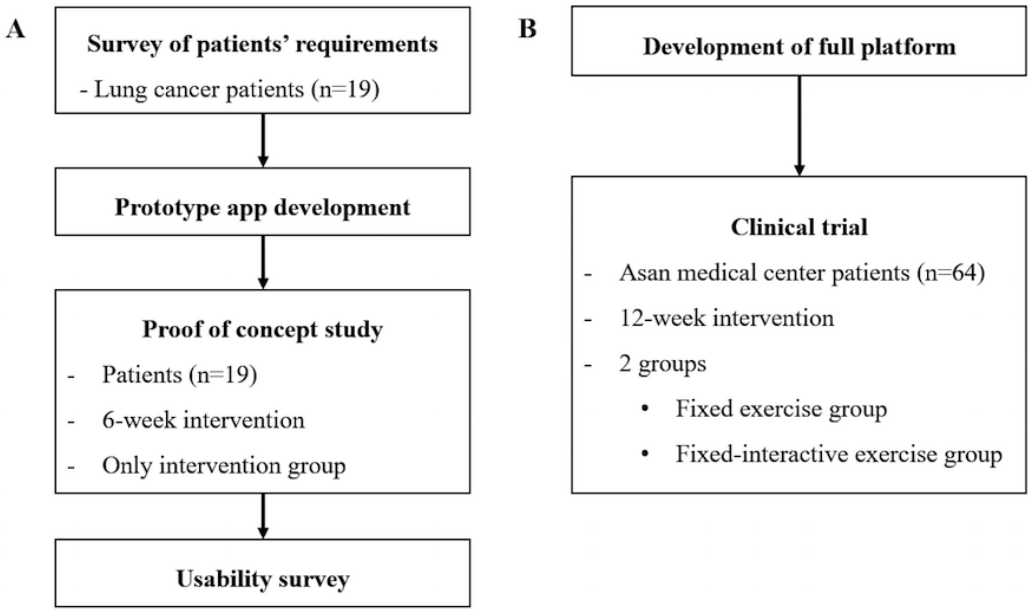
The aim of this study was to examine the outcome of home-based PR regarding exercise capacity, dyspnea symptoms, and QoL in adult patients who were being treated for non–small cell lung cancer (NSCLC). To achieve this goal, we developed efil breath, a mobile health (mHealth) rehabilitation platform for patients with NSCLC that custom tailors symptom management and exercise programs to individual patient's health conditions. The platform comprises 2 apps for pulmonary rehabilitation and a patient monitoring website. Patient compliance and progress were monitored for timely intervention by health care professionals. We conducted a prospective clinical trial to investigate the effectiveness of the platform. To our knowledge, this was the first clinical trial of an mHealth PR service for patients in South Korea with lung cancer.

Methods

Study Design

We conducted the study in 2 phases: (1) a proof of concept phase, and (2) full platform development and a clinical trial (Figure 1). In the second phase, we developed 2 apps: one with a fixed exercise program, and the other with interactive exercise programs that conformed to the individual patient's physical capacities. We also developed a patient monitoring website to collect data from the apps in real time. Although the main goal of the clinical trial was to evaluate the effects of home-based PR in the entire study population, we also evaluated the effect of PR according to the 2 apps.

Figure 1. Study design. (A) Flow of the proof-of-concept study; (B) flow of the clinical trial.



Development of a Personalized Mobile Health–Based Pulmonary Rehabilitation Platform

Apps

Use of the apps and patient monitoring website within the mHealth rehabilitation platform are described elsewhere [28]. The platform was developed as a comprehensive rehabilitation-focused platform for patients with chronic obstructive pulmonary disease [28], lung cancer, breast cancer, and others. Figures 2 and 3 show screenshots of the app for the fixed exercise program. The contents of the app were translated from Korean to English by HK.

Figure 2. App screenshots showing (A) the 6-minute walk test, (B) exercise routines, and (C) a summary of the results.

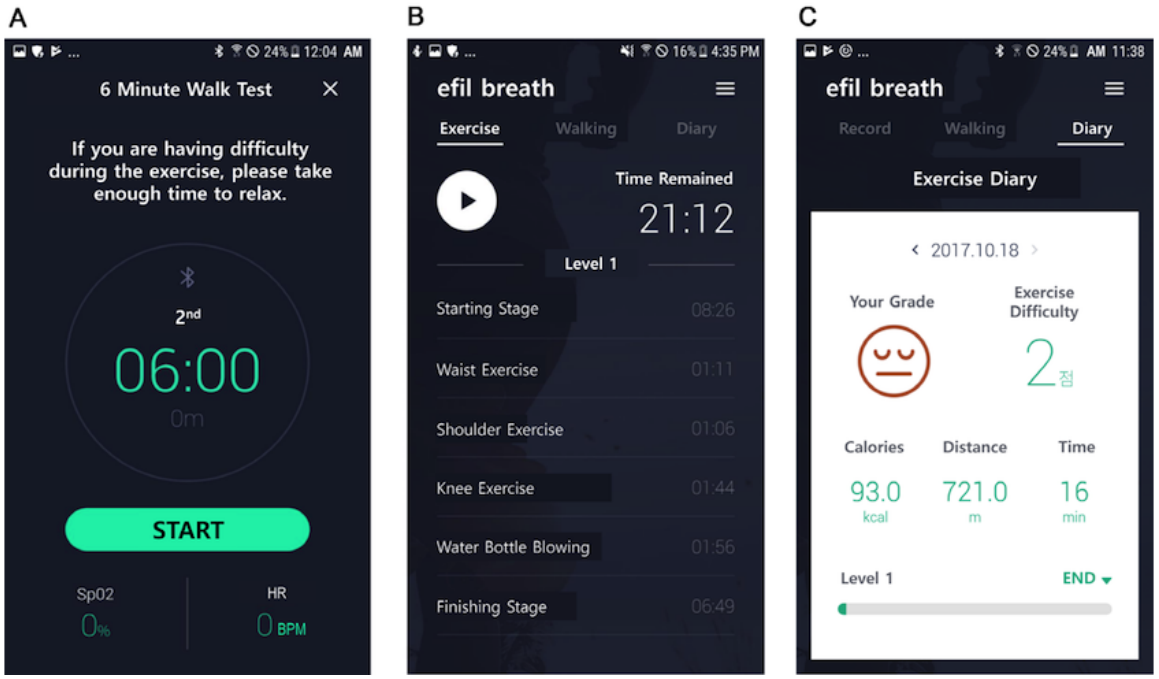
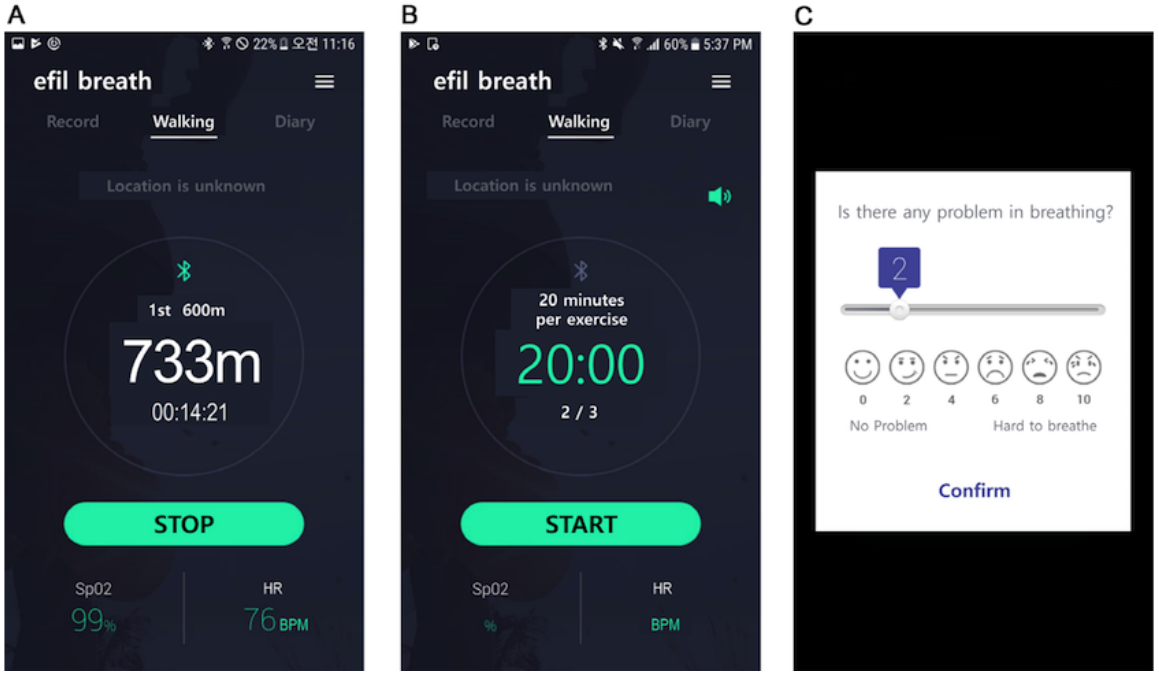


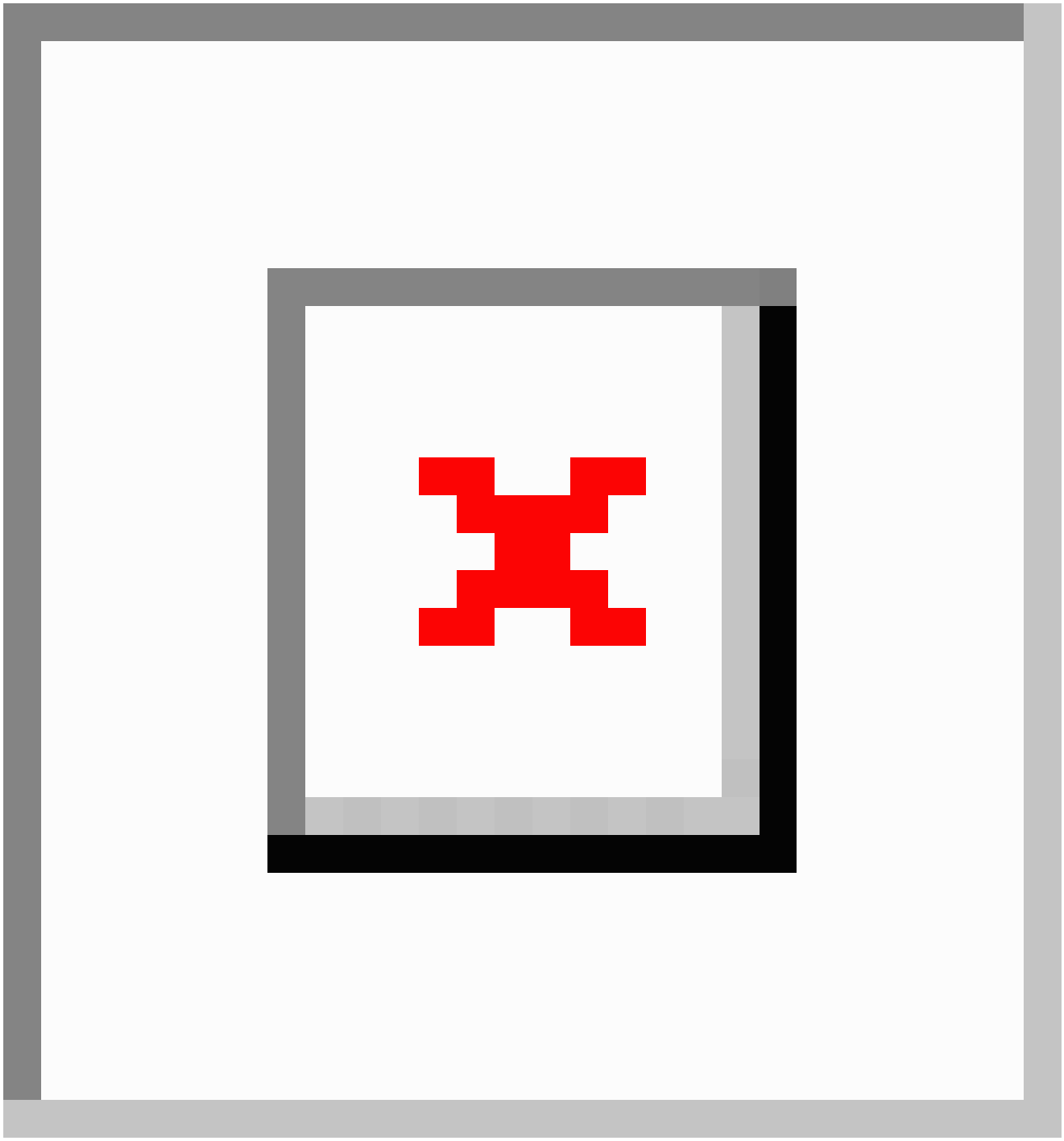
Figure 3. App screenshots for the walking exercise showing (A) the operation screen, (B) standby screen, and (C) measurement of dyspnea.



Patient Monitoring Website

All patient-generated health data, such as 6-minute walk test (6MWT) results, rehabilitation exercise progress, heart rate, and breathing difficulty levels, are sent from the apps to a central database. Health care professionals, such as lung cancer specialists and nurses, can access the database through the monitoring website but are permitted to view only the patients' data. When an individual patient is selected, the patient's health status is displayed, as Figure 4 shows.

Figure 4. Patient health data screen showing detailed information about exercise and health measurements for a fictional patient.



Clinical Trial

We recruited participants from May 2017 to December 2017 from the outpatient clinics of Asan Medical Center, a tertiary hospital in South Korea. This study was approved by the institutional review board of Asan Medical Center. The NSCLC participant inclusion criteria were as follows: (1) adults aged 20 to 80 years, (2) patients with histologically confirmed NSCLC, (3) the presence of obstructive ventilatory disturbance, defined by a forced expiratory volume in the first second of inspiration of less than 80% in spirometry, (4) the ability to walk more than 150 m in a 6MWT, (5) possession of an Android mobile phone, and (6) patients who agreed to provide informed written consent before the study. Patients who met the following criteria were excluded: (1) having severe cerebrovascular or musculoskeletal disease and being unable to follow the exercise regimen, (2) being illiterate or having limited communication ability, (3) having a major disability that can cause death within 1 year, or (4) declining to sign the consent form.

At the initial visit, all the participants were randomly allocated to 1 of 2 groups for 12 weeks: a fixed exercise group and a fixed-interactive exercise group. The fixed exercise group used only the fixed exercise program during the 12 weeks, whereas the fixed-interactive exercise group received the app with the fixed exercise regimen for the first 6 weeks and then switched to an app with an interactive exercise regimen for the remaining 6 weeks. We based this sequence on the expectation that the interactive exercise program could provide more effective

personalized exercise based on the participant's PR record for the previous 6 weeks. All participants performed a 6MWT, which is a standard indicator of physical activity in PR. We measured subjective dyspnea (modified Medical Research Council [mMRC] symptom score) at baseline (visit [V] 1), 6 weeks (V2), and 12 weeks (V3). Participants were provided a wearable pulse oximeter (Checkme O2, Viatom, China) to monitor their pulse rate and oxygen saturation during exercise. QoL was measured by the EuroQol 5 dimensions questionnaire (EQ-5D), which is a standardized instrument for measuring general health status and has been widely used in clinical studies and for routine outcome measurement of operational health care [29].

The primary end points were the pulmonary function parameters 6-minute walk distance (6MWD) and mMRC grade of dyspnea at 12 weeks. The secondary outcomes were QoL (EuroQol-visual analog scale [EQ-VAS] and EQ-5D) and participant satisfaction (Patient Global Assessment [PGA]) based on a 5-point Likert scale (1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=strongly agree) measured at V3 (12 weeks). After confirming the change in outcome values in the entire study group, we also analyzed the difference between the fixed and fixed-interactive group.

Measurement and Data Analysis

We performed all statistical analyses using SAS version 9.4 (SAS Institute Inc). Paired *t* test was used to evaluate the changes in pulmonary parameters in all participants and to compare the 2 exercise program groups. We considered $P<.05$ to be statistically significant. Based on previous studies of 6MWD improvement in related trials [12], we hypothesized that participants would have a mean improvement in 6MWD of 30 m at 12 weeks. Assuming an SD of 60, 2-sided test at an alpha level of .05, power of 80%, and dropout rate of 20%, a sample size of 60 patients (30 per group) was required for the primary analysis.

Results

Baseline Participant Characteristics

We screened a total of 158 patients. Of these, 64 patients (40.5%) met the inclusion criteria and were enrolled into the 2 groups (Figure 5). Table 1 shows participants' demographics and baseline characteristics.

Figure 5. Flow of the study groups. (a) Participants using the fixed-regimen app; (b) participants using the interactive-regimen app. IC: informed consent.

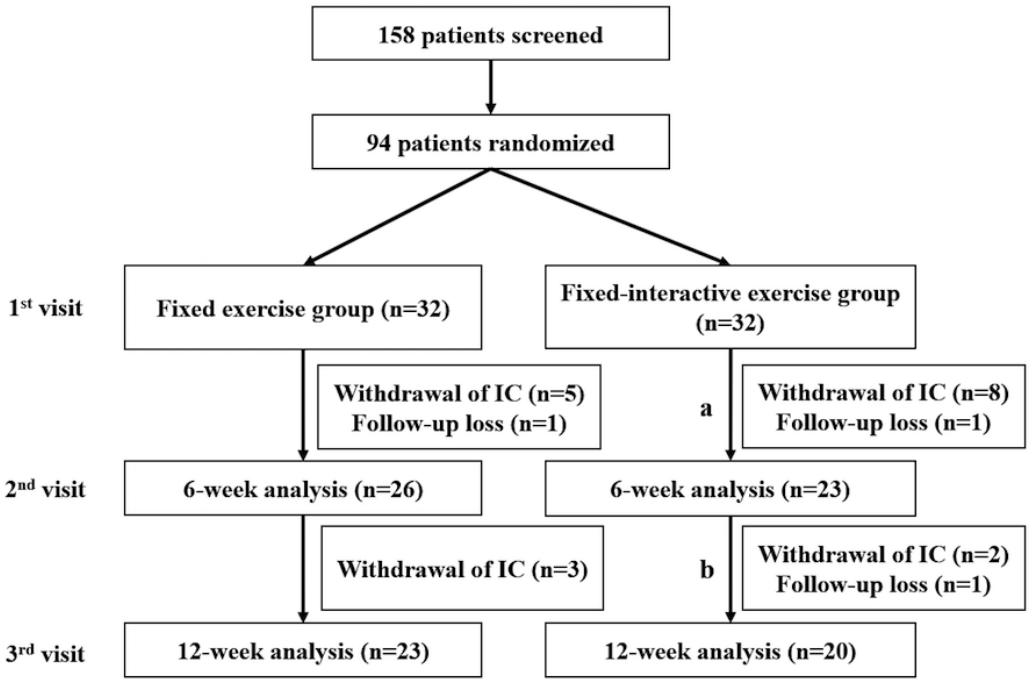


Table 1. Baseline participant characteristics (N=64).

Variables	Fixed exercise group (n=32)		Fixed-interactive exercise group (n=32)		*P* value
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Age (years)		57.97 (9.868)		60.50 (10.198)	.32
Sex: male	21 (66)		24 (75)		.41
Body mass index (kg/m^2)		23.939 (3.332)		23.458 (3.452)	.57
Smoking status					.59
Never	13 (41)		10 (31)		
Former smoker	11 (34)		15 (47)		
Current smoker	8 (25)		7 (22)		
Cormobidities					
Diabetes mellitus	7 (22)		5 (16)		.52
Hypertension	9 (28)		11 (34)		.59
Cerebrovascular disease	0 (0)		0 (0)		
Tuberculosis	4 (13)		7 (22)		.32
Other	20 (63)		14 (44)		.27
Previous surgery	28 (88)		21 (66)		.04
Stage of lung cancer					.32
I	13 (42)		7 (22)		
II	3 (10)		5 (16)		
IIIA	6 (19)		7 (22)		
IIIB	1 (3)		0 (0)		
IV	8 (26)		13 (41)		
Histologic type					>.99
Adenocarcinoma	27 (84)		27 (84)		
Squamous	5 (16)		4 (13)		
NSCLC[a] and others	0 (0)		1 (3)		
Baseline pulmonary function tests					
FEV_1[b] (% predicted)		71.31 (18.220)		68.19 (10.772)	.41
FVC[c] (% predicted)		75.03 (17.453)		74.03 (11.232)	.79
FEV_1 / FVC[d] (% predicted)		0.96 (0.150)		0.93 (0.112)	.37
6-minute walk distance (m)		433.13 (55.790)		427.13 (79.034)	.73
Modified Medical Research Council dyspnea scale score					.07
0	11 (35)		3 (10)		
1	17 (53)		25 (78)		
2	3 (10)		2 (6)		
3	1 (3)		2 (6)		
EQ-5D[e]		7.50 (1.951)		7.50 (2.356)	>.99

[a]NSCLC: non–small cell lung cancer.

[b]FEV_1: forced expiratory volume in the first second of inspiration; normal value ≥80%.

[c]FVC: forced vital capacity; normal value ≥80%.

[d]FEV_1 / FVC: normal value ≥70%.

[e]EQ-5D: EuroQol 5 dimensions questionnaire.

Comparison Between Preintervention and Postintervention Changes in Primary and Secondary Outcomes by Participant

For all participants in both groups, the 6MWD improved significantly from V1 to V3 (from mean 433.43 m [SD 65.60] to 471.25m [SD 75.69]; *P*=.001). Similarly, mMRC score improved from V1 to V3 (from 0.94, SD 0.66 to 0.61, SD 0.82; *P*=.02; Table 2). As a measure of QoL, EQ-VAS score improved significantly from 76.05 (SD 12.37) at V1 to 82.09, (SD 13.67) at V3 (*P*=.002). The mean value of EQ-5D was not significantly different between time points (Table 2). PGA scores measured at V3 showed significant improvement over PGA scores at V2 (from 13.77, SD 3.68 to 15.08, SD 3.99; *P*=.01).

Table 2. Results of all participants' outcome variables according to time point.

Variables	Visit 1 (baseline, N=49)		Visit 2 (6 weeks, N=49)			Visit 3 (12 weeks, N=43)		
	n (%)	mean (SD)	n (%)	mean (SD)	*P* value	n (%)	mean (SD)	*P* value
6MWD[a] (m)	49 (100)	433.429 (65.595)	42 (86)	448.095 (76.421)	.14	40 (93)	471.250 (75.691)	.001
mMRC[b] score	49 (100)	0.939 (0.659)	45 (92)	0.733 (0.618)	.04	43 (100)	0.605 (0.821)	.02
EQ-VAS[c] score	43 (88)	76.047 (12.371)	N/A[d]	N/A	N/A	43 (100)	82.093 (13.674)	.002
EQ-5D[e] score	43 (88)	7.535 (1.817)	N/A	N/A	N/A	43 (100)	6.930 (2.849)	.17
PGA[f] score	N/A	N/A	39 (80)	13.769 (3.681)	N/A	39 (91)	15.077 (3.989)	.01

[a]6MWD: 6-minute walk distance.

[b]mMRC: modified Medical Research Council.

[c]EQ-VAS: EuroQol-visual analog scale.

[d]N/A: not applicable.

[e]EQ-5D: EuroQol 5 dimensions questionnaire.

[f]PGA: Patient Global Assessment.

Table 3. Preintervention and postintervention changes in outcome variables between the 2 groups.

Variables	Fixed exercise group (N=23)		Fixed-interactive exercise group (N=20)		*P* value
	n (%)	Mean (SD)	n (%)	Mean (SD)	
6MWD[a] (m)	21 (91)	58.095 (73.663)	19 (95)	25.368 (66.640)	.30
mMRC[b] score	23 (100)	–0.435 (0.945)	20 (100)	–0.250 (0.910)	.68
EQ-VAS[c] score	23 (100)	6.304 (9.073)	20 (100)	5.750 (14.825)	.99
EQ-5D[d] score	23 (100)	–0.957 (1.745)	20 (100)	–0.200 (3.722)	.50

[a]6MWD: 6-minute walk distance.

[b]mMRC: modified Medical Research Council.

[c]EQ-VAS: EuroQol-visual analog scale.

[d]EQ-5D: EuroQol 5 dimensions questionnaire.

Comparison Between Preintervention and Postintervention Changes in Primary and Secondary Outcomes by Exercise Program

Unlike the results of all participants, there were no statistical differences in the primary and secondary outcome indicators between the fixed and fixed-interactive exercise groups (Table 3).

Discussion

Principal Findings

We developed a personalized mHealth PR platform to provide rehabilitation treatment for patients with lung cancer. Our mobile app-based PR improved exercise capacity (as measured by 6MWD), dyspnea grade, and QoL. This suggests that a mobile-based home rehabilitation program might be an alternative method of conventional PR for patients with lung cancer, especially under the limitations of the clinical environment in South Korea, where many PR programs are inadequate. We demonstrated that personalized PR using mobile technology significantly improved exercise capacity and QoL for patients with lung cancer regardless of disease status.

Previous studies that reported the effects of PR in patients with lung cancer focused mostly on perioperative periods [10-17,23], whereas only a few studies assessed patients with advanced lung cancer [19,21,24,27,30]. In this study, we analyzed the effect of PR on both early- and advanced-stage lung cancer patients, and found improvements in exercise capacity (6MWD) and dyspnea, consistent with previous studies that reported improvements in exercise capacity [10-13,16,21,24,27] or dyspnea grade [21,27].

Only a few studies have reported on the effect of PR on the QoL of patients with lung cancer [11,24,26,27], and evidence is still insufficient. Although Stigt et al [11] reported that PR did not improve patients' QoL after lung cancer surgery, other studies supported the positive effect of PR [26,27]. Our study found that EQ-VAS was significantly improved at the end of PR, compared with baseline, which is consistent with previous studies [24] reporting significant improvement in daily physical activity and anxiety scores. However, we did not find a significant improvement in EQ-5D, which measures QoL in 5 dimensions: mobility, self-care, daily activity, pain/discomfort, and anxiety/depression. Each item is scored from 1 to 5; the higher the score, the lower the QoL. On the other hand, EQ-VAS scores reflect patients' subjective perception of QoL on a continuous scale from 0 to 100 points. A score of 100 indicates the best health condition, whereas a score of 0 indicates the worst condition. In our study, the EQ-VAS scores showed a more widely distributed pattern than the EQ-5D scores; we suggest that this difference could be responsible for the difference in the measures' statistical significance. Importantly, these 2 indicators have consistently shown that PR can improve the QoL of patients with lung cancer.

Limitations

This study had a few limitations. First, the proportion of patients with advanced lung cancer was relatively small. About 70% of the patients in this study had surgical treatments prior to PR, and about 33% were in an advanced stage of the disease. Considering that lung cancer is diagnosed at advanced stages in about two-thirds of patients [1], further studies involving more patients with advanced-stage lung cancer are needed to better confirm the effects of mobile-based PR.

Second, this study reported the effects of 12-week PR on patients with lung cancer, but it did not assess long-term effects. Most studies of PR for patients with lung cancer have reported only a 12-week follow-up, similar to this study. Future studies are required to confirm the long-term benefit of PR platforms for both patients and health care professionals. Also, this mobile-based personalized rehabilitation platform should be studied in patients with lung cancer who live in medically deprived areas, such as remote islands with small populations.

Third, only 40.5% of the screened patients were enrolled in the study. As Table 1 shows, the participants' mean age was around 60 years; thus, many patients had difficulty with handling the technology or did not own an adequate device for this study. Some patients declined to use this app-based PR service. Together, these suggest that technology limitations might affect the application of home-based PR service among the elderly. Therefore, further research is warranted to study how to best provide effective PR service using smart devices to older patients.

Fourth, we found no significant difference between the fixed and the fixed-interactive exercise regimen groups. To our knowledge, this was the first clinical study to confirm the effectiveness of home-based PR using app-based exercise programs in patients with lung cancer. Although the main goal of this study was to evaluate the effects of home-based PR, we expected that the interactive exercise program could provide more effective personalized exercise based on the participant's PR record for the previous 6 weeks. We suggest that the 6-week duration was likely too short to create a meaningful difference between the 2 groups. Further prospective studies are needed to evaluate the effects of long-term exercise according to whether participants followed a fixed or fixed-interactive exercise program.

Conclusions

Personalized mHealth-based PR in patients with NSCLC can supplement traditional health care center–based rehabilitation programs. Our clinical trial results support the use of this technology to improve exercise capacity, dyspnea symptoms, and QoL. mHealth technology is now a robust and supplementary tool for self-management, remote monitoring, and rehabilitation for patients who require chronic care. The relevance of mHealth-based rehabilitation is enhanced only when its rehabilitative efficacy is corroborated by evidence-based clinical results. Evidence-based PR for lung cancer enables prescription of exercise regimens commensurate with a patient's individual physical status. To our knowledge, efil breath is the first attempt in South Korea at developing an mHealth management platform for patients with lung cancer. The results of this study may form the foundation of other mHealth-based PR endeavors to support optimal home-based self-management.

Acknowledgments

This work was supported by the Creative Industrial Technology Development Program (10053249, Development of Personalized Healthcare System Exploiting User Life-log and Open Government Data for Business Service Model Proof on Whole Life Cycle Care) funded by the Ministry of Trade, Industry & Energy (South Korea). We would like to thank Enago for the English language review.

Conflicts of Interest

None declared.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018 Jan;68(1):7-30 [FREE Full text] [doi: 10.3322/caac.21442] [Medline: 29313949]

2. Morrison EJ, Novotny PJ, Sloan JA, Yang P, Patten CA, Ruddy KJ, et al. Emotional problems, quality of life, and symptom burden in patients with lung cancer. Clin Lung Cancer 2017 Dec;18(5):497-503. [doi: 10.1016/j.cllc.2017.02.008] [Medline: 28412094]
3. Iyer S, Roughley A, Rider A, Taylor-Stokes G. The symptom burden of non-small cell lung cancer in the USA: a real-world cross-sectional study. Support Care Cancer 2014 Jan;22(1):181-187. [doi: 10.1007/s00520-013-1959-4] [Medline: 24026981]
4. Cheville AL, Novotny PJ, Sloan JA, Basford JR, Wampfler JA, Garces YI, et al. Fatigue, dyspnea, and cough comprise a persistent symptom cluster up to five years after diagnosis with lung cancer. J Pain Symptom Manage 2011 Aug;42(2):202-212 [FREE Full text] [doi: 10.1016/j.jpainsymman.2010.10.257] [Medline: 21398090]
5. Koczywas M, Williams AC, Cristea M, Reckamp K, Grannis FW, Tiep BL, et al. Longitudinal changes in function, symptom burden, and quality of life in patients with early-stage lung cancer. Ann Surg Oncol 2013 Jun;20(6):1788-1797 [FREE Full text] [doi: 10.1245/s10434-012-2741-4] [Medline: 23143593]
6. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA, et al. Early palliative care for patients with metastatic non-small-cell lung cancer. N Engl J Med 2010 Aug 19;363(8):733-742. [doi: 10.1056/NEJMoa1000678] [Medline: 20818875]
7. Nici L, Donner C, Wouters E, Zuwallack R, Ambrosino N, Bourbeau J, ATS/ERS Pulmonary Rehabilitation Writing Committee. American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation. Am J Respir Crit Care Med 2006 Jun 15;173(12):1390-1413. [doi: 10.1164/rccm.200508-1211ST] [Medline: 16760357]
8. Zeng Y, Jiang F, Chen Y, Chen P, Cai S. Exercise assessments and trainings of pulmonary rehabilitation in COPD: a literature review. Int J Chron Obstruct Pulmon Dis 2018;13:2013-2023 [FREE Full text] [doi: 10.2147/COPD.S167098] [Medline: 29983556]
9. Varadi RG, Goldstein RS. Pulmonary rehabilitation for restrictive lung diseases. Chest 2010 Feb;137(2):247-248. [doi: 10.1378/chest.09-1857] [Medline: 20133284]
10. Bobbio A, Chetta A, Ampollini L, Primomo GL, Internullo E, Carbognani P, et al. Preoperative pulmonary rehabilitation in patients undergoing lung resection for non-small cell lung cancer. Eur J Cardiothorac Surg 2008 Jan;33(1):95-98. [doi: 10.1016/j.ejcts.2007.10.003] [Medline: 18006327]
11. Stigt JA, Uil SM, van Riesen SJH, Simons FJNA, Denekamp M, Shahin GM, et al. A randomized controlled trial of postthoracotomy pulmonary rehabilitation in patients with resectable lung cancer. J Thorac Oncol 2013 Feb;8(2):214-221 [FREE Full text] [doi: 10.1097/JTO.0b013e318279d52a] [Medline: 23238118]
12. Lai Y, Su J, Qiu P, Wang M, Zhou K, Tang Y, et al. Systematic short-term pulmonary rehabilitation before lung cancer lobectomy: a randomized trial. Interact Cardiovasc Thorac Surg 2017 Dec 01;25(3):476-483. [doi: 10.1093/icvts/ivx141] [Medline: 28520962]
13. Bradley A, Marshall A, Stonehewer L, Reaper L, Parker K, Bevan-Smith E, et al. Pulmonary rehabilitation programme for patients undergoing curative lung cancer surgery. Eur J Cardiothorac Surg 2013 Oct;44(4):e266-e271. [doi: 10.1093/ejcts/ezt381] [Medline: 23959742]
14. Maeda K, Higashimoto Y, Honda N, Shiraishi M, Hirohata T, Minami K, et al. Effect of a postoperative outpatient pulmonary rehabilitation program on physical activity in patients who underwent pulmonary resection for lung cancer. Geriatr Gerontol Int 2016 May;16(5):550-555. [doi: 10.1111/ggi.12505] [Medline: 25953128]
15. Saito H, Hatakeyama K, Konno H, Matsunaga T, Shimada Y, Minamiya Y. Impact of pulmonary rehabilitation on postoperative complications in patients with lung cancer and chronic obstructive pulmonary disease. Thorac Cancer 2017 Dec;8(5):451-460 [FREE Full text] [doi: 10.1111/1759-7714.12466] [Medline: 28696575]
16. Morano MT, Araújo AS, Nascimento FB, da Silva GF, Mesquita R, Pinto JS, et al. Preoperative pulmonary rehabilitation versus chest physical therapy in patients undergoing lung cancer resection: a pilot randomized controlled trial. Arch Phys Med Rehabil 2013 Jan;94(1):53-58. [doi: 10.1016/j.apmr.2012.08.206] [Medline: 22926460]
17. Benzo R, Wigle D, Novotny P, Wetzstein M, Nichols F, Shen RK, et al. Preoperative pulmonary rehabilitation before lung cancer resection: results from two randomized studies. Lung Cancer 2011 Dec;74(3):441-445 [FREE Full text] [doi: 10.1016/j.lungcan.2011.05.011] [Medline: 21663994]
18. Janssen SMJ, Abbink JJ, Lindeboom R, Vliet Vlieland TPM. Outcomes of pulmonary rehabilitation after treatment for non-small cell lung cancer stages I to IIIa: an observational study. J Cardiopulm Rehabil Prev 2017 Jan;37(1):65-71. [doi: 10.1097/HCR.0000000000000227] [Medline: 28005681]
19. Tarumi S, Yokomise H, Gotoh M, Kasai Y, Matsuura N, Chang SS, et al. Pulmonary rehabilitation during induction chemoradiotherapy for lung cancer improves pulmonary function. J Thorac Cardiovasc Surg 2015 Feb;149(2):569-573 [FREE Full text] [doi: 10.1016/j.jtcvs.2014.09.123] [Medline: 25451483]
20. Pasqua F, D'Angelillo R, Mattei F, Bonassi S, Biscione GL, Geraneo K, et al. Pulmonary rehabilitation following radical chemo-radiation in locally advanced non surgical NSCLC: preliminary evidences. Lung Cancer 2012 May;76(2):258-259. [doi: 10.1016/j.lungcan.2012.01.015] [Medline: 22386840]
21. Jastrzębski D, Maksymiak M, Kostorz S, Bezubka B, Osmanska I, Młynczak T, et al. Pulmonary rehabilitation in advanced lung cancer patients during chemotherapy. Adv Exp Med Biol 2015;861:57-64. [doi: 10.1007/5584_2015_134] [Medline: 26017725]

22. Nwosu AC, Bayly JL, Gaunt KE, Mayland CR. Lung cancer and rehabilitation--what are the barriers? Results of a questionnaire survey and the development of regional lung cancer rehabilitation standards and guidelines. Support Care Cancer 2012 Dec;20(12):3247-3254. [doi: 10.1007/s00520-012-1472-1] [Medline: 22539052]
23. Sommer MS, Trier K, Vibe-Petersen J, Missel M, Christensen M, Larsen KR, et al. Perioperative Rehabilitation in Operable Lung Cancer Patients (PROLUCA): a feasibility study. Integr Cancer Ther 2016 Dec;15(4):455-466 [FREE Full text] [doi: 10.1177/1534735416635741] [Medline: 27151595]
24. Olivier C, Grosbois JM, Cortot AB, Peres S, Heron C, Delourme J, et al. Real-life feasibility of home-based pulmonary rehabilitation in chemotherapy-treated patients with thoracic cancers: a pilot study. BMC Cancer 2018 Dec 13;18(1):178 [FREE Full text] [doi: 10.1186/s12885-018-4102-6] [Medline: 29433474]
25. Driessen EJ, Peeters ME, Bongers BC, Maas HA, Bootsma GP, van Meeteren NL, et al. Effects of prehabilitation and rehabilitation including a home-based component on physical fitness, adherence, treatment tolerance, and recovery in patients with non-small cell lung cancer: a systematic review. Crit Rev Oncol Hematol 2017 Jun;114:63-76. [doi: 10.1016/j.critrevonc.2017.03.031] [Medline: 28477748]
26. Hoffman AJ, Brintnall RA, von Eye A, Jones LW, Alderink G, Patzelt LH, et al. Home-based exercise: promising rehabilitation for symptom relief, improved functional status and quality of life for post-surgical lung cancer patients. J Thorac Dis 2014 Jun;6(6):632-640 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2014.06.08] [Medline: 24976984]
27. Hoffman AJ, Brintnall RA, von Eye A, Jones LW, Alderink G, Patzelt LH, et al. A rehabilitation program for lung cancer patients during postthoracotomy chemotherapy. Onco Targets Ther 2014;7:415-423 [FREE Full text] [doi: 10.2147/OTT.S57262] [Medline: 24648745]
28. Kwon H, Lee S, Jung EJ, Kim S, Lee J, Kim DK, et al. An mHealth management platform for patients with chronic obstructive pulmonary disease (efil breath): randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 24;6(8):e10502 [FREE Full text] [doi: 10.2196/10502] [Medline: 30143475]
29. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
30. Park YS, Lee J, Keum B, Oh BM. Feasibility of an eight-week outpatient-based pulmonary rehabilitation program for advanced lung cancer patients undergoing cytotoxic chemotherapy in Korea. Thorac Cancer 2018 Aug;9(8):1069-1073 [FREE Full text] [doi: 10.1111/1759-7714.12788] [Medline: 29932301]

Abbreviations

6MWD: 6-minute walk distance
6MWT: 6-minute walk test
EQ-5D: EuroQol 5 dimensions questionnaire
EQ-VAS: EuroQol-visual analog scale
mHealth: mobile health
mMRC: modified Medical Research Council
NSCLC: non–small cell lung cancer
PGA: Patient Global Assessment
PR: pulmonary rehabilitation
QoL: quality of life
V: visit

Edited by G Eysenbach; submitted 30.10.18; peer-reviewed by C Smeets, J Lingeman, N Ocal, V Sambandam; comments to author 05.01.19; revised version received 30.03.19; accepted 05.05.19; published 21.06.19.

Please cite as:
Ji W, Kwon H, Lee S, Kim S, Hong JS, Park YR, Kim HR, Lee JC, Jung EJ, Kim D, Choi CM
Mobile Health Management Platform–Based Pulmonary Rehabilitation for Patients With Non–Small Cell Lung Cancer: Prospective Clinical Trial
JMIR Mhealth Uhealth 2019;7(6):e12645
URL: http://mhealth.jmir.org/2019/6/e12645/
doi:10.2196/12645
PMID:31228180

Original Paper

Accuracy of Fitbit Wristbands in Measuring Sleep Stage Transitions and the Effect of User-Specific Factors

Zilu Liang[1,2], PhD; Mario Alberto Chapa-Martell[3], PhD

[1]School of Engineering, Kyoto University of Advanced Science, Kyoto, Japan

[2]Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[3]Advanced Technology Division, CAC Corporation, Tokyo, Japan

Corresponding Author:
Zilu Liang, PhD
School of Engineering
Kyoto University of Advanced Science
18 Yamanouchi Gotanda-Cho
Kyoto, 6158577
Japan
Phone: 81 8040866433
Email: z.liang@cnl.t.u-tokyo.ac.jp

Abstract

Background: It has become possible for the new generation of consumer wristbands to classify sleep stages based on multisensory data. Several studies have validated the accuracy of one of the latest models, that is, Fitbit Charge 2, in measuring polysomnographic parameters, including total sleep time, wake time, sleep efficiency (SE), and the ratio of each sleep stage. Nevertheless, its accuracy in measuring sleep stage transitions remains unknown.

Objective: This study aimed to examine the accuracy of Fitbit Charge 2 in measuring transition probabilities among wake, light sleep, deep sleep, and rapid eye movement (REM) sleep under free-living conditions. The secondary goal was to investigate the effect of user-specific factors, including demographic information and sleep pattern on measurement accuracy.

Methods: A Fitbit Charge 2 and a medical device were used concurrently to measure a whole night's sleep in participants' homes. Sleep stage transition probabilities were derived from sleep hypnograms. Measurement errors were obtained by comparing the data obtained by Fitbit with those obtained by the medical device. Paired 2-tailed *t* test and Bland-Altman plots were used to examine the agreement of Fitbit to the medical device. Wilcoxon signed–rank test was performed to investigate the effect of user-specific factors.

Results: Sleep data were collected from 23 participants. Sleep stage transition probabilities measured by Fitbit Charge 2 significantly deviated from those measured by the medical device, except for the transition probability from deep sleep to wake, from light sleep to REM sleep, and the probability of staying in REM sleep. Bland-Altman plots demonstrated that systematic bias ranged from 0% to 60%. Fitbit had the tendency of overestimating the probability of staying in a sleep stage while underestimating the probability of transiting to another stage. SE>90% (P=.047) was associated with significant increase in measurement error. Pittsburgh sleep quality index (PSQI)<5 and wake after sleep onset (WASO)<30 min could be associated to significantly decreased or increased errors, depending on the outcome sleep metrics.

Conclusions: Our analysis shows that Fitbit Charge 2 underestimated sleep stage transition dynamics compared with the medical device. Device accuracy may be significantly affected by perceived sleep quality (PSQI), WASO, and SE.

(JMIR Mhealth Uhealth 2019;7(6):e13384) doi:10.2196/13384

KEYWORDS

wearable electronic devices; sleep; validation studies

Introduction

Importance of Consumer Sleep Tracking Devices

Having enough restorative sleep is essential for physical and mental health [1]. In recent years, consumer sleep-monitoring wristbands and associated mobile phone apps have created an effective way for individuals to understand personal sleep patterns or improve sleep quality in daily settings [2]. These devices are relatively affordable, easy to use, and ready to purchase in the consumer market. Most of the consumer wristbands rely on a similar mechanism of clinical actigraphy that infers wake and sleep cycles from limb movement [2]. Newly launched models also incorporate other streams of biosignals, such as heart rate to measure sleep stages. Users can visualize a whole night's sleep hypnogram (the temporal sequence of sleep stages) and the aggregated sleep parameters, such as total sleep time (TST) and the ratio of each sleep stage on a dashboard [3]. There is increasing evidence that consumer sleep-monitoring wristbands raise awareness of sleep health and have a positive impact on personal sleep hygiene [4-6], though the long-term impact of these technologies has not been elucidated [7]. In the meantime, researchers and clinicians are increasingly adopting consumer wristbands, such as Fitbit devices, as outcome measurement tools in research studies [6,8-14]. Compared with traditional polysomnography (PSG), Fitbit devices significantly reduce the time and monetary cost for longitudinal sleep data collection, and they could provide rich information that was not possible to collect outside sleep laboratories or clinics in the past. Participants can use the devices under free-living conditions, without the need of constant technical support. The new generation of Fitbit devices could also possibly outperform clinical actigraphy, as they leverage multiple streams of biosignals for sleep staging, whereas actigraphy is only able to detect wake and sleep on the basis of limb movement [15].

Accuracy of Consumer Sleep Tracking Devices

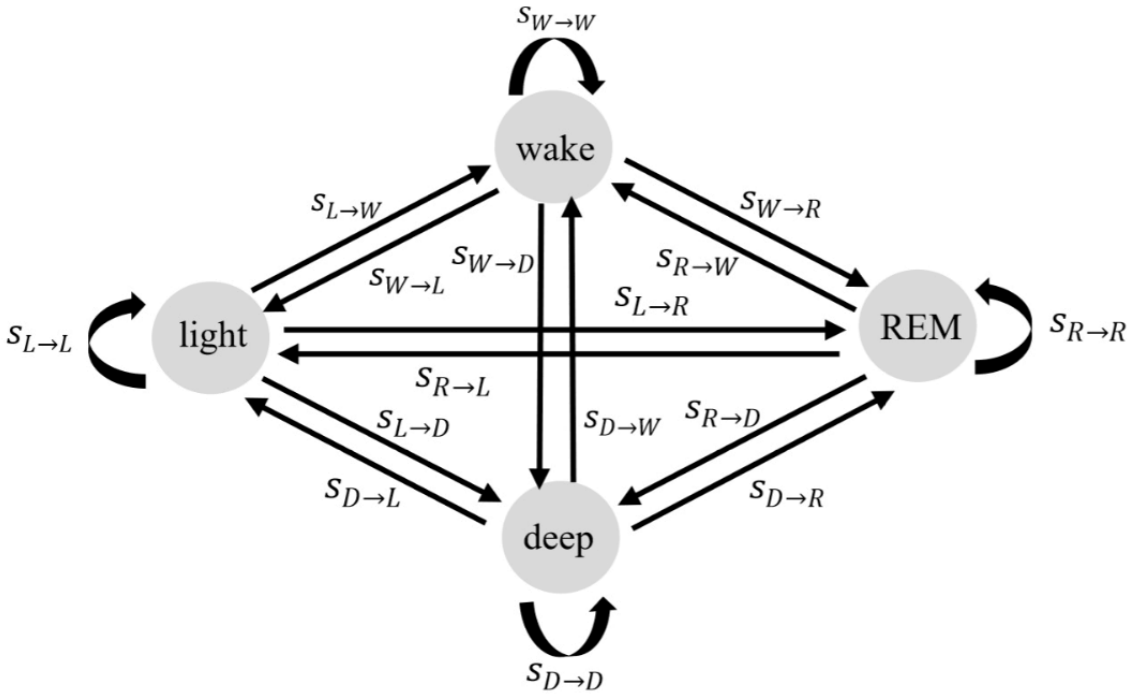
As consumer sleep-monitoring wristbands continue to gain popularity, their limitation in measurement accuracy raised wide concerns on the quality of data collected using these devices [7,16,17]. Data of low quality may mislead users to arrive at wrong conclusions of their sleep. In addition, data quality is of top priority for researchers who intend to use these devices in scientific studies. Therefore, understanding the validity of consumer sleep trackers has practical benefit for both individual users and for the research community. In response to this need, many studies have examined the accuracy of popular sleep trackers compared with medical devices in terms of aggregated sleep metrics, including TST, wake after sleep onset (WASO), sleep efficiency (SE), and sleep stages, that is, light sleep, deep sleep, and rapid eye movement (REM) sleep [18-24]. These studies show that the previous models of consumer wristbands have a common problem of overestimating sleep and underestimating wake [18-20]. Recent models, such as Fitbit Charge 2, that rely on multistreams of biosignals have satisfying performance in measuring TST and SE but fail to produce accurate results in classifying sleep stages [21,24].

Although the main body of validation studies has been dominantly focused on polysomnographic metrics (eg, TST, WASO, sensitivity, and specificity) [2,13,24-27], the performance of consumer wristbands in measuring sleep stage transitions remains unknown. Sleep research has shown that sleep stage transition probabilities comprise rich information of sleep patterns, which have been considered more effective than polysomnographic parameters in characterizing sleep stability [28-37]. Sleep stage transition abnormality is an important indicator of sleep disorders [28,32,33,38-43]. Some studies also relied on sleep stage transition probabilities to assess the effect of treatment [44]. The clinical significance of sleep stage transition dynamics suggests the necessity of including relevant metrics (sleep stage transition probabilities) as outcome sleep parameters in validation studies. In Figure 1, a visualization of sleep stage transition dynamics is presented. The total transition probability from a single state to other states (including staying in the same state) is always 1. The $s_{X \to Y}$ represents the transition probability from sleep stage X to Y. The { X, Y } are derived from { W, L, D, R }, which are abbreviations for wake, light sleep, deep sleep, and REM sleep. For example, $s_{W \to R}$ denotes the transition probability from wake to REM sleep, and $s_{W \to W}$ denotes the probability of staying in wake.

Significance of This Study

This study aimed to examine whether it would accurately measure sleep stage transitions (the transition probabilities among waking, light, deep, and REM sleep) using Fitbit Charge 2. Despite the abundant validation studies, the accuracy of consumer wristbands in measuring sleep stage transition has not been investigated. We also examined the factors that are associated with the measurement errors on sleep stage transition probabilities. Previous validation studies on other types of wearable devices found that device accuracy could vary as a function of the underlying sleep patterns, the population studied, and even how the measurand was defined [45-48]. Along the same line, we selected a set of independent variables (possible predictors), including demographic characteristics of participants, subjective sleep quality measured by Pittsburgh Sleep Quality Index (PSQI) [49], and objective sleep quality derived from medical data. The dependent variables were the absolute percent errors of Fitbit Charge 2 on sleep stage transition probabilities compared with the medical device. The outcomes of this study complement previous validation studies and contribute to the establishment of a holistic view of the capacity of consumer wristbands in measuring sleep structure under free-living conditions. This study also establishes a preliminary reference for researchers who intend to use Fitbit to measure sleep stage transitions and for individual users who rely on Fitbit sleep data to make health decisions.

Figure 1. Sleep stage transition dynamics. The W, L, D, R in the subscripts denotes the abbreviation of wake, light sleep, deep sleep, and rapid eye movement sleep.



Methods

Recruitment

We recruited participants by distributing posters around the campus of The University of Tokyo. In total, 38 people registered interest through a Web-based form, of whom 28 (74%) were eligible to participate in the study. The inclusion criteria required that the participants were adults (age>18 years), were free of diagnosed chronic conditions, and were able to attend a briefing before the data collection phase. This research was approved by the ethical committee of the University of Tokyo. All participants provided informed consent.

Study Procedures

A face-to-face briefing was held with each participant individually before the data collection phase. In this meeting, we installed the Fitbit app on participants' mobile phones and provided verbal instructions on how to use the devices and how to synchronize the Fitbit device with its mobile phone app. Participants were provided with the following items for data collection: a Fitbit Charge 2, a medical device named Sleep Scope, electrodes, chargers, and manuals. At the end of the briefing, participants were asked to fill in a PSQI questionnaire [49] to measure their perceived sleep quality. The PSQI is a widely used instrument for assessing subjective sleep quality averaged over the past 1 month, and a PSQI≥5 is indicative of perceived poor sleep. We collected the PSQI, as it may associate to the measurement accuracy of Fitbit. More details on potential association factors of measurement accuracy will be provided in the next section.

After the briefing, participants measured their sleep using both devices for 3 consecutive nights in their homes to ensure that Fitbit Charge 2 was evaluated in an ecologically valid setting. They were asked to wear the Fitbit on the nondominant wrist during data collection. All participants received a monetary reward when they returned the devices after data collection.

Data Collection

In this study, we collected sleep data concurrently using Fitbit Charge 2 and a medical device. Fitbit Charge 2 (Fitbit Inc) is a wearable activity wristband with an embedded triaxial accelerometer. It estimates sleep stages for each 30 second period by integrating a user's movement and heart rate data. With advances in software and hardware, Fitbit Charge 2 has overcome some problems of previous models, and it is able to measure TST and SE with good accuracy [21,24]. A medical sleep monitor named Sleep Scope (Sleep Well Co) was used to obtain the ground truth on sleep hypnograms. Sleep Scope is a clinical-grade single-channel electroencephalogram (Japanese Medical Device Certification 225ADBZX00020000), which was validated against PSG (agreement=86.9%, average Cohen Kappa value =0.75) [50,51]. Sleep Scope was chosen over PSG as it enabled data collection in participants' homes rather than in a sleep laboratory. This ensures that Fitbit Charge 2 was evaluated in an ecologically valid setting; this also ensures minimalizing the possible disruption of sleep by unfamiliar environment.

In the data collection phase, participants tracked their sleep for 3 consecutive nights in their homes. Following the common practice in sleep science, we analyzed the second night for each participant to remove *the first night effect* [52,53]. If the data of the second night were not valid, then the data of the third night were analyzed. The data of the first night were only selected when neither the second night nor the third night was valid.

Fitbit sleep data were retrieved through the application program interface (API) of Fitbit. Fitbit Charge 2 provides sleep data at 2 levels through public API. The *stage level* data comprise sleep

stage levels, including wake, light sleep, deep sleep, and REM sleep. These data are aggregated at 30-second granularity, which complies with the standard sleep staging in the clinical setting. If the *stage level* data are not available, the *classic level* data will be provided as an alternative. *Classic level* data comprise sleep pattern levels, including asleep, restless, and awake, and they are aggregated at a coarser granularity of 60 seconds. In this study, we were interested in the *stage level* sleep data, and the *classic level* data were discarded, as they contained no information on deep sleep, light sleep, and REM sleep.

The data of the medical device were analyzed by the Sleep Well Company, using proprietary automatic scoring algorithms, followed by epoch-by-epoch visual inspection by specialists on the basis of established standards [54], and corrections were added if needed. Fitbit data and medical data were synchronized to make sure that the start time was aligned.

To examine the effect of user-specific factors on measurement accuracy, we also collected data on the factors listed in Table 1. Age and sex were based on self-report, and PSQI was measured by the PSQI questionnaire [49]. Sleep quality metrics were all derived from the medical data.

Table 1. A full list of user-specific factors.

Factors	Data type	Data collection method	Cut-off threshold
Age (years)	Ordinal	Self-reported	25
Sex	Nominal	Self-reported	Female or male
PSQI[a]	Ordinal	PSQI questionnaire	5
TST[b] (min)	Continuous	Sleep scope (medical device)	360
WASO[c] (min)	Continuous	Sleep scope	30
SOL[d] (min)	Continuous	Sleep scope	30
SE[e], %	Continuous	Sleep scope	90.0
Light sleep, %	Continuous	Sleep scope	65.0
SWS[f], %	Continuous	Sleep scope	20.0
REM[g], %	Continuous	Sleep scope	20.0
T_{avg}[h] (min)	Continuous	Sleep scope	90

[a]PSQI: Pittsburgh Sleep Quality Index.

[b]TST: total sleep time.

[c]WASO: wake after sleep onset.

[d]SOL: sleep onset latency.

[e]SE: sleep efficiency.

[f]SWS: slow wave sleep.

[g]REM: rapid eye movement sleep.

[h]T_{avg}: average sleep cycle.

Statistical Analysis

The overall goal of the analysis was two-fold. We aimed to examine the accuracy of Fitbit Charge 2 in measuring sleep stage transitions compared with a medical device. We were also interested in the associations of user-specific factors with the measurement accuracy of Fitbit Charge 2. All statistical significance levels reported were 2 sided, and statistical analysis was performed using R statistical software version 3.5.3 (The R Foundation)[55].

First, descriptive statics of sleep parameters were derived from the medical data. Paired 2-tailed *t* test was used to probe if there were statistically significant differences on sleep patterns between men and women, as well as between participants below 25 years of age and above 25 years of age. Second, sleep stage transition probabilities were calculated by dividing the number of transitions from a specific sleep state to a specific sleep state by the total number of transitions from that specific state to all sleep states (including staying in the same state). As shown in Figure 2, { *X*, *Y*, and *B* } are derived from { *W*, *L*, *D*, and *R* } and $n_{X \to Y}$ is the number of transitions from sleep stage *X* to *Y* during a whole night's sleep. The *W*, *L*, *D*, and *R* are the abbreviations for wake, light sleep, deep sleep, and REM sleep. Sleep stage transition probabilities were calculated from Fitbit data and medical data for each participant and then averaged over the whole cohort to obtain the average sleep stage transition probabilities. Systematic difference between the 2 devices was assessed by applying paired *t* test on the sleep stage transition probabilities. A *P* value below .05 was considered statistically significant. The level of agreement between 2 devices was examined using the Bland-Altman plots [56].

Figure 2. The calculation of sleep stage transition probabilities.

$$s_{X\to Y} = \frac{n_{X\to Y}}{\sum_B n_{X\to B}} \times 100$$

Figure 3. The calculation of absolute percent error.

$$e_{X\to Y} = \left| \frac{s^F_{X\to Y} - s^M_{X\to Y}}{s^M_{X\to Y}} \right| \times 100$$

The absolute percent error $e_{X\to Y}$ was calculated using the equation in Figure 3, where { *X*, *Y*, and *B* } are derived from { *W*, *L*, *D*, and *R* }, $s^F_{X\to Y}$ and $s^M_{X\to Y}$ are the transition probability from sleep stage *X* to *Y*, derived from Fitbit data and medical data.

To examine the effect of user-specific factors on absolute percent error, the dataset was divided into 2 subsets according to the cut-off threshold values listed in Table 1. Wilcoxon signed–rank test was conducted to examine if there were significant differences between the 2 subsets in terms of the outcome sleep metrics (sleep stage transition probabilities). The selection of cut-off threshold values was in line with literature in sleep science [49,57].

Results

Descriptive Statistics

A total of 28 young adults without chronic diseases participated in the study. A total of 5 participants were excluded from analysis because of failure to obtain *stage level* sleep data with Fitbit. That is, only *classic level* sleep data were obtained from these participants; the data had no information on light, deep, and REM sleep. Therefore, it was not possible to calculate sleep stage transition probabilities for these participants. The final dataset thus comprises sleep data from 23 participants (men:women=14:9). This number of participants is comparable with other validation studies [20,27,58-61]. All the participants were university students between 21 to 30 years old (mean 24.3, SD 2.7). A total of 8 out of the 23 participants had a PSQI higher than 5, which was indicative of unsatisfied sleep quality. Statistically significant differences were found between men and women in terms of wake time (women: 9.7 min; men: 22.8 min; *P*=.02) and the ratio of sleep stage 1 (women: 7.7(%); men: 14.3(%); *P*=.02). We also compared the sleep patterns between participants below and above 25 years. Statistically significant differences were found in terms of TST (below 25 years: 308.7 min; above 25 years: 396.8 min; *P*=.03), transition probability from deep sleep to light sleep (below 25 years: 5.5%; above 25 years: 1.5%; *P*=.02), and the probability of staying in light sleep (below 25 years: 85.3(%); above 25 years: 94.8(%); *P*=.008).

Systematic Differences

Table 2 presents the estimated sleep stage transition probabilities derived from medical data and Fitbit data, as well as the results of paired *t* test. We calculated sleep stage transition probabilities individually for each participant and then averaged results across the whole cohort. It is shown that the following transitions rarely occurred: deep sleep to REM sleep and wake, light sleep to REM sleep, REM sleep to deep sleep, and REM sleep to light sleep. The *t* test results indicated that there were significant differences between the sleep stage transition probabilities measured by Fitbit and those measured by the medical device. Fitbit deviated from the medical device on all the transition probabilities except for the transition probability from light sleep to REM sleep ($s^F_{L\to R}$ = 0.9%; $s^M_{L\to R}$ =1.7%), the transition probability from deep sleep to wake ($s^F_{D\to W} = s^M_{D\to W}$ =0.2%), and the probability of staying in REM sleep stage ($s^F_{R\to R} = s^M_{R\to R}$ =96.9%). In general, Fitbit underestimated sleep stage transition dynamics. The probabilities of staying in a specific sleep stage were significantly overestimated, whereas the probabilities of transitions from a specific stage to a different stage were mostly underestimated.

Table 2. Average sleep stage transition probabilities (%) and results of paired *t* test. Data are displayed as mean and ±95% CI.

Sleep stage	Wake	Light	Deep	REM[a]
Wake				
Medical	53.7 (44.0-63.3)	43.6 (33.8-53.4)	0.2 (0.0-0.4)	2.6 (1.5-3.7)
Fitbit	89.8 (81.2-98.3)	5.5 (4.3-6.7)	0.2 (0.0-0.5)	0.2 (0.0-0.5)
P value	<.001	<.001	.83	<.001
Light				
Medical	2.6 (2.0-3.3)	92.6 (90.9-94.4)	3.9 (2.1-5.8)	0.8 (0.7-0.9)
Fitbit	0.5 (0.3, 0.6)	97.8 (97.6-98.1)	1.1 (0.9-1.3)	0.5 (0.3-0.7)
P value	<.001	<.001	.005	.02
Deep				
Medical	2.5 (0.7-4.3)	57.7 (43.8-71.6)	35.5 (22.6-48.4)	0.0 (0.0-0.0)
Fitbit	0.2 (0-1.8)	3.8 (2.9-4.6)	94.9 (93.4-96.4)	1.1 (0.4-1.8)
P value	.02	<.001	<.001	.002
REM				
Medical	2.0 (1.6-2.4)	0.9 (0.7-1.2)	0.0 (0.0-0.0)	96.9 (96.5-97.5)
Fitbit	0.1 (0.0-0.2)	1.7 (0.7-2.6)	1.2 (0.3-2.2)	96.9 (96.0-98.0)
P value	<.001	.14	.01	>.99

[a]REM: rapid eye movement.

Level of Agreement and Correlations

Figures 4-6 show the Bland-Altman plots comparing Fitbit Charge 2 with the medical device. Device discrepancies for sleep outcomes are plotted as a function of the medical outcomes for each individual. The mean bias ranged from 0% ($s_{R\to R}$ and $s_{D\to W}$) to approximately 60% ($s_{L\to D}$). No more than 2 participants were situated outside the lower limit of agreement or the upper limit of agreement.

In line with previous studies [62,63], we defined the acceptable error range as $e_i \leq 5\%$, as this approximates a widely acceptable standard for statistical significance in literature [64]. On the basis of this criterion, no systematic bias was found between Fitbit and the medical device in measuring $s_{W\to L}$, $s_{W\to R}$, $s_{L\to R}$, $s_{D\to W}$, $s_{R\to L}$, $s_{R\to D}$, and $s_{R\to R}$.

Figure 4 shows that no trend was found between the difference and the mean of $s_{R\to L}$, $s_{L\to R}$ and $s_{R\to R}$. In contrast, Figure 5 and Figure 6 show clear trends that the measurement differences were greater for lower $s_{L\to L}$, $s_{D\to D}$, and $s_{W\to W}$, and the differences were greater for higher $s_{W\to L,}s_{W\to R}$, $s_{W\to D}$, $s_{L\to W}$, $s_{L\to D}$, $s_{D\to W}$, $s_{D\to L}$, $s_{D\to R}$, $s_{R\to W}$, and $s_{R\to D}$. These findings suggest that the accuracy of Fitbit Charge 2 in measuring sleep stage transitions could be deteriorated as sleep became more dynamic (more transitions between different sleep stages).

Figure 4. Bland-Altman plots assessing the level and limits of agreement between Fitbit Charge 2 and medical device on the transition probabilities from rapid eye movement (REM) sleep to light sleep, from light sleep to REM sleep, and the probability of staying in REM sleep. The dashed line in the middle represents the mean difference, whereas the upper and lower dashed lines represent the upper limit of agreement and the lower limit of agreement.

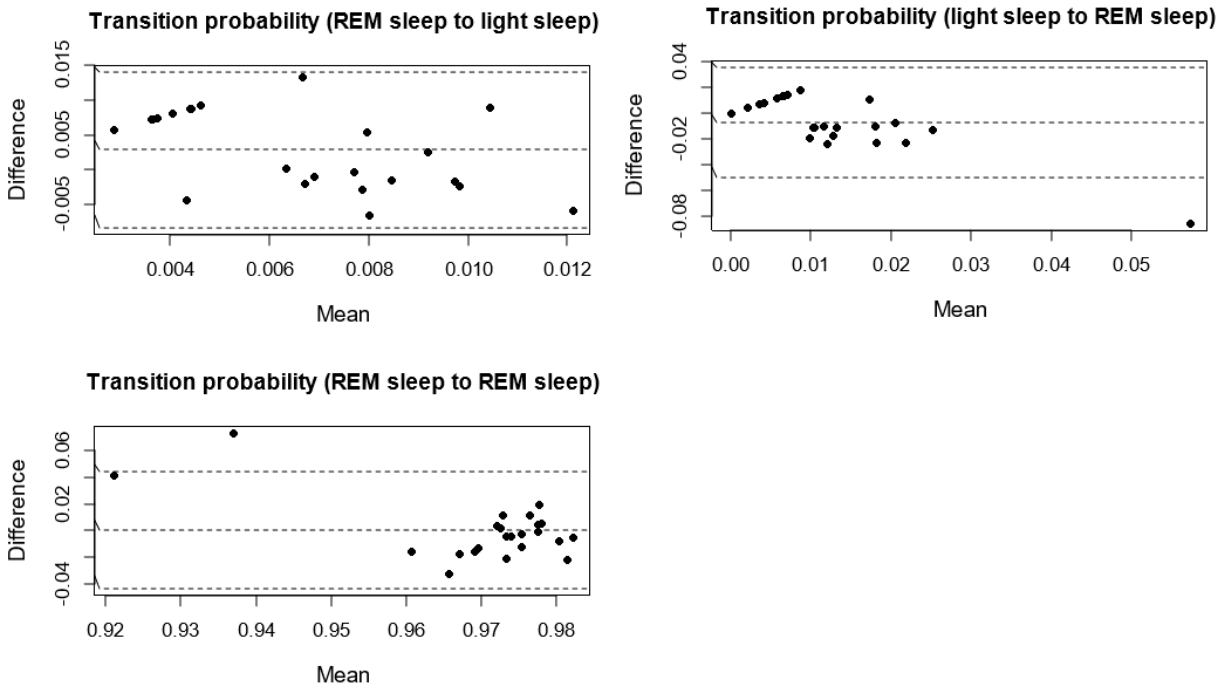


Figure 5. Bland-Altman plots assessing the level and limits of agreement between Fitbit Charge 2 and medical device on the probability of staying in light sleep, in deep sleep, and in wake.

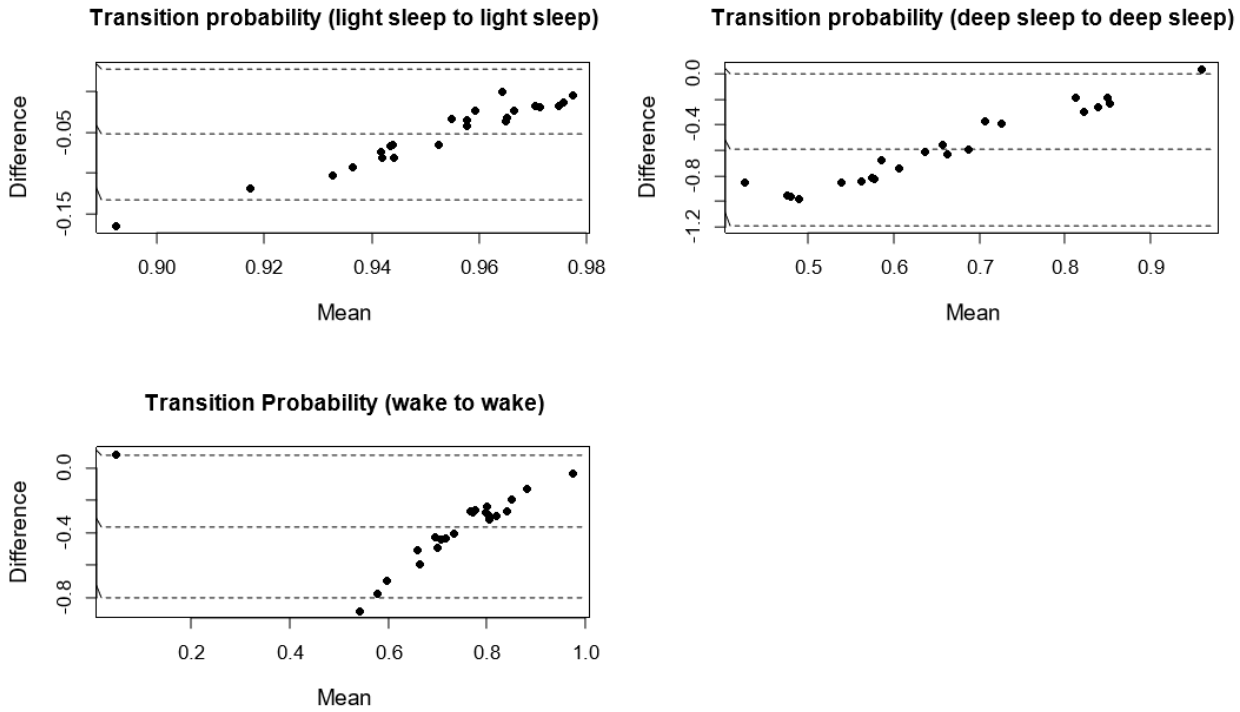
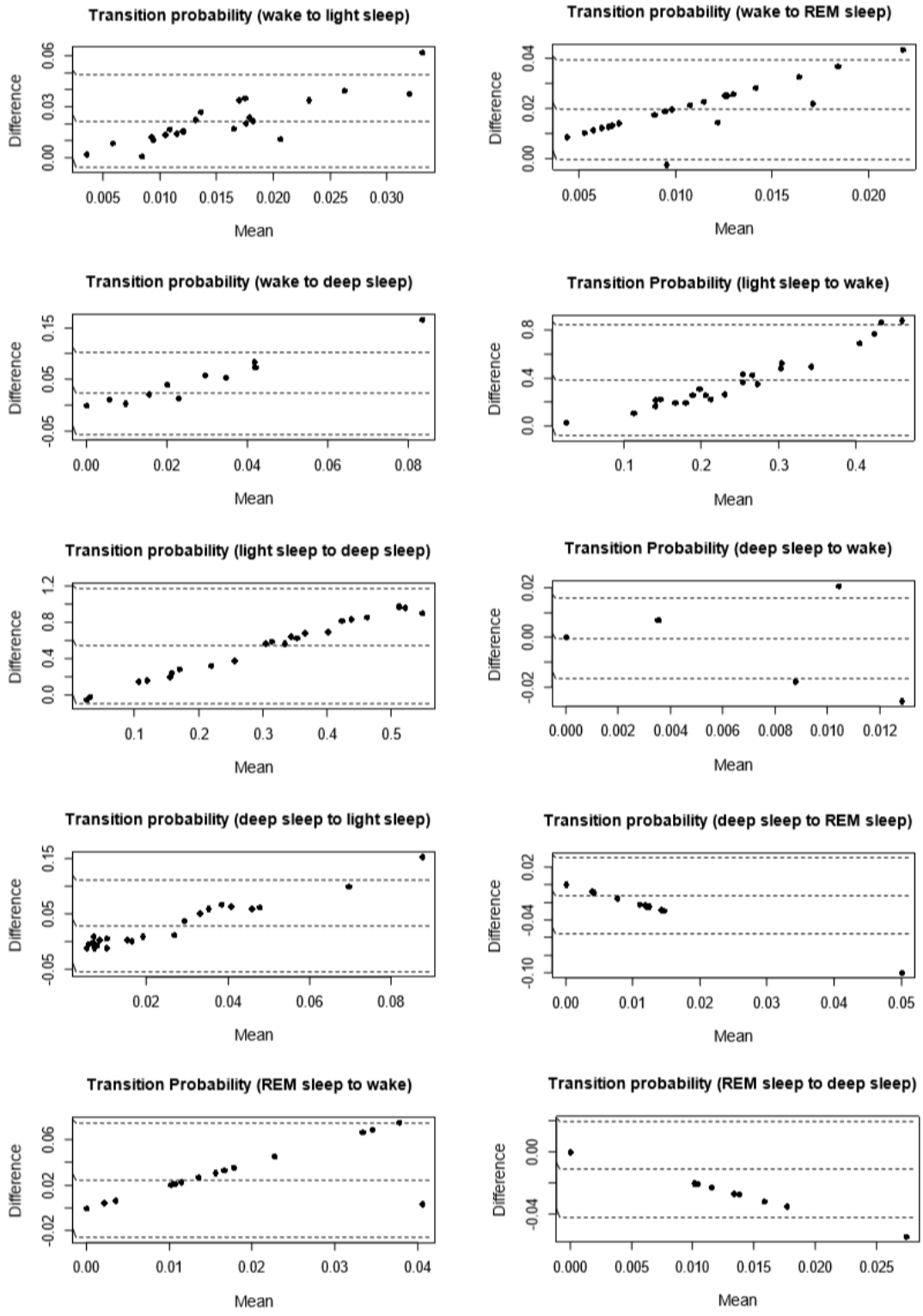


Figure 6. Bland-Altman plots assessing the level and limits of agreement between Fitbit Charge 2 and medical device on the transition probabilities from wake to light sleep, from wake to rapid eye movement (REM) sleep, from wake to deep sleep, from light sleep to wake, from light sleep to deep sleep, from deep sleep to wake, from deep sleep to light sleep, from deep sleep to REM sleep, from REM sleep to wake, and from REM sleep to deep sleep.



Effect of User-Specific Factors

The results of Wilcoxon signed–rank test showed that good subjective sleep quality indicated by PSQI as lower than 5 was associated with decreased errors in the probability of staying in deep sleep stage (PSQI<5, 132.1±173.1%; PSQI≥5, 346.8±250.0%; *P*=.04), but it was associated with increased errors in transition probability from waking to REM sleep (PSQI<5, 100.0±0.0%; PSQI≥5, 85.1±25.5%; *P*=.02).

Wake time longer than 30 min was associated with increased errors in transition probability from light sleep to REM sleep (WASO≥30, 265.8±176.5; WASO<30, 103.9±49.1%; *P*=.02), but it was associated with decreased errors in transition

probability from light sleep to wake (WASO≥30, 78.6±10.2%; WASO<30, 86.7±8.6%; *P*=.049), as well as the probability of staying in wake (WASO≥30, 117.3±269.5%; WASO<30, 125.2±103.6%; *P*=.006).

SE above 90% was associated with increased measurement errors in transition probability from REM sleep to light sleep (SE>90, 107.1±53.2%; SE≤90%, 55.9±40.4%; *P*=.047).

In addition, age below 25 years (age<25, 7.9±5.4%; age≥25, 3.1±2.3%; *P*=.01), sleep onset latency (SOL) shorter than 30 min (SOL<30, 8.6±5.8%; SOL≥30, 4.1±3.4%; *P*=.02), and deep sleep ratio above 20% (slow wave sleep; SWS<20%, 3.9±3.5%; SWS≥20, 9.5±5.2; *P*=.007) were associated with slight increased measurement error in the probability of staying in light sleep stage. Nevertheless, the average errors were no more than 10% in all the corresponding cases.

No significant associations were found between measurement errors of Fitbit and other factors, including sex, TST, SOL, light sleep ratio, REM sleep ratio, and T_{avg}.

Discussion

Principal Findings

We have demonstrated a numerical comparison on sleep stage transition probabilities between Fitbit Charge 2 and the medical device. The level and limits of agreement between the 2 types of devices were illustrated using Bland-Altman plots. The results of Wilcoxon signed–rank test were presented to demonstrate the associations between user-specific factors and measurement errors. This study generated 2 main findings. First, we found that Fitbit Charge 2 underestimated sleep stage transition dynamics compared with the medical device. Second, device accuracy was mainly associated with 3 user-specific factors: subjective sleep quality measured by PSQI, WASO, and SE.

Sleep stage transition analysis has been used to characterize sleep continuity and the temporal stability of non-REM and REM bouts in sleep science [28-30,32,40,44]. In this study, the sleep stage transition probabilities derived from the medical data demonstrated interesting patterns. As expected, the probability for any sleep stage to stay in the same stage was constantly higher than that for this stage to change to a different stage. Direct transition between deep sleep and REM sleep rarely happened. The probability of transitions from wake to deep sleep or from wake to REM sleep was low. Similarly, the probability of transition from deep sleep to wake was also low. These characteristics were consistent with findings reported in previous sleep studies on sleep stage transition patterns in healthy people [31,44].

Sleep stage transition is the result of complex interactions among many brain regions. Not being able to detect markers in brainwaves, such as k-complexes [54], consumer wristbands have limited performance in classifying sleep stages. Previous studies show that Fitbit Charge 2 devices significantly overestimated light sleep and underestimated deep sleep when validated in lab settings [21], whereas they underestimated deep sleep and overestimated light and REM sleep when validated under free-living conditions [24]. This study complements previous findings and contributes new insights into Fitbit's capacity in capturing sleep stage transitions. Overall, we observed that Fitbit Charge 2 significantly deviated from the medical device in measuring sleep stage transition dynamics. Notably, the average probabilities of staying in wake stage and deep stage measured by Fitbit were significantly higher than those measured by the medical device. In contrast, Fitbit underestimated the probabilities of stage transitions from light sleep to wake and from light sleep to deep sleep. This is probably because of the misclassification of wake and deep sleep epochs to light sleep [21]. Systematic bias (between 40% and 60%) was illustrated in the Bland-Altman plots on these sleep stage transition probabilities. On the other hand, no systematic bias and mean difference were observed in measuring the probability of staying in REM sleep stage. This result provides complementary evidence to the finding in the study by De Zambotti et al [21] that Fitbit Charge 2 agreed well to medical devices in detecting REM sleep.

A unique aspect of this study is that we also examined the effect of user-specific factors and found multiple associations. Our analysis showed that subjective sleep quality measured by PSQI, wake after WASO, and SE were significantly strong predictors of measurement errors in sleep stage transition probabilities. Age, SOL, and deep sleep ratio were significant but weak predictors, whereas sex, TST, light sleep ratio, REM sleep ratio, and average sleep cycle were not associated with the measurement errors of Fitbit.

Despite the finding from previous validation studies that poor sleep quality is associated with deteriorated performance of sleep monitoring devices in measuring polysomnographic sleep metrics [21,25,65], this study reveals that the relationship is more complicated between sleep quality and device accuracy in measuring sleep stage transitions. Indeed, we found that good subjective sleep quality (PSQI<5) was associated with decreased measurement error in the probability of staying in deep sleep stage, and less fragmented sleep (WASO<30 min) was associated with decreased errors in transition probability from light sleep to REM sleep. Nevertheless, it is also found that good sleep characterized by quick sleep onset (SOL<30 min), high ratio of deep sleep (SWS>20%), good subjective feeling (PSQI<5), short awakenings (WASO<30 min), and high SE (SE>90%) were associated with increased measurement errors in different outcome transition probabilities. This result contradicts previous findings on actigraphy that deteriorated sleep (eg, long WASO and SOL) increased measurement errors [21,25,65]. This disparity suggests that findings related to clinical actigraphy should not be generalized to consumer wristbands without further validation.

In addition, age was found to be a significant but weak predictor of measurement errors. Participants in the age range of 25 to 30 had decreased measurement errors in the probability of staying in light sleep stage compared with those younger than the age of 25. As age has been widely recognized as a significant factor that alters sleep patterns [43,57], the effect of age may also be traced back to the difference in underlying sleep patterns. The medical sleep data showed that younger participants generally had shorter sleep and higher sleep stage transition dynamics (transition from deep sleep to light sleep), which may

account for the increase in measurement errors. Nevertheless, this finding should not be generalized to a wide range of age groups because of the restricted sampling of age in this study. Further studies are needed to systematically examine the effect of age on device accuracy.

Our findings complement those of previous validation studies on consumer wristbands for sleep tracking in general. Fitbit Charge 2 has demonstrated satisfying performance in measuring TST and SE, but it remains incapable of classifying sleep stages with good accuracy [21,24]. Our findings show that Fitbit Charge 2 may also underestimate sleep transition dynamics, and it should thus be used with caution. This study establishes a preliminary reference for researchers who intend to use the Fitbit device to measure sleep stage transitions in scientific studies, and this study suggests that both perceived and objective sleep patterns may need to be considered when choosing sleep monitoring tools.

Limitations

This study is subject to the following limitations. First, the participants represent a young healthy population that was free of sleep disorders or chronic diseases. Therefore, the results cannot be generalized to older or clinical populations. Second, the data collection phase was not longitudinal in nature, and only 1 night of sleep from each participant was analyzed. Thus, the results may fail to count intrapersonal variations. Third, the list of potential affecting factors investigated in this study was not exhaustive, and it may be affected by restricted sampling. Further research should address these limitations by including a diverse population, extending data collection duration, and examining the effect of other potential predictors of device accuracy.

Conclusions

We have demonstrated that Fitbit Charge 2 significantly underestimated sleep stage transition dynamics compared with the medical device and that measurement accuracy could be mainly affected by perceived sleep quality, sleep continuity, and SE. Despite the positive trend of enhanced accuracy for the latest consumer wearable sleep trackers, the limitation of these devices in detecting sleep stage transition dynamics needs to be recognized. As an outcome measurement tool, Fitbit Charge 2 may not be suited for research studies related to sleep stage transitions or for health care decision making. Further research should focus on enhancing the accuracy of these consumer wristbands in measuring not only polysomnographic parameters but also sleep stage transition dynamics.

Acknowledgments

This study was sponsored by a JSPS KAKENHI Grant-in-Aid for Research Activity Start-up (Grant Number 16H07469) and a JSPS KAKENHI Grant-in-Aid for Early Career Scientists (Grant Number 19K20141).

Conflicts of Interest

None declared.

References

1. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 01;37(1):9-17 [FREE Full text] [doi: 10.5665/sleep.3298] [Medline: 24470692]
2. Kolla BP, Mansukhani S, Mansukhani MP. Consumer sleep tracking devices: a review of mechanisms, validity and utility. Expert Rev Med Devices 2016 May;13(5):497-506. [doi: 10.1586/17434440.2016.1171708] [Medline: 27043070]
3. 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] [Medline: 28807889]
4. Shelgikar AV, Anderson PF, Stephens MR. Sleep tracking, wearable technology, and opportunities for research and clinical care. Chest 2016 Dec;150(3):732-743. [doi: 10.1016/j.chest.2016.04.016] [Medline: 27132701]
5. Liu W, Ploderer B, Hoang T. In bed with technology: challenges and opportunities for sleep tracking. New York, NY: ACM; 2015 Presented at: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction ; , Parkville, Australia; December 07-10, 2015; Parkville, VIC, Australia p. 142-151 URL: https://dl.acm.org/citation.cfm?id=2838739.2838742 [doi: 10.1145/2838739.2838742]
6. Liang Z, Ploderer B, Liu W, Nagata Y, Bailey J, Kulik L, et al. SleepExplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. Pers Ubiquit Comput 2016 Sep 16;20(6):985-1000. [doi: 10.1007/s00779-016-0960-6]
7. Liang Z, Ploderer B. Sleep tracking in the real world: a qualitative study into barriers for improving sleep. New York, NY: ACM; 2016 Presented at: Proceedings of the 28th Australian Conference on Computer-Human Interaction ;, Tasmania, Australia; 2016; Launceston, Tasmania, Australia p. 537-541 URL: https://dl.acm.org/citation.cfm?id=3010988&dl=ACM&coll=DL [doi: 10.1145/3010915.3010988]
8. Yang R, Shin E, Newman M, Ackerman M. When fitness trackers don't 'fit'nd-user difficulties in the assessment of personal tracking device accuracy. In: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, NY: ACM; 2015 Presented at: International Joint Conference on Pervasive and Ubiquitous Computing; September 07-11, 2015; Osaka, Japan p. 623-634 URL: https://dl.acm.org/citation.cfm?id=2804269 [doi: 10.1145/2750858.2804269]

9. Liang Z, Chapa-Martell M, Nishimura T. A personalized approach for detecting unusual sleep from time series sleep-tracking data. In: Proceedings of the IEEE International Conference on Health Informatics (ICHI).: IEEE; 2016 Presented at: IEEE International Conference on Health Informatics (ICHI); 2016; Chicago, US URL: https://ieeexplore.ieee.org/document/7776322 [doi: 10.1109/ICHI.2016.99]
10. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: a comparison against polysomnography and wrist-worn actigraphy. J Affect Disord 2017 Aug 01;217:299-305. [doi: 10.1016/j.jad.2017.04.030] [Medline: 28448949]
11. Weatherall J, Paprocki Y, Meyer T, Kudel I, Witt E. Sleep tracking and exercise in patients with type 2 diabetes mellitus (Step-D): pilot study to determine correlations between Fitbit data and patient-reported outcomes. JMIR Mhealth Uhealth 2018 Jun 05;6(6):e131 [FREE Full text] [doi: 10.2196/mhealth.8122] [Medline: 29871856]
12. Bian J, Guo Y, Xie M, Parish AE, Wardlaw I, Brown R, et al. Exploring the association between self-reported asthma impact and Fitbit-derived sleep quality and physical activity measures in adolescents. JMIR Mhealth Uhealth 2017 Jul 25;5(7):e105 [FREE Full text] [doi: 10.2196/mhealth.7346] [Medline: 28743679]
13. Baron K, Duffecy J, Berendsen M, Cheung Mason I, Lattie E, Manalo N. Feeling validated yet? A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep. Sleep Med Rev 2018 Dec;40:151-159. [doi: 10.1016/j.smrv.2017.12.002] [Medline: 29395985]
14. Kelly JM, Strecker RE, Bianchi MT. Recent developments in home sleep-monitoring devices. ISRN Neurol 2012;2012:768794-768710 [FREE Full text] [doi: 10.5402/2012/768794] [Medline: 23097718]
15. Goldstone A, Baker F, de Zambotti M. Actigraphy in the digital health revolution: still asleep? Sleep 2018 Sep 01;41(9):-. [doi: 10.1093/sleep/zsy120] [Medline: 29982834]
16. West P, Van Kleek M, Giordano R, Weal M, Shadbolt N. Information quality challenges of patient-generated data in clinical practice. Front Public Health 2017;5:284 [FREE Full text] [doi: 10.3389/fpubh.2017.00284] [Medline: 29209601]
17. Liang Z, Ploderer B, Chapa-Martell M. Is Fitbit fit for sleep-tracking? Sources of measurement errors and proposed countermeasures. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare. New York, NY: ACM; 2017 Presented at: International Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2017; Barcelona, Spain p. 476-479 URL: https://dl.acm.org/citation.cfm?id=3154897 [doi: 10.1145/3154862.3154897]
18. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015;32(7):1024-1028 [FREE Full text] [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]
19. de Zambotti M, Baker FC, Willoughby AR, Godino JG, Wing D, Patrick K, et al. Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. Physiol Behav 2016 May 01;158:143-149 [FREE Full text] [doi: 10.1016/j.physbeh.2016.03.006] [Medline: 26969518]
20. de Zambotti M, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015 Sep 01;38(9):1461-1468 [FREE Full text] [doi: 10.5665/sleep.4990] [Medline: 26158896]
21. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. Chronobiol Int 2018 Apr;35(4):465-476. [doi: 10.1080/07420528.2017.1413578] [Medline: 29235907]
22. 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;38(8):1323-1330 [FREE Full text] [doi: 10.5665/sleep.4918] [Medline: 26118555]
23. Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. Sleep Breath 2012 Sep;16(3):913-917. [doi: 10.1007/s11325-011-0585-y] [Medline: 21971963]
24. Liang Z, Chapa Martell MA. Validity of consumer activity wristbands and wearable EEG for measuring overall sleep parameters and sleep structure in free-living conditions. J Healthc Inform Res 2018 Apr 20;2(1-2):152-178. [doi: 10.1007/s41666-018-0013-1]
25. Sánchez-Ortuño MM, Edinger JD, Means MK, Almirall D. Home is where sleep is: an ecological approach to test the validity of actigraphy for the assessment of insomnia. J Clin Sleep Med 2010 Feb 15;6(1):21-29 [FREE Full text] [Medline: 20191934]
26. Kang S, Kang JM, Ko K, Park S, Mariani S, Weng J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. J Psychosom Res 2017 Jun;97:38-44. [doi: 10.1016/j.jpsychores.2017.03.009] [Medline: 28606497]
27. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
28. Kishi A, Struzik ZR, Natelson BH, Togo F, Yamamoto Y. Dynamics of sleep stage transitions in healthy humans and patients with chronic fatigue syndrome. Am J Physiol Regul Integr Comp Physiol 2008 Jun;294(6):R1980-R1987 [FREE Full text] [doi: 10.1152/ajpregu.00925.2007] [Medline: 18417644]

29. Kishi A, Yamaguchi I, Togo F, Yamamoto Y. Markov modeling of sleep stage transitions and ultradian REM sleep rhythm. Physiol Meas 2018 Dec 31;39(8):084005. [doi: 10.1088/1361-6579/aad900] [Medline: 30089099]
30. Kishi A, Yasuda H, Matsumoto T, Inami Y, Horiguchi J, Tamaki M, et al. NREM sleep stage transitions control ultradian REM sleep rhythm. Sleep 2011 Oct 01;34(10):1423-1432 [FREE Full text] [doi: 10.5665/SLEEP.1292] [Medline: 21966074]
31. Lo C, Bartsch RP, Ivanov PC. Asymmetry and basic pathways in sleep-stage transitions. Europhys Lett 2013 Apr 01;102(1):10008 [FREE Full text] [doi: 10.1209/0295-5075/102/10008] [Medline: 24653582]
32. Bianchi MT, Cash SS, Mietus J, Peng C, Thomas R. Obstructive sleep apnea alters sleep stage transition dynamics. PLoS One 2010 Jun 28;5(6):e11356 [FREE Full text] [doi: 10.1371/journal.pone.0011356] [Medline: 20596541]
33. Burns JW, Crofford LJ, Chervin RD. Sleep stage dynamics in fibromyalgia patients and controls. Sleep Med 2008 Aug;9(6):689-696. [doi: 10.1016/j.sleep.2007.10.022] [Medline: 18314389]
34. Kemp B, Kamphuisen HA. Simulation of human hypnograms using a Markov chain model. Sleep 1986;9(3):405-414. [doi: 10.1093/sleep/9.3.405] [Medline: 3764288]
35. Yang MC, Hursch CJ. The use of a semi-Markov model for describing sleep patterns. Biometrics 1973 Dec;29(4):667-676. [Medline: 4785232]
36. Zung W, Naylor T, Gianturco D, Wilson W. Computer simulation of sleep EEG patterns with a Markov chain model. Recent Adv Biol Psychiatry 1965;8:335-355. [Medline: 5871725]
37. Comte JC, Ravassard P, Salin PA. Sleep dynamics: a self-organized critical system. Phys Rev E Stat Nonlin Soft Matter Phys 2006 May;73(5 Pt 2):056127. [doi: 10.1103/PhysRevE.73.056127] [Medline: 16803018]
38. Penzel T, Kantelhardt JW, Lo C, Voigt K, Vogelmeier C. Dynamics of heart rate and sleep stages in normals and patients with sleep apnea. Neuropsychopharmacology 2003 Jul;28 Suppl 1:S48-S53 [FREE Full text] [doi: 10.1038/sj.npp.1300146] [Medline: 12827144]
39. Klerman EB, Davis JB, Duffy JF, Dijk D, Kronauer RE. Older people awaken more frequently but fall back asleep at the same rate as younger people. Sleep 2004 Jun 15;27(4):793-798. [Medline: 15283016]
40. Langrock R, Swihart BJ, Caffo BS, Punjabi NM, Crainiceanu CM. Combining hidden Markov models for comparing the dynamics of multiple sleep electroencephalograms. Stat Med 2013 Aug 30;32(19):3342-3356 [FREE Full text] [doi: 10.1002/sim.5747] [Medline: 23348835]
41. Schlemmer A, Parlitz U, Luther S, Wessel N, Penzel T. Changes of sleep-stage transitions due to ageing and sleep disorder. Philos Trans A Math Phys Eng Sci 2015 Feb 13;373(2034):-. [doi: 10.1098/rsta.2014.0093] [Medline: 25548271]
42. Ferri R, Pizza F, Vandi S, Iloti M, Plazzi G. Decreased sleep stage transition pattern complexity in narcolepsy type 1. Clin Neurophysiol 2016 Dec;127(8):2812-2819. [doi: 10.1016/j.clinph.2016.05.364] [Medline: 27417057]
43. Zhang X, Kantelhardt JW, Dong XS, Krefting D, Li J, Yan H, et al. Nocturnal dynamics of sleep-wake transitions in patients with narcolepsy. Sleep 2017 Feb 01;40(2). [doi: 10.1093/sleep/zsw050] [Medline: 28364512]
44. Kim JW, Lee J, Robinson PA, Jeong D. Markov analysis of sleep dynamics. Phys Rev Lett 2009 May 01;102(17):178104. [doi: 10.1103/PhysRevLett.102.178104] [Medline: 19518839]
45. Van de Water AT, Holmes A, Hurley D. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography: a systematic review. J Sleep Res 2011 Mar;20(1 Pt 2):183-200 [FREE Full text] [doi: 10.1111/j.1365-2869.2009.00814.x] [Medline: 20374444]
46. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. Sleep 2007 Oct;30(10):1362-1369 [FREE Full text] [Medline: 17969470]
47. Modave F, Guo Y, Bian J, Gurka MJ, Parish A, Smith MD, et al. Mobile device accuracy for step counting across age groups. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e88 [FREE Full text] [doi: 10.2196/mhealth.7870] [Medline: 28659255]
48. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. Sleep 1992 Oct;15(5):461-469. [doi: 10.1093/sleep/15.5.461] [Medline: 1455130]
49. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [Medline: 2748771]
50. Matsuo M, Masuda F, Sumi Y, Takahashi M, Yamada N, Ohira MH, et al. Comparisons of portable sleep monitors of different modalities: potential as naturalistic sleep recorders. Front Neurol 2016;7:110 [FREE Full text] [doi: 10.3389/fneur.2016.00110] [Medline: 27471489]
51. Yoshida M, Shinohara H, Kodama H. Assessment of nocturnal sleep architecture by actigraphy and one-channel electroencephalography in early infancy. Early Hum Dev 2015 Sep;91(9):519-526. [doi: 10.1016/j.earlhumdev.2015.06.005] [Medline: 26140905]
52. McCall C, McCall WV. Objective vs subjective measurements of sleep in depressed insomniacs: first night effect or reverse first night effect? J Clin Sleep Med 2012 Feb 15;8(1):59-65 [FREE Full text] [doi: 10.5664/jcsm.1664] [Medline: 22334811]
53. Ahmadi N, Shapiro G, Chung S, Shapiro C. Clinical diagnosis of sleep apnea based on single night of polysomnography vs two nights of polysomnography. Sleep Breath 2009 Aug;13(3):221-226. [doi: 10.1007/s11325-008-0234-2] [Medline: 19067010]
54. Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, et al. AASM Scoring Manual Updates for 2017 (Version 2.4). J Clin Sleep Med 2017 Dec 15;13(5):665-666 [FREE Full text] [doi: 10.5664/jcsm.6576] [Medline: 28416048]

55. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
56. Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. Br Med J 1994 Jun 11;308(6943):1552 [FREE Full text] [doi: 10.1136/bmj.308.6943.1552] [Medline: 8019315]
57. Ohayon M, Wickwire E, Hirshkowitz M, Albert S, Avidan A, Daly F, et al. National Sleep Foundation's sleep quality recommendations: first report. Sleep Health 2017 Dec;3(1):6-19. [doi: 10.1016/j.sleh.2016.11.006] [Medline: 28346153]
58. Taibi DM, Landis CA, Vitiello MV. Concordance of polysomnographic and actigraphic measurement of sleep and wake in older women with insomnia. J Clin Sleep Med 2013 Mar 15;9(3):217-225 [FREE Full text] [doi: 10.5664/jcsm.2482] [Medline: 23493815]
59. O'Hare E, Flanagan D, Penzel T, Garcia C, Frohberg D, Heneghan C. A comparison of radio-frequency biomotion sensors and actigraphy versus polysomnography for the assessment of sleep in normal subjects. Sleep Breath 2015 Mar;19(1):91-98. [doi: 10.1007/s11325-014-0967-z] [Medline: 24614968]
60. Bellone GJ, Plano SA, Cardinali DP, Chada DP, Vigo DE, Golombek DA. Comparative analysis of actigraphy performance in healthy young subjects. Sleep Sci 2016;9(4):272-279 [FREE Full text] [doi: 10.1016/j.slsci.2016.05.004] [Medline: 28154740]
61. Lucey B, Mcleland JS, Toedebusch C, Boyd J, Morris J, Landsness E, et al. Comparison of a single-channel EEG sleep study to polysomnography. J Sleep Res 2016 Dec;25(6):625-635 [FREE Full text] [doi: 10.1111/jsr.12417] [Medline: 27252090]
62. Meltzer L, Walsh C, Traylor J, Westin A. Direct comparison of two new actigraphs and polysomnography in children and adolescents. Sleep 2012 Jan 01;35(1):159-166 [FREE Full text] [doi: 10.5665/sleep.1608] [Medline: 22215930]
63. Werner H, Molinari L, Guyer C, Jenni O. Agreement rates between actigraphy, diary, and questionnaire for children's sleep patterns. Arch Pediatr Adolesc Med 2008 Apr;162(4):350-358. [doi: 10.1001/archpedi.162.4.350] [Medline: 18391144]
64. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. 24 hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc 2015 Oct 17. [doi: 10.1249/MSS.0000000000000778] [Medline: 26484953]
65. Kushida C, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement W. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. Sleep Med 2001 Sep;2(5):389-396. [Medline: 14592388]

Abbreviations

API: application programming interface
PSG: polysomnography
PSQI: Pittsburgh Sleep Quality Index
REM: rapid eye movement
SE: sleep efficiency
SOL: sleep onset latency
SWS: slow wave sleep
TST: total sleep time
WASO: wake after sleep onset

Edited by G Eysenbach; submitted 19.01.19; peer-reviewed by F Modave, K Ng, S Berrouiguet; comments to author 28.03.19; revised version received 04.04.19; accepted 23.04.19; published 06.06.19.

Please cite as:
Liang Z, Chapa-Martell MA
Accuracy of Fitbit Wristbands in Measuring Sleep Stage Transitions and the Effect of User-Specific Factors
JMIR Mhealth Uhealth 2019;7(6):e13384
URL: https://mhealth.jmir.org/2019/6/e13384/
doi:10.2196/13384
PMID:31172956

Original Paper

Development of a Sensor-Based Behavioral Monitoring Solution to Support Dementia Care

Julia Rosemary Thorpe[1], PhD; Birgitte Hysse Forchhammer[2], PhD; Anja M Maier[1], PhD

[1]Engineering Systems Group, DTU Management, Technical University of Denmark, Kongens Lyngby, Denmark

[2]Department of Neurology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Corresponding Author:
Julia Rosemary Thorpe, PhD
Engineering Systems Group
DTU Management
Technical University of Denmark
Diplomvej, Building 371
Kongens Lyngby, 2800
Denmark
Phone: 45 30113332
Email: jrth@dtu.dk

Abstract

Background: Mobile and wearable technology presents exciting opportunities for monitoring behavior using widely available sensor data. This could support clinical research and practice aimed at improving quality of life among the growing number of people with dementia. However, it requires suitable tools for measuring behavior in a natural real-life setting that can be easily implemented by others.

Objective: The objectives of this study were to develop and test a set of algorithms for measuring mobility and activity and to describe a technical setup for collecting the sensor data that these algorithms require using off-the-shelf devices.

Methods: A mobility measurement module was developed to extract travel trajectories and home location from raw GPS (global positioning system) data and to use this information to calculate a set of spatial, temporal, and count-based mobility metrics. Activity measurement comprises activity bout extraction from recognized activity data and daily step counts. Location, activity, and step count data were collected using smartwatches and mobile phones, relying on open-source resources as far as possible for accessing data from device sensors. The behavioral monitoring solution was evaluated among 5 healthy subjects who simultaneously logged their movements for 1 week.

Results: The evaluation showed that the behavioral monitoring solution successfully measures travel trajectories and mobility metrics from location data and extracts multimodal activity bouts during travel between locations. While step count could be used to indicate overall daily activity level, a concern was raised regarding device validity for step count measurement, which was substantially higher from the smartwatches than the mobile phones.

Conclusions: This study contributes to clinical research and practice by providing a comprehensive behavioral monitoring solution for use in a real-life setting that can be replicated for a range of applications where knowledge about individual mobility and activity is relevant.

(JMIR Mhealth Uhealth 2019;7(6):e12013) doi:10.2196/12013

KEYWORDS

ambulatory monitoring; patient-centered care; physical activity; dementia; wearable electronics devices; activity trackers; mHealth; human behavior; system design

Introduction

Background

The aging of the population and consequent rise in prevalence of conditions such as dementia present a great challenge to society [1]. New care approaches are needed to overcome the increasing disparity between available resources and demands on our health care systems. Mobile and wearable devices, and the rich health-related data these generate, present exciting opportunities to broaden access to care, while enabling

predictive, preventive, personalized, and participatory (P4) health care interventions [2]. One interesting avenue is the use of sensor data to monitor behavior. Individual mobility and activity is highly meaningful with regard to independence and quality of life among people with dementia [3] and could be measured using data recorded from smartphones and wearables, such as location, activity, and step count.

Relevance of Behavioral Monitoring for Dementia Care

Life-space mobility (also referred to as *out-of-home* or *global* mobility) describes the extent to which an individual moves within their environment by any means. Both increasing age and cognitive impairment are associated with reduced mobility [4,5]. This is of great concern with regard to quality of life for people with dementia, as mobility is intrinsically linked to social engagement, functional capacity, affective state and caregiver burden, and a decisive factor for active aging [5-8]. Several factors may be at play when cognitive impairment leads to reduced mobility, such as concerns about safety, usual activities becoming too cognitively demanding, and depressive symptoms (eg, reclusiveness and apathy). Reduced mobility can present a dangerous feedback loop by inhibiting social engagement and stimulation, thereby aggravating the cognitive decline and depressive symptoms that contribute to further mobility reduction. This underscores the importance of maintaining mobility among the elderly and especially the cognitively impaired. Activity can include physical activity levels or activity states or types. Although physical activity is directly linked to mobility, in that a person's functional capacity contributes to their ability to move in their environment, its measurement also complements out-of-home mobility measures by informing how active a person is while home (or other locations). Monitoring activity is relevant among people with dementia in several ways. For rehabilitation, activity monitoring can guide strategies for increasing engagement in meaningful activities [9] and provide insight into how structured an individual's daily routines are. Activity levels also provide a useful indicator for functional capacity loss with cognitive impairment. Some studies have even shown a possible association between physical activity and reduced risk of dementia or dementia progression [10].

Related Work and Open Challenges

Measurement of mobility and physical activity has traditionally been performed using surveys. This approach is limited by its reliance on patients' memory and subjective perceptions of values such as the distances they cover or time spent active each day, which is especially problematic among people with cognitive impairment. Surveys require input from both patients and health care professionals and thus tend to be restricted to discrete measurements at widely spaced intervals with no information about changes that occur daily or even weekly or monthly. The last decade has seen significant progress toward sensor-based behavior measurement, including among the elderly and cognitively impaired. Mobility and activity features have been calculated using specialized global positioning system (GPS) kits and ankle-worn accelerometers [4,8,11,12]. Although these works offer valuable contributions toward sensor-based behavioral monitoring, the use of specialized systems or strict protocols regarding device placement to measure behavior under experimental conditions is unrealistic for long-term everyday use, therefore difficult to replicate in a real-world setting. This motivates a growing interest in leveraging the wide availability and acceptance of today's personal devices [8,13-15]. Smartphones and wearables have successfully been applied to measure activity among older adults under free-living conditions [13], daily step count and distance covered among people with dementia [14], and life space among people with Parkinson disease and mild-to-moderate Alzheimer disease [8,15]. System design considerations for real-world use are addressed in a previous study [14], which demonstrates how adequate data are recorded over an extended period (5 months) to reveal behavioral patterns. In some studies [8,13,15], the sensor-based approach is evaluated by comparing measures between experimental and control groups, indicating that significant change in sensor-based behavioral measures might be detected with disease onset/progression; however, no comparison is made with manually reported data. The behavioral measures used vary as follows: activity measures range from daily steps to more detailed descriptions of active and sedentary states; and life space measures range from basic distances to trips or time frames away from home but without extraction of travel trajectories in their estimation. Instead, a threshold distance from home is used to determine whether points are at or away from home. A high threshold (such as 500m in one study [15]) may not detect trips within the subject's neighborhood. Even with a lower threshold (such as 25m in another study [8]), it is not possible to infer how many places the subject visited if they did not travel home between places, or whether they are continuously moving (eg, going for a long walk) compared with staying at a single location (eg, visiting a friend, in hospital). Without reference data such as self-reports, it is difficult to evaluate the performance of such methods.

This study therefore aimed to advance progress toward mobile/wearable technology–based behavioral monitoring by building upon noted strengths regarding real-world suitability, extending mobility measurement to incorporate GPS trajectory extraction, and providing evidence comparing sensor-derived measures with reference data in the form of self-reports.

Objectives

The main objective of this study was to develop and test a complete set of tools for measuring mobility and activity using widely available data from off-the-shelf devices. Furthermore, we have described a generic setup for collecting the required data. Together, these are intended to fulfil the purpose of monitoring behavior on an individual level to observe patterns or changes among community-dwelling older adults, such as people with early-stage dementia. Two important goals include

- Transferability: Others should be able to implement the solution at minimal additional effort or expense.
- Real-life suitability: The solution should be developed for long-term, unsupervised, everyday use rather than for laboratory test conditions.

Through fulfilling these objectives, this study contributes a comprehensive behavioral monitoring solution to advance clinical research and practice and enable P4 health care systems.

Methods

Extracting Features and Metrics From Behavioral Data

Here we have described the methods used to measure mobility and activity from sensor data, including the algorithms that were adapted and developed for this purpose. We have chosen to focus on 3 types of data: location, step count, and recognized activities, as these are both widely available and highly relevant for measuring mobility and activity. All algorithms and calculations described in this section were implemented in the R programming environment and are available on request.

Mobility Measurement

Mobility measurement has been described in a number of works on chronic diseases, mental health, and among the elderly [16-19]. These works describe a variety of metrics that we categorize here as spatial, temporal, or frequency-based. Spatial measurements, often termed *life space*, include geographical areas or distances covered, such as the area of the minimum convex polygon (MCP) enveloping a specified quantile of GPS coordinates, distances from home (action range), or total distance covered. Temporal measurements include time spent at or out of home or visiting places of interest. Frequency-based measures include counts within a given period, such as the number of trips out of home or places visited per day or week. Certain life-space measures can be calculated from raw GPS data, whereas other metrics require knowledge of a home location, and some require knowledge of GPS trajectories, that is, the series of stays and moves within the location data.

The raw location data for each user comprises merged watch and phone data, including time stamp, latitude, longitude, and accuracy in meters. The sampling frequency is irregular, and besides any periods of missing data, readings are spaced between a few milliseconds and approximately 5 min apart. The only preprocessing applied was to filter the data according to accuracy with an upper limit of 25 meters. The inputs required to calculate the mobility metrics include the set of coordinate pairs, GPS trajectories (a series of stays and moves), and a known home location. The following sections describe how these inputs are obtained, followed by a description of the metrics calculations.

Extraction of Travel Trajectories and Home Centroid

Many temporal and frequency-based measures require GPS trajectories describing how the person stays at or moves between locations, also referred to as *mobility traces* [17]. This requires analysis of the raw location data to extract *stay* (or *stop*, *visit*) and *move* (or *go*) events and identification of geolocations (or *points of interest*, *hotspots*) in the dataset. The identification of trajectories follows a similar approach to those described in the literature [20-22], whereby the data are first split into *stays* and *moves* using time and/or distance thresholds and then clustered to identify geolocations in the dataset. We further included a filtering step to merge temporally close stays or moves that likely belong to the same event. An overview is provided in Figure 1.

Figure 1. Overview of the trajectory extraction algorithm. DBSCAN: density-based spatial clustering of applications with noise; GPS: global positioning system.

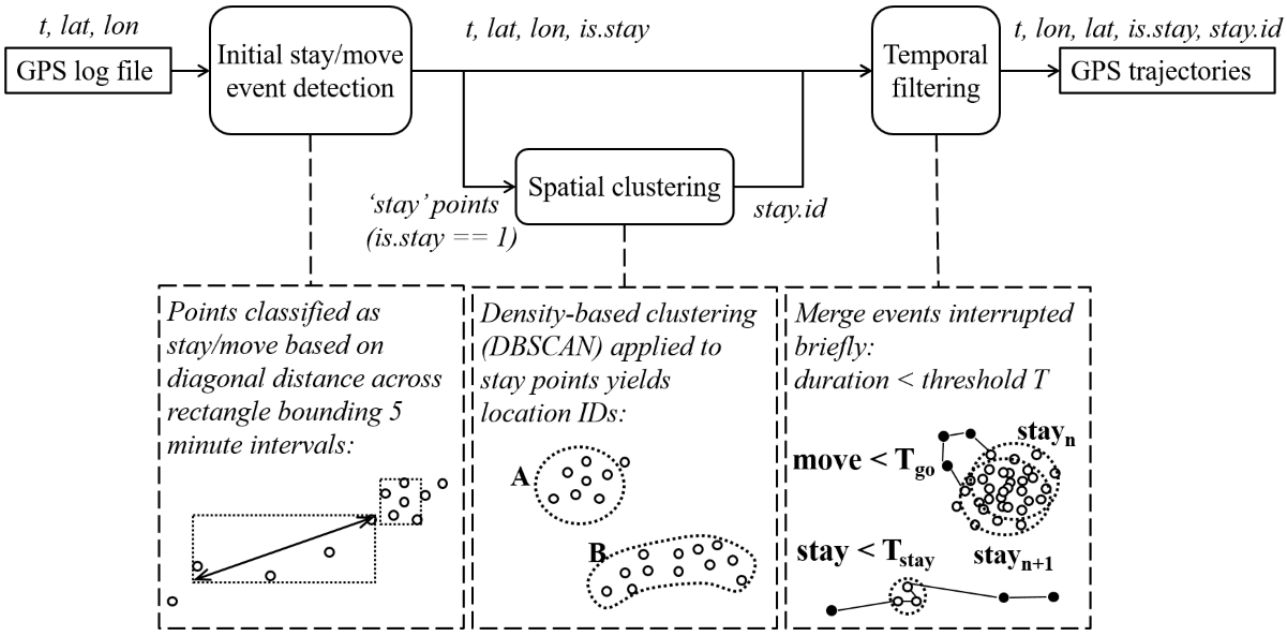
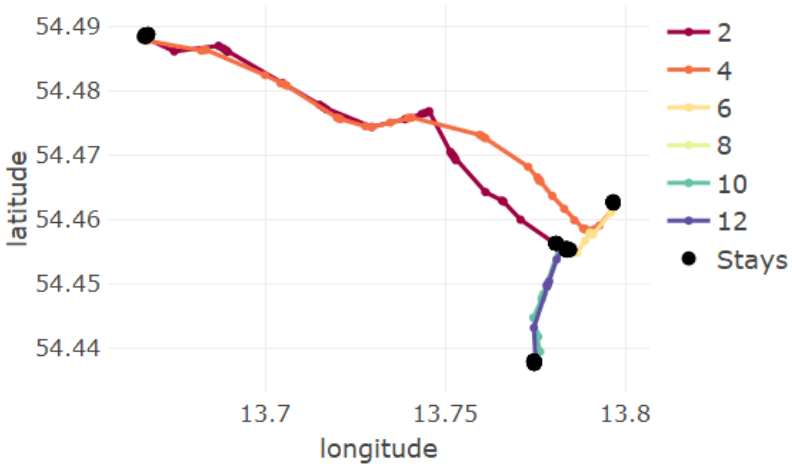


Figure 2. Travel trajectory showing the sequence of stay/move events. All stays are shown as black discs, and each move event shown as a colored line named with its chronological order. The "stays" make up trajectory events 1, 3, 5, 7, 9, 11, and 13. Note: actual GPS data is displaced and the map background removed for anonymity.



The initial stay detection step combines both time and distance information (as in a previous study [22]) to accommodate irregularly sampled location data without the need for further preprocessing. For each location, a rectangle is calculated bounding data from intervals of 5-min ahead. If the diagonal distance across the rectangle exceeds a predefined threshold, the point belongs to a *move*; if below the threshold, all points in the interval are classified as *stay* points. The stay points are then clustered into distinct locations using density-based spatial clustering of applications with noise (DBSCAN), as this method is well-suited to clusters of varying shapes, does not require a priori knowledge about the number of clusters, can handle outliers, and has been successfully applied previously for this purpose [20]. Any outliers not belonging to stay locations are reclassified as move points. In the third (final) step, a time threshold is used to filter very short stays or moves, which is effectively the same as merging move segments or stay events at the same location that are very close together in time.

The algorithm assigns indices to locations without inferring any further information about the nature of the location, with the exception of the subject's home. The home location is estimated in 2 steps. First, the statistical mode of all GPS points is calculated as *home*. Once all stay locations are extracted from the dataset, those close to home (within a specified threshold) are classified as being at home. An updated home location is then calculated as the centroid of the subset of all points classified as stays at home.

In summary, the GPS log data (time stamp, latitude, and longitude) are used to calculate the following information for each point: whether it is a stay/move; which event it belongs to in a chronologically ordered sequence per day; and for stay points, a location index and whether it is home. An example trajectory is shown in Figure 2.

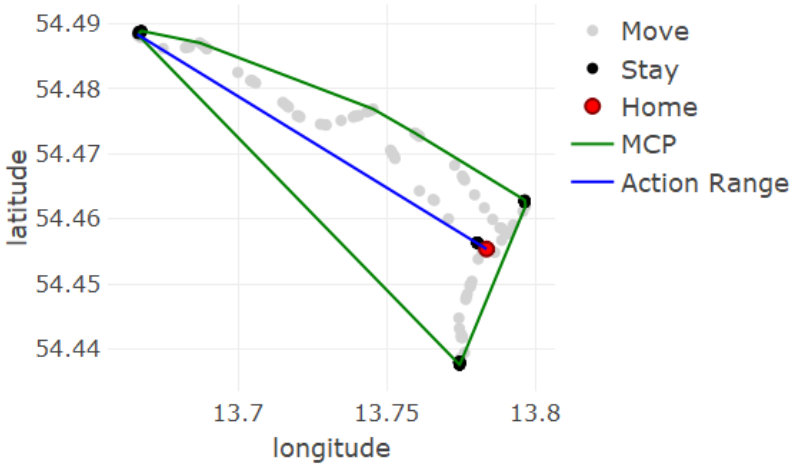
Calculation of Mobility Metrics

The set of mobility metrics was selected by combining various other selections used in the literature for similar purposes [3,8,17,23,24], ensuring that different types of measures are included (ie, spatial, temporal, and frequency-based). The set of mobility metrics includes the following calculations:

- *MCP:* Area of the smallest possible convex polygon constructed around the data, also referred to as the *mobility envelope*. This is calculated by applying the R function, *chull*, to a subset of points for which the distance to the centroid falls within a 99% quantile.
- *Action range:* Straight-line distance between home and the most distal point of a journey, sometimes referred to as *home range*. The geodesic distance is calculated between the home centroid and all other points in the dataset. For each stay and move event in the GPS trajectory, action range is calculated as the maximum of these distances.
- *Distance covered:* Sum of all geodesic distances between consecutive stays centroids.
- *Time spent out:* Sum of durations for all events excluding stays at home.
- *Time spent moving between locations:* Sum of durations for all move events.
- *Number of places visited:* Count of unique places visited (including home). This requires location identifications (IDs) so that a single place is only counted once even when visited multiple times per day.
- *Number of trips:* Count of all moves in the GPS trajectories.

An example of 2 of the spatial measures, MCP and action range, is shown in Figure 3, where these are calculated for the GPS trajectory in Figure 2.

Figure 3. Visual representation of the mobility metrics minimum convex polygon (MCP) and action range, overlaid over global positioning system (GPS) data for one subject for a single day (same trajectory as in Figure 2).



Activity Measurement

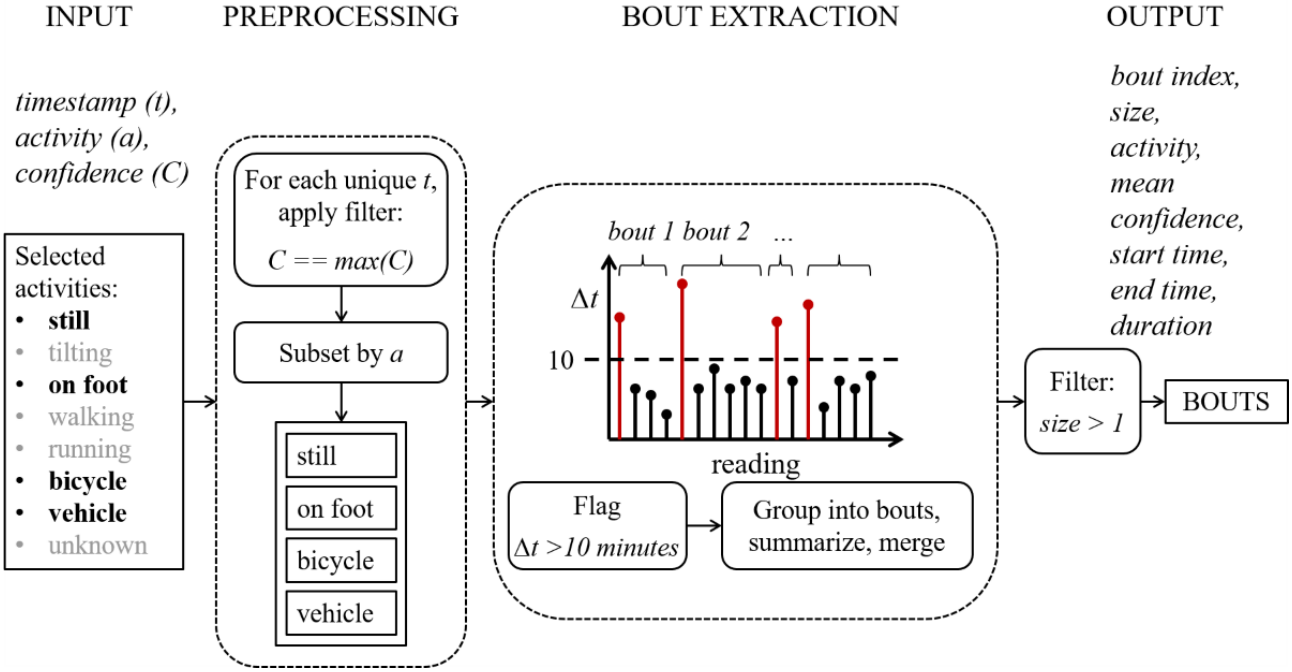
Within the context of behavioral monitoring for people with dementia, activity measurement is used to gauge how active a person's everyday life is generally rather than to provide detailed information about physical exercise. Examples of activity measurement in related works (eg, those investigating mobility and activity among similar target groups) include measurements such as active/walking time, number and duration of walking bouts, and total steps per day [7,8,25]. Here we proposed extending the measurement of activity bouts beyond walking (or active) to include other modes of transport, as these offer insight into a person's everyday routines and preferences, for which any gradual or sudden change could be telling regarding changes in health status. This section describes the methods used to extract these activity features (activity bouts and steps) from sensor data, including recognized activities and step count, respectively.

Activity Bout Detection

Activity bouts are detected using data obtained from Google's activity recognition application programming interface (API: ActivityRecognitionClient). This includes the following types of movement: *still*, *tilting*, *on foot*, *walking*, *running*, *on bicycle*, *in vehicle*, and *unknown*, where *running* and *walking* are both subsets of *on foot*. Each instance of an activity is recorded with a time stamp and confidence level for its recognition. A number of activities can be recorded at the same time instance. The sampling period is typically approximately 5 min.

The data are first preprocessed to keep only those activities of interest: *still*, *on foot*, *bicycle*, and *vehicle*. *On foot* is kept in place of both walking and running in accordance with the target group and purpose. The dataset is then reduced further by keeping only those readings with maximum confidence within each distinct time stamp. Where multiple activity types show equal and maximum confidence, all are included. A time threshold is used to split each activity type subset into bouts of continuous activity. Occurrences of the same activity type within 10 min of one another are grouped into the same bout. Bouts comprising only a single reading are filtered out. The bouts are then summarized yielding the following set of measures: bout number (chronologically ordered per day), activity, start time, end time, duration, and number of readings. An overview of the bout extraction algorithm is provided in Figure 4.

Figure 4. Bout extraction from activity data.



Step Count

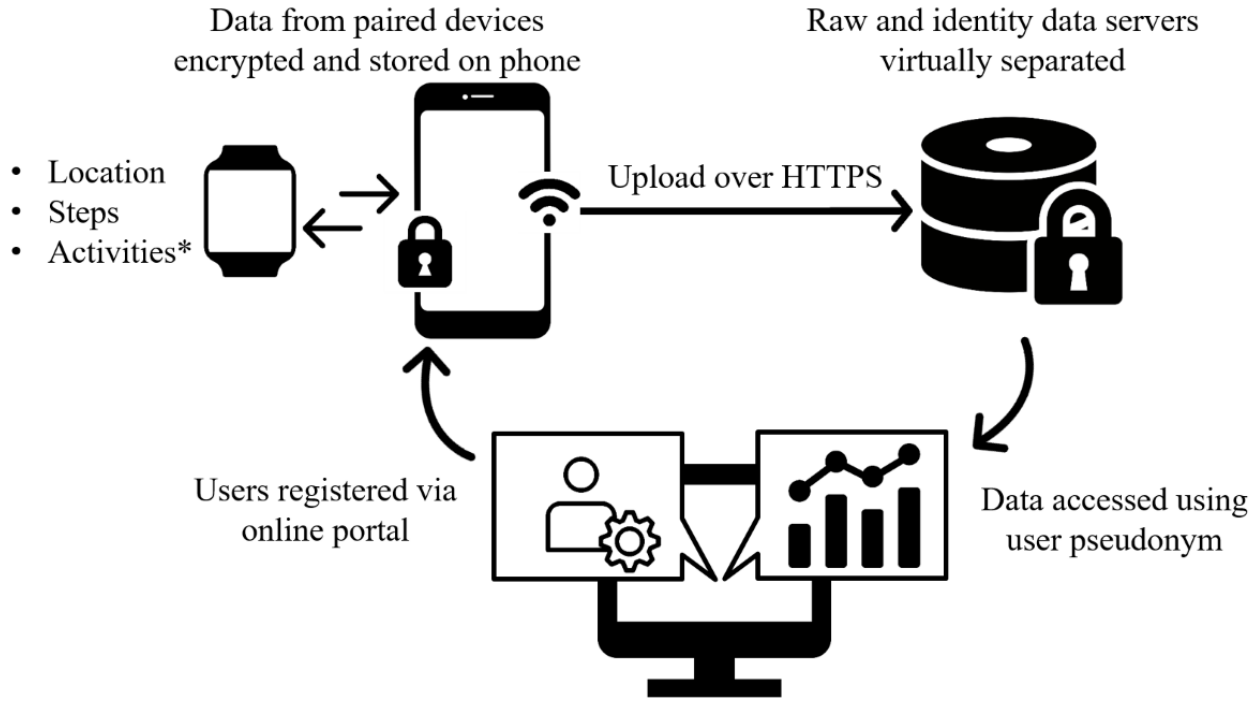
Step count data are used in a straightforward manner to calculate total daily steps. The data are first restructured to obtain a cumulative sum of steps over each day, rather than increasing until the device restarts/reboots. Daily step count from a worn device can potentially account for short ad hoc bouts of activity (such as performing household chores) that take place over and above other exercise regimes or broader movement between geolocations.

Data Collection Setup: Devices, Sensors, and Apps

We have presented a collection of algorithms and metrics to describe individual mobility and activity. Applying these tools to monitor behavior requires infrastructure to gather the necessary data inputs. In this section, we describe such data collection setup, including devices, sensors, and applications. We have sought to compile a setup that can be replicated by others by using off-the-shelf devices and open-source resources as far as possible.

The 3 types of data included are location, activity, and step count. Location can be recorded using GPS sensors on board most smartphones and a number of wearables (smartwatches and activity trackers) currently in the market. Recognized activities are calculated primarily from accelerometer data (and in combination with pedometers, gyroscopes, and barometers where available). Step count is typically recorded using activity trackers and smartphones, either from on board pedometers or derived from accelerometer data. The devices used in this study are Google Nexus 5 smartphones and Sony SmartWatch 3 smartwatches, running on Android operating system v6.0.1 and Android Wear operating system, respectively. Android devices offer unrestrictive platforms for development and have been shown to be comparable with ActiGraph for physical activity estimates [26]. Both devices record location and step count, whereas activity types are recorded on the phone only. A custom-built application was used to securely collect, store, and transfer data. The app is an adapted version of that described by Stopczynski et al [27], which is publicly available under the *OpenSensing* GitHub organization. The app uses Google APIs to access sensor data (*LocationListener*, *ActivityRecognitionApi*, and *SensorManager*). New users are registered through a Web portal where they create a front-end username and password and are assigned a back-end pseudonym. The custom app is then accessed from Google Play Store and installed on both paired devices. Watch data are uploaded to the phone, where all data are continuously collected, encrypted, and stored locally on the phone. Only when a Wi-Fi connection is available are encrypted data files uploaded to a server over https. To ensure security, 2 virtually separate servers are employed, an anonymous raw data server and an identity server with user pseudonyms. Data are then decrypted and transferred to a database for analysis as required by an authorized user (eg, a researcher with administrator rights). An overview of the data collection setup is provided in Figure 5.

Figure 5. Data collection setup: data is collected using sensors on-board a smartwatch and smartphone (*activities collected using the phone only), encrypted and stored locally on the phone then transferred securely to a server from where it is accessed by an administrator.



Through combining the data analysis methods presented in the previous section together with the data collection setup presented here, we provide the necessary tools for a complete behavioral monitoring solution to be applied directly in clinical research and practice. This is evaluated, and the results presented in the following sections.

Evaluation

The behavioral monitoring solution was implemented in a pilot feasibility study with 5 healthy volunteers (3 female and 2 male), aged between 31 and 40 years. The purpose of the evaluation study was to test the setup under free-living conditions to obtain real-world behavioral data with which to examine the performance of the feature extraction algorithms. Although the intended target group is ultimately people with dementia, the solution is tested among adults with no cognitive impairment to ensure reliable self-reporting, before carrying out any further testing among the target group in future.

Equipment, Material, and Methods

Participants were provided with the behavioral monitoring setup, including a smartphone and smartwatch, to use for a period of 1 week. They were instructed to try to wear the watch during the day and charge it at night as required. All participants continued to use their own smartphone during the study and did not interact with the study phone besides taking it with them when going from place to place in everyday life and keeping it charged and paired with the smartwatch. Participants were also provided with a set of log sheets in which they had to record their movements daily. The log sheets comprised 15-min intervals with columns for stay, move, and the stay location or mode of travel.

Data Analysis

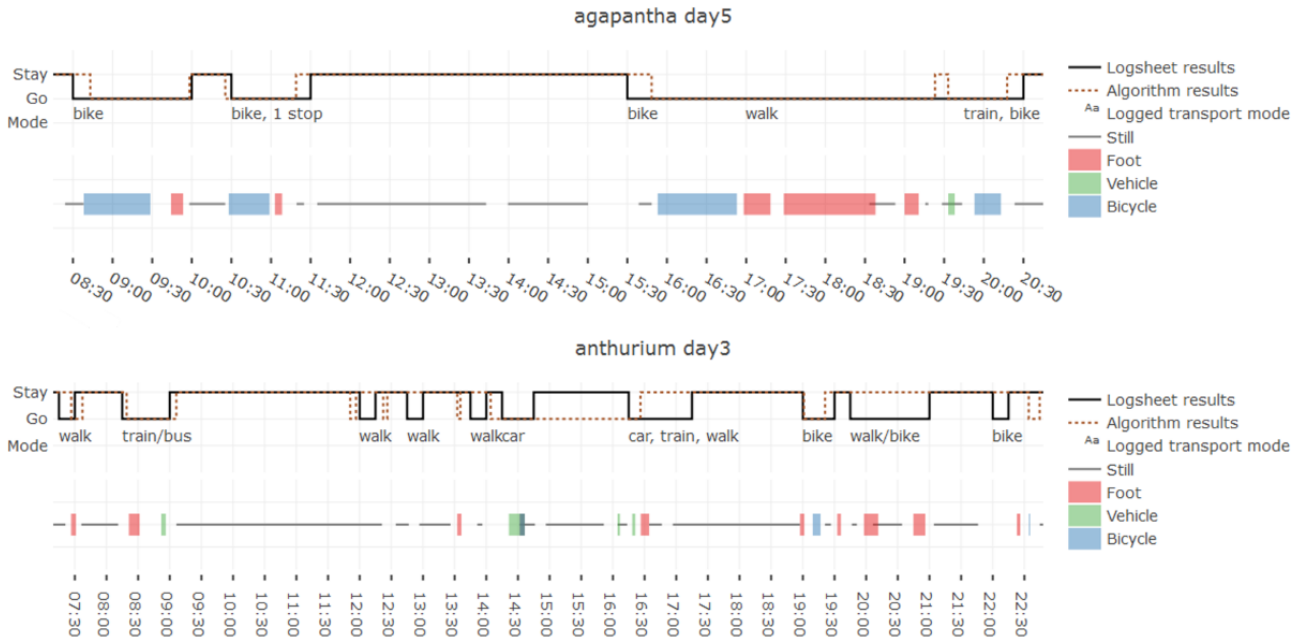
Mobility Measurement

Time charts were created to compare participants' logs with algorithm results. Results of the trajectory extraction algorithm were mapped to daily time charts as a signal indicating when the participant was in a *stay* or *move* state. Participants' log sheets were then captured and processed to create matching time charts for their travel trajectories. These were assessed to determine the level of agreement between the logged and reported travel, in terms of both whether trajectory events are detected or not and their timing. The mobility metrics *time outside of home* and *number of places visited* were calculated using both log sheets and the algorithm results and compared. These metrics were selected as they cover different types of information (time and count) that requires both durations and location identification and could feasibly be calculated from participant-reported data.

Activity Measurement

Automatically extracted activity bouts were plotted on the daily time charts (alongside stay/move signals). Each *move* epoch from the trajectories was annotated with participants' reported transport modes for comparison (see Figure 6). These combined time charts were assessed to gauge the level of agreement between sensor-derived and logged transport modes, identify recurrent errors, and infer strengths or weakness of the approach. Step count data were collected from both the smartwatch and smartphone to compare the 2 sources in terms of daily totals and cumulative step count signals.

Figure 6. Example time charts for one day from two participants comparing log sheets and algorithm results for trajectory extraction and activity bout detection.



Results

Mobility Measurement

Participants' log sheets were compared with results obtained automatically from sensor data. The trajectory extraction algorithm achieved 92% sensitivity for detecting a *move* event. As participants reported that log sheet times were often only approximate estimates (the log sheets could not always be filled at the time of the move), a window of 30 min surrounding the logged *move* was used to determine the correct detection (true positive) by the algorithm.

The results for estimation of mobility metrics *number of unique places visited* and *total time spent out* are compared in Figure 7 and Figure 8. Residuals between the algorithm- and log sheet–based estimates offer insight into the performance of the algorithm in relation to manual reporting. The root mean square deviation (RMSD) indicates that the overall difference between algorithm and log sheets is <1 place visited and <1 hour away from home per day (Table 1). The residual means indicate that the algorithm tends slightly toward more places than logged and tends slightly toward less time at home.

Example time charts comparing reported and extracted travel trajectories are shown in Figure 6. These provide further insight into where and how the algorithm deviates from the log sheets. Time offsets between the 2 signals for move start/end times and durations tend to be below 15 min, which is the time resolution of the log sheet data. Certain moves are interrupted by a short stay in either one of the signals and not the other, which could be attributed to, for example, waiting for or changing between transport modes, or stopping to look in a shop. In very few cases, the disagreement between algorithm results and logged data corresponds to longer periods or certain events are missed entirely (eg, in Figure 6, *anthurium day3*, where a 1.5-hour stay is missed). This could be because of signal loss when the participant goes indoors. Conversely, certain moves are detected that are not logged, which may be because of the participant forgetting to report a move or moving very near to their stay location (eg, in their garden).

Figure 7. Comparison between the numbers of unique places detected from sensor data and reported by study participants in logsheets.

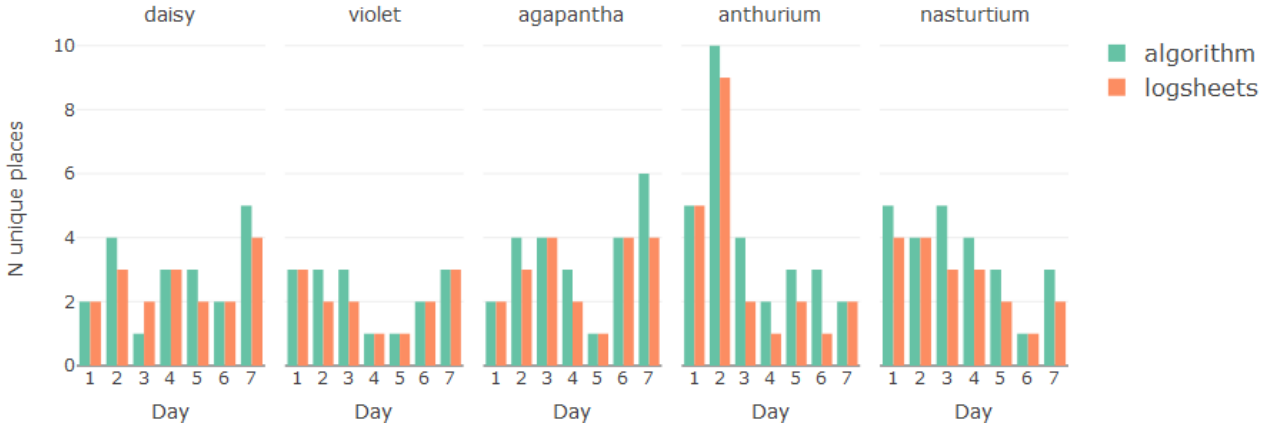


Figure 8. Comparison between the time spent out of home detected from sensor data and reported by study participants in logsheets.

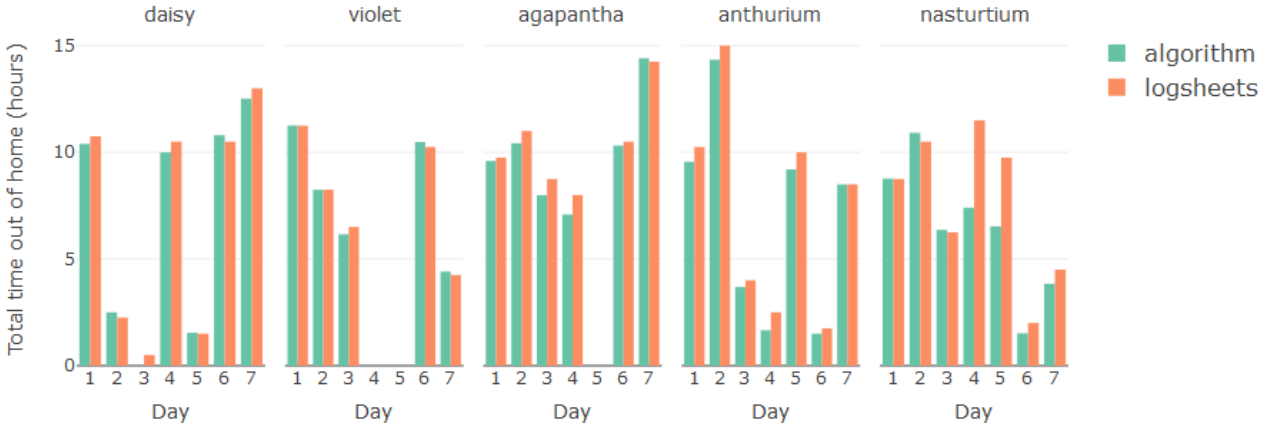


Table 1. Comparison between metrics calculated from log sheet and algorithm trajectories. Residuals are calculated by subtraction of log sheet results from algorithm results, and are used to calculate root mean square deviation (RMSD) and mean (SD).

Residuals (algorithm-logs)	RMSD	Mean (SD)
Number of places per day	0.94	0.60 (0.74)
Time out of home per day (min)	58.66	−25.87 (53.42)

Activity Measurement

Activity Bout Detection

The detected activity bouts were plotted together with trajectories in the time charts and compared with reported transport modes. Time charts were initially reviewed to assess overall performance and pinpoint recurring errors to investigate in further detail. The time charts showed that the algorithm generally detects activity bouts during moves and is able to detect multimodal travel, with only limited examples where no bouts are detected during reported travel. One evident limitation is that the detected activity bouts tend not to fill the entire duration of the travel. Instead, each journey comprises one or more shorter bouts along with still periods and gaps in the signal. A recurring error was confusion between travel by bicycle and by vehicle, for which the classification accuracy is determined by Google's activity recognition. This was investigated further, particularly as the difference between travel by bicycle and vehicle has implications for measuring physical activity or illness-related changes in transport preferences (eg, compared with confusion between bicycle and on foot). Figure 9 shows a classification matrix subset to include only classification between bicycle and vehicle, which indicates accuracy of 85% for classification between the 2 transport modes.

A further observation was a lack of walking bouts during stays, which might be expected as part of an ordinary workday. One contributor may be the filtering of single-reading bouts, presenting an apparent trade-off between detecting more actual activity bouts and introducing additional, incorrect activities during others. Another likely cause is that the data were collected from the smartphone, which may not have been carried on the participant's person for shorter walking trips around the home or workplace, highlighting the importance of investigation within real-life settings. This is examined more closely in the step count data, which is available for both the phone and watch.

Figure 9. Classification matrix showing confusion between detection of bicycle and vehicle activities in number of occurrences.

Logged \ Detected	Bicycle	Vehicle
Bicycle	47	11
Vehicle	2	25

Step Count

Step count is compared between the watch and phone for all days where both sources are available. Owing to technical failures in accessing the watch data, it is available for selected participant-days only (agapantha=4, anthurium=5, daisy=0, nasturtium=7, and violet=3). Step count measured from the watches was substantially higher than from the phones. The cause of the disparity between counts appears to be twofold. There are periods where only the watch step count increases and not the phone (Figure 10, section B), indicating that the watch may be worn while the phone is placed still. There are periods where both increase simultaneously, where the watch steps increase at a faster rate than the phone (Figure 10, section C). The substantial difference between daily total steps measured from the watch and phone is demonstrated in Figure 11.

Figure 10. Comparison of step count accumulated over the day between watch and phone. Outlined sections show A) missing updates in watch step count, B) watch records steps while phone does not, and C) watch step count increases at a faster rate than phone.

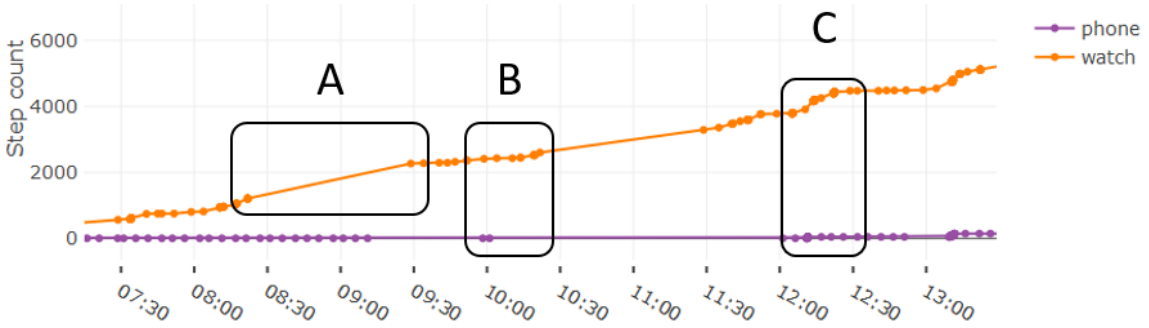
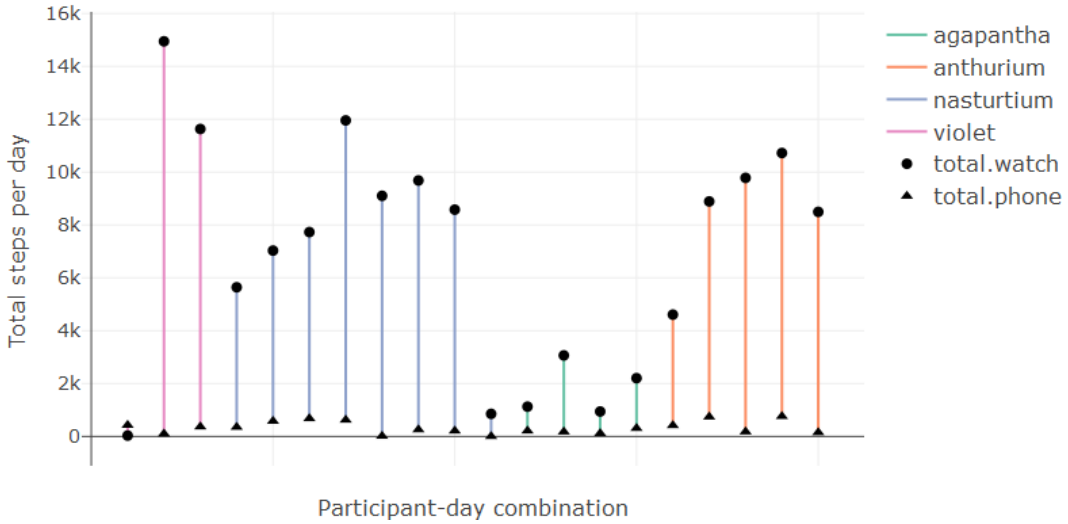


Figure 11. Error bars demonstrating the large difference between total daily step count between watch and phone for each available day.



An important limitation evident in the step count data is regarding sampling irregularities. Long periods (over an hour) without any readings are common in the watch step count (Figure 10, section A). Although the steps appear to be logged during these periods (and therefore not expected to influence daily total step counts), this does not support bout detection from the step count data, as it is impossible to infer the distribution of walking bouts over the course of the day.

Discussion

Principal Findings

This study has presented a behavioral monitoring solution that leverages widely available data (location, activity, and step count) from mobile and wearable devices used in everyday life. A set of tools was developed to measure mobility and activity, and a generic data collection setup was described to support its implementation. Evaluation of the behavioral monitoring solution showed that it is capable of estimating participants' travel trajectories from raw GPS data to calculate spatial, temporal, and count-based mobility metrics. An activity bout detection algorithm was shown to successfully extract bouts of activity during travel, including for multimodal transport; however, it appears not to capture their full duration. Step count was shown to be more reliable from a wearable device than from a smartphone under real-world conditions, in which case it is preferable over activity bouts for estimating general daily activity, including during stays such as at work or home.

These results build upon research toward sensor-based monitoring of mobility and activity among older adult populations in [8,13-15] by incorporating trajectory extraction into the mobility measurement, as well as information about modes of transport in activity monitoring, and by providing evidence comparing the sensor-based approach with self-reports. An important goal has been to describe a solution that can be replicated by others for monitoring behavior to support clinical research and practice in a range of applications. The following sections discuss the extent to which the setup is transferable in terms of technical and practical considerations, and of clinical utility.

Technical Feasibility and Real-World Considerations

A core contribution of this study is the set of algorithms developed to translate sensor data into meaningful behavioral insights. For these to be replicable and useful, we first consider what data are required as inputs and how this is acquired.

The mobility monitoring uses GPS data, which can be acquired using a wide range of available mobile phones. The sampling frequency need not be regular and depends on anticipated trajectory event duration, with a trade-off regarding power consumption. In this study, the mode of sampling periods for location data is approximately 5 min, as it is not of interest to

capture stay/move events of shorter duration. Participants were instructed to charge the device overnight, which was sufficient in this study (although it is noted that the devices were not used for other purposes than for data capture). Requirements for location accuracy depend on the anticipated geographical span of stay locations or moves of interest, which in this study was restricted to 25 meters. The activity bout detection relies on Google's activity recognition (though this could also be implemented using output from another activity recognition algorithm). Step count data are used only for daily totals, which can be obtained from a range of available devices; however, the results in this study indicate that devices vary substantially in step count estimations.

This study has used devices running Android and Google APIs to access the sensor data these generate. This can be replicated using any Android device with the same sensors. The infrastructure for storing and accessing data is based on open-source software from the OpenSensing GitHub organization described in [27] (for Apple products, a similar approach could be feasibly implemented using frameworks such as Apple Research Kit).

This study has endeavored to capture real-world data from everyday life, which leads to important considerations regarding data availability and quality. Data availability from the phone was high, with readings recorded approximately every 5 min throughout the recording period. However, data availability from the watch was poor, with extensive periods (up to days) of missing data. This could be partly attributed to a lost connection between the phone and watch, as the collection of watch data relies on an additional step (compared with phone sensors), whereby data are transferred over Bluetooth to the phone. With regard to data quality, it is not possible to control device placement or wear time, which can lead to missed steps, activities, or trajectory events. This is particularly relevant for counting steps during stay events, for example, to estimate short bouts of physical activity during work, for which a wearable device is recommended.

Clinical Utility

The transferability and modular structure of the behavioral monitoring approach offers broad potential for supporting clinical research and practice, for example, to evaluate the effect of an intervention on behavior, to monitor behavioral change as an indicator of disease progression, or to advance understanding of how different factors or conditions are associated with changes in behavior.

To better gauge clinical utility, algorithm performance is discussed in relation to intended purpose. The first algorithm uses location data to derive metrics for monitoring life space mobility. Existing methods for assessing life space mobility include a questionnaire-based approach that asks about travel within a set of ranges from home, the frequency (in days per week) at which the subject moves in each range, and about independence level (ie, support from equipment or people) [28]. The algorithm is not able to detect the lowest 2 mobility levels included in the questionnaire: movement within the room in which the subject sleeps and movement to other rooms in their place of residence. It is also not possible to infer the level of support using the algorithm-based approach. However, compared with the questionnaire, the algorithm-based approach offers other types of mobility measures besides distance from home (such as time spent at home or unique locations visited), does not rely on the subject's ability to recall events, and is more scalable as it does not rely on manual reporting.

Algorithm performance holds implications for the detectable increments of change in behavior. Although there is limited evidence available quantifying change in life space mobility with disease progression among people with dementia, significant differences in mobility between older adults with and without mild-to-moderate Alzheimer disease are documented in [8] using a similar approach. These include differences between group means for an area of approximately 50 km^2, perimeter of approximately 15 km, distance from home of approximately 1 km, and time away from home of approximately 50 min (6.6% with a reported mean recording time of 7.5 hours per day), where the first 3 are significant and the latter marginally significant. For the spatial measures, such increments should be detectable given the availability, accuracy and sampling frequency of the location data, and trajectory detection presented. For time away from home, the algorithm results deviated from subject reports by >50 min; however, as reporting errors may be a contributing factor, further investigation is necessary to determine whether the algorithm is adequately sensitive to detect changes under 50 min.

The sensor-based activity monitoring approach can be used to gain insight into how a person travels around and to estimate total daily steps as an indicator of overall activity level. Owing to the large discrepancy between device step counts in this study, it is recommended that a device with validated step count is used, or at least that the same device is used throughout any monitoring period. An existing instrument for assessing physical activity is the Global Physical Activity Questionnaire (GPAQ) [29]. Compared with the GPAQ, the sensor-based approach described in this study does not distinguish between activity intensities nor precisely measures time spent in activity states but offers daily step count as an alternative measure of overall daily activity. The GPAQ asks about modes of transport in days per week and time for those days to infer activity performed while moving from place to place. The sensor-based approach is similarly able to measure transport habits; however, as the duration is not precisely measured, this may be better suited to monitoring everyday life habits to detect a change in lifestyle (eg, increase in vehicle use and decrease in walking), rather than, say, adherence to an exercise program.

On the basis of these characteristics with regard to both mobility and activity, suitable target groups could include older adults who live in the community at a baseline functional level sufficient for independent journeys beyond the home. The algorithm-based approach is particularly beneficial where self-reporting is challenging or inaccurate, such as for cognitively impaired individuals. Other relevant application areas in which mobility- and activity-related behavior is highly meaningful include depression or other mental illnesses, active aging, rehabilitation, and lifestyle-diseases (or risk thereof).

Implications for Advancing Predictive, Preventive, Personalized, and Participatory Health Care

By relying on common, personal devices and allowing for real-world data, the behavioral monitoring solution is geared to enable P4 health care approaches by generating information within the scale and context that this requires. Continuous behavioral monitoring in a home setting can reveal patterns in behavior and fuel the development of predictive models to anticipate disease trajectories or adverse events. This also allows health care professionals to proactively prevent problems earlier than would be possible with prescheduled visits months apart. Knowledge about behavior and lifestyle can inform personalized interventions that take into account which aspects of quality of life are most important to the individual. Data describing behavioral patterns also offer a valuable resource for sharing information about patient status between patients and health care providers in participatory care approaches.

Conclusions

This study has described a novel sensor-based approach to behavioral monitoring for use among people with dementia. We have presented a set of algorithms to measure mobility and activity from sensor data, including location, recognized activity, and step count data and a technical setup for collecting these data inputs.

An evaluation of the behavioral monitoring solution among 5 participants for 1 week showed that the setup was capable of extracting travel trajectories, mobility features, activity bouts, and daily step count using a smartphone and smartwatch in a natural setting. Each set of results provides related yet distinct information about a person's daily life: mobility describes the extent to which the persons goes out, activity bouts describe *how* they go out, and step count supplements this with information about how active they are generally, including periods at home or work. Combining these measures provides insights into daily rhythms or longer temporal patterns. This could support clinical applications involving patient groups for whom mobility and activity behavior is closely tied to intervention outcomes, such as among the elderly and people with dementia, and to advance P4 health care.

Acknowledgments

The authors would like to acknowledge the Dementia and Memory Clinic at Rigshospitalet-Glostrup, Radu Gatej from Kl7, Luna S Hansen, the Copenhagen Centre for Health Technology, and Videncenter for velfærdsteknologi (VihTek), all of who have supported and assisted with this research.

Conflicts of Interest

None declared.

References

1. Response to the growing dementia burden must be faster. The Lancet Neurology 2018 Aug;17(8):651. [doi: 10.1016/S1474-4422(18)30256-4] [Medline: 30033050]
2. Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E. Cognitive testing in people at increased risk of dementia using a smartphone app: the iVitality proof-of-principle study. JMIR Mhealth Uhealth 2017 May 25;5(5):e68 [FREE Full text] [doi: 10.2196/mhealth.6939] [Medline: 28546139]
3. Thorpe JR, Forchhammer BH, Maier AM. Sensing behaviour in healthcare design. : Design Society; 2017 Presented at: 21st International Conference on Engineering Design (ICED17); August 21-25, 2017; Vancouver p. 171-181 URL: https://www.designsociety.org/publication/39622/Sensing+behaviour+in+healthcare+design
4. Shoval N, Wahl HW, Auslander G, Isaacson M, Oswald F, Edry T, et al. Use of the global positioning system to measure the out-of-home mobility of older adults with differing cognitive functioning. Ageing and Society 2011 May 16;31(05):849-869. [doi: 10.1017/S0144686X10001455]
5. Rosso AL, Taylor JA, Tabb LP, Michael YL. Mobility, disability, and social engagement in older adults. J Aging Health 2013 Jun;25(4):617-637 [FREE Full text] [doi: 10.1177/0898264313482489] [Medline: 23548944]
6. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. Gerontologist 2010 Aug;50(4):443-450. [doi: 10.1093/geront/gnq013] [Medline: 20145017]
7. Werner S, Auslander GK, Shoval N, Gitlitz T, Landau R, Heinik J. Caregiving burden and out-of-home mobility of cognitively impaired care-recipients based on GPS tracking. Int Psychogeriatr 2012 Nov;24(11):1836-1845. [doi: 10.1017/S1041610212001135] [Medline: 22874772]
8. Tung JY, Rose RV, Gammada E, Lam I, Roy EA, Black SE, et al. Measuring life space in older adults with mild-to-moderate Alzheimer's disease using mobile phone GPS. Gerontology 2014;60(2):154-162. [doi: 10.1159/000355669] [Medline: 24356464]
9. Clare L, Bayer A, Burns A, Corbett A, Jones R, Knapp M, et al. Goal-oriented cognitive rehabilitation in early-stage dementia: study protocol for a multi-centre single-blind randomised controlled trial (GREAT). Trials 2013 May 27;14:152 [FREE Full text] [doi: 10.1186/1745-6215-14-152] [Medline: 23710796]

10. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
11. Wettstein M, Wahl H, Shoval N, Oswald F, Voss E, Seidl U, et al. Out-of-home behavior and cognitive impairment in older adults: findings of the SenTra Project. J Appl Gerontol 2015 Feb;34(1):3-25. [doi: 10.1177/0733464812459373] [Medline: 25548086]
12. Tung J, Semple J, Woo W, Hsu W. Ambulatory assessment of lifestyle factors for alzheimer's disease and related dementias. : AAAI; 2011 Presented at: AAAI Spring Symposium: Computational Physiology; 2011; Stanford, California, USA p. 50-54.
13. Rye HC, Kwon Y, Aung T, Whittington J, High RR, Goulding EH, et al. Mobile phone-based measures of activity, step count, and gait speed: results from a study of older ambulatory adults in a naturalistic setting. JMIR Mhealth Uhealth 2017 Oct 03;5(10):e104 [FREE Full text] [doi: 10.2196/mhealth.5090] [Medline: 28974482]
14. Zylstra B, Netscher G, Jacquemot J, Schaffer M, Shen G, Bowhay A, 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 Dec 15;300:59-67 [FREE Full text] [doi: 10.1016/j.jneumeth.2017.08.034] [Medline: 28865985]
15. Liddle J, Ireland D, McBride SJ, Brauer SG, Hall LM, Ding H, et al. Measuring the lifespace of people with Parkinson's disease using smartphones: proof of principle. JMIR Mhealth Uhealth 2014;2(1):e13 [FREE Full text] [doi: 10.2196/mhealth.2799] [Medline: 25100206]
16. Allet L, Knols RH, Shirato K, de Bruin ED. Wearable systems for monitoring mobility-related activities in chronic disease: a systematic review. Sensors (Basel) 2010;10(10):9026-9052 [FREE Full text] [doi: 10.3390/s101009026] [Medline: 22163393]
17. Canzian L, Musolesi M. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, New York, USA: ACM Press; 2015 Sep 07 Presented at: UbiComp; September 07-11, 2015; Osaka, Japan p. 1293-1304. [doi: 10.1145/2750858.2805845]
18. Boissy P, Brière S, Hamel M, Jog M, Speechley M, Karelis A, EMAP Group. Wireless inertial measurement unit with GPS (WIMU-GPS)--wearable monitoring platform for ecological assessment of lifespace and mobility in aging and disease. Conf Proc IEEE Eng Med Biol Soc 2011;2011:5815-5819. [doi: 10.1109/IEMBS.2011.6091439] [Medline: 22255662]
19. Hirsch JA, Winters M, Clarke P, McKay H. Generating GPS activity spaces that shed light upon the mobility habits of older adults: a descriptive analysis. Int J Health Geogr 2014 Dec 12;13:51 [FREE Full text] [doi: 10.1186/1476-072X-13-51] [Medline: 25495710]
20. Cuttone A, Lehmann S, Larsen JE. Inferring human mobility from sparse low accuracy mobile sensing data. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp '14; September 13-17, 2014; Seattle, WA, United States p. 995-1004. [doi: 10.1145/2638728.2641283]
21. Zheng Y, Zhang L, Xie X, Ma WY. Mining interesting locations and travel sequences from GPS trajectories. In: Proceedings of the 18th international conference on world wide web. New York, NY, USA: ACM; 2009 Presented at: WWW '09; April 20-24, 2009; Madrid, Spain p. 791-800. [doi: 10.1145/1526709.1526816]
22. Andrienko G, Andrienko N, Fuchs G, Olteanu RA, Symanzik J, Ziemlicki C. Extracting Semantics of Individual Places from Movement Data by Analyzing Temporal Patterns of Visits. In: Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place. New York, New York, USA: ACM Press; 2013 Presented at: COMP '13; November 05-08, 2013; Orlando FL, USA p. 9. [doi: 10.1145/2534848.2534851]
23. Ireland D, McBride S, Liddle J, Chenery H. Towards quantifying the impact of Parkinson's disease using GPS and lifespace assessment. 2014 Presented at: 6th International Conference on Biomedical Engineering and Informatics; 16-18 Dec. 2013; Hangzhou, China p. 564. [doi: 10.1109/BMEI.2013.6747004]
24. Wettstein M, Wahl HW, Shoval N, Auslander G, Oswald F, Heinik J. Cognitive status moderates the relationship between out-of-home behavior (OOHB), environmental mastery and affect. Arch Gerontol Geriatr 2014;59(1):113-121. [doi: 10.1016/j.archger.2014.03.015] [Medline: 24768430]
25. Wettstein M, Wahl HW, Shoval N, Auslander G, Oswald F, Heinik J. Identifying mobility types in cognitively heterogeneous older adults based on GPS-tracking: what discriminates best? J Appl Gerontol 2015 Dec;34(8):1001-1027. [doi: 10.1177/0733464813512897] [Medline: 24652916]
26. Hekler EB, Buman MP, Grieco L, Rosenberger M, Winter SJ, Haskell W, et al. Validation of physical activity tracking via android smartphones compared to ActiGraph accelerometer: laboratory-based and free-living validation studies. JMIR Mhealth Uhealth 2015 Apr 15;3(2):e36 [FREE Full text] [doi: 10.2196/mhealth.3505] [Medline: 25881662]
27. Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen MM, Larsen JE, et al. Measuring large-scale social networks with high resolution. PLoS One 2014;9(4):e95978 [FREE Full text] [doi: 10.1371/journal.pone.0095978] [Medline: 24770359]
28. Baker PS, Bodner EV, Allman RM. Measuring life-space mobility in community-dwelling older adults. J Am Geriatr Soc 2003 Nov;51(11):1610-1614. [Medline: 14687391]

29. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. J Phys Act Health 2009 Nov;6(6):790-804. [Medline: 20101923]

Abbreviations

API: application programming interface
DBSCAN: density-based spatial clustering of applications with noise
GPAQ: Global Physical Activity Questionnaire
GPS: global positioning system
MCP: minimum convex polygon
P4: predictive, preventive, personalized, and participatory
RMSD: root mean square deviation

Edited by G Eysenbach; submitted 23.08.18; peer-reviewed by M Karunanithi, J Tung; comments to author 05.11.18; revised version received 30.03.19; accepted 02.04.19; published 30.05.19.

Please cite as:
Thorpe JR, Forchhammer BH, Maier AM
Development of a Sensor-Based Behavioral Monitoring Solution to Support Dementia Care
JMIR Mhealth Uhealth 2019;7(6):e12013
URL: https://mhealth.jmir.org/2019/6/e12013/
doi:10.2196/12013
PMID:31199304

Original Paper

Wearable Finger Pulse Oximetry for Continuous Oxygen Saturation Measurements During Daily Home Routines of Patients With Chronic Obstructive Pulmonary Disease (COPD) Over One Week: Observational Study

Joren Buekers[1,2], MSc; Jan Theunis[1], MSc; Patrick De Boever[1,3], PhD; Anouk W Vaes[4], PhD; Maud Koopman[4], MSc; Eefje VM Janssen[4], MSc; Emiel FM Wouters[4,5], PhD; Martijn A Spruit[4,5,6,7], PhD; Jean-Marie Aerts[2], PhD

[1]Health Unit, Flemish Institute for Technological Research (VITO), Mol, Belgium

[2]Measure, Model & Manage Bioresponses, Department of Biosystems, KU Leuven, Leuven, Belgium

[3]Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium

[4]Department of Research and Education, Centre of Expertise for Chronic Organ Failure (CIRO), Horn, Netherlands

[5]Department of Respiratory Medicine, Maastricht University Medical Centre, Maastricht, Netherlands

[6]Rehabilitation Research Center (REVAL), Biomedical Research Institute (BIOMED), Faculty of Rehabilitation Sciences, Hasselt University, Diepenbeek, Belgium

[7]School of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University Medical Centre, Maastricht, Netherlands

Corresponding Author:
Jean-Marie Aerts, PhD
Measure, Model & Manage Bioresponses
Department of Biosystems
KU Leuven
Kasteelpark Arenberg 30
Leuven, 3000
Belgium
Phone: 32 16321434
Email: jean-marie.aerts@kuleuven.be

Abstract

Background: Chronic obstructive pulmonary disease (COPD) patients can suffer from low blood oxygen concentrations. Peripheral blood oxygen saturation (SpO_2), as assessed by pulse oximetry, is commonly measured during the day using a spot check, or continuously during one or two nights to estimate nocturnal desaturation. Sampling at this frequency may overlook natural fluctuations in SpO_2.

Objective: This study used wearable finger pulse oximeters to continuously measure SpO_2 during daily home routines of COPD patients and assess natural SpO_2 fluctuations.

Methods: A total of 20 COPD patients wore a $WristOx_2$ pulse oximeter for 1 week to collect continuous SpO_2 measurements. A SenseWear Armband simultaneously collected actigraphy measurements to provide contextual information. SpO_2 time series were preprocessed and data quality was assessed afterward. Mean SpO_2, SpO_2 SD, and cumulative time spent with SpO_2 below 90% (CT90) were calculated for every (1) day, (2) day in rest, and (3) night to assess SpO_2 fluctuations.

Results: A high percentage of valid SpO_2 data (daytime: 93.27%; nocturnal: 99.31%) could be obtained during a 7-day monitoring period, except during moderate-to-vigorous physical activity (MVPA) (67.86%). Mean nocturnal SpO_2 (89.9%, SD 3.4) was lower than mean daytime SpO_2 in rest (92.1%, SD 2.9; P<.001). On average, SpO_2 in rest ranged over 10.8% (SD 4.4) within one day. Highly varying CT90 values between different nights led to 50% (10/20) of the included patients changing categories between desaturator and nondesaturator over the course of 1 week.

Conclusions: Continuous SpO_2 measurements with wearable finger pulse oximeters identified significant SpO_2 fluctuations between and within multiple days and nights of patients with COPD. Continuous SpO_2 measurements during daily home routines of patients with COPD generally had high amounts of valid data, except for motion artifacts during MVPA. The identified

fluctuations can have implications for telemonitoring applications that are based on daily SpO_2 spot checks. CT90 values can vary greatly from night to night in patients with a nocturnal mean SpO_2 around 90%, indicating that these patients cannot be consistently categorized as desaturators or nondesaturators. We recommend using wearable sensors for continuous SpO_2 measurements over longer time periods to determine the clinical relevance of the identified SpO_2 fluctuations.

(JMIR Mhealth Uhealth 2019;7(6):e12866) doi:10.2196/12866

KEYWORDS

COPD; oxygen saturation; finger pulse oximeter; wearable sensor; nocturnal desaturation; telemonitoring

Introduction

Chronic obstructive pulmonary disease (COPD) is a highly prevalent lung disease that is characterized by persistent airflow limitation due to a mixture of obstructive bronchiolitis and emphysema [1]. Morbidity and mortality of COPD are high and still increasing [2], leading COPD to become the third-leading cause of death worldwide by 2030 [3]. COPD patients can suffer from low blood oxygen concentrations due to gas exchange abnormalities [1]. Hypoxemia during the night (ie, nocturnal desaturation) is also common in patients with COPD [4-6] due to nocturnal alveolar hypoventilation and ventilation-perfusion mismatching [7]. Hypoxemia can worsen with increasing disease severity [8]. Furthermore, it has been shown that hypoxemia is associated with lower exercise tolerance; decreased quality of life; increased risk for exacerbations, defined as "a sustained worsening of the patient's condition, from the stable state and beyond normal day-to-day variations" [9]; and higher risks of death [8].

The presence of hypoxemia can be assessed by invasive blood gas analyses. A noninvasive method to assess peripheral blood oxygen saturation (SpO_2) is pulse oximetry. Spot check SpO_2 measurements with thresholds of 88%-92% have been suggested for the detection of hypoxemia [1,10,11]. In telemonitoring applications, daily SpO_2 spot checks are used to raise alerts for exacerbations when SpO_2 spot check values drop below predefined SpO_2 thresholds [12]. Nocturnal desaturation is usually defined as having an SpO_2 value below 90% for more than 30% of the time in bed, measured during one [6,13-16] or two nights [5,17-20].

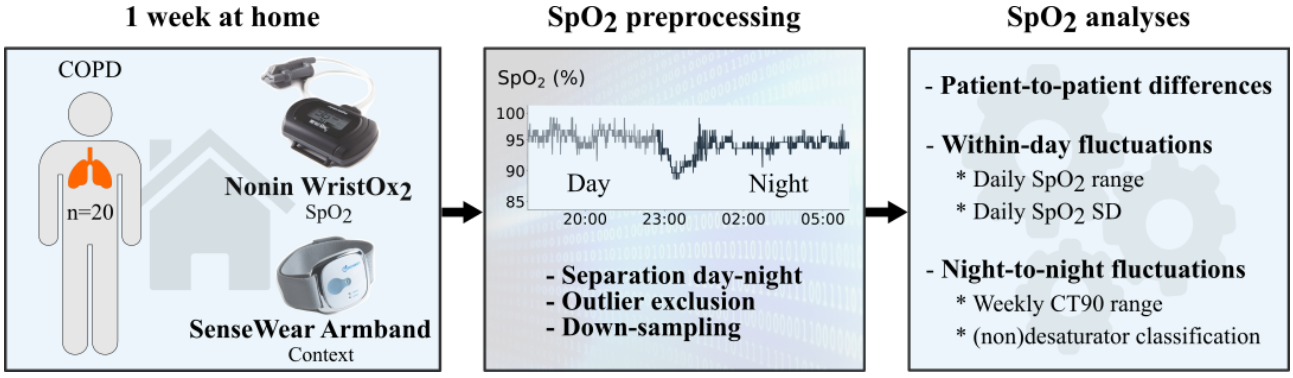
The current SpO_2 monitoring strategies do not take into account natural fluctuations in SpO_2. Sampling SpO_2 with such a low time frequency or during such a short time period may thus lead to classification errors (ie, hypoxemic or nonhypoxemic and nocturnal desaturator or nondesaturator) or false alerts in telemonitoring applications [12]. Wearable finger pulse oximeters provide the possibility to collect SpO_2 data at higher time frequencies and over longer time periods. This makes it possible to assess and account for oxygen saturation fluctuations in patients with COPD. Therefore, the objective of this study was to use wearable finger pulse oximeters to continuously measure SpO_2 during daily home routines of COPD patients and assess natural SpO_2 fluctuations. We hypothesized that significant natural SpO_2 fluctuations are present within and between multiple days and nights, which may lead to classification errors (ie, nocturnal desaturator or nondesaturator) or false alerts in telemonitoring applications.

Methods

Study Design and Participants

Figure 1 provides a general overview of the methods applied in this single-center, 1-week observational study. COPD patients at Global Initiative for Chronic Obstructive Lung Disease stages II-IV (GOLD II-IV) were recruited at the Centre of Expertise for Chronic Organ Failure (CIRO), a COPD treatment center located in Horn, the Netherlands, during a standard baseline assessment prior to pulmonary rehabilitation [21]. The target sample size was set a priori to 20 patients. COPD patients were eligible to enroll in the study based on the following criteria: (1) clinically stable (ie, no exacerbation in the past 4 weeks), (2) no rollator use, and (3) no long-term oxygen therapy. Patients that were interested in participating were called a few days after the baseline assessment to schedule a home visit for the delivery of the wearable sensors. During this visit, the functioning of the sensors and data acquisition protocol were explained, written informed consent was obtained, and a new visit was planned for collection of the sensors at the end of the study period. During the 7-day study period, which took place before the start of pulmonary rehabilitation, two phone calls were made to resolve potential technical difficulties. Demographics, oxygen partial pressure in arterial blood (PaO_2), postbronchodilator pulmonary function data (ie, forced expiratory volume in 1 second [FEV1], forced vital capacity [FVC], and transfer factor for carbon monoxide [TLCO]), 6-minute walking distance (6MWD), and COPD assessment test (CAT) results were collected during the standard baseline assessment at CIRO. The study was approved by the Medical Research Ethics Committees United (MEC-U) (study approval number: NL58079.100.16) in the Netherlands and executed between December 2016 and April 2018.

Figure 1. General overview of the applied methods. COPD: chronic obstructive pulmonary disease; SpO_2: peripheral blood oxygen saturation; CT90: cumulative time spent with SpO_2 below 90%.



Wearable Sensors and Protocol

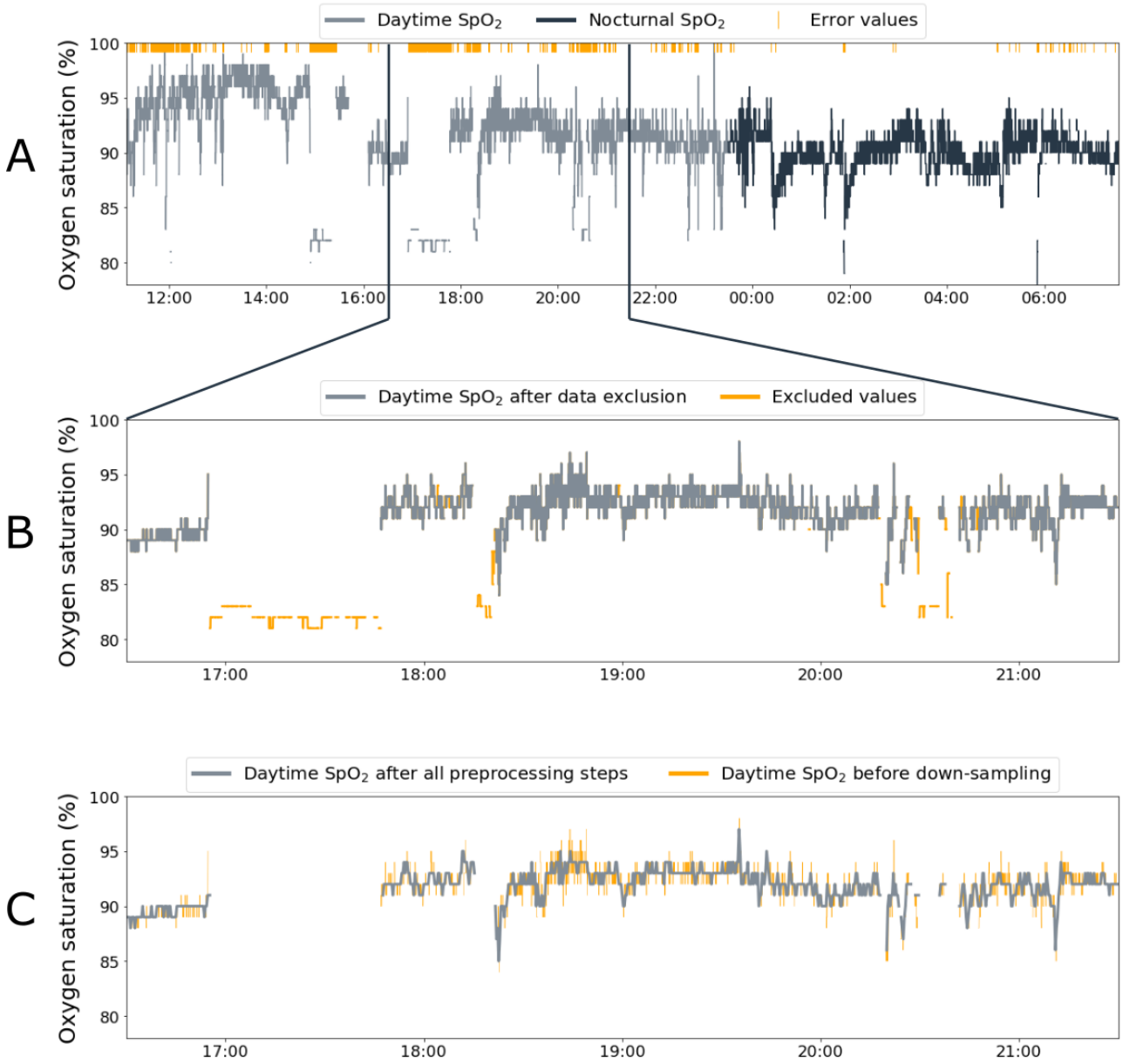
Continuous SpO_2 measurements were performed for 1 week using a wearable finger pulse oximeter: $WristOx_2$ 3150 (Nonin Medical). Nonin oximeters have frequently been used for home monitoring of patients with COPD [12] and the $WristOx_2$ 3150 model complies with the International Organization for Standardization (ISO) standards ISO 10993-1 and ISO 80601-2-61. The manufacturer reports an accuracy of ±2% for SpO_2 measurements [22]. Sampling frequency was fixed at 1 Hz. Only the cumulative measurement time was visible for the participants. SpO_2 values were not shown on the $WristOx_2$ display to prevent patients from changing their behavior when deviating SpO_2 values would occur. Every participant received three $WristOx_2$ devices to deal with the limited battery life (ie, 48 hours of continuous measurements with one $WristOx_2$). Participants were instructed to wear the $WristOx_2$ on the index finger of their nondominant hand every night and as much as possible during the day, depending on their daily routines and comfort of wearing the finger clip. Raw data were stored on the internal memory of the $WristOx_2$ and downloaded at the end of the week using nVISION software, version 6.4.0.10 (Nonin Medical).

Simultaneously, actigraphy was performed with the SenseWear Armband (BodyMedia) for obtaining contextual information about physical activity levels and when the participants were lying down and/or asleep. The SenseWear Armband is a multisensory triaxial accelerometer, combining accelerometry with measurements of heat flux, galvanic skin response, and skin temperature. Based on these measurements, the armband provides information about, for example, energy expenditure (EE), expressed as metabolic equivalent of task (MET), or steps taken, while also indicating when the wearer is lying down and sleeping, at a standard sampling time of 1 sample per minute. Measurements of MET were used for classifying physical activity levels, whereas the indications of lying down and sleeping were used to separate daytime measurements from nocturnal measurements, as further described below. The SenseWear Armband has been shown to be accurate for measurements of both physical activity [23,24], except when using a rollator [24], and sleep estimations [25]. Patients were asked to wear the armband on the left upper arm, except when there was contact with water (eg, when taking a shower). Battery life of one SenseWear Armband was sufficient for continuous 24-hour measurements with a 1-minute sampling time during the whole week. Data were stored on the internal memory and downloaded at the end of the week using the BodyMedia SenseWear 8.1 software (BodyMedia).

Data Preprocessing

Figure 2 visualizes the preprocessing steps. The SenseWear Armband indications about lying down and sleep were used to determine the time of going to bed in the evening and the time of getting out of bed in the morning. These time stamps were used to divide SpO_2 data into nocturnal and daytime data (see Figure 2A). Only full-night nocturnal SpO_2 measurements were retained for further analyses and daytime SpO_2 measurements were retained if there was at least one hour of SpO_2 measurements during that day. One hour of daytime measurements was considered sufficient to examine whether significant fluctuations occurred in SpO_2 values during the day. Previous studies examining intraday fluctuations only examined measurements of one hour or less [26-29]. No days or nights had to be excluded due to lacking SenseWear Armband measurements. Days with at least one hour of measurements in both the afternoon and evening were used for the comparison between afternoon (ie, 13:00-18:00) and evening (ie, 18:00-going to bed) SpO_2 values. No comparison was performed between SpO_2 values in the morning (ie, before 13:00) and the afternoon and evening, because SpO_2 measurements were often not performed before 13:00 (see Multimedia Appendix 1).

Figure 2. Visualization of the different preprocessing steps. Panel A shows all original data, containing error values, that are divided into daytime and nocturnal data. Panel B zooms in on the effect of data exclusion on a specific part of daytime peripheral blood oxygen saturation (SpO_2) data. Panel C zooms in on the effect of down-sampling and interpolating on the same part of daytime SpO_2 data.



The raw SpO_2 data contained error values (ie, the number 500 was provided in the raw data file) when measurements were considered invalid by the algorithms of the $WristOx_2$ manufacturer (ie, orange dashes in Figure 2A). However, close inspection of the time series indicated that invalid data (eg, sudden low values) surrounded these error values. A data-cleaning algorithm was developed to exclude these invalid SpO_2 values as follows. First, small blocks of data (ie, less than 20 samples) between error values were excluded. Second, bigger blocks of data (ie, between 20 and 100 samples) between error values were excluded, only when the mean SpO_2 value of this data block was deviating more than 6% from the mean SpO_2 value of the full day or night under consideration: both steps are shown in Figure 2B. Thorough visual and raw data examination of both valid and invalid data led to the choice of this 6% threshold for excluding invalid blocks of data (eg, Figure 2B: around 17:00-18:00). Third, remaining outliers were excluded by down-sampling the data (see Figure 2C). Autocorrelation analyses indicated that a sampling time of 20 seconds was appropriate (see Multimedia Appendix 2) and all data were thus down-sampled by taking the median SpO_2 value of each consecutive 20-second block. By taking the median value, the effect of outliers was excluded. As a last step, small gaps of SpO_2 data (ie, 3 or fewer missing samples) were filled using linear interpolation (see Figure 2C). After application of the data-cleaning algorithm, all data were visually checked to ensure valid SpO_2 values were retained, while invalid values were removed.

Data Analyses

The SenseWear Armband indications about lying down and sleep were used to calculate the total night sleeping time (TNST) (ie, sum of all minutes indicated as sleep), wake time after sleep

onset (WASO) (ie, sum of all minutes spent awake during the time in bed, after the first onset of sleep), and sleep efficiency (Seff) (ie, the ratio between TNST and time in bed) for every separate night. Weekly averages of TNST, WASO, and Seff were calculated to describe the sleep quality of the included patients.

Daytime SpO_2 data were divided into daytime data during rest (EE≤1.5 MET, while the patient was still awake), during low-intensity physical activities (LIPAs) (1.5 MET<EE≤3 MET), and during moderate-to-vigorous intensity physical activities (MVPAs) (EE>3 MET). Data quality of the continuous SpO_2 measurements was assessed for both nocturnal and daytime data based on the amount of valid data (ie, excluding error values and cleaned values). Furthermore, the effect of physical activity on the amount of valid data was examined by comparing data quality during MVPA with data quality during rest and LIPA (EE≤3 MET).

Mean SpO_2, SpO_2 SD, and cumulative time spent with SpO_2 below 90% (CT90) were calculated for every separate (1) day, (2) day in rest, and (3) night. Hereafter, the weekly average and weekly range (ie, the difference between the maximum and minimum value over the different days or nights of the same patient) of these features were calculated for every patient.

Intraday and intranight fluctuations were quantified as the weekly average of the standard deviation of the SpO_2 measurements. The range of SpO_2 values during the day in rest, calculated as the difference between the maximum and minimum SpO_2 values of that day in rest, was determined to indicate how much spot-check values in telemonitoring applications could differ depending on the moment of the measurement. Only daytime SpO_2 in rest was considered because only at these moments could spot checks have been performed in a telemonitoring application. In addition, the difference between mean SpO_2 values in rest in the afternoon (ie, between 13:00 and 18:00) and in the evening (ie, between 18:00 and going to bed) was calculated for every day separately to examine differences in SpO_2 baseline levels during the day.

Night-to-night and day-to-day SpO_2 fluctuations were quantified as the weekly ranges of the three features (ie, mean SpO_2, SpO_2 SD, and CT90). Furthermore, we examined how many of the included patients changed category between nocturnal desaturator and nondesaturator.

Statistical Analysis

Patient characteristics, weekly averages, and weekly ranges were summarized for all patients as mean and SD. Paired-sample *t* tests were used to test for differences between weekly averages of nocturnal and daytime SpO_2 in rest. Pearson correlations assessed the relationship between mean SpO_2 and patient characteristics, intraday SpO_2 fluctuations, or intranight SpO_2 fluctuations. A P value of <.05 was considered statistically significant. All analyses were carried out in Jupyter Notebooks (Project Jupyter) [30] using the Python 3.5 programming language (Python Software Foundation) [31].

Results

Participants

A total of 21 out of 41 patients that were approached accepted study participation. One patient suffered from an exacerbation before the start of the 7-day study period and was excluded. This finally led to the inclusion of 20 patients (14 males, 70%; 6 females, 30%) with moderate (8/20, 40%), severe (10/20, 50%), or very severe (2/20, 10%) COPD. General demographics, postbronchodilator lung function, resting arterial blood gases, 6MWD, CAT score, and sleep-quality characteristics are summarized in Table 1.

Table 1. Characteristics of the 20 included patients with moderate-to-very severe COPD[a].

Characteristics	Mean (SD)
Age (years)	63 (8)
Body mass index (kg/m²)	26 (4)
Forced expiratory volume in 1 second (L)	1.4 (0.5)
Forced expiratory volume in 1 second (% predicted)	48 (15)
Forced vital capacity (L)	3.8 (1.1)
Forced vital capacity (% predicted)	97 (20)
Transfer factor for carbon monoxide (% predicted)	50 (16)
Partial pressure of oxygen (kPa)	8.7 (1.7)
Partial pressure of carbon dioxide (kPa)	5.2 (0.6)
6-minute walking distance (m)	411 (59)
COPD assessment test score	21 (5)
Total night sleeping time (min)	411 (72)
Wake time after sleep onset (min)	74 (36)
Sleep efficiency (%)	83 (8)

[a]COPD: chronic obstructive pulmonary disease.

Continuous SpO_2 Measurements, Preprocessing, and Data Quality

An overview of the amount of SpO_2 measurements that were included for further analyses is provided in Table 2 and Multimedia Appendix 1. A total of 2 days and 2 nights of measurements from patient 4 were missing due to battery issues. A total of 3 days and 3 nights of measurements from patient 11 were excluded because correct time stamps were missing, due to the patient unintentionally resetting the time indication by taking out the batteries. The last day and night of patient 11's measurements could not be analyzed because no clear distinction could be made between daytime and nocturnal data. Still, 3 days and 3 nights of patient 11's measurements were used for further analyses.

The SpO_2 dataset contained 1.83% error values and an additional 1.49% were excluded during the first two steps of the data-cleaning algorithm (see Figure 3), resulting in 96.68% of valid data. Nocturnal data had 99.31% of valid data (0.45% error data and 0.24% cleaned) compared to 93.27% of valid data during the day (3.62% error data and 3.11% cleaned). This was similar when only considering daytime data during rest and LIPA (EE≤3 MET; 93.90% of valid data: 3.25% error data and 2.85% cleaned). However, during MVPA (EE>3 MET), the amount of valid data decreased to 67.86% (18.34% error data and 13.80% cleaned).

Table 2. Summary of the amount of peripheral blood oxygen saturation (SpO_2) measurements that were included for further analyses[a].

Measurements included for further analyses	Mean (SD)
Days per patient	5.3 (1.8)
Nights per patient	5.9 (1.2)
Hours per day	
Total	7.8 (3.9)
During rest (EE[b]≤1.5 MET[c])	6.1 (3.1)
During LIPA[d] (1.5 MET<EE≤3 MET)	1.5 (2.7)
During MVPA[e] (EE>3 MET)	0.2 (0.2)
Hours per night	8.0 (0.9)
Data samples per patient	
Total	296,533 (95,580)
During the day	149,665 (88,909)
During the night	169,318 (35,488)

[a]All 20 patients performed nocturnal measurements and 17 patients performed daytime measurements. Therefore, averages were calculated with 17 patients for daytime measurements and 20 patients for nocturnal and total measurements.

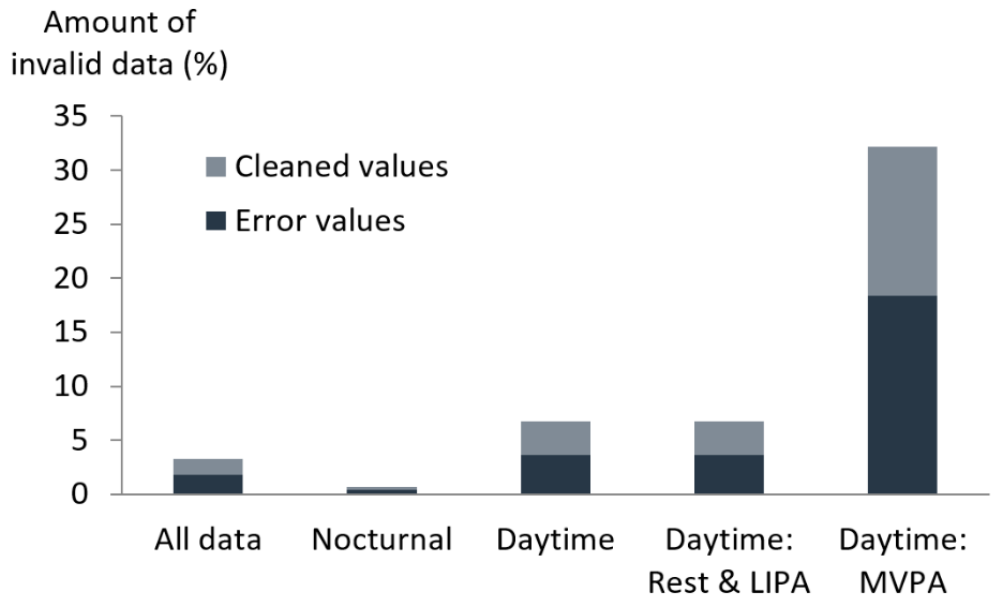
[b]EE: energy expenditure.

[c]MET: metabolic equivalent of task.

[d]LIPA: low-intensity physical activity.

[e]MVPA: moderate-to-vigorous physical activity.

Figure 3. The amount of invalid data of continuous peripheral blood oxygen saturation (SpO_2) measurements for different activities. LIPA: low-intensity physical activity; MVPA: moderate-to-vigorous physical activity.



SpO_2 Features and Fluctuations

The weekly averages of mean SpO_2 were 89.9% (SD 3.4), 91.9% (SD 3.1), and 92.1% (SD 2.9) for nocturnal (n=20), daytime (n=17), and daytime-in-rest (n=17) measurements, respectively. Figure 4 shows that mean nocturnal SpO_2 was lower than mean daytime SpO_2 in rest (P<.001). Weekly averages of mean SpO_2 were between 84.7% and 96.0% for daytime SpO_2 in rest and between 80.3% and 94.3% for nocturnal SpO_2 (see Figure 4). The weekly averages of CT90 were 35% (SD 34), 20% (SD 29), and 18% (SD 29) for nocturnal, daytime, and daytime-in-rest measurements, respectively. No significant correlations were found between patient characteristics and weekly averages of the SpO_2 features.

Intraday and intranight SpO_2 fluctuations were quantified as the weekly average of SpO_2 SD for nocturnal (1.6%, SD 0.6), daytime (1.8%, SD 0.7), and daytime-in-rest (1.6%, SD 0.6) measurements. Nocturnal, daytime, and daytime-in-rest SpO_2 SD values were inversely correlated with nocturnal, daytime, and daytime-in-rest mean SpO_2 values (all: P<.001, R^2>0.4), respectively. On average, SpO_2 in rest values ranged over 10.8% (SD 4.4) within one day. There was a significant difference between mean SpO_2 in the afternoon and in the evening (P<.001) (see Multimedia Appendix 3), with an average difference of 0.8% (SD 0.7) for those days where patients performed at least one hour of measurements during both the afternoon and evening. For this comparison, a total of 46 days from 12 different patients were used (median of 4 days per patient).

Night-to-night and day-to-day SpO_2 fluctuations were quantified as the weekly ranges of mean SpO_2 (2.0%, SD 1.1; 1.9%, SD 1.5; 1.9%, SD 1.5) and CT90 (28%, SD 25; 19%, SD 20; 18%, SD 21) for nocturnal, daytime, and daytime-in-rest measurements, respectively. A more detailed analysis of the night-to-night changes in CT90 is shown in Figure 5. When considering the definition of a nocturnal desaturator, three types of patients can be distinguished: consistent desaturators (5/20, 25%), consistent nondesaturators (5/20, 25%), and occasional desaturators that changed category over the 7 nights (10/20, 50%). A total of 6 occasional desaturators (6/10, 60%) were desaturators or nondesaturators in the first two nights, but changed category when SpO_2 measurements of another night were considered. The weekly average of mean SpO_2 for occasional desaturators was between 89.5% and 91.9%, compared to 80.3%-88.8% and 91.8%-94.3% for consistent desaturators and consistent nondesaturators, respectively.

Figure 4. Mean nocturnal peripheral blood oxygen saturation (SpO_2) compared to mean daytime SpO_2 in rest. The dots indicate weekly averages of mean SpO_2 for every patient, lines indicate the standard deviation of the mean SpO_2 values over different days and nights, and the orange line is the line of equality.

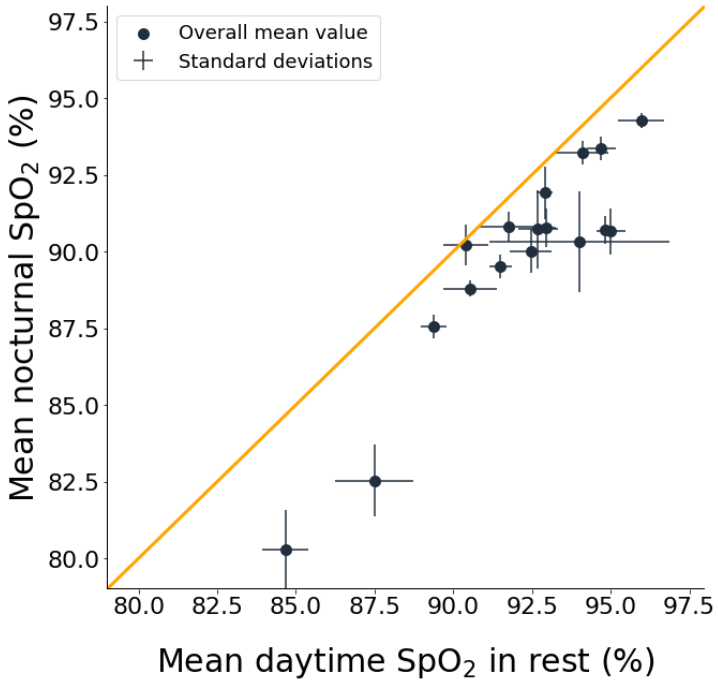
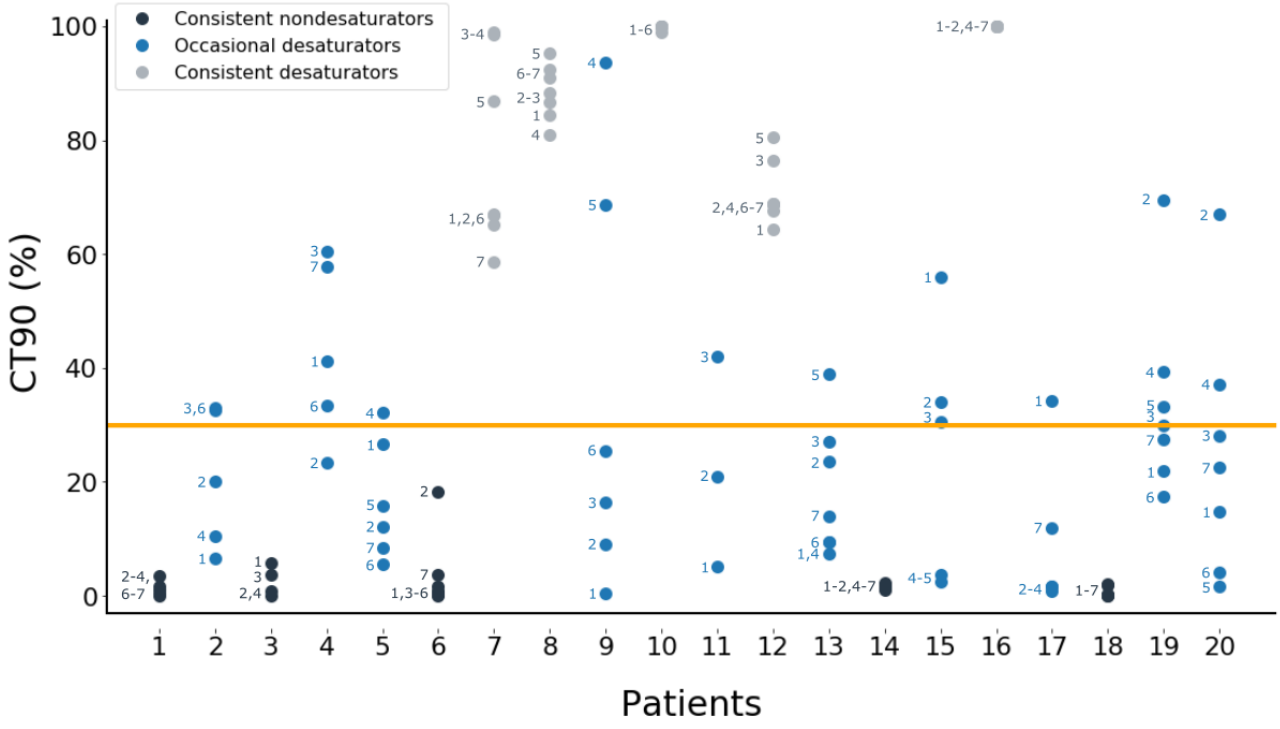


Figure 5. Cumulative time spent with peripheral blood oxygen saturation (SpO_2) below 90% (CT90) values for every night of every patient showed that 50% (10/20) of the included patients changed category between desaturator and nondesaturator. The number next to each dot indicates the corresponding night of the measuring week and the orange line indicates the threshold of CT90 (30%) that divides nights with and without desaturation. A missing night number indicates that no data was available for that night.



Discussion

Continuous SpO_2 measurements with wearable pulse oximeters identified significant SpO_2 fluctuations between and within multiple days and nights of patients with COPD. COPD patients are known to be a heterogeneous group [32,33], which leads to large differences in mean SpO_2 values between patients. Day-to-day SpO_2 fluctuations have not yet received a lot of attention in scientific literature and night-to-night fluctuations have not yet been studied over more than two nights in patients

with COPD. Some studies examined intraday fluctuations, however, these studies only examined short-term SpO_2 fluctuations (ie, measurements of one hour or less) in healthy individuals [29] or infants [26-28]. Our results showed that significant SpO_2 fluctuations occurred between and within multiple days and nights for the included patients. This is often not taken into account in telemonitoring applications [12]. When using the current definition of nocturnal desaturation, 50% (10/20) of the included patients changed category between desaturator and nondesaturator over the course of 1 week.

Continuous SpO_2 measurements with high amounts of valid data could be obtained during a 7-day monitoring period, except during MVPA (EE>3 MET). Nocturnal measurements were well-perceived by the patients, leading to the high compliance. These measurements only had a small amount of data that was considered invalid (0.69%). This indicates that it is feasible to perform multi-night continuous SpO_2 measurements. Most patients also performed daytime measurements during rest and LIPA (ie, 17 patients measured, on average, 7.6 hours per day). These had a limited amount of invalid data (6.10%). However, very few measurements were performed during MVPA (ie, 0.2 hours per day) and almost one-third of these measurements produced invalid data due to motion artifacts. Therefore, it might not be recommended to analyze continuous SpO_2 measurements during periods of MVPA.

Two types of preprocessing steps for continuous SpO_2 data have been reported: (1) removing deviating samples or (2) down-sampling the data. Deviating samples have been removed visually [29,34] by using an unspecified data-cleaning algorithm [35] or by removing samples with a sample-to-sample deviation of more than 4% [36] or 8% [37]. The latter method, however, does not remove artifacts that last over longer periods of time (eg, see Figure 2B: around 17:00-18:00). Therefore, artifacts in this study were removed by removing blocks of deviating values instead of separate samples. After sample removal, Morillo et al excluded remaining artifacts by down-sampling the data from 8 Hz to 1 Hz [37]. However, autocorrelation analyses in our study showed that a sampling time of 1 sample per second still led to oversampling of the data (see Multimedia Appendix 2). SpO_2 could be down-sampled to 1 sample per 20 seconds for improved artifact removal, without losing information about the SpO_2 dynamics.

Mean nocturnal SpO_2 was lower than mean daytime SpO_2 in rest. By taking into account the SpO_2 fluctuations over different days and nights, the results of this study generalize the findings of Soguel Schenkel et al, who only performed SpO_2 measurements during a single day and night that were up to five days apart [38]. Other studies used *awake* measurements of SpO_2 to predict nocturnal mean SpO_2 or CT90. They used the first 15-30 minutes of the nocturnal measurements to calculate the mean *awake* SpO_2 values [15,39-41]. For these studies, it should be noted that slightly lower SpO_2 values were observed in the evening compared to the afternoon, on average a 0.8% difference, and *awake* SpO_2 values depend on the time of day when the measurements are performed. It is unlikely that this difference is solely due to the specified oximeter accuracy of ±2%, as no consistent bias has been reported for the $WristOx_2$ when measuring over a longer time period.

The observed patient-to-patient, day-to-day, and intraday differences can have implications for telemonitoring applications that are based on daily SpO_2 spot checks [12]. In several telemonitoring setups, alerts were raised when daily spot checks of SpO_2 dropped below a generic threshold value (eg, 90% for all patients) [12]. However, the high patient-to-patient differences point out the shortcomings of these generic thresholds. Other telemonitoring setups used personalized thresholds, but these were still fixed on one specific threshold value for every patient [12]. For these setups, alerts can be merely a consequence of natural day-to-day and intraday fluctuations, instead of being triggered by the onset of an exacerbation. A recent paper better dealt with the day-to-day fluctuations by suggesting a day-to-day decrease of more than 4% to alert for exacerbations, however, a more thorough examination of this method is needed [42]. Intraday SpO_2 values during rest ranged over 10.8% (SD 4.4) within one day. Moreover, statistical interpretation of weekly average of SpO_2 SD (1.6%, SD 0.6), which is a measure of intraday fluctuations, indicates that during 5% of the day in rest, SpO_2 fluctuates more than 3.2% (ie, 2 times SpO_2 SD) beyond the mean daytime-in-rest value. In comparison, previous studies only reported a 1%-2% decrease in SpO_2 spot checks around exacerbation onset compared to stable periods [43-45]. These natural intraday fluctuations can thus easily result in false alerts for exacerbations. Patients with lower SpO_2 values might experience a higher number of false alerts as, similar to healthy individuals [29], intraday fluctuations (ie, SpO_2 SD) increased with decreasing mean SpO_2. Altogether, personalized alerts based on intelligent algorithms will be necessary to cope with all of these natural fluctuations in daytime SpO_2 in rest. Preferably, these alerts should be based on continuous measurements over longer time periods, in contrast to the currently used daily spot checks, to account for the identified SpO_2 fluctuations and to exclude the potential effect of the ±2% oximeter accuracy. In addition, the slightly lower SpO_2 values in the evening compared to the afternoon show that the daily measurements should always be performed at the same moment of the day. A clearly defined measurement protocol, which is often not specified [12], can thus further attempt to limit the effect of natural SpO_2 fluctuations in these telemonitoring applications.

The night-to-night differences of nocturnal mean SpO_2 (ie, average weekly range of 2.0% over the different nights) led to highly varying CT90 values over the different nights of the measurement week (ie, average range of 28%). These high variations in CT90 resulted in 50% (10/20) of the included patients changing category between desaturator and nondesaturator, due to the fact that these occasional desaturators all had a mean SpO_2 value around the threshold of 90%. A similar finding has been reported by Lewis et al, who concluded that 35% of the included patients changed category over two consecutive nights of measurements [20]. Later studies then tackled this problem by performing measurements over two

nights, categorizing a patient as a desaturator if desaturation occurred in at least one of both nights [5,17-19]. However, the results of this study indicate that even two nights are insufficient to make a consistent separation between desaturators and nondesaturators, as 6 out of 10 occasional desaturators (60%) were desaturators or nondesaturators in the first two nights and only changed category afterward. Our results suggest that it might be impossible to consistently categorize COPD patients with a mean nocturnal SpO_2 value around 90% as desaturator or nondesaturator. Based on measurements over multiple nights, these patients could thus be referred to as occasional desaturators. Further research is needed to assess the clinical relevance of identifying these three different nocturnal desaturation profiles.

Some limitations should be taken into consideration when interpreting the results of this study. The main limitation, as is often the case in similar studies, is the small sample size. Consequently, no comparison could be made between mild, moderate, and severe hypoxemic patients. Nevertheless, the increasing SpO_2 SD with decreasing mean SpO_2 suggests an increase in SpO_2 fluctuations for more hypoxemic patients. No control group was included because this study aimed to perform continuous SpO_2 measurements in COPD patients for identification of SpO_2 fluctuations that could affect SpO_2 applications, rather than comparing SpO_2 between COPD patients and healthy controls. In addition, daily home routines can greatly differ between patients and healthy controls [46], impeding a proper comparison. As shown in Multimedia Appendix 1, daytime measurements were only seldom performed during full days due to the impracticalities of wearing the finger clip during activities that require manual actions (eg, during morning routines). This, in combination with the high amount of invalid data during MVPA, suggests that continuous SpO_2 measurements with a finger clip will have more profound limitations in a population that is more physically active than the target population of this study. Reliable wearable oximeters that do not require a finger clip could thus increase compliance. Moreover, the limited battery life and the inability of real-time data transmission of the used oximeter can further complicate the integration of continuous SpO_2 measurements into practice. Technological advances are thus needed to allow for long-term, continuous monitoring of SpO_2. This study, however, mainly intended to show the potential of prolonged continuous measurements to identify SpO_2 fluctuations. Therefore, a certified wearable pulse oximeter with finger clip was preferred over more user-friendly, watch-type oximeters that have not yet been proven to be accurate. The resulting, more fragmented, daytime measurements were sufficient to identify large natural fluctuations occurring within one day (ie, SpO_2 in rest ranged over 10.8% [SD 4.4] within one day), confirming the a priori posed hypothesis. The limited battery life was addressed by providing the patients with multiple sensors to cover the 7-day monitoring period.

This study was the first to use wearable finger pulse oximeters for prolonged continuous SpO_2 measurements in COPD patients, as opposed to only performing spot checks or continuous measurements during one or two nights. These measurements showed that spot checks or one- or two-night measurements should be interpreted with caution, as the conclusions based on these measurements might change depending on the moment of the measurement. Measurements were performed at home during daily life routines of COPD patients, which provides a more natural SpO_2 profile compared to supervised measurements. By adding actigraphy measurements, the necessary contextual information could be gathered for more accurate analyses of the continuous SpO_2 measurements.

In conclusion, continuous SpO_2 measurements with wearable pulse oximeters identified significant SpO_2 fluctuations between and within multiple days and nights of patients with COPD. Continuous SpO_2 measurements during the daily home routine of patients with COPD generally had high amounts of valid data, except for motion artifacts during MVPA. The continuous measurements showed that mean nocturnal SpO_2 was lower than mean daytime SpO_2 in rest, and significant SpO_2 fluctuations occurred between and within multiple days and nights. The large fluctuations of daytime SpO_2 in rest indicate that clear measurement protocols and personalized alerts, based on intelligent algorithms, will be needed to increase the performance of telemonitoring applications that make use of daily SpO_2 spot checks. Lastly, it was shown that CT90 values can vary greatly from night to night in patients with a nocturnal mean SpO_2 around 90%, indicating that these patients cannot be consistently categorized as desaturators or nondesaturators. We recommend using wearable sensors for performing continuous SpO_2 measurements over longer time periods to determine the clinical relevance of the identified SpO_2 fluctuations.

Acknowledgments

This research is part of a PhD research project funded by the Flemish Institute for Technological Research (VITO), Mol, Belgium.

Conflicts of Interest

MAS discloses receipt of renumeration for consultancy and/or lectures from Boehringer Ingelheim, GlaxoSmithKline, and AstraZeneca outside the scope of this work. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript, apart from those disclosed.

Multimedia Appendix 1

Visual overview of the continuous peripheral blood oxygen saturation (SpO2) measurements that were included for analyses.

[PDF File (Adobe PDF File), 32KB - mhealth_v7i5e12866_app1.pdf]

Multimedia Appendix 2

Autocorrelation analyses.

[PDF File (Adobe PDF File), 95KB - mhealth_v7i6e12866_app2.pdf]

Multimedia Appendix 3

Differences between mean peripheral blood oxygen saturation (SpO2) in the afternoon and in the evening. Grey dots indicate the differences within the same day for the same patient and orange dots indicate the average of these differences for every patient. The number of grey dots per patient indicate the amount of days with measurements in both the afternoon and evening.

[PDF File (Adobe PDF File), 22KB - mhealth_v7i6e12866_app3.pdf]

References

1. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: 2019 Report. Fontana, WI: Global Initiative for Chronic Obstructive Lung Disease (GOLD); 2019. URL: https://goldcopd.org/wp-content/uploads/2018/11/GOLD-2019-v1.7-FINAL-14Nov2018-WMS.pdf [accessed 2019-04-16] [WebCite Cache ID 77g4jvX6Y]
2. López-Campos JL, Tan W, Soriano JB. Global burden of COPD. Respirology 2016 Jan;21(1):14-23 [FREE Full text] [doi: 10.1111/resp.12660] [Medline: 26494423]
3. World Health Organization. 2019. Chronic obstructive pulmonary disease (COPD) URL: http://www.who.int/respiratory/copd/en/ [accessed 2018-11-13] [WebCite Cache ID 73tsne72z]
4. Casanova C, Hernández MC, Sánchez A, García-Talavera I, de Torres JP, Abreu J, et al. Twenty-four-hour ambulatory oximetry monitoring in COPD patients with moderate hypoxemia. Respir Care 2006 Dec;51(12):1416-1423 [FREE Full text] [Medline: 17134522]
5. Lacasse Y, Sériès F, Vujovic-Zotovic N, Goldstein R, Bourbeau J, Lecours R, et al. Evaluating nocturnal oxygen desaturation in COPD--revised. Respir Med 2011 Sep;105(9):1331-1337 [FREE Full text] [doi: 10.1016/j.rmed.2011.04.003] [Medline: 21561753]
6. Chaouat A, Weitzenblum E, Kessler R, Charpentier C, Ehrhart M, Levi-Valensi P, et al. Sleep-related O2 desaturation and daytime pulmonary haemodynamics in COPD patients with mild hypoxaemia. Eur Respir J 1997 Aug;10(8):1730-1735 [FREE Full text] [Medline: 9272911]
7. Weitzenblum E, Chaouat A. Sleep and chronic obstructive pulmonary disease. Sleep Med Rev 2004 Aug;8(4):281-294. [doi: 10.1016/j.smrv.2004.03.006] [Medline: 15233956]
8. Kent BD, Mitchell PD, McNicholas WT. Hypoxemia in patients with COPD: Cause, effects, and disease progression. Int J Chron Obstruct Pulmon Dis 2011;6:199-208 [FREE Full text] [doi: 10.2147/COPD.S10611] [Medline: 21660297]
9. Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. Chest 2000 May;117(5 Suppl 2):398S-401S. [doi: 10.1378/chest.117.5_suppl_2.398s] [Medline: 10843984]
10. Kelly AM, McAlpine R, Kyle E. How accurate are pulse oximeters in patients with acute exacerbations of chronic obstructive airways disease? Respir Med 2001 May;95(5):336-340 [FREE Full text] [doi: 10.1053/rmed.2001.1046] [Medline: 11392573]
11. Garcia-Gutierrez S, Unzurrunzaga A, Arostegui I, Quintana JM, Pulido E, Gallardo MS, IRYSS-COPD group. The use of pulse oximetry to determine hypoxemia in acute exacerbations of COPD. COPD 2015;12(6):613-620. [doi: 10.3109/15412555.2014.995291] [Medline: 25774875]
12. Buekers J, De Boever P, Vaes AW, Aerts J, Wouters EFM, Spruit MA, et al. Oxygen saturation measurements in telemonitoring of patients with COPD: A systematic review. Expert Rev Respir Med 2018 Dec;12(2):113-123. [doi: 10.1080/17476348.2018.1417842] [Medline: 29241369]
13. Toraldo DM, Nicolardi G, De Nuccio F, Lorenzo R, Ambrosino N. Pattern of variables describing desaturator COPD patients, as revealed by cluster analysis. Chest 2005 Dec;128(6):3828-3837. [doi: 10.1378/chest.128.6.3828] [Medline: 16354851]
14. Scott AS, Baltzman MA, Chan R, Wolkove N. Oxygen desaturation during a 6 min walk test is a sign of nocturnal hypoxemia. Can Respir J 2011;18(6):333-337 [FREE Full text] [doi: 10.1155/2011/242636] [Medline: 22187689]
15. Levi-Valensi P, Weitzenblum E, Rida Z, Aubry P, Braghiroli A, Donner C, et al. Sleep-related oxygen desaturation and daytime pulmonary haemodynamics in COPD patients. Eur Respir J 1992 Mar;5(3):301-307. [Medline: 1572442]
16. Iliaz S, Cagatay T, Bingol Z, Okumus G, Iliaz R, Kuran G, et al. Does the 6-minute walk test predict nocturnal oxygen desaturation in patients with moderate to severe COPD? Chron Respir Dis 2015 Feb;12(1):61-68. [doi: 10.1177/1479972314562406] [Medline: 25480424]

17. Nisbet M, Eaton T, Lewis C, Fergusson W, Kolbe J. Overnight prescription of oxygen in long term oxygen therapy: Time to reconsider the guidelines? Thorax 2006 Sep;61(9):779-782 [FREE Full text] [doi: 10.1136/thx.2005.056119] [Medline: 16769716]
18. Lacasse Y, Bernard S, Sériès F, Nguyen VH, Bourbeau J, Aaron S, International Nocturnal Oxygen (INOX) Research Group. Multi-center, randomized, placebo-controlled trial of nocturnal oxygen therapy in chronic obstructive pulmonary disease: A study protocol for the INOX trial. BMC Pulm Med 2017 Dec 09;17(1):8. [doi: 10.1186/s12890-016-0343-9] [Medline: 28069009]
19. Lewis CA, Fergusson W, Eaton T, Zeng I, Kolbe J. Isolated nocturnal desaturation in COPD: Prevalence and impact on quality of life and sleep. Thorax 2009 Feb;64(2):133-138. [doi: 10.1136/thx.2007.088930] [Medline: 18390630]
20. Lewis CA, Eaton TE, Fergusson W, Whyte KF, Garrett JE, Kolbe J. Home overnight pulse oximetry in patients with COPD: More than one recording may be needed. Chest 2003 Apr;123(4):1127-1133. [doi: 10.1378/chest.123.4.1127] [Medline: 12684303]
21. Spruit MA, Vanderhoven-Augustin I, Janssen PP, Wouters EFM. Integration of pulmonary rehabilitation in COPD. Lancet 2008 Jan 05;371(9606):12-13. [doi: 10.1016/S0140-6736(08)60048-3] [Medline: 18177760]
22. Nonin Medical. WristOx2 ® Model 3150 USB: Wrist-worn pulse oximeter URL: https://www.nonin.com/wp-content/uploads/2018/09/Spec-Sheet-WristOx2-Model-3150-USB.pdf [accessed 2019-04-04] [WebCite Cache ID 77Ni6Sp3k]
23. Van Remoortel H, Raste Y, Louvaris Z, Giavedoni S, Burtin C, Langer D, PROactive consortium. Validity of six activity monitors in chronic obstructive pulmonary disease: A comparison with indirect calorimetry. PLoS One 2012;7(6):e39198 [FREE Full text] [doi: 10.1371/journal.pone.0039198] [Medline: 22745715]
24. Hill K, Dolmage TE, Woon L, Goldstein R, Brooks D. Measurement properties of the SenseWear armband in adults with chronic obstructive pulmonary disease. Thorax 2010 Jun;65(6):486-491. [doi: 10.1136/thx.2009.128702] [Medline: 20522844]
25. Sharif MM, Bahammam AS. Sleep estimation using BodyMedia's SenseWear™ armband in patients with obstructive sleep apnea. Ann Thorac Med 2013 Jan;8(1):53-57 [FREE Full text] [doi: 10.4103/1817-1737.105720] [Medline: 23440703]
26. de Jesus JA, Tristao RM, Storm H, da Rocha AF, Campos D. Heart rate, oxygen saturation, and skin conductance: A comparison study of acute pain in Brazilian newborns. Conf Proc IEEE Eng Med Biol Soc 2011;2011:1875-1879. [doi: 10.1109/IEMBS.2011.6090532] [Medline: 22254696]
27. de Oliveira MV, de Jesus JA, Tristao RM. Psychophysical parameters of a multidimensional pain scale in newborns. Physiol Meas 2012 Jan;33(1):39-49. [doi: 10.1088/0967-3334/33/1/39] [Medline: 22155897]
28. Garde A, Zhou G, Raihana S, Dunsmuir D, Karlen W, Dekhordi P, et al. Respiratory rate and pulse oximetry derived information as predictors of hospital admission in young children in Bangladesh: A prospective observational study. BMJ Open 2016 Dec 17;6(8):e011094 [FREE Full text] [doi: 10.1136/bmjopen-2016-011094] [Medline: 27534987]
29. Bhogal AS, Mani AR. Pattern analysis of oxygen saturation variability in healthy individuals: Entropy of pulse oximetry signals carries information about mean oxygen saturation. Front Physiol 2017;8:555 [FREE Full text] [doi: 10.3389/fphys.2017.00555] [Medline: 28824451]
30. Jupyter. URL: https://jupyter.org/ [accessed 2019-05-21] [WebCite Cache ID 78XYVWOKv]
31. Python. URL: https://www.python.org/ [accessed 2019-05-21] [WebCite Cache ID 78XYTIZtV]
32. Sillen MJ, Franssen FM, Delbressine JM, Uszko-Lencer NH, Vanfleteren LE, Rutten EP, et al. Heterogeneity in clinical characteristics and co-morbidities in dyspneic individuals with COPD GOLD D: Findings of the DICES trial. Respir Med 2013 Aug;107(8):1186-1194 [FREE Full text] [doi: 10.1016/j.rmed.2013.04.020] [Medline: 23706780]
33. Agusti A, Calverley PM, Celli B, Coxson HO, Edwards LD, Lomas DA, Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) investigators. Characterisation of COPD heterogeneity in the ECLIPSE cohort. Respir Res 2010 Sep 10;11:122 [FREE Full text] [doi: 10.1186/1465-9921-11-122] [Medline: 20831787]
34. Fussell KM, Ayo DS, Branca P, Rogers JT, Rodriguez M, Light RW. Assessing need for long-term oxygen therapy: A comparison of conventional evaluation and measures of ambulatory oximetry monitoring. Respir Care 2003 Feb;48(2):115-119 [FREE Full text] [Medline: 12556251]
35. Álvarez D, Hornero R, Marcos JV, Del Campo F. Feature selection from nocturnal oximetry using genetic algorithms to assist in obstructive sleep apnoea diagnosis. Med Eng Phys 2012 Oct;34(8):1049-1057. [doi: 10.1016/j.medengphy.2011.11.009] [Medline: 22154238]
36. Marcos JV, Hornero R, Nabney IT, Álvarez D, Gutiérrez-Tobal GC, del Campo F. Regularity analysis of nocturnal oximetry recordings to assist in the diagnosis of sleep apnoea syndrome. Med Eng Phys 2016 Mar;38(3):216-224. [doi: 10.1016/j.medengphy.2015.11.010] [Medline: 26719242]
37. Morillo DS, Gross N. Probabilistic neural network approach for the detection of SAHS from overnight pulse oximetry. Med Biol Eng Comput 2013 Mar;51(3):305-315. [doi: 10.1007/s11517-012-0995-4] [Medline: 23160897]
38. Soguel Schenkel N, Burdet L, de Muralt B, Fitting JW. Oxygen saturation during daily activities in chronic obstructive pulmonary disease. Eur Respir J 1996 Dec;9(12):2584-2589 [FREE Full text] [Medline: 8980973]
39. Mulloy E, Fitzpatrick M, Bourke S, O'Regan A, McNicholas WT. Oxygen desaturation during sleep and exercise in patients with severe chronic obstructive pulmonary disease. Respir Med 1995 Mar;89(3):193-198 [FREE Full text] [Medline: 7746912]

40. Heijdra YF, Dekhuijzen PN, van Herwaarden CL, Folgering HT. Nocturnal saturation and respiratory muscle function in patients with chronic obstructive pulmonary disease. Thorax 1995 Jun;50(6):610-612 [FREE Full text] [doi: 10.1136/thx.50.6.610] [Medline: 7638800]
41. Mohsenin V, Guffanti EE, Hilbert J, Ferranti R. Daytime oxygen saturation does not predict nocturnal oxygen desaturation in patients with chronic obstructive pulmonary disease. Arch Phys Med Rehabil 1994 Mar;75(3):285-289. [Medline: 8129580]
42. Miłkowska-Dymanowska J, Białas AJ, Obrębski W, Górski P, Piotrowski WJ. A pilot study of daily telemonitoring to predict acute exacerbation in chronic obstructive pulmonary disease. Int J Med Inform 2018 Dec;116:46-51 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.04.013] [Medline: 29887234]
43. Martín-Lesende I, Orruño E, Bilbao A, Vergara I, Cairo MC, Bayón JC, et al. Impact of telemonitoring home care patients with heart failure or chronic lung disease from primary care on healthcare resource use (the TELBIL study randomised controlled trial). BMC Health Serv Res 2013 Mar 28;13:118 [FREE Full text] [doi: 10.1186/1472-6963-13-118] [Medline: 23537332]
44. Burton C, Pinnock H, McKinstry B. Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease. J Telemed Telecare 2015 Jan;21(1):29-36. [doi: 10.1177/1357633X14562733] [Medline: 25475218]
45. Shah SA, Velardo C, Farmer A, Tarassenko L. Exacerbations in chronic obstructive pulmonary disease: Identification and prediction using a digital health system. J Med Internet Res 2017 Dec 07;19(3):e69 [FREE Full text] [doi: 10.2196/jmir.7207] [Medline: 28270380]
46. Troosters T, Sciurba F, Battaglia S, Langer D, Valluri SR, Martino L, et al. Physical inactivity in patients with COPD, a controlled multi-center pilot-study. Respir Med 2010 Jul;104(7):1005-1011 [FREE Full text] [doi: 10.1016/j.rmed.2010.01.012] [Medline: 20167463]

Abbreviations

6MWD: 6-minute walking distance
CAT: chronic obstructive pulmonary disease assessment test
CIRO: Centre of Expertise for Chronic Organ Failure
COPD: chronic obstructive pulmonary disease
CT90: cumulative time spent with SpO_2 below 90%
EE: energy expenditure
FEV1: forced expiratory volume in 1 second
FVC: forced vital capacity
GOLD II-IV: Global Initiative for Chronic Obstructive Lung Disease stages II-IV
ISO: International Organization for Standardization
LIPA: low-intensity physical activities
MEC-U: Medical Research Ethics Committees United
MET: metabolic equivalent of task
MVPA: moderate-to-vigorous physical activities
PaO_2: oxygen partial pressure in arterial blood
Seff: sleep efficiency
SpO_2: peripheral blood oxygen saturation
TLCO: transfer factor for carbon monoxide
TNST: total night sleeping time
VITO: Flemish Institute for Technological Research
WASO: wake time after sleep onset

Edited by G Eysenbach; submitted 21.11.18; peer-reviewed by N Hernandes, YR Park, A Farmer, D Dunsmuir; comments to author 31.03.19; revised version received 16.04.19; accepted 27.04.19; published 06.06.19.

Please cite as:
Buekers J, Theunis J, De Boever P, Vaes AW, Koopman M, Janssen EVM, Wouters EFM, Spruit MA, Aerts JM
Wearable Finger Pulse Oximetry for Continuous Oxygen Saturation Measurements During Daily Home Routines of Patients With Chronic Obstructive Pulmonary Disease (COPD) Over One Week: Observational Study
JMIR Mhealth Uhealth 2019;7(6):e12866
URL: https://mhealth.jmir.org/2019/6/e12866/
doi:10.2196/12866
PMID:31199331

Original Paper

Deep Learning Approaches to Detect Atrial Fibrillation Using Photoplethysmographic Signals: Algorithms Development Study

Soonil Kwon[1*], MD; Joonki Hong[2*], MS; Eue-Keun Choi[1], MD, PhD; Euijae Lee[1], MD; David Earl Hostallero[2], BS; Wan Ju Kang[2], BS; Byunghwan Lee[3], MS; Eui-Rim Jeong[4], PhD; Bon-Kwon Koo[1], MD, PhD; Seil Oh[1], MD, PhD, FHRS, FESC; Yung Yi[2], PhD

[1]Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea

[2]School of Electrical Engineering, KAIST, Daejeon, Republic of Korea

[3]Sky Labs Inc, Seongnam, Republic of Korea

[4]Department of Information and Communication Engineering, Hanbat National University, Daejeon, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Eue-Keun Choi, MD, PhD
Department of Internal Medicine
Seoul National University Hospital
101 Daehang-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 2 2072 0688
Fax: 82 2 762 9662
Email: choiek417@gmail.com

Abstract

Background: Wearable devices have evolved as screening tools for atrial fibrillation (AF). A photoplethysmographic (PPG) AF detection algorithm was developed and applied to a convenient smartphone-based device with good accuracy. However, patients with paroxysmal AF frequently exhibit premature atrial complexes (PACs), which result in poor unmanned AF detection, mainly because of rule-based or handcrafted machine learning techniques that are limited in terms of diagnostic accuracy and reliability.

Objective: This study aimed to develop deep learning (DL) classifiers using PPG data to detect AF from the sinus rhythm (SR) in the presence of PACs after successful cardioversion.

Methods: We examined 75 patients with AF who underwent successful elective direct-current cardioversion (DCC). Electrocardiogram and pulse oximetry data over a 15-min period were obtained before and after DCC and labeled as AF or SR. A 1-dimensional convolutional neural network (1D-CNN) and recurrent neural network (RNN) were chosen as the 2 DL architectures. The PAC indicator estimated the burden of PACs on the PPG dataset. We defined a metric called the confidence level (CL) of AF or SR diagnosis and compared the CLs of true and false diagnoses. We also compared the diagnostic performance of 1D-CNN and RNN with previously developed AF detectors (support vector machine with root-mean-square of successive difference of RR intervals and Shannon entropy, autocorrelation, and ensemble by combining 2 previous methods) using 10 5-fold cross-validation processes.

Results: Among the 14,298 training samples containing PPG data, 7157 samples were obtained during the post-DCC period. The PAC indicator estimated 29.79% (2132/7157) of post-DCC samples had PACs. The diagnostic accuracy of AF versus SR was 99.32% (70,925/71,410) versus 95.85% (68,602/71,570) in 1D-CNN and 98.27% (70,176/71,410) versus 96.04% (68,736/71,570) in RNN methods. The area under receiver operating characteristic curves of the 2 DL classifiers was 0.998 (95% CI 0.995-1.000) for 1D-CNN and 0.996 (95% CI 0.993-0.998) for RNN, which were significantly higher than other AF detectors (P<.001). If we assumed that the dataset could emulate a sufficient number of patients in training, both DL classifiers improved their diagnostic performances even further especially for the samples with a high burden of PACs. The average CLs for true versus false classification were 98.56% versus 78.75% for 1D-CNN and 98.37% versus 82.57% for RNN (P<.001 for all cases).

Conclusions: New DL classifiers could detect AF using PPG monitoring signals with high diagnostic accuracy even with frequent PACs and could outperform previously developed AF detectors. Although diagnostic performance decreased as the

burden of PACs increased, performance improved when samples from more patients were trained. Moreover, the reliability of the diagnosis could be indicated by the CL. Wearable devices sensing PPG signals with DL classifiers should be validated as tools to screen for AF.

(JMIR Mhealth Uhealth 2019;7(6):e12770) doi:10.2196/12770

KEYWORDS

atrial fibrillation; deep learning; photoplethysmography; pulse oximetry; diagnosis

Introduction

Background

Atrial fibrillation (AF) is the most common cardiac arrhythmia in clinical practice [1]. The prevalence and incidence of AF have risen over the years with the aging population [2]. The gold standard used to diagnose AF is the electrocardiogram (ECG) [3]. However, many patients with AF present paroxysmal symptoms or are asymptomatic; thus, the limited accessibility of the ECG during the symptom could lower the detection rate of AF. It is important to detect AF regardless of symptoms because asymptomatic patients with AF could present with stroke at their first manifestation [4]. AF is one of the major causes of stroke and its severity is worse in patients with AF than those without [5]. Therefore, it is important to detect AF in a potentially high-risk population who might benefit from stroke prevention with adequate anticoagulation control [6-9].

Recently, photoplethysmography (PPG) has been studied for long-term monitoring of AF, because of the ease of use and their utilization using wearable and mobile devices [10-12]. PPG monitoring incorporated into wearable technologies might permit improvements in the detection rate of AF in high-risk patients. Past studies using PPG for AF detection relied on predetermined feature extractions, for example, root-mean-square of successive difference of RR intervals (RMSSD) with Shannon entropy (ShE) and machine learning techniques such as the support vector machine (SVM), artificial neural network, and k-nearest neighbor [11,13]. The accuracy, sensitivity, and specificity for differentiating AF from a *clean* sinus rhythm (SR; ie, SR without any premature atrial complexes or PACs) of these methods were promising [14-18]. However, PACs were frequently exhibited in patients with paroxysmal AF or in those after successful cardioversion [19-21], which rendered AF detection using PPG from *unclean* SRs less practical. Low accuracy in differentiating AF from SR with PACs in the feature extraction step was a major limitation of previous approaches [22,23]. More sophisticated AF detection algorithms should be designed to render PPG monitoring more pragmatic.

Objective

We aimed to develop deep learning (DL) classifiers using PPG as an input to distinguish AF from SR in the presence of PACs. We also suggested a method to compute a confidence level (CL) [24,25] for each decision in tested samples so that physicians could quantify the reliability of the results from the DL classifiers.

Methods

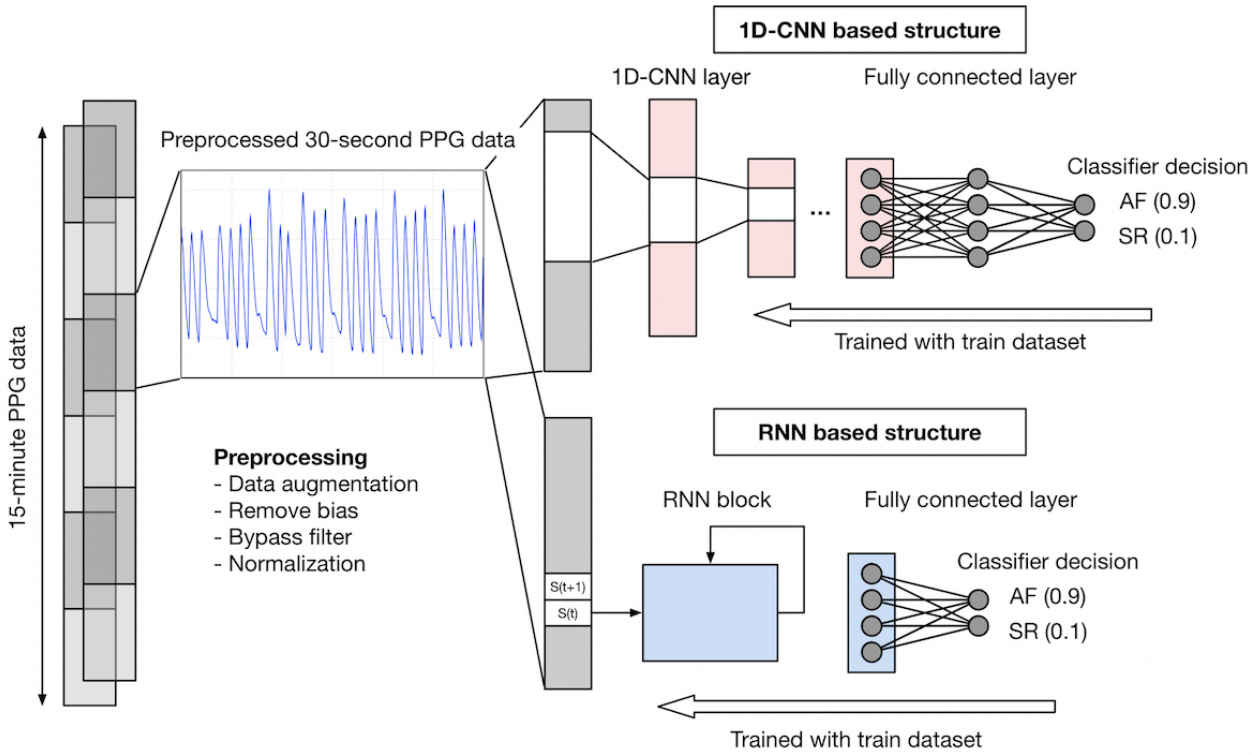
Study Population and Data Acquisition

This was a prospective, single-center study including patients with persistent AF admitted for elective DCC from September 2017 to April 2018. A total of 81 consecutive patients were enrolled. After verifying AF with 12-lead ECG, baseline PPG signals were collected over a 15-min interval by attaching a pulse oximeter to the patient's nondominant arm's index finger in the supine position. In addition, a single-lead ECG signal was acquired simultaneously to confirm the rhythm and was used as the gold standard. DCC (biphasic 100 to 200 J) was performed under light sedation after the baseline recording. Among 81 patients with DCC, 5 patients could not be converted to SR and 1 patient had improper data acquisition because of inappropriate bandwidth filters and sampling rate. In total, 75 patients with successful DCC underwent post-DCC PPG and ECG recording for over 15 min using the same methods. PACs were also monitored during the post-DCC recording period. During both periods of the measurements, the subject was required to rest on the bed with a supine position such that potential motion artifacts could be minimized. In total, 3 cardiologists interpreted the single-lead rhythm strips and verified the PACs and other atrial tachyarrhythmia. If there was a discrepancy between readings, then the senior electrophysiologists (EKC and EL) interpreted the rhythm and determined the final conclusion for the rhythms. We applied bandwidth filters (0.2 to 18 Hz) on both PPG and ECG data and then exported them in XML format for the DL training. The study protocol was approved by the Seoul National University Hospital Institutional Review Board and adhered to the Declaration of Helsinki.

Dataset Manipulation and Deep Learning Framework

We constructed PPG samples for training and testing from the PPG recording data of 75 patients. Each patient's 15-min pre-DCC and post-DCC data were divided into multiple 30-second fragments with 20-second overlaps. We used a data augmentation technique to increase the number of samples [24]. Each sample was labeled as *AF* if it was generated before the DCC and as *SR* if generated after the successful DCC. The detailed dataset manipulation method is presented in Multimedia Appendix 1.

Figure 1. Flowchart illustrating the deep learning process. For each subject, 15-min PPG data during pre- and post- direct-current cardioversion periods were obtained. Each 15-min sample was preprocessed by removing bias, applying bypass filters, and normalization. Then, each sample was subdivided into 30-second samples with 20-second overlaps for data augmentation. The 30-second samples were trained and tested by 1-dimensional convolutional neural network (1D-CNN) and recurrent neural network (RNN) methods. Each sample was labeled as atrial fibrillation (AF) or sinus rhythm (SR). The number in parenthesis shows an example of the corresponding confidence level. In this example, the confidence level for diagnosing AF was 0.9 in 1D-CNN and RNN models, whereas 0.1 in SR for both models. 1D-CNN: 1-dimensional convolutional neural network; AF: atrial fibrillation; PPG: photoplethysmography; RNN: recurrent neural network; SR: sinus rhythm.



The entire DL framework of AF diagnosis is described in Figure 1. We used a 1-dimensional convolutional neural network (1D-CNN) and a recurrent neural network (RNN) as 2 DL architectures and compared their diagnostic performance. The DL process was divided into 2 phases: training and testing. The DL classifier consisted of several layers of artificial neurons, forming a kind of function approximator, simulating the neural connections of human brains. The neural network (NN) was trained to approximate a target function, which was the function of the 30-second-long PPG sample as input and the diagnostic decision as output. In the training phase, we trained the NN based on a supervised method that minimized differences between the AF and SR true labels and the NN outputs using a back-propagation algorithm. In the testing phase, we evaluated the trained NN classifiers with data that were not used in the training phase.

Training and Testing Dataset Design

We used a 5-fold cross-validation approach to compare and evaluate 2 DL classifiers. The training phase included randomized features to initialize weight parameters between the NN nodes (weight and node corresponded to synapse and neuron, respectively) and the application of a rule to update weight parameters. Our 5-fold cross-validation process was as follows. In Scenario A, patients were randomly assigned to 5 groups, in which 4 were the training dataset, whereas the remainder was the testing dataset. Given that there were 5 groups of patients, 5 different combinations of the training and testing datasets were analyzed. In this scheme, no patients belonged to both the training and testing dataset at the same time, that is, DL classifiers always faced new patients during testing. We repeated the validation process over 10 times for each combination, and the final results were obtained by averaging the total of 50 validations. In Scenario B, we have performed 5-fold validation with random choices of samples, not in patients. Therefore, randomly chosen, 80% of the entire samples were assigned to training, whereas the remaining 20% were assigned to testing. Unlike Scenario A, Scenario B permits to allocate samples from the same patient into both training and testing datasets. We compared the performance with previous well-known AF detectors including SVM with RMSSD+ShE [15], SVM with autocorrelation [11], and an ensemble method by combining the 2 previous methods. Linear-kernel SVM was used in our study.

Diagnostic Performance Using Different Algorithms

We compared the diagnostic performance of the different methods by generating receiver operating characteristic (ROC) curves. The area under the ROC curve (AUC) and the 95% CI for each method was calculated and compared using the DeLong test [26]. Statistical analysis was performed as a 2-sided test, and a *P* value less than .05 was considered statistically significant.

We also analyzed the accuracy (total number of true diagnosis of AF or SR divided by total number of test samples), sensitivity

(the number of true diagnosis of AF divided by total number of AF-labeled test samples), specificity (the number of true diagnosis of SR divided by total number of SR-labeled test samples), positive predictive value (the proportion of AF-labeled test samples among the samples diagnosed as AF), and negative predictive value (the proportion of SR-labeled test samples among the samples diagnosed as SR). Each value was averaged over 50 validation processes for the 2 DL classifiers.

Diagnostic Performance According to the Premature Atrial Complex Burden

We further analyzed the specificity of the DL classifier over the PAC burden in post-DCC rhythms. In Scenario A, the trained DL classifier faced new patients during the validation. Therefore, it is highly likely to encounter unknown samples during the testing phase. However, if the number of patients in the training set grows, even though the patients in the validation are new to the DL classifier, they will likely be similar to the patients seen in the training set. Scenario B emulates such a circumstance by making the sample distribution of the testing dataset similar to that of the training dataset. This could be achieved because the samples from the same patient could appear in both datasets. For each PAC burden and scenario, we compared a specificity by different algorithms. Then we evaluated whether DL classifiers outperformed previous algorithms over various PAC burdens and how they are improved by an assumption of the same sample distribution in both datasets (Scenario B).

Confidence Levels

In this study, we defined the metric CL to measure the reliability of the diagnosis by a certain DL classifier. We refer readers to Multimedia Appendix 1 for a detailed description of the CL. The true CL represented the confidence of the classifier's output when it has correctly diagnosed the patient (ie, AF as AF and SR as SR), whereas a false CL indicated the confidence when AF was diagnosed as SR or SR diagnosed as AF. The minimum CL was 50%, meaning that the diagnosis was randomly AF or SR, whereas the maximum CL was 100%, indicating that the diagnosis could be unquestionably AF or SR irrespective of the number of times the test was repeated.

Results

Baseline Characteristics of Study Population

A total of 75 patients (men 68/75, 91%; mean age 63 years, SD 7.8) were enrolled. Clinical characteristics of the study population are summarized in Table 1. A total of 18 patients (18/75, 24%) were long-standing persistent AF (AF history of >1 year). The median value of the CHA_2DS_2-VASc score was 1.

Characteristics of the Study Dataset

A total of 14,298 samples consisting of 30-second-long PPG were generated from the 75 patients. Each 30-second PPG sample was synchronized with a single-lead ECG to be diagnosed as AF or SR. Figure 2 shows examples of AF and SR determined by PPG recordings.

We developed a PAC indicator that could automatically detect the number of PACs in each post-DCC PPG sample to quantify the PAC burden of the study dataset. We applied a simple criterion: consider a beat as a PAC when the interval with the previous beat was less than 85% of the average interval. Multimedia Appendix 2 illustrates the PAC detection results of the indicator for a post-DCC PPG sample with 5 PAC episodes and the corresponding ECG signal. The proposed PAC indicator's result was verified based on the cardiologist's decision with a matched ECG signal. Both the PPG and the matched ECG samples were reviewed and the sensitivity and specificity of the PAC indicator after the review were 92.55% and 98.18%, respectively. The PAC burden of each post-DCC PPG sample was calculated by dividing the number of PAC with the number of beats in the sample (Multimedia Appendix 3). Inspection of the PAC indicator showed that 29.79% (2132/7157) of post-DCC samples contained PACs in their data.

The deep learning method may require considerable computational power and time during the training. However, because our NN structure was relatively lightweight because of a small number of layers (6 layers in 1D-CNN and 2 layers in RNN), the training phase in our study took about 10 min under the computational environment with a single GPU system using NVIDIA TITAN Xp graphics card.

Comparison of the Performance of Deep Learning Classifiers

The diagnostic performance of the DL classifiers with previous well-known AF detecting algorithms is summarized in Table 2. The results were obtained under Scenario A. The 1D-CNN and RNN showed high accuracy (97.58% [139,527/142,980] and 97.15% [138,912/142,980], respectively), sensitivity (99.32% [70,925/71,410] and 98.27% [70,176/71,410], respectively), and specificity (95.85% [68,602/71,570] and 96.04% [68,736/71,570], respectively). In addition, both methods showed high positive predictive values (95.98% [70,925/73,893] and 96.12% [70,176/73,010], respectively) and negative predictive values (99.30% [68,602/69,087] and 98.24% [68,736/69,970], respectively). We derived ROCs for all the algorithms and compared the different AUCs (Figure 3). The AUCs for 1D-CNN and RNN were 0.998 (95% CI 0.995-1.000) and 0.996 (95% CI 0.993-0.998), respectively, which were significantly higher than previous methods. The 1D-CNN and RNN showed larger AUCs compared with the SVM with RMSSD+ShE, SVM with an autocorrelation or ensemble method (all $P<.001$, compared using the DeLong test). There were no significant differences in the AUC for the 1D-CNN and RNN algorithms ($P=.12$).

In Scenario B, the performances of the DL classifiers improved in overall aspects but more in specificity and positive predictive value. Both 1D-CNN and RNN had improved sensitivity (99.71% [71,206/71,410] and 99.51% [71,059/71,410], respectively), specificity (99.35% [71,104/71,570] and 99.29% [71,062/71,570], respectively), positive predictive value (99.35% [71,206/71,672] and 99.29% [71,059/71,567], respectively), negative predictive value (99.71% [71,104/71,308] and 99.51% [71,062/71,413], respectively), and accuracy (99.53%

[142,310/142,980] and 99.40% [142,121/142,980], respectively). Both classifiers had 1.000 of AUCs (95% CI 1.000-1.000).

Table 1. The clinical characteristics of the study population (N=75).

Variables	Values
Demographics	
Age (years), mean (SD)	63 (7.8)
Male, n (%)	68 (91)
Body mass index (kg/m^2), mean (SD)	25.2 (2.9)
Body surface area (m^2), mean (SD)	1.83 (0.16)
Types of atrial fibrillation (AF), n (%)	
Persistent[a]	57 (76)
Long-standing persistent[b]	18 (24)
CHA_2 DS_2-VASc score[c]	1 (1,2)
Comorbidity, n (%)	
Congestive heart failure	5 (7)
Hypertension	38 (51)
Diabetes mellitus	10 (13)
Stroke or transient ischemic attack	8 (11)
Myocardial infarction	1 (1)
Valvular heart disease	5 (7)
Dyslipidemia	19 (25)
Chronic renal failure	3 (4)
Chronic obstructive pulmonary disease	0
Hyperthyroidism	6 (8)
Previous AF ablation history, n (%)	3 (5)
Antiarrhythmic agents, n (%)	
Propafenone	19 (25)
Flecainide	9 (12)
Pilsicainide	4 (5)
Sotalol	1 (1)
Amiodarone	40 (53)
Beta blocker	18 (24)
Calcium channel blocker[d]	12 (16)
Digoxin	1 (1)
Anticoagulant, n (%)	
Aspirin	2 (3)
Warfarin	18 (24)
Nonvitamin K oral anticoagulant	55 (73)
Other medications	
Angiotensin converting enzyme inhibitor	0
Angiotensin II receptor blocker	16 (21)
Diuretics	7 (9)
Statin	13 (17)

[a]AF history more than 1 month and less than 1 year.

[b]AF history more than 1 year.

[c]The value is expressed as both median and interquartile range.

[d]Nondihydropyridine class.

Figure 2. Typical examples of 15-second-long photoplethysmography and corresponding synchronized electrocardiogram samples for (A) stereotypic normal sinus rhythm and (B) atrial fibrillation with suggested confidence level using the 1-dimensional convolutional neural network algorithm. AF: atrial fibrillation; CL: confidence level; ECG: electrocardiogram; PPG: photoplethysmography; SR: sinus rhythm.

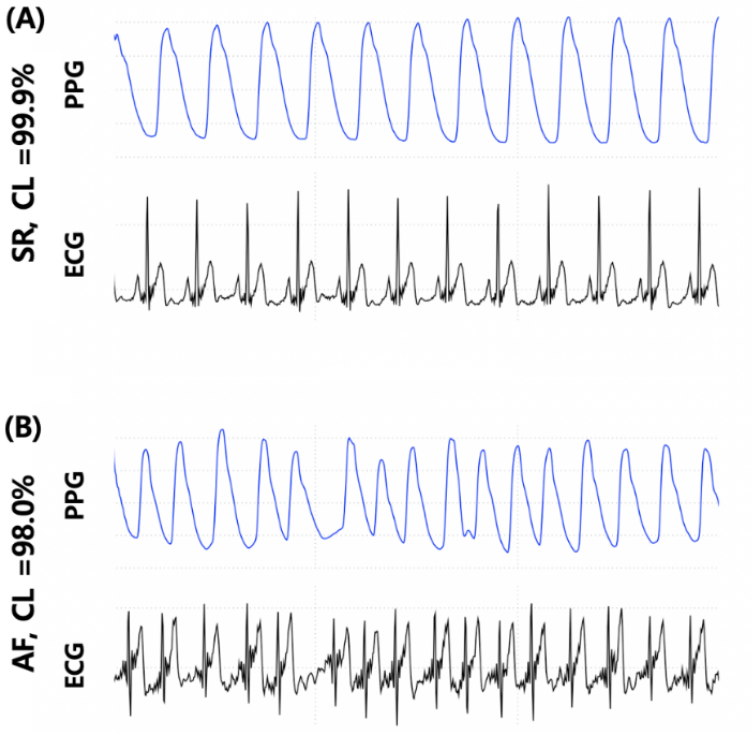


Table 2. The diagnostic performance of various algorithms for classifying photoplethysmography samples of atrial fibrillation and sinus rhythm after electrically cardioverted patients.

Algorithms	Accuracy	Mean sensitivity (%)	Mean specificity (%)	Mean positive predictive value (%)	Mean negative predictive value (%)	AUC[a]	95% CI	True mean confidence level (%)	False CL
1-Dimensional convolutional neural network	97.58	99.32	95.85	95.98	99.30	0.998	(0.995-1.000)	98.56	78.75
Recurrent neural network	97.15	98.27	96.04	96.12	98.24	0.996	(0.993-0.998)	98.37	82.57
Support vector machine, root-mean square of the successive differences of RR intervals + ShE[b]	86.82	89.13	84.50	85.16	88.63	0.868	(0.854-0.881)	—[c]	—
SVM, autocorrelation[d]	91.43	93.26	89.60	89.94	93.02	0.977	(0.972-0.982)	—	—
SVM, ensemble[e]	90.72	88.57	92.87	92.53	89.07	0.976	(0.970-0.981)	—	—

[a]AUC: mean area under the receiver operating characteristic curves. The standard errors by binomial exact test were all <0.01 except SVM with ensemble (0.01).

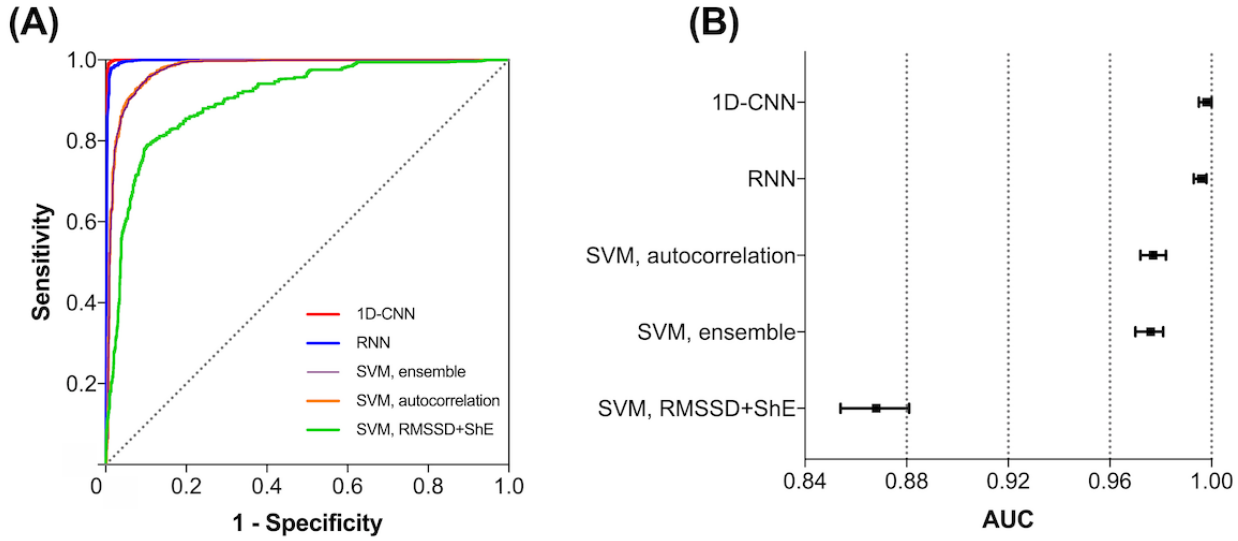
[b]SVM using RMSSD and ShE as a feature.

[c]Not applicable.

[d]SVM using autocorrelation method.

[e]SVM using RMSSD, ShE and autocorrelation.

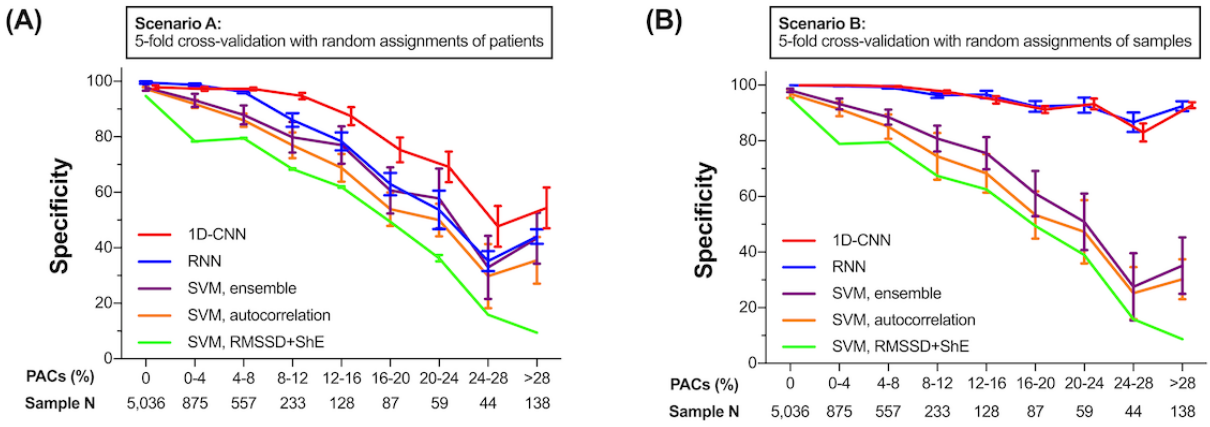
Figure 3. The receiver operating characteristic (ROC) curves of 2 deep learning classifiers (1-dimensional convolutional neural network, 1D-CNN and recurrent neural network, RNN) compared with other previous high-end atrial fibrillation (AF) detectors. (A) A Comparison of several ROC curves by different AF-detection algorithms. (B) The area under the curve and corresponding 95% CI by different algorithms. Both 1D-CNN and RNN methods showed significantly better diagnostic performance than previous detectors. 1D-CNN: 1-dimensional convolutional neural network; RMSSD: root-mean square of the successive differences of RR intervals; RNN: recurrent neural network; ShE: Shannon entropy; SVM: support vector machine.



Performance of the Deep Learning Classifier and Burden of Premature Atrial Complexes

Figure 4 shows the results with Scenario A and B. The burden of PACs was calculated for each 30-second PPG sample as the ratio of the number of PAC beats to the number of normal beats. The Sample N in Figure 4 shows the PAC burden distributions of the entire (both training and testing) dataset. In Scenario A, all the algorithms showed a decreasing tendency in specificity as the burden of PACs increased. Among the previous algorithms, the SVM with autocorrelation or ensemble maintained a higher specificity than with RMSSD+ShE. However, both 1D-CNN and RNN had a significantly higher specificity than the SVM with autocorrelation or ensemble (1D-CNN versus SVM with autocorrelation or ensemble, $P<.001$; RNN versus SVM with autocorrelation or ensemble, $P<.001$; all P values were calculated using a Student t test). Interestingly, the 1D-CNN maintained a significantly higher specificity than RNN in Scenario A (1D-CNN versus RNN, $P=.02$). In Scenario B, both DL classifiers improved significantly in specificity compared with Scenario A. The DL classifier could maintain 91.1% (1D-CNN) and 91.5% (RNN) specificity even for samples with a PAC burden ≥20%. Therefore, if the DL classifiers were trained with a sufficiently large dataset, they would maintain higher specificity even with a high PAC burden of and would outperform previous AF detectors.

Figure 4. Comparison of performances of deep learning classifiers and previous state-of-the-art atrial fibrillation detectors by premature atrial complexes (PACs) burden. The performance of classifying photoplethysmography samples during post- direct-current cardioversion period as sinus rhythm by each algorithm was measured by specificity. (A) Scenario A was obtained by the 5-fold cross-validation with random assignment of patients. In this case, each algorithm faced new patient's data during testing. (B) Scenario B was obtained by the 5-fold cross-validation with random assignment of samples. This approach assumed that the training distribution could emulate the test distribution. Regardless of the method, both 1-dimensional convolutional neural network and recurrent neural network maintained higher specificity over burden of PACs. Both DL classifiers showed higher specificity in Scenario B than Scenario A. 1D-CNN: 1-dimensional convolutional neural network; PAC: premature atrial complex; PPG: photoplethysmography; RNN: recurrent neural network root-mean square of successive difference of RR intervals.

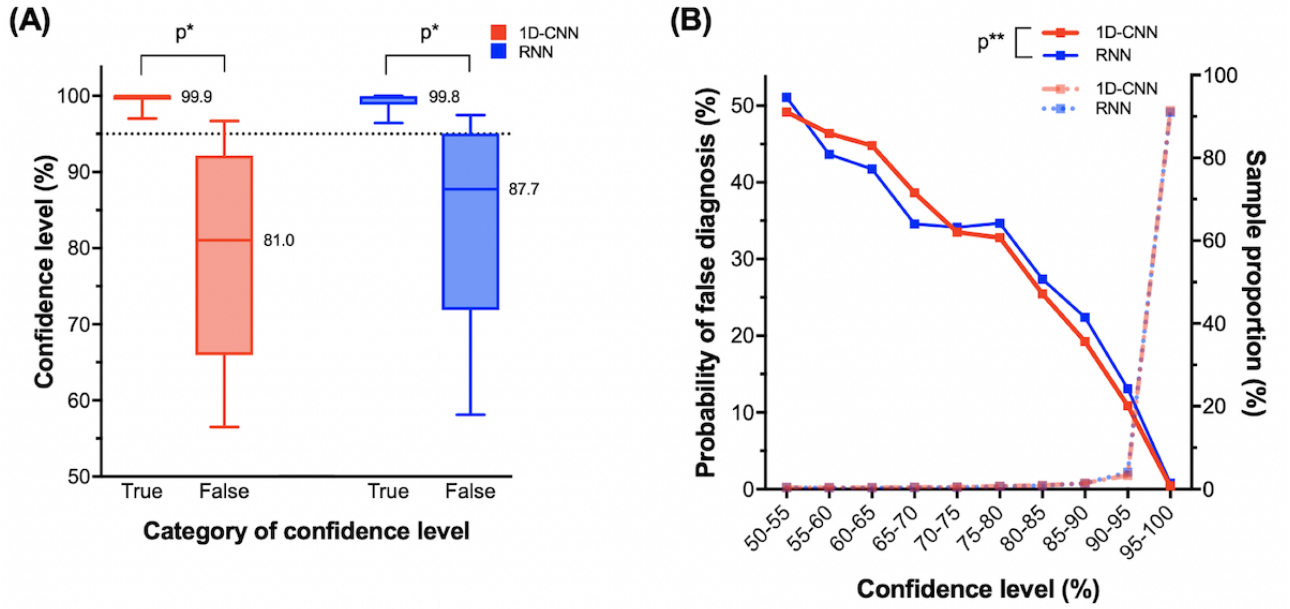


Confidence Level of the Deep Learning Classifier

Table 2 also shows true and false CLs of the 2 DL classifiers. The mean true and false CLs of the 1D-CNN classifier were 98.56% and 78.75%, respectively. With the RNN classifier, the true and false CLs were 98.37% and 82.57%, respectively. Therefore, significantly low CL values could indicate potential misdiagnoses. A further evaluation of the distribution of CLs is presented in Figure 5. For the 1D-CNN, the median values of true or false CL were 99.98% and 81.04%, respectively. Similarly, the median values for RNN were 99.81% and 87.74%, respectively. Comparison of the distribution of true or false CLs showed there were no significant differences in either the 1D-CNN or RNN methods (P<.001, calculated using a Student t test). If we set the cut-off level of CL to be 95%, the diagnostic accuracies were 99.58% (130,138/130,688) for 1D-CNN and 99.21% (129,053/130,082) for RNN. Therefore, a diagnosis with a CL ≥95% may be regarded as confident.

Even though DL classifiers were able to output a helpful metric, the classifier would become useless if most of the CL output were lower than 95%. However, 91.40% (130,688/142,980) of the tested samples had a CL >95% and the probability of a misdiagnosis for such sample was only 0.42% (550/130,688) for 1D-CNN. For RNN, likewise, 90.98% (130,082/142,980) of the tested samples had a CL >95% and the probability was 0.79% (1029/130,082). Therefore, most of the diagnosis made by the DL classifiers were confident. The differences in the probability of false diagnoses according to CL were not significantly different for both 1D-CNN and RNN (P=.98, calculated using a Student t test).

Figure 5. The characteristics of confidence level (CL) calculated by deep learning (DL) classifiers. Data were obtained by repeating the 5-fold cross-validation test over 10 times. (A) Comparison of true and false CLs of 1-dimensional convolutional neural network (1D-CNN) and recurrent neural network (RNN) methods by Box-and-Whiskers plot. True CLs indicate the cases where the diagnosis of a DL classifier was correct. Conversely, false CLs indicate cases where a DL classifier was incorrect. In the both 1D-CNN and RNN methods, the distributions of true or false CLs were significantly different (P<.001) for both 1D-CNN and RNN methods. If cut-off level of CL is to be 95% (dashed line), the diagnostic accuracy was 99.6% in 1D-CNN and 99.2% in RNN. Therefore, a diagnosis with a CL ≥95% can be regarded as certain. (B) The association between the probability of misdiagnosis and sample proportions and the respective CLs. Because 91% of the tested samples showed CL ≥95%, most diagnoses made by DL classifiers were valid. The probability of false diagnoses decreases from 50% to 0% as the CL increases from 50% to 100%. Comparing 1D-CNN and RNN, there was no significant difference in CLs (P=.98). *P<.001, calculated by Student t test. **P=.98, calculated by Student t test. 1D-CNN: 1-dimensional convolutional neural network; RNN: recurrent neural network.



Discussion

In this study, we developed a DL-based algorithm to diagnose AF using PPG data with better performance than previous algorithms. We found that (1) both 1D-CNN and RNN showed high diagnostic performance (AUC=0.998 and 0.996 for 1D-CNN and RNN, respectively); (2) both DL-based algorithms showed a better diagnostic performance than previous well-known AF detection algorithms, even under a high PAC burden, and had the potential to improve as more samples were allowed to be trained; and (3) most diagnoses by the DL classifiers were confident and the respective calculated CLs provided an easily interpreted reliability of the diagnosis.

Previous Photoplethysmographic Signals–Based Atrial Fibrillation Detectors

Before the development of DL, several well-performing PPG-based AF detectors with algorithms, such as RMSSD+ShE with Poincaré plots [16], SVM with autocorrelation or with RMSSD+ShE [11,15] had been described that could detect the irregularity of intervals between each PPG pulse by utilizing explicit rules or features. A recent study showed that, among the selected non-DL algorithms, SVM performed better than any others including Poincaré plots [27]. However, most of these previous algorithms were based on explicit features regarding peak-to-peak intervals but not the information such as amplitudes or waveforms. This loss of information during feature extraction steps was a limitation of these algorithms. In

this study, DL-based algorithms utilized the entire training data without any loss of information.

In contrast to the P-wave in ECG, the PPG has no markers for atrial contraction and this hinders the interpretation of cardiac rhythm from PPG alone. As a result, most previously developed algorithms for detecting AF from PPG relied heavily on the irregularity of peak-to-peak intervals. However, in the real-world setting, PACs are frequently observed in cardioverted AF patients and can simulate AF recurrence in these patients. Therefore, the diagnostic accuracy of PPG for detecting AF could be underpowered if the algorithm is predefined using a handcrafted approach. Therefore, more sophisticated methods to detect AF from PPG are needed.

Novelty of Deep Learning Classifiers for Detection of Atrial Fibrillation

With regard to the outperformance of DL compared with traditional machine learning (ML), the following explanations should be considered. To solve a classification problem, as in our AF detection method, traditional approaches mainly rely on algorithms that are rule-based or handcrafted ML-based. However, discriminating AF from SR using PPG becomes much more challenging under the presence of high burden of PACs, as there are no P-waves in the PPG. Furthermore, the complexity increases as the frequency of the PACs grows, which often holds in practice. Although previous ML approaches rely on handcrafted features, which are extracted mainly based on human intuition, the DL analyzes all the characteristics of the trained data, which is not limited only to peak-to-peak intervals, but also contains waveform characteristics such as amplitude, frequency, and wave morphology, and then automatically and implicitly quantifies their significance. The DL is composed of multiple layers of NNs and their smart connections have proven to be a powerful and efficient tool to handle complicated problems via automatic feature extraction of data and an in-depth understanding of their correlations. The 1D-CNN and RNN are the NN architectures specialized in handling the sequential data. The 1D-CNN network is a set of learnable kernels to extract the specific features or patterns in the sequential data. The kernels are convolved across the time axis of the input sequence to compute more compressed output sequence. Multiple layers of 1D-CNN are stacked to get the fully compressed features from the input sequence. The RNN compresses the input sequence by performing the same feedforward operation for every input token with the output being dependent on previous operations. The information abstracted from the previous feedforwards is accumulated to the last operation so that the last output contains the fully compressed features. Unlike the 1D-CNN, we used only a single layer of RNN. This would explain how DL outperforms previously described algorithms. In addition to the higher detection accuracy achieved by DL, it is also able to provide an output of the Softmax probability for each decision, which is used to quantify the probability of a correct decision, that is, of the CL. These advantages of DL for higher detection accuracy and provision of the CL were significant for detecting AF using PPG as a dataset.

Comparing Deep Learning Classifiers to Previous Algorithms

Recently, Poh et al reported that SVM showed the best performance among several non-DL algorithms and the deep convolutional NN was superior to SVM [27]. In our study, we compared 2 DL classifiers (1D-CNN and RNN) to previous well-known non-DL AF detection algorithms. The result showed that both DL classifiers had significantly better ROCs than previous methods including SVM and there were no significant differences between the ROCs of the 2 DL classifiers (Figure 3).

It should be noted that compared with previous studies, our results of both DL and non-DL algorithms showed fewer performances than expected. One may argue that our study should have had better results because we used a medical-grade pulse oximeter whereas some of the previous literature used sensors from smartphones, which would have a poorer signal-to-noise ratio [10,11,13,15,16]. However, such paradoxically lower results in our study may be explained by much higher PAC burdens in our samples (29.8%). Unlike other studies, we were able to graphically describe how diagnostic performances of various AF detectors degraded by PAC burdens (Figure 4). This implies any AF detector would give poorer results with more difficult samples, that is, the samples with more PAC burdens.

For DL classifiers, Poh et al reported better performance of CNN compared with this study (100.0% sensitivity and 99.6% specificity in the study of Poh et al but 99.3% and 95.9%, respectively, in our analysis) [27]. However, this can be explained by the different datasets used for training and testing. Our dataset was obtained from AF patients who underwent DCC. This scenario defines post-DCC rhythms with much more frequent PACs. Subsequently, the proportion of non-AF arrhythmia in the study sample was 6.4% in the study of Poh et al and 29.8% in our analysis. In other words, our training and testing dataset may be considered to be much more challenging in terms of AF or SR classification than those used in Poh's study. Non-DL classifiers in our study also had lower performances than previous reports, but this may be again because of a much higher burden of PAC in our study. For example, with an SVM-based approach, Chan reported 92.9% sensitivity and 97.7% specificity for diagnosing AF from SR. However, the proportion of samples with PAC to those with non-AF was only 2.8% [11]. Therefore, the testing results of any algorithm heavily depend on the characteristics of the sample dataset. Nevertheless, as one can observe from Figure 4, the coherent superiority of the DL classifier's performance over other algorithms for any PAC burden supports the advantages of using DL approaches to detect AF using PPG rather than previous non-DL algorithms.

Finally, there were other attempts to diagnose PAC from a normal SR or AF [16,28]. By incorporating RMSSD, ShE, and Poincaré plots, PACs were successfully diagnosed with 100% sensitivity and 97.8% specificity, yet information on how those results vary according to PAC burdens is lacking [28]. Similarly, such diagnoses would be performed by DL classifiers, provided that an investigator additionally labels the samples with PAC.

Future research is needed to observe how DL classifiers are different from other algorithms when diagnosing not only AF but also PACs.

Probability of Atrial Fibrillation Diagnosis

Besides the superior performance, the CL derived from DL classifiers was a useful metric to distinguish potentially mistaken decisions from correct decisions (Figure 5 and Multimedia Appendix 4). Multimedia Appendix 4 shows 2 exemplary cases where DL falsely diagnosed an AF but managed to generate low CLs (ie, 63.6% and 78.2%). However, according to Figure 5, such low CLs can be interpreted as *the diagnosis would possibly be incorrect* because the values are outliers to the lower 25th percentile of the distribution of true CLs. Thus, the possibility that the diagnosis in the examples would be correct would be approximately 35% to 45%. In such circumstances, the physician may attempt to confirm the cardiac rhythm using 12-lead ECG to validate the DL's diagnosis. More simply, one may only accept the diagnosis by DL classifiers when the CL is ≥95%. We have described above how a diagnosis having such a high CL may be regarded as confident with an extremely low probability of misdiagnosis. An easy interpretation using the CL may be that *X%* of CL has an *X%* chance of true diagnosis by DL classifiers by observing the linearly decreasing probability of false diagnosis according to the CL (Figure 5).

In summary, our new DL classifiers showed not only better diagnostic performance of detecting AF compared with previous algorithms especially under the high burden of PACs, but were also helpful to physicians by suggesting the reliability of the diagnosis through CLs.

Usefulness of Photoplethysmographic Signals for Atrial Fibrillation Detection

Most AF detectors rely on the ECG rather than on PPG because the interpretation of arrhythmia is more accurate and easier given the existence of the P-wave. However, in the clinical setting, monitoring the PPG rather than the ECG is easier and simpler. Furthermore, the emergence of wearable devices and mobile technologies enable clinicians to monitor the PPGs of a patient over long-term periods. Thus, the combination of AF detectors using PPG and wearable technology has the synergistic potential to screen AF recurrence and potentially to help prevent stroke in patients with AF. Though much effort has been made to utilize PPG to detect AF from SR, the presence of intermittent disruptions of regularity by frequent PACs hinder the application of PPG-monitoring for AF patients in real-life clinical situations. Future studies are needed to demonstrate that wearable devices sensing PPG to detect AF have clinical benefits in the prevention of stroke.

Limitations

First, we did not diagnose other arrhythmias such as ventricular premature complex, atrial tachycardia, and sinus arrhythmia, but focused specifically only on PAC. As a result, DL classifiers designed in our study may not be applied to patients with other arrhythmias. However, this can be justified as most electrically cardioverted AF patients have more frequent PACs than other arrhythmias. Also, PACs have importance as a high burden is associated with AF recurrence [29]. Second, the training and testing datasets of this study consisted of PPG data from 75 patients. The total number of samples generated by data augmentation was 14,298, which was sufficient to train DL. Despite this number, 75 patients may be insufficient to capture AFs presenting different characteristics and PACs' burden. Applying these trained DL classifiers to data from new patients may result in poorer performance. However, our analysis comparing scenario A and B (Figure 4) suggested that a sufficient number of samples from additional patients may result in improved performance. Finally, motion artifacts can arise during the measurement and cause performance degradation in a wearable device. To reduce the influence of motion artifact on the recording, we performed the signal in the supine position with minimal movement in our study. DL classifiers showed satisfactory performances without correction or removal of samples because of motion artifacts. However, we could not guarantee the performance of the developed deep learning algorithm in ambulatory patients would be as satisfactory as our results. Future works are needed to develop an algorithm that covers deleterious effects from motion artifacts.

Conclusions

In this study, we developed DL classifiers to detect AF from SR under the presence of PACs using 30-second PPG samples during pre- and post-DCC periods from patients with AF. The diagnostic performances of 1D-CNN and RNN models were significantly superior to other previously well-known AF detectors. The diagnostic performance of this model was still better than previous algorithms even under frequent PACs, and we observed the potential for further improvement in performance when sufficiently larger samples could be trained. As a metric representing the reliability of diagnosis, the CLs can be calculated using DL classifiers and a diagnosis with a CL ≥95% may be considered as confident. Besides, most diagnoses determined by DL classifiers were found to be confident. Taken together, these advantages indicate that implementing DL classifiers to wearable devices sensing PPG signals can be useful for AF screening among patients with AF.

Acknowledgments

This study was supported by a Korea National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2014R1A1A2A16055218), and sponsored by Sky Labs Inc., Seongnam, Republic of Korea.

Authors' Contributions

The first 2 authors contributed equally to this work.

Conflicts of Interest

SK, JH, EL, DEH, WK, BL, and SO have nothing to disclose. EKC and YY received research grants from Sky Labs. EKC also had grants from Daiichi-Sankyo, BMS/Pfizer, and Biosense Webster, and is a stockholder of Sky Labs. KBK is a stockholder of Sky Labs.

Multimedia Appendix 1

Detailed description of dataset manipulation, deep learning framework, and confidence level.

[PDF File (Adobe PDF File), 61KB - mhealth_v7i6e12770_app1.pdf]

Multimedia Appendix 2

Exemplary comparison between single-lead electrocardiogram (ECG; upper panel) and simultaneous photoplethysmography (PPG; lower panel). Premature atrial complexes (PACs) in PPG found by PAC indicator (red dots) were well corresponded to true PACs observed in ECG. The green dots were normal peaks of PPG pulse calculated by the indicator.

[PPTX File, 1MB - mhealth_v7i6e12770_app2.pptx]

Multimedia Appendix 3

The distribution of the post- direct-current cardioversion (DCC) photoplethysmography (PPG) samples used in the study subdivided by premature atrial complex (PAC) burden inspected by the PAC indicator devised in the study. Overall, 2132 out of total 7157 post-DCC PPG samples (29.79%) presented a PACs burden.

[PPTX File, 213KB - mhealth_v7i6e12770_app3.pptx]

Multimedia Appendix 4

Exceptional examples of both photoplethysmography (PPG) and corresponding to synchronized electrocardiogram samples with frequent premature atrial complexes (PACs) mimicking atrial fibrillation (AF) and 1-dimensional convolutional neural network (1D-CNN) misdiagnosed as AF but able to generate low confidence level (CL) values. All samples were obtained during the post- direct-current cardioversion period. Asterisks denote PACs. (A) The case in which frequent PACs occurred and the CL was determined as 63.6%. (B) The presence of frequent PACs resulting in a couplet and with a CL of 78.2%.

[PPTX File, 1MB - mhealth_v7i6e12770_app4.pptx]

References

1. Williams BA, Honushefsky AM, Berger PB. Temporal trends in the incidence, prevalence, and survival of patients with atrial fibrillation from 2004 to 2016. Am J Cardiol 2017 Dec 1;120(11):1961-1965. [doi: 10.1016/j.amjcard.2017.08.014] [Medline: 29033050]
2. Lee S, Choi E, Han K, Cha M, Oh S. Trends in the incidence and prevalence of atrial fibrillation and estimated thromboembolic risk using the CHADS-VASc score in the entire Korean population. Int J Cardiol 2017 Jun 1;236:226-231. [doi: 10.1016/j.ijcard.2017.02.039] [Medline: 28233629]
3. Camm AJ, Corbucci G, Padeletti L. Usefulness of continuous electrocardiographic monitoring for atrial fibrillation. Am J Cardiol 2012 Jul 15;110(2):270-276. [doi: 10.1016/j.amjcard.2012.03.021] [Medline: 22503584]
4. Lubitz SA, Yin X, McManus DD, Weng L, Aparicio HJ, Walkey AJ, et al. Stroke as the initial manifestation of atrial fibrillation: the Framingham Heart Study. Stroke 2017 Dec;48(2):490-492 [FREE Full text] [doi: 10.1161/STROKEAHA.116.015071] [Medline: 28082669]
5. Steger C, Pratter A, Martinek-Bregel M, Avanzini M, Valentin A, Slany J, et al. Stroke patients with atrial fibrillation have a worse prognosis than patients without: data from the Austrian Stroke registry. Eur Heart J 2004 Oct;25(19):1734-1740. [doi: 10.1016/j.ehj.2004.06.030] [Medline: 15451152]
6. Carrazco C, Golyan D, Kahen M, Black K, Libman RB, Katz JM. Prevalence and risk factors for paroxysmal atrial fibrillation and flutter detection after cryptogenic ischemic stroke. J Stroke Cerebrovasc Dis 2018 Jan;27(1):203-209. [doi: 10.1016/j.jstrokecerebrovasdis.2017.08.022] [Medline: 29032886]
7. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. J Am Med Assoc 2001 May 9;285(18):2370-2375. [Medline: 11343485]
8. Otite FO, Khandelwal P, Chaturvedi S, Romano JG, Sacco RL, Malik AM. Increasing atrial fibrillation prevalence in acute ischemic stroke and TIA. Neurology 2016 Nov 8;87(19):2034-2042 [FREE Full text] [doi: 10.1212/WNL.0000000000003321] [Medline: 27733570]

9. Charitos EI, Ziegler PD, Stierle U, Robinson DR, Graf B, Sievers H, et al. How often should we monitor for reliable detection of atrial fibrillation recurrence? Efficiency considerations and implications for study design. PLoS One 2014;9(2):e89022 [FREE Full text] [doi: 10.1371/journal.pone.0089022] [Medline: 24563690]
10. Yan BP, Lai WH, Chan CK, Chan SC, Chan L, Lam K, et al. Contact-free screening of atrial fibrillation by a smartphone using facial pulsatile photoplethysmographic signals. J Am Heart Assoc 2018 Apr 5;7(8) [FREE Full text] [doi: 10.1161/JAHA.118.008585] [Medline: 29622592]
11. Chan P, Wong C, Poh YC, Pun L, Leung WW, Wong Y, et al. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. J Am Heart Assoc 2016 Dec 21;5(7) [FREE Full text] [doi: 10.1161/JAHA.116.003428] [Medline: 27444506]
12. Conroy T, Guzman JH, Hall B, Tsouri G, Couderc J. Detection of atrial fibrillation using an earlobe photoplethysmographic sensor. Physiol Meas 2017 Sep 26;38(10):1906-1918. [doi: 10.1088/1361-6579/aa8830] [Medline: 28836507]
13. McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, et al. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart Rhythm 2013 Mar;10(3):315-319 [FREE Full text] [doi: 10.1016/j.hrthm.2012.12.001] [Medline: 23220686]
14. Berkaya SK, Uysal AK, Gunal ES, Ergin S, Gunal S, Gulmezoglu MB. A survey on ECG analysis. Biomed Signal Process Contr 2018 May;43:216-235. [doi: 10.1016/j.bspc.2018.03.003]
15. Lee K, Choi HO, Min SD, Lee J, Gupta BB, Nam Y. A comparative evaluation of atrial fibrillation detection methods in Koreans based on optical recordings using a smartphone. IEEE Access 2017;5:11437-11443. [doi: 10.1109/ACCESS.2017.2700488]
16. Mc MD, Chong JW, Soni A, Saczynski JS, Esa N, Napolitano C, et al. PULSE-SMART: pulse-based arrhythmia discrimination using a novel smartphone application. J Cardiovasc Electrophysiol 2016 Jan;27(1):51-57. [doi: 10.1111/jce.12842] [Medline: 26391728]
17. Nemati S, Ghassemi MM, Ambai V, Isakadze N, Levantsevych O, Shah A, et al. Monitoring and detecting atrial fibrillation using wearable technology. Conf Proc IEEE Eng Med Biol Soc 2016 Aug;2016:3394-3397. [doi: 10.1109/EMBC.2016.7591456] [Medline: 28269032]
18. Shan SM, Tang SC, Huang PW, Lin YM, Huang WH, Lai DM, et al. Reliable PPG-based algorithm in atrial fibrillation detection. 2017 Jan 26 Presented at: IEEE Biomedical Circuits and Systems Conference (BioCAS); October 19-21, 2016; Shanghai p. 340-343. [doi: 10.1109/BioCAS.2016.7833801]
19. Timmermans C, Rodriguez LM, Smeets JL, Wellens HJ. Immediate reinitiation of atrial fibrillation following internal atrial defibrillation. J Cardiovasc Electrophysiol 1998 Feb;9(2):122-128. [Medline: 9511886]
20. Wallmann D, Tüller D, Kucher N, Fuhrer J, Arnold M, Delacretaz E. Frequent atrial premature contractions as a surrogate marker for paroxysmal atrial fibrillation in patients with acute ischaemic stroke. Heart 2003 Oct;89(10):1247-1248 [FREE Full text] [Medline: 12975433]
21. Wallmann D, Tüller D, Wustmann K, Meier P, Isenegger J, Arnold M, et al. Frequent atrial premature beats predict paroxysmal atrial fibrillation in stroke patients: an opportunity for a new diagnostic strategy. Stroke 2007 Aug;38(8):2292-2294. [doi: 10.1161/STROKEAHA.107.485110] [Medline: 17585079]
22. Bae MH, Lee JH, Yang DH, Park HS, Cho Y, Chae SC, et al. Erroneous computer electrocardiogram interpretation of atrial fibrillation and its clinical consequences. Clin Cardiol 2012 Jun;35(6):348-353 [FREE Full text] [doi: 10.1002/clc.22000] [Medline: 22644921]
23. Poon K, Okin PM, Kligfield P. Diagnostic performance of a computer-based ECG rhythm algorithm. J Electrocardiol 2005 Jul;38(3):235-238. [Medline: 16003708]
24. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017 May 24;60(6):84-90. [doi: 10.1145/3065386]
25. Chung J, Gulcehre C, Cho K, Bengio Y. arXiv. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling URL: https://arxiv.org/pdf/1412.3555 [accessed 2019-05-22] [WebCite Cache ID 78YpHJMcS]
26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [Medline: 3203132]
27. Poh M, Poh YC, Chan P, Wong C, Pun L, Leung WW, et al. Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. Heart 2018 May 31:1921-1928. [doi: 10.1136/heartjnl-2018-313147] [Medline: 29853485]
28. Chong JW, Esa N, McManus DD, Chon KH. Arrhythmia discrimination using a smart phone. IEEE J Biomed Health Inform 2015 May;19(3):815-824. [doi: 10.1109/JBHI.2015.2418195] [Medline: 25838530]
29. Haïssaguerre M, Jaïs P, Shah DC, Takahashi A, Hocini M, Quiniou G, et al. Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. N Engl J Med 1998 Sep 3;339(10):659-666. [doi: 10.1056/NEJM199809033391003] [Medline: 9725923]

Abbreviations

1D-CNN: 1-dimensional convolutional neural network

AF: atrial fibrillation
AUC: area under the curve
CL: confidence level
DCC: direct-current cardioversion
DL: deep learning
ECG: electrocardiogram
ML: machine learning
NN: neural network
PAC: premature atrial complex
PPG: photoplethysmography
RMSSD: root-mean square of successive difference of RR intervals
RNN: recurrent neural network
ROC: receiver operating characteristic
ShE: Shannon entropy
SR: sinus rhythm
SVM: support vector machine

Edited by G Eysenbach; submitted 09.11.18; peer-reviewed by A Soni, K Chon; comments to author 20.12.18; revised version received 25.03.19; accepted 02.05.19; published 06.06.19.

Please cite as:
Kwon S, Hong J, Choi EK, Lee E, Hostallero DE, Kang WJ, Lee B, Jeong ER, Koo BK, Oh S, Yi Y
Deep Learning Approaches to Detect Atrial Fibrillation Using Photoplethysmographic Signals: Algorithms Development Study
JMIR Mhealth Uhealth 2019;7(6):e12770
URL: http://mhealth.jmir.org/2019/6/e12770/
doi:10.2196/12770
PMID:31199302

Original Paper

Estimating Maximal Oxygen Uptake From Daily Activity Data Measured by a Watch-Type Fitness Tracker: Cross-Sectional Study

Soon Bin Kwon[1], BSc; Joong Woo Ahn[2], PhD; Seung Min Lee[3], MA; Joonnyong Lee[1], PhD; Dongheon Lee[1], MA; Jeeyoung Hong[2,4], PhD; Hee Chan Kim[1,2,3], PhD; Hyung-Jin Yoon[1,2,3,4], MD, PhD

[1]Interdisciplinary Program in Bioengineering, Seoul National University, Seoul, Republic of Korea

[2]Institute of Medical & Biological Engineering, Medical Research Center, Seoul National University, Seoul, Republic of Korea

[3]Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, Republic of Korea

[4]Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

Corresponding Author:
Hyung-Jin Yoon, MD, PhD
Interdisciplinary Program in Bioengineering
Seoul National University
103 Daehak-ro, Jongro-gu
Seoul, 03080
Republic of Korea
Phone: 82 2 740 8596
Fax: 82 2 745 7870
Email: hjyoon@snu.ac.kr

Abstract

Background: Cardiorespiratory fitness (CRF), an important index of physical fitness, is the ability to inhale and provide oxygen to the exercising muscle. However, despite its importance, the current gold standard for measuring CRF is impractical, requiring maximal exercise from the participants.

Objective: This study aimed to develop a convenient and practical estimation model for CRF using data collected from daily life with a wristwatch-type device.

Methods: A total of 191 subjects, aged 20 to 65 years, participated in this study. Maximal oxygen uptake (VO_2 max), a standard measure of CRF, was measured with a maximal exercise test. Heart rate (HR) and physical activity data were collected using a commercial wristwatch-type fitness tracker (Fitbit; Fitbit Charge; Fitbit) for 3 consecutive days. Maximal activity energy expenditure (aEEmax) and slope between HR and physical activity were calculated using a linear regression. A VO_2 max estimation model was built using multiple linear regression with data on age, sex, height, percent body fat, aEEmax, and the slope. The result was validated with 2 different cross-validation methods.

Results: aEEmax showed a moderate correlation with VO_2 max (r=0.50). The correlation coefficient for the multiple linear regression model was 0.81, and the SE of estimate (SEE) was 3.518 mL/kg/min. The regression model was cross-validated through the predicted residual error sum of square (PRESS). The PRESS correlation coefficient was 0.79, and the PRESS SEE was 3.667 mL/kg/min. The model was further validated by dividing it into different subgroups and calculating the constant error (CE) where a low CE showed that the model does not significantly overestimate or underestimate VO_2 max.

Conclusions: This study proposes a CRF estimation method using data collected by a wristwatch-type fitness tracker without any specific protocol for a wide range of the population.

(JMIR Mhealth Uhealth 2019;7(6):e13327) doi:10.2196/13327

KEYWORDS

cardiorespiratory fitness; oxygen consumption; fitness tracker

Introduction

Cardiorespiratory fitness (CRF) is an important component of physical fitness, representing the body's ability to take oxygen in and deliver this oxygen to muscle cells throughout the body during physical activity. Previous studies have emphasized the importance of CRF, providing convincing evidence that CRF is closely related to all-cause mortality [1,2]. In addition, CRF is known to be correlated with various physiological factors, such as body composition and blood pressure, and psychological factors, such as depression [1-4]. Erikssen et al [5] have reported that a change in physical fitness is a strong predictor of mortality. They found that a small improvement in physical fitness can significantly lower the risk of death.

Maximal oxygen uptake (VO_2 max) is regarded as a representative feature of CRF. The current gold standard for measuring VO_2 max is a metabolic gas analysis during a maximal graded exercise test (GXT) on a treadmill or other equipment, such as cycle ergometer. Even though the maximal exercise test provides an accurate measurement of VO_2 max, there are several limitations. The maximal exercise requires a high level of motivation from the subject and should be performed under medical supervision for older or high-risk subjects who need this test the most [6]. Furthermore, the gas analysis requires expensive equipment and a trained technician to operate the process [7]. In addition, because of the high cost and inconvenience, it is impractical to repeat the maximal exercise test to regularly monitor VO_2 max.

Several estimation models have been developed to estimate VO_2 max. Some of these models have developed a submaximal exercise protocol in an attempt to overcome the limitation of the maximal test [8-10]. Submaximal models obtain exercise-related data through a specified exercise protocol, such as shuttle run, and build estimation models along with other anthropometric features. Although submaximal models have overcome some of the limitations of the GXT, they still require trained personnel to conduct the submaximal test, and familiarity with the exercise protocol could affect the results of the test [11], making it unsuitable for regular VO_2 max monitoring. There are other estimation models that do not involve an exercise protocol [12,13]. These models estimate VO_2 max by collecting data from physical activity and heart rate (HR) from daily life and using the relationship between the collected data and VO_2 max. Although these methods are more suitable for regular VO_2 max measurement, they are time-consuming (requiring a week of data collection) and use multiple devices, making it uncomfortable for application in daily life.

In our previous study [14], we developed a nonexercise VO_2 max estimation model using a new feature, maximum activity energy expenditure (aEEmax), which was calculated using activity energy expenditure and HR. Using aEEmax, we were able to build an accurate estimation model. However, aEEmax and our previous model were validated only in homogenous subjects, young Asian males. Furthermore, the device was worn on the chest, which might cause discomfort when used in daily life.

The aim of this study was to overcome the limitations of our previous study by using a wristwatch-type fitness tracker with various groups in terms of age and sex. We also sought to develop a new VO_2 max estimation model using aEEmax and the slope between physical activity and HR as new features, which could be applied to daily life data collected from a single convenient device worn on a wrist with a relatively short estimation time.

Methods

Participants

A total of 240 participants were recruited for this study. All participants completed the Physical Activity Readiness Questionnaire and health evaluation, including medical history related to cardiovascular disease, hypertension, and/or diabetes. Only participants without such medical history were included for this study. There were a total of 6 groups, divided according to age (20 to 35, 36 to 50, and 51 to 65 years) and sex, and there were 40 subjects for each group. Subjects who failed to achieve VO_2 max were excluded from the study. The achievement of VO_2 max was defined by accomplishing at least 2 of the following 3 criteria: a respiratory exchange ratio reaching >1.2, plateau of VO_2 despite increasing work load, or self-reported volitional fatigue [15]. Participants who did not wear the device for 3 days or participants with data loss were also excluded. A total of 49 subjects were excluded because of failure to achieve VO_2 max or unappropriated data collection. The characteristics of the participants are shown in Table 1.

Table 1. Subject characteristics.

Characteristics	Male, mean (SD)			Female, mean (SD)		
	20-35 years (n=34)	36-50 years (n=26)	51-65 years (n=30)	20-35 years (n=36)	36-50 years (n=35)	51-65 years (n=30)
Height (cm)	174.3 (5.6)	172.6 (6.1)	167.1 (4.9)	161.9 (5.4)	160.0 (5.2)	155.3 (4.8)
Weight (kg)	73.9 (8.0)	74.7 (9.7)	67.0 (6.1)	55.8 (7.3)	60.6 (6.1)	56.2 (6.1)
Percent body fat	20.4 (5.2)	24.8 (4.9)	24.2 (5.1)	29.4 (6.7)	33.9 (4.9)	33.8 (5.4)
aEEmax[a] (kcal/kg/h)	141.0 (14.6)	123.7 (15.1)	111.2 (11.0)	112.5 (15.8)	107.7 (12.1)	102.0 (10.3)
Slope (kcal/kg/h/bpm)	1.10 (0.14)	1.04 (0.13)	1.04 (0.12)	0.92 (0.17)	0.98 (0.14)	0.96 (0.13)
VO_2 max[b] (mL/kg/min)	42.3 (3.6)	39.9 (3.5)	38.1 (4.6)	35.3 (3.5)	31.4 (4.1)	30.5 (3.9)

[a]aEEmax: maximal activity energy expenditure.

[b]VO_2 max: maximal oxygen uptake.

Anthropometrics

Body mass and height were measured using a medical scale with a stadiometer (BSM330; InBody). Body mass was measured to the nearest 0.1 kg, and height was measured to the nearest 0.1 cm. Percent body fat was measured using a bioimpedance analysis (InBody720; InBody) to the nearest 0.1%.

Measurement of Maximal Oxygen Uptake

The reference VO_2 max value was measured using the modified Bruce protocol. The equipment used in the modified Bruce protocol includes a respiration gas analyzer (Vmax Encore System; CareFusion) and an aerobic exercise test system (CASE v6.61; GE Healthcare). Standard 12-lead electrocardiogram (ECG), blood oxygen saturation, and blood pressure were measured throughout the procedure.

Before performing the modified Bruce protocol, the baseline physiologic measures for all devices used were measured in a resting state for 5 min and subsequently in a standing position. The modified Bruce protocol was performed immediately after the baseline measurement. The treadmill's velocity and slope increased at 3-min intervals until the subject reached VO_2 max.

Experimental Methods

Participants wore a Fitbit (Fitbit Charge; Fitbit) on the left wrist for 3 consecutive days. From our previous study, we have shown that a minimum of 15 hours of physical activity data are required to acquire aEEmax [14]. In total, 3 consecutive days, regardless of weekday or weekend days, were shown to be enough to obtain the data needed based on the study. The Fitbit simultaneously measured HR and daily physical activity in terms of metabolic equivalent (kcal/kg/h), which was an expression of energy expenditure of activities.

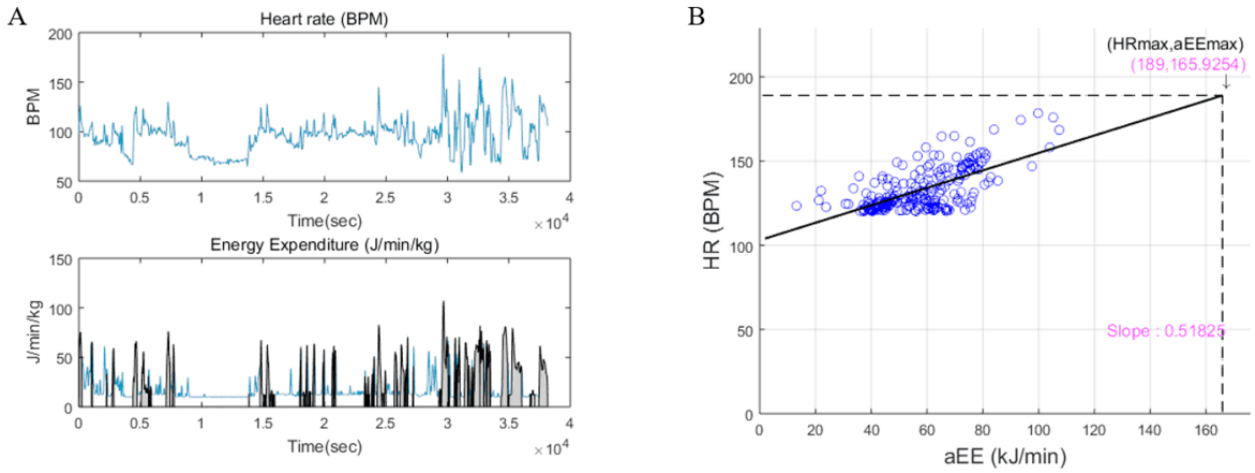
The participants removed the sensor during sleep or showering. Data obtained from the sensor were retrieved after the experiment via the internet. The Fitbit returned HR and metabolic equivalent data for every 1 min.

Signal Processing

Moving average filter was applied to both HR and physical activity data. After filtering the data, only data points at which both HR and physical activity data increased were selected as a period of physical activity for further processing. This was done by differentiating the data and selecting where both differentiated data were positive.

Figure 1 shows the HR (upper graph) and physical activity (lower graph) for a representative participant's filtered data over time. The shaded area under the physical activity curve represents the periods of increasing HR and physical activity. The scatter plot for HR versus physical activity is shown in Figure 1. A simple linear regression was performed between HR and physical activity to estimate aEEmax. The part of data where HR was greater than 120 beats/min was selected to remove the data in which the relationship between HR and physical activity was nonlinear [16]. The maximal HR was calculated as 200–age×0.67 for women and 216–age×0.93 for men [17]. The physical activity value of the point of intersection between the maximum HR and the regression line was defined as the aEEmax. The slope of the regression line was also used as a feature to estimate VO_2 max; hereafter, it will be referred as the *slope*. After the calculation of aEEmax and the slope, a multiple linear regression model was developed with aEEmax and anthropometric values to estimate VO_2 max.

Figure 1. (A) HR and aEE data from Fitbit. The shaded area indicates the period of data where both HR and aEE are increasing. (B) Scatter plot between selected periods of aEE versus HR of a representative subject. Data where HR was less than 120 bpm were removed to select the data where HR and aEE had a linear relationship. aEEmax is defined as the intersection between the interoperation line and HRmax, and the slope is the slope of the interpolation line. aEE: activity energy expenditure; aEEmax: maximal activity energy expenditure; BPM: beat per minute; HR: heart rate; HRmax: maximum heart rate.



Statistical Analysis

Pearson correlation coefficient was calculated between the independent variables (age, percent body fat, height, gender, aEEmax, and slope) and the measured VO_2 max. The regression model for estimating VO_2 max was evaluated with the coefficients of determination (adjusted R^2) and absolute SE of the estimate (SEE). The predicted residual error sum of squares (PRESS) statistic method was selected for cross-validation of the model [18]. The PRESS statistic is a cross-validation method calculating the error for each case by excluding a case each time from generating the estimation model and applying the model to the excluded case. The PRESS adjusted R^2 (R^2p) and the PRESS SEE (SEEp) were calculated as 1 – (PRESS/SS) and . The model was further validated by dividing it into different subgroups and calculating the constant error (CE) for each group. The standard for recruiting participants was to retain diversity. However, we wanted to observe CE based on age, sex, and VO2max level. The median value for age and VO2max was chosen to divide the groups. All signal processing, cross-validation, and statistical analyses were performed using MATLAB (MATLAB2017a; MathWorks).

Ethics Statement

This study protocol was reviewed and approved by the Institutional Review Board of the Seoul National University Hospital (IRB No. 1505-022-669). Written informed consent was submitted by all subjects when they were enrolled. This study followed the Helsinki Declaration.

Results

The general characteristics of all subjects are summarized in Table 1. The average value for age, weight, height, and percent body fat for excluded male subjects were 42.0, 72.8 kg, 171.0 cm, and 25.0% respectively. For excluded female subjects, the average values were 41.7, 56.5 kg, 158.4 cm, and 31.6%, respectively. A student *t* test was performed to compare the *P* value for age, weight, height, and percent body fat between included and excluded subjects. For male subjects, the *P* values were .646, .937, .883, and .026, respectively. For female subjects, the *P* values were .936, .415, .663, and .497 respectively.

The Pearson correlations between VO_2 max and selected features are shown in Table 2. The Pearson correlations between selected features are also summarized. The correlations between VO_2 max and independent variables were all statistically significant ($P<.001$ for all). The independent variable that showed the highest correlation was sex, with a correlation of .675. The lowest correlation for the model was the *slope*, with a correlation of .237.

The multiple linear regression analysis for the model is shown in Table 3. The scatter plot for measured VO_2 max versus predicted VO_2 max is shown in Figure 1. The R^2 for the model was 0.651, and the SEE was 3.518 mL/kg/min. As shown in Table 3, the decrease in R^2 and the increase in SEE were small for the cross-validation result of the PRESS method. R^2 decreased by 0.032 and SEE increased by 0.148 mL/kg/min. The scatter plot of the multiple linear regression for the model is shown in Figure 2.

The model was further validated by dividing the groups into various subgroups and calculating the CE and SD. Each subgroup was divided into 2 groups according to age, sex, and measured VO_2 max. The results are shown in Table 4. The CEs were positive for the younger group and negative for the older groups. The CE for the younger group was 0.024 mL/kg/min. The CE for the older group was −0.021 mL/kg/min. The CEs were all positive for the subgroups divided according to sex. The CE was positive for individuals with high VO_2 max and negative for individuals with low VO_2 max. However, as shown in Table 4, CE values were low for all subgroups, indicating that our model does not overestimate or underestimate VO_2 max.

Table 2. Correlation matrix between VO_2 max and independent variables.

Independent variables	VO_2 max[a]	Age	Height	Sex	Percent fat	aEEmax[b]
Age (years)	−0.372[c]	—	—	—	—	—
Height	0.372[c]	−0.329	—	—	—	—
Sex	0.675[c]	−0.015	0.517	—	—	—
Percent fat	−0.652[c]	0.202	−0.423	−0.620	—	—
aEEmax	0.503[c]	−0.473	0.397	0.496	−0.340	—
Slope	0.237[c]	−0.016	0.193	0.369	−0.130	0.830

[a]VO_2 max: maximal oxygen uptake.

[b]aEEmax: maximal activity energy expenditure.

[c]$P<.001$.

Table 3. A multiple regression nonexercise model for estimating VO_2 max (maximal oxygen uptake; mL/kg/min).

Independent variables	Fitbit model	
	Coefficient	Beta[a]
Constant	63.262	—[b]
aEEmax[c] (kcal/kg/h)	0.027	.082
Slope (kcal/kg/h/bpm)	−1.776	−.045
Percent body fat	−0.242	−.296
Age (years)	−0.150	−.321
Sex	3.264	.548
Height (cm)	−0.09	−.166
R[d]	0.807	—
SEE[e] (mL/kg/min)	3.518	—
R_P[f]	0.787	—
SEE_P[f] (mL/kg/min)	3.667	—

[a]Beta is the normalized coefficient of the model.

[b]—: not applicable.

[c]aEEmax: maximal activity energy expenditure.

[d]R is the Pearson correlation.

[e]SEE: SE of estimate.

[f]R_P and SEE_P are the cross-validated results of the model.

Figure 2. The correlation between estimated and measured VO2max value for all subjects (N=191). The red solid line is the identity line of the measured and estimated VO2max. VO2max: maximal oxygen uptake.

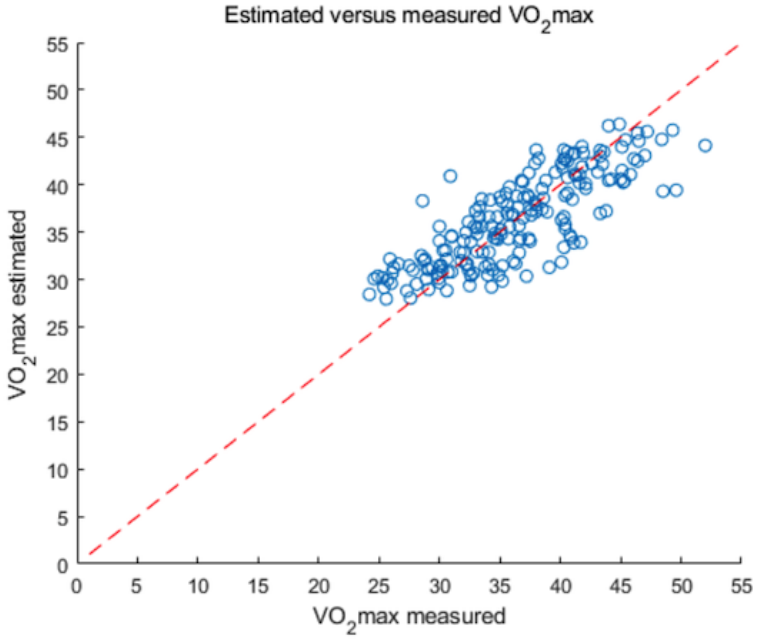


Table 4. Constant error and SD for each subgroup. Subjects aged older than 40 years are considered the old group, and subjects with a VO_2 max higher than 36 mL/kg/min are considered as the high VO_2 max group.

Group	n (%)	CE[a]	SD
Female	101 (52.9)	−4.397e-15	4.341
Male	90 (47.1)	3.000e-15	4.484
Old	101 (52.9)	−0.021	5.587
Young	90 (47.1)	0.024	5.180
High VO_2 max[b]	77 (40.3)	1.535	3.565
Low VO_2 max	114 (59.7)	−1.488	3.232

[a]CE: constant error.

[b]VO_2 max: maximal oxygen uptake.

Discussion

In this study, we developed VO_2 max estimation models with data collected from the commercially available device, Fitbit, which could estimate VO_2 max conveniently from daily life. The commercial device used in this study provided data on physical activity, along with its physiological response (HR). Even though some previous studies have reported that Fitbit does not always generate accurate data [19,20], our method for estimating VO_2 max does not depend on the absolute value of each data point. It rather depends on the trend of a large set of data points and would balance the inaccuracy of a single data point. Even though Fitbit data were biased under certain conditions, a large number of data collected from daily life would minimize this bias. This characteristic of our 2 novel features would only need calibration to be applied with other hardware and processing methods. Previous studies reported that lower HR was not linearly correlated with physical activity [16]. It is also known that heterogeneous recovery of HR after physical exercise does not have a clear correlation with VO_2 max [21]. Therefore, in an attempt to select the data during physical activity, we have selected periods of data where HR is greater than 120 beats/min and where both HR and physical activity increased at the same time to calculate aEEmax and the slope. By calculating 2 features from the linear relationship between HR and energy expenditure, our model would estimate VO2max accurately without any protocol or training.

In our previous study [14], aEEmax and our estimation model were validated with a homogenous group of young healthy Asian males using aEEmax and BMI (VO2max=0.192 x aEEmax – 0.708 x BMI). In this study, we have validated aEEmax and the model with a large pool of subjects, including both men and women aged from 20 to 65 years. The validation of the model was performed by cross-validation with the PRESS method and by calculating the CE of the subgroups. The cross-validated result shows that the validation sample fitted well with the model with little error. Furthermore, the CE for the model shows that our methodology did not significantly

overestimate nor underestimate VO_2 max for all subgroups, whereas other studies [12,22] have reported significant overestimation and underestimation with both highly and poorly fitted individuals.

To the best of our knowledge, this study is the first to provide a VO_2 max estimation model with a commercially available device on a wrist without any specific protocol. Tönis et al used a submaximal exercise protocol to estimate VO_2 max [23]. Polar Electro Oy Inc developed a nonexercise protocol, Polar Fitness Test, and devices to estimate VO_2 max. However, Esco et al [24] reported that the Polar Fitness Test had low accuracy for estimating VO_2 max when it was tested with one of its own products. Altini et al [25] developed a nonexercise estimation model for VO_2 max with data collected from daily life; however, they used an ECG necklace with a wet electrode attached to the chest and stomach, which could cause an inconvenience when used in daily life. Our method allows individuals to measure their VO_2 max on a wrist without the requirement for any electrode attachment. In addition, although other smartwatches, including Fitbit, require a specific protocol, such as running for at least 10 min on flat terrain, our protocol does not require any protocol and allows for easy monitoring of physical fitness.

Instead of using the absolute value of HR or physical activity data, we developed 2 new features to portray physical fitness. The change in HR for a given physical activity differs depending on the physical fitness of a subject [8]. Thus, the slope between HR and physical activity would be smaller and aEEmax would be larger for a subject with higher VO_2 max. As shown in Table 2, aEEmax has a moderate correlation with VO_2 max, supporting our hypothesis. Features based on physical activity can easily fluctuate depending on change in the short-term lifestyle of the subject during the period of data collection. Those features would be vulnerable to a sudden increase or decrease in the amount of physical activity. On the contrary, aEEmax and the slope represent the relationship between physical activity and HR and thus would be less affected by a sudden short-term change in physical activity. In addition, aEEmax and slope have been shown to be applicable to a wide range of subject ages and different genders. Other studies [26-28] have been validated with a relatively homogenous group compared with this study. The percent body fat used in this study was obtained using a professional bioimpedance analyzer. However, we have used general percent body fat which did not necessarily require a professional analyzer. There are products available, such as an AURA device, and ongoing studies about measuring body fat percentage from the wrist [29]. These efforts will make the measurement of percent body fat more accessible to the public.

There are limitations to this study. First, our methodology needs to be validated with more devices worn on a wrist. There are many commercially available devices that provide physical activity and HR data. To provide a more generalized VO_2 max estimation method, it is important to prove device independency of our method. In addition, this study was conducted with healthy subjects who were not taking any medication that might affect the HR. Another limitation of this study was error with maximal HR calculated from a basic population-derived formula. A more accurate method for calculating maximal HR could increase the accuracy of our model. A future study could include subjects who are on cardiac-related medication. Additionally, in our previous study, we have shown that 900 min of data were enough to calculate aEEmax; however, it would be worthwhile to observe change in the correlation coefficient of aEEmax for a longer period of time.

In summary, we have developed a new estimation model for VO_2 max using novel features, aEEmax and the slope between physical activity and HR, along with other anthropometric variables. The new features represent the relationship between physical activity and its physiological response. The high correlation between VO_2 max and aEEmax is in agreement with our previous study and supports our hypothesis. Our model requires data from only 3 days of daily life, without any specific exercise protocol. This hypothesis was validated with a diverse and large number of participants based on age and sex. Furthermore, all material required for our study is available in the conventional market as fully built products. The result of this study allows individuals to measure their VO_2 max conveniently in their daily life without any burden of an exercise protocol and allows them to easily monitor physical fitness.

Acknowledgments

This study was supported by the Samsung Research Funding Center of Samsung Electronics under Project Number SRFC-IT1402-04. This study was also supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF), funded by the Korean government (Ministry of Science, ICT, and Future Planning [MSIP]; No. 2016M3A9F1939646). The authors would like to thank all participants.

Conflicts of Interest

None declared.

References

1. Blair SN, Kampert JB, Kohl HW, Barlow CE, Macera CA, Paffenbarger RS, et al. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. J Am Med Assoc 1996 Jul 17;276(3):205-210. [Medline: 8667564]

2. Lee DC, Artero EG, Sui X, Blair SN. Mortality trends in the general population: the importance of cardiorespiratory fitness. J Psychopharmacol 2010 Nov;24(4 Suppl):27-35 [FREE Full text] [doi: 10.1177/1359786810382057] [Medline: 20923918]
3. Gerber M, Lindwall M, Lindegård A, Börjesson M, Jonsdottir IH. Cardiorespiratory fitness protects against stress-related symptoms of burnout and depression. Patient Educ Couns 2013 Oct;93(1):146-152. [doi: 10.1016/j.pec.2013.03.021] [Medline: 23623176]
4. Kaminsky LA, Arena R, Beckie TM, Brubaker PH, Church TS, Forman DE, et al. The importance of cardiorespiratory fitness in the United States: the need for a national registry: a policy statement from the American Heart Association. Circulation 2013;127(5):652-662. [Medline: 23295916]
5. Erikssen G, Liestøl K, Bjørnholt J, Thaulow E, Sandvik L, Erikssen J. Changes in physical fitness and changes in mortality. Lancet 1998 Sep 05;352(9130):759-762. [doi: 10.1016/S0140-6736(98)02268-5] [Medline: 9737279]
6. Akay MF, Inan C, Bradshaw DI, George JD. Support vector regression and multilayer feed forward neural networks for non-exercise prediction of VO2max. Expert Syst Appl 2009 Aug;36(6):10112-10119. [doi: 10.1016/j.eswa.2009.01.009]
7. Ernesto C, Martins da Silva F, Pereira L, de Melo G. Cross Validation of Different Equations to Predict Aerobic Fitness by the Shuttle Run 20 Meters Test in Brazilian Students. Journal of Exercise Physiology Online 2015;18(1).
8. Ekblom-Bak E, Björkman F, Hellenius ML, Ekblom B. A new submaximal cycle ergometer test for prediction of VO2max. Scand J Med Sci Sports 2014 Apr;24(2):319-326. [doi: 10.1111/sms.12014] [Medline: 23126417]
9. Larsen GE, George JD, Alexander JL, Fellingham GW, Aldana SG, Parcell AC. Prediction of maximum oxygen consumption from walking, jogging, or running. Res Q Exerc Sport 2002 Mar;73(1):66-72. [doi: 10.1080/02701367.2002.10608993] [Medline: 11926486]
10. Matabuena M, Vidal JC, Hayes PR, Huelin Trillo F. A 6-minute sub-maximal run test to predict VO max. J Sports Sci 2018 Nov;36(22):2531-2536. [doi: 10.1080/02640414.2018.1468149] [Medline: 29688149]
11. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther 2000 Aug;80(8):782-807. [Medline: 10911416]
12. Cao ZB, Miyatake N, Higuchi M, Miyachi M, Ishikawa-Takata K, Tabata I. Predicting VO2max with an objectively measured physical activity in Japanese women. Med Sci Sports Exerc 2010 Jan;42(1):179-186. [doi: 10.1249/MSS.0b013e3181af238d] [Medline: 20010115]
13. Plasqui G, Westerterp KR. Accelerometry and heart rate as a measure of physical fitness: proof of concept. Med Sci Sports Exerc 2005 May;37(5):872-876. [Medline: 15870644]
14. Ahn JW, Hwang SH, Yoon C, Lee J, Kim HC, Yoon HJ. Unobtrusive estimation of cardiorespiratory fitness with daily activity in healthy young men. J Korean Med Sci 2017 Dec;32(12):1947-1952 [FREE Full text] [doi: 10.3346/jkms.2017.32.12.1947] [Medline: 29115075]
15. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. Sports Med 2007;37(12):1019-1028. [doi: 10.2165/00007256-200737120-00002] [Medline: 18027991]
16. Ainslie PN, Reilly T, Westerterp KR. Estimating human energy expenditure: a review of techniques with particular reference to doubly labelled water. Sports Med 2003;33(9):683-698. [doi: 10.2165/00007256-200333090-00004] [Medline: 12846591]
17. Casteel B. American college of cardiology. 2014. The Heart Responds Differently to Exercise in Men vs Women URL: https://www.acc.org/about-acc/press-releases/2014/03/27/12/29/allison-peak-hr-pr [accessed 2019-06-03] [WebCite Cache ID 78qdZcV8g]
18. Holiday DB, Ballard JE, McKeown BC. PRESS-related statistics: regression tools for cross-validation and case diagnostics. Med Sci Sports Exerc 1995 Apr;27(4):612-620. [Medline: 7791595]
19. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
20. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2017 Dec 01;2(1):104-106. [doi: 10.1001/jamacardio.2016.3340] [Medline: 27732703]
21. Sartor F, Vernillo G, de Morree HM, Bonomi AG, La Torre A, Kubis HP, et al. Estimation of maximal oxygen uptake via submaximal exercise testing in sports, clinical, and home settings. Sports Med 2013 Sep;43(9):865-873. [doi: 10.1007/s40279-013-0068-3] [Medline: 23821468]
22. Jackson AS, Blair SN, Mahar MT, Wier LT, Ross RM, Stuteville JE. Prediction of functional aerobic capacity without exercise testing. Med Sci Sports Exerc 1990 Dec;22(6):863-870. [Medline: 2287267]
23. Tönis TM, Gorter K, Vollenbroek-Hutten MM, Hermens H. Comparing VO2max determined by using the relation between heart rate and accelerometry with submaximal estimated VO2max. J Sports Med Phys Fitness 2012 Aug;52(4):337-343. [Medline: 22828454]
24. Astorino T, Baker J, Brock S, Dalleck L, Goulet E, Gotshall R, et al. Cross-Validation of the Polar Fitness Test TM via the Polar F11 Heart Rate Monitor in Predicting VO 2 Max. J Exerc Physiol 2011;14(5) [FREE Full text]
25. Altini M, Casale P, Penders J, Amft O. Cardiorespiratory fitness estimation in free-living using wearable sensors. Artif Intell Med 2016 Mar;68:37-46. [doi: 10.1016/j.artmed.2016.02.002] [Medline: 26948954]

26. Fudge BW, Wilson J, Easton C, Irwin L, Clark J, Haddow O, et al. Estimation of oxygen uptake during fast running using accelerometry and heart rate. Med Sci Sports Exerc 2007 Jan;39(1):192-198. [doi: 10.1249/01.mss.0000235884.71487.21] [Medline: 17218902]
27. Malek MH, Housh TJ, Berger DE, Coburn JW, Beck TW. A new nonexercise-based VO2(max) equation for aerobically trained females. Med Sci Sports Exerc 2004 Oct;36(10):1804-1810. [Medline: 15595304]
28. Bradshaw DI, George JD, Hyde A, LaMonte MJ, Vehrs PR, Hager RL, et al. An accurate VO2max nonexercise regression model for 18-65-year-old adults. Res Q Exerc Sport 2005 Dec;76(4):426-432. [doi: 10.1080/02701367.2005.10599315] [Medline: 16739680]
29. Jung M, Namkoong K, Lee Y, Koh YJ, Eom K, Jang H, et al. Wrist-wearable bioelectrical impedance analyzer with contact resistance compensation function. IEEE Sensors 2016:1-3. [doi: 10.1109/ICSENS.2016.7808916]

Abbreviations

aEEmax: maximal activity energy expenditure
CE: constant error
CRF: cardiorespiratory fitness
ECG: electrocardiogram
GXT: graded exercise test
HR: heart rate
PRESS: predicted residual error sum of square
SEE: SE of estimate
VO2max: maximal oxygen uptake

Edited by G Eysenbach; submitted 08.01.19; peer-reviewed by G Signorelli, M Altini, K Lu; comments to author 28.03.19; revised version received 22.04.19; accepted 17.05.19; published 13.06.19.

Please cite as:
Kwon SB, Ahn JW, Lee SM, Lee J, Lee D, Hong J, Kim HC, Yoon HJ
Estimating Maximal Oxygen Uptake From Daily Activity Data Measured by a Watch-Type Fitness Tracker: Cross-Sectional Study
JMIR Mhealth Uhealth 2019;7(6):e13327
URL: https://mhealth.jmir.org/2019/6/e13327/
doi:10.2196/13327
PMID:31199336

Original Paper

Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort

Salvatore Tedesco[1], MSc; Marco Sica[1], MSc; Andrea Ancillao[1], PhD; Suzanne Timmons[2], MD; John Barton[1], MEng; Brendan O'Flynn[1], PhD

[1]Tyndall National Institute, University College Cork, Cork, Ireland

[2]Centre for Gerontology and Rehabilitation, University College Cork, Cork, Ireland

Corresponding Author:
Salvatore Tedesco, MSc
Tyndall National Institute, University College Cork
Lee Maltings, Prospect Row
Cork, T12R5CP
Ireland
Phone: 353 212346286
Email: salvatore.tedesco@tyndall.ie

Abstract

Background: Few studies have investigated the validity of mainstream wrist-based activity trackers in healthy older adults in real life, as opposed to laboratory settings.

Objective: This study explored the performance of two wrist-worn trackers (Fitbit Charge 2 and Garmin vivosmart HR+) in estimating steps, energy expenditure, moderate-to-vigorous physical activity (MVPA) levels, and sleep parameters (total sleep time [TST] and wake after sleep onset [WASO]) against gold-standard technologies in a cohort of healthy older adults in a free-living environment.

Methods: Overall, 20 participants (>65 years) took part in the study. The devices were worn by the participants for 24 hours, and the results were compared against validated technology (ActiGraph and New-Lifestyles NL-2000i). Mean error, mean percentage error (MPE), mean absolute percentage error (MAPE), intraclass correlation (ICC), and Bland-Altman plots were computed for all the parameters considered.

Results: For step counting, all trackers were highly correlated with one another (ICCs>0.89). Although the Fitbit tended to overcount steps (MPE=12.36%), the Garmin and ActiGraph undercounted (MPE 9.36% and 11.53%, respectively). The Garmin had poor ICC values when energy expenditure was compared against the criterion. The Fitbit had moderate-to-good ICCs in comparison to the other activity trackers, and showed the best results (MAPE=12.25%), although it underestimated calories burned. For MVPA levels estimation, the wristband trackers were highly correlated (ICC=0.96); however, they were moderately correlated against the criterion and they overestimated MVPA activity minutes. For the sleep parameters, the ICCs were poor for all cases, except when comparing the Fitbit with the criterion, which showed moderate agreement. The TST was slightly overestimated with the Fitbit, although it provided good results with an average MAPE equal to 10.13%. Conversely, WASO estimation was poorer and was overestimated by the Fitbit but underestimated by the Garmin. Again, the Fitbit was the most accurate, with an average MAPE of 49.7%.

Conclusions: The tested well-known devices could be adopted to estimate steps, energy expenditure, and sleep duration with an acceptable level of accuracy in the population of interest, although clinicians should be cautious in considering other parameters for clinical and research purposes.

(JMIR Mhealth Uhealth 2019;7(6):e13084) doi:10.2196/13084

KEYWORDS

aging; fitness trackers; wristbands; older adults; wearable activity trackers; Fitbit; Garmin; energy expenditure; physical activity; sleep

Introduction

Fitness trackers are popular devices used by athletes and the general public to monitor their physical activity levels, sport performance, and even their general health status in real time, with the latter having the potential to also predict the person's future health status [1]. Compared to other body positions, the wrist has been identified as the most suitable location for enhancing user acceptability and the user-friendliness of the device [2]. Common consumer-level, wrist-worn devices typically provide data on step count, distance traveled, number of floors climbed, and minutes of physical activity, as well as sport-related activity recognition, physiological measurements, energy expenditure, and sleep patterns. This information can promote a healthier lifestyle or an optimal training program through user-friendly visual feedback of current status or performance compared to set targets.

Fitness trackers based on motion sensors are also being used for monitoring biomechanical quantities of clinical interest, such as gait analysis applications [3,4], indirect estimation of ground reaction forces, and posture in general [5,6].

Although the number of studies investigating the validity and reliability of different fitness trackers is growing, the majority of the evidence is limited to young and middle-aged adult populations, mostly in good health [7-9]. Considering the multiple applications of wrist-based technology and its potential adoption in health care, and with an aging population, it is important to investigate the use of these devices in different populations, such as older people [10]. Although older adults perceive commercial trackers as useful and acceptable [11,12], older person-specific activity trackers are still limited [13].

Few studies have investigated the validity of mainstream wrist-based activity trackers in healthy older adults [14,15]. However, such investigations mainly involved a protocol structured around a number of daily activities simulated or recreated in a laboratory environment. Studies that investigated fitness trackers' performance when used by older people in their home environment, where older adults can perform their real daily routine, are scarce and mainly limited to step-counting features [16]. This study reviewed the validity and reliability of consumer-grade activity trackers in older community-dwelling adults through seven observational studies, of which only five studied free-living settings for a monitoring period of between 3 and 7 days.

For example, Paul et al [17] reported that the average steps per day measured over 7 days in a community-dwelling older adult population with a Fitbit and an ActiGraph showed excellent agreement, with the ActiGraph undercounting steps compared against participants' physical activity logs.

In another study, a Fitbit Flex and an ActiGraph were worn by a cohort of cardiac patients and their family members to measure steps and moderate-to-vigorous physical activity (MVPA) levels for 4 days [18]. It showed a significant correlation for step counts but lower values for MVPA, with the Fitbit Flex slightly overestimating both parameters.

Boeselt et al [19] compared a Polar A300 with a BodyMedia SenseWear in a cohort of patients with chronic obstructive pulmonary disease (mean age 66.4 years). Participants used the devices for three consecutive days, measuring steps, calories burned, daily activity time, and metabolic equivalents. The study showed a high correlation for step count and calories burned.

Farina and Lowry [20] compared the accuracy of step counts from two consumer-level activity monitors (Misfit Shine on both wrist and waist, and Fitbit Charge HR on the wrist) against two waist-worn reference devices (ActiGraph GT3X+ and New Lifestyle NL2000i) in healthy, community-dwelling older adults in free-living conditions over seven consecutive days. All consumer-level activity monitors positively correlated with reference devices. Compared to the ActiGraph GT3X+, the waist-worn Misfit Shine displayed the highest agreement, whereas the wrist-worn devices showed poorer performances.

Finally, Burton et al [21] reported good reliability and validity for the Fitbit Flex and Fitbit Charge HR compared against a GENEactiv accelerometer in a free-living environment over 14 days. Step count, distance traveled, MVPA minutes, and sleep were measured. Good strength of agreement was found for total distance and steps (obtained with the fitness tracker) and the MVPA estimated by the GENEactiv.

It is evident that a comparative analysis of mainstream trackers worn by healthy older people in a more ecologically valid environment is needed. This study aims to investigate the reliability and accuracy of the wrist-based Fitbit Charge 2 and the Garmin vivosmart HR+ activity trackers in the estimation of daily step count, total calorie expenditure, MVPA, and sleep parameters within a home environment in a cohort of older adults.

Methods

Participants

This study was based on a sample of 20 healthy older people (9 males, 11 females). Volunteers were recruited via a general invitation email, posters, and word of mouth to exstaff at University College Cork (Cork, Ireland) and their relatives, and also through local social and voluntary groups that had older adults as members. They were informed of the study by the Centre for Gerontology and Rehabilitation in University College Cork.

For the cohort, the inclusion criteria were age 65 years and older, with no history of neurological or other disorders or disability that could affect the participant's movements, and in good general health. Before participation, volunteers received an oral and written explanation of the study protocol, and written consent was obtained. Sociodemographic information was collected on gender, age, weight, height, and dominant arm. The study received approval by the Clinical Research Ethics Committee at the University College Cork. Demographic information on the participants who completed the study protocol is presented in Table 1.

Table 1. Participant characteristics (N=20).

Characteristic	Males (n=9)	Females (n=11)
Age (years), mean (SD)	70.2 (2.9)	71.1 (3.1)
Height (cm), mean (SD)	176.8 (4.7)	161.1 (6.9)
Weight (kg), mean (SD)	81.6 (12.5)	65.9 (9.3)
People with right dominant arm, n	9	10
People with left dominant arm, n	0	1

Equipment

The following consumer-level and research-grade devices were selected for comparison:

1. Fitbit Charge 2 (Fitbit Inc, San Francisco, CA, USA): a wrist-based device with a large organic light-emitting diode screen featuring heartrate monitoring and tracking of steps, distance, calories burned, floors climbed, active minutes, and sleep duration.
2. Garmin vivosmart HR+ (Garmin, Olathe, KS, USA): a wrist-based device that monitors heart rate, calories burned, intensity of fitness activities, distance, time, and pace for indoor or outdoor activities.
3. ActiGraph GT9X-BT (ActiGraph LLC, Pensacola, FL, USA): a research-grade activity monitor based on motion sensors that provide raw data and numerous activity and sleep measures (eg, activity counts, energy expenditure, steps taken, activity/sedentary bouts, sleep latency/efficiency) via publicly available validated algorithms. The device can be worn on the waist, hip, wrist, ankle, or thigh.
4. New-Lifestyles NL-2000i Activity Monitor (New-Lifestyles Inc, Lee's Summit, MO, USA): a research-grade measurement tool for data collection based on a three-dimensional piezoelectric accelerometer. The device can provide details on steps, MVPA, total/active calories, and distance, and can store data for 7 to 14 days. Examples of studies that have validated the pedometer are available [20,22-24].

Experimental Protocol

Two consumer-based wrist-mounted brands were tested (Fitbit and Garmin), worn on the nondominant arm. The trackers' position on the wrist was randomized. The dominant side of the waist (midaxillary line) and the dominant wrist are reported to be optimal for monitoring energy expenditure, MVPA, and sleep in older adults [25-28]; therefore, two ActiGraph monitors were located in these positions as a reference for those parameters. Energy expenditure, MVPA levels, and steps were measured with the ActiGraph on the waist; sleep parameters were extrapolated from the ActiGraph on the wrist. The New-Lifestyles NL-2000i tracker was also worn on the dominant waist (midaxillary line) and was considered as a reference for step counting. Figure 1 illustrates the body positions of the different devices on a participant.

The algorithm adopted by the ActiGraph for estimating energy expenditure was based on the method designed by Crouter et al [29] and also considered in Patterson et al [26]. This method provides estimations expressed in metabolic equivalent (MET), which are later converted into total calories per day. Likewise, the algorithms of Troiano et al [30] and Cole-Kripke et al [31] were considered for estimating MVPA levels and sleep parameters, respectively, through the ActiGraph accelerometer [28,32]. By definition, MVPA level is the amount of time spent performing any activities requiring more than 3 METs, which according to Troiano et al [30] is defined by the ActiGraph by at least 2020 counts per minute. Finally, to guarantee a fair comparison for the different trackers, only the sleep parameters measured by both the Fitbit and Garmin were analyzed. Those parameters were the total sleep time (TST), and the wake after sleep onset (WASO). Participants were asked to complete a sleep diary as well, and the in-bed and out-bed information required by the Cole-Kripke method were input manually according to the values reported in the sleep diary.

The devices were attached on to the person in the morning for data collection and were returned to the researchers the following morning. All trackers were removed by the participants during bathing, whereas only the trackers on the waist were removed during sleep.

Nonwear periods were defined as 90 minutes or more with no activity counts [33]. A valid day was defined as 10 wearing hours or more in a 24-hour period [30].

Statistical Analysis

Descriptive statistics were run on the computed parameters. The following indicators were computed for each parameter and device: mean estimated value with related standard deviation (SD), mean bias with standard deviation, mean percentage error (MPE) with standard deviation, and mean absolute percentage error (MAPE). Intraclass correlation (ICC[2,1]) was performed for each tracker compared against all other devices and the criterion as well. The related 95% confidence intervals (CIs) were also computed. ICC values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively [34]. Bland-Altman plots were also obtained for every parameter comparing all the possible permutations of trackers and the criterion. All statistical analyses were performed using MATLAB (MathWorks, Natick, MA, USA).

Figure 1. Placement of devices on a participant.

Results

Overall, 20 participants took part in the data collection. All participants were white of Irish and British ancestry. Data collection was carried out at the Tyndall National Institute between April 2018 and August 2018.

Table 2 shows the mean values measured, mean error, mean percentage error, and MAPE with related standard deviations for the activity monitors and each parameter. Likewise, Table 3 shows the related ICCs with the 95% CI for each tracker and every parameter. Figure 2 displays the MAPE.

The average wear time for each tracker among all participants was mean 963 (SD 102) minutes per day; thus, every monitored day for each participant was deemed a valid test day above the 10 wearing hours threshold. All monitored days were on weekdays.

For step counting, all the trackers were highly correlated with one another (ICCs>0.89). Although the Fitbit tended to overcount steps (MPE=12.36%), the Garmin and ActiGraph undercounted with a MPE of 9.36% and 11.53%, respectively. For the MAPE, the Garmin and ActiGraph were slightly more accurate with mean MAPEs of 12.89% and 14.23%, respectively. Therefore, all the considered activity trackers can accurately capture steps when worn on the nondominant wrist.

However, the Garmin had poor ICC values when comparing energy expenditure against the New-Lifestyles NL-2000i and the criterion. Likewise, similar results were shown when comparing the New-Lifestyles NL-2000i and the criterion. Conversely, the Fitbit had moderate-to-good ICCs compared against the other activity trackers. All the tested activity monitors underestimated the amount of calories burned, with the Fitbit showing the best results with a MAPE of 12.25%.

For MVPA level estimation, the Fitbit and Garmin were highly correlated (ICC=0.96), whereas the New-Lifestyles NL-2000i showed poor correlation with these two devices. However, all monitors were moderately correlated against the criterion. The Fitbit and Garmin overestimated the minutes of MVPA activity (mean 12.63, SD 28.31 and mean 13.8, SD 35.7 minutes per day), whereas New-Lifestyles-2000i underestimated, although it showed the best MAPE results, at 45.45%, suggesting that consumer-grade activity trackers may not be reliable in estimating MVPA in older adults.

When analyzing the sleep parameters, the ICCs were poor for all cases, except when comparing the Fitbit to the criterion, which showed a moderate agreement. The TST was slightly overestimated with the Fitbit (mean 5.72, SD 49.11 minutes), although it provided good results with a mean MAPE equal to 10.13%. Conversely, the WASO estimation was poorer; it was overestimated by the Fitbit but underestimated by the Garmin. Again, the Fitbit was the most accurate, with a mean MAPE of 49.7%.

The Bland-Altman plots are shown in Figures 3-6 for steps, energy expenditure, MVPA, and sleep parameters, respectively, and summarized in Table 4.

Table 2. Mean values measured, mean error, mean percentagage error, and mean absolute percentage error with related standard deviation for each parameter and tracker (N=20).

Parameter	Mean (SD)	Mean error (SD)	MPE[a] (SD)	MAPE[b] (SD)
Steps				
Fitbit	10088.45 (5067.32)	698.60 (1491.27)	12.36 (18.26)	17.05 (13.71)
Garmin	8834.85 (5067.07)	–555.00 (928.11)	–9.36 (18.55)	12.90 (16.16)
ActiGraph	8375.70 (4716.39)	–1014.15 (846.81)	–11.53 (11.55)	14.23 (7.76)
Energy expenditure (cal)				
Fitbit	2324.25 (547.02)	–175.70 (293.18)	–6.98 (12.67)	12.26 (7.33)
Garmin	2434.65 (804.12)	–65.30 (682.94)	–2.09 (27.27)	20.05 (18.04)
NL2000i	2102.30 (256.93)	–397.65 (306.13)	–14.61 (10.41)	16.70 (6.3)
MVPA[c] (min)				
Fitbit	44.32 (45.83)	12.63 (28.31)	31.78 (103.51)	75.74 (75.32)
Garmin	39.10 (57.18)	13.80 (35.70)	6.42 (109.75)	91.98 (47.13)
NL2000i	20.30 (19.03)	–11.70 (18.45)	–41.43 (35.37)	45.45 (29.67)
TST[d] (min)				
Fitbit	389.83 (59.33)	5.72 (49.11)	2.27 (13.66)	10.14 (9.12)
Garmin	442.83 (48.64)	55.39 (48.07)	15.85 (14.56)	16.86 (13.3)
WASO[e] (min)				
Fitbit	49.21 (16.5)	0.21 (22.15)	25.02 (84.72)	49.73 (72.03)
Garmin	13.14 (8.93)	–35.86 (22.13)	–66.89 (28.61)	66.89 (28.61)

[a]MPE: mean percentage error.

[b]MAPE: mean absolute percentage error.

[c]MVPA: moderate-to-vigorous physical activity.

[d]TST: total sleep time.

[e]WASO: wake after sleep onset.

Table 3. Intraclass correlation (ICC) and 95% CI for each parameter and tracker (N=20).

Parameter	ICC (95% CI)				
	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Step					
Fitbit	__[a]	0.93 (0.65, 0.98)	0.89 (0.39, 0.97)	—	0.95 (0.87, 0.98)
Garmin	—	—	0.98 (0.94, 0.99)	—	0.98 (0.93, 0.99)
ActiGraph (waist)	—	—	—	—	0.97 (0.59, 0.99)
Energy expenditure					
Fitbit	—	0.64 (0.29, 0.84)	—	0.66 (0.21, 0.86)	0.80 (0.48, 0.92)
Garmin	—	—	—	0.32 (–0.07, 0.65)	0.48 (0.05, 0.76)
NL2000i	—	—	—	—	0.45 (–0.10, 0.78)
MVPA[b]					
Fitbit	—	0.96 (0.86, 0.99)	—	0.41 (–0.01, 0.72)	0.69 (0.35, 0.87)
Garmin	—	—		0.46 (–0.11, 0.82)	0.68 (0.18, 0.91)
NL2000i	—	—	—	—	0.60 (0.18, 0.83)
TST[c]					
Fitbit	—	0.43 (–0.10, 0.77)	—	—	0.67 (0.30, 0.86)
Garmin	—	—	—	—	0.42 (–0.10, 0.76)
WASO[d]					
Fitbit	—	<0.01 (–0.08, 0.18)	—	—	0.32 (–0.28, 0.72)
Garmin	—	—	—	—	0.01 (–0.10, 0.25)

[a]Not applicable.

[b]MVPA: moderate-to-vigorous physical activity.

[c]TST: total sleep time.

[d]WASO: wake after sleep onset.

Figure 2. Mean absolute percentage error (MAPE) with standard deviation for each parameter and tracker. EE: energy expenditure; MVPA: moderate-to-vigorous physical activity; TST: total sleep time; WASO: wake after sleep onset.

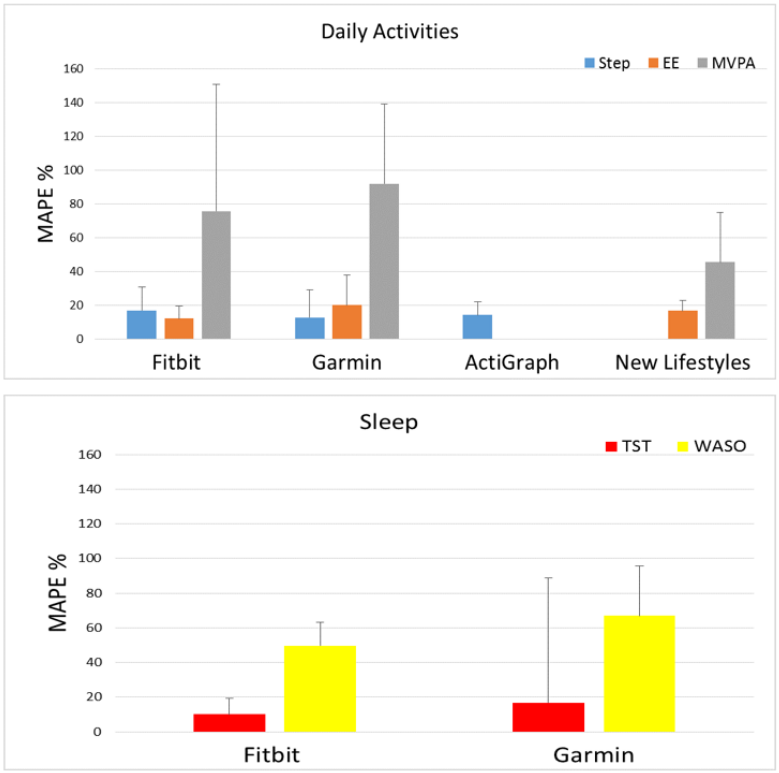


Figure 3. Bland-Altman plots for steps.

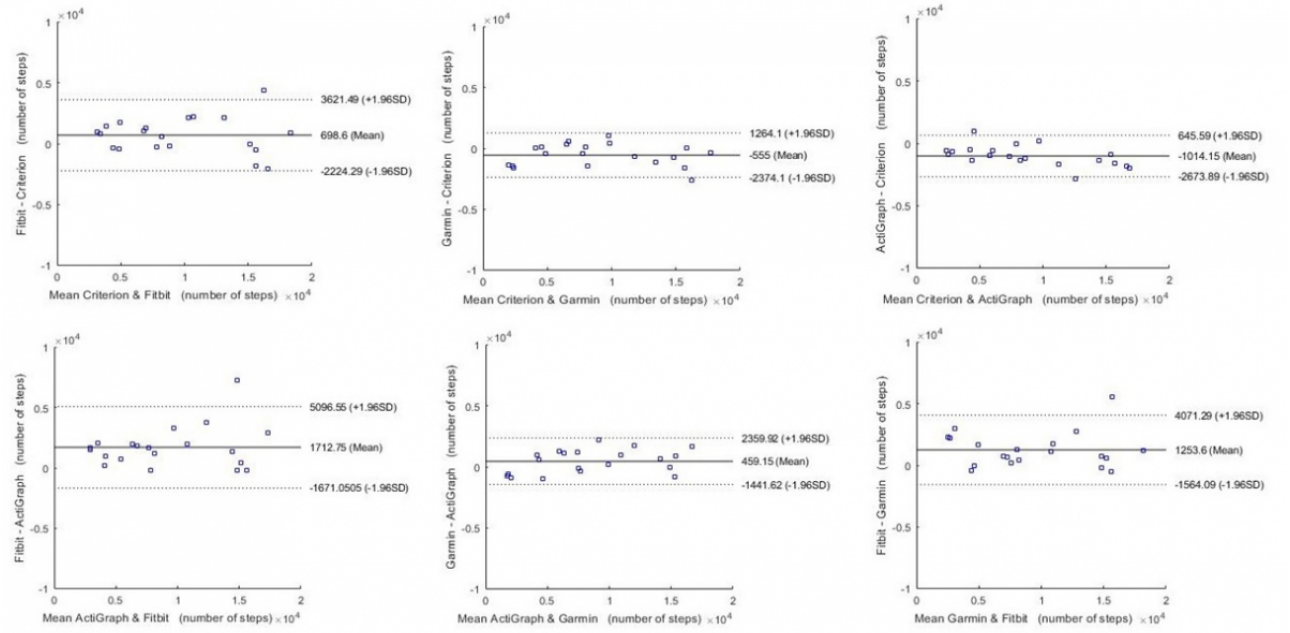


Figure 4. Bland-Altman plots for energy expenditure.

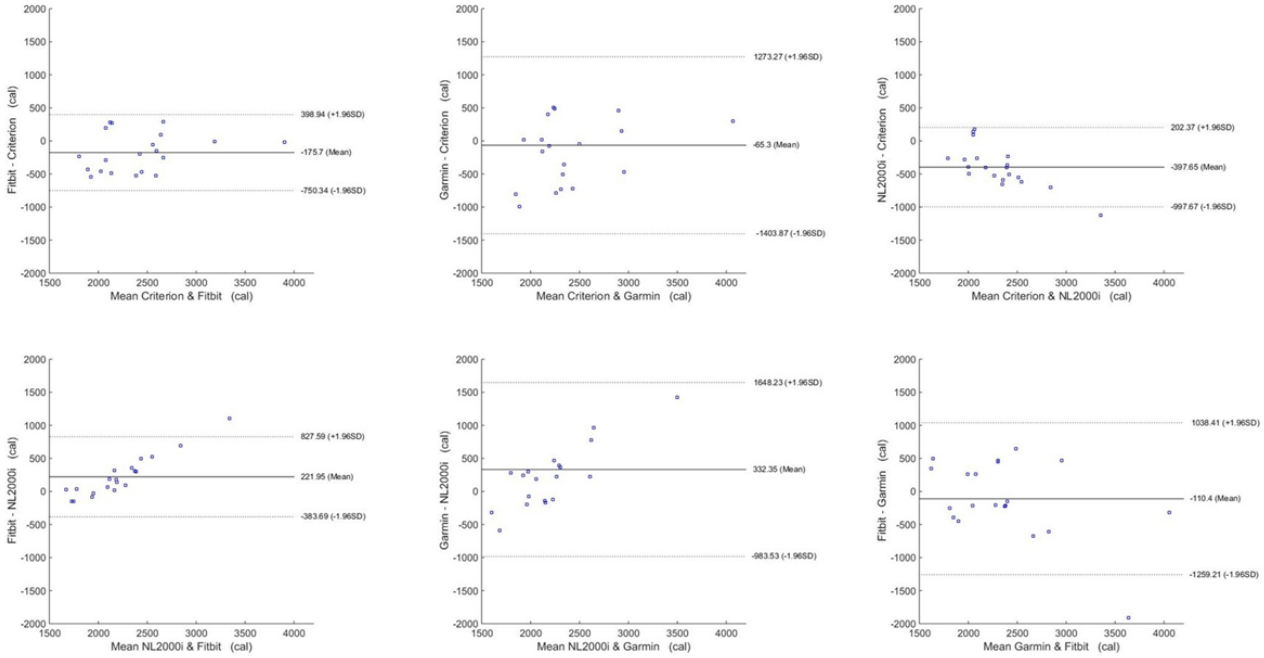


Figure 5. Bland-Altman plots for moderate-to-vigorous physical activity.

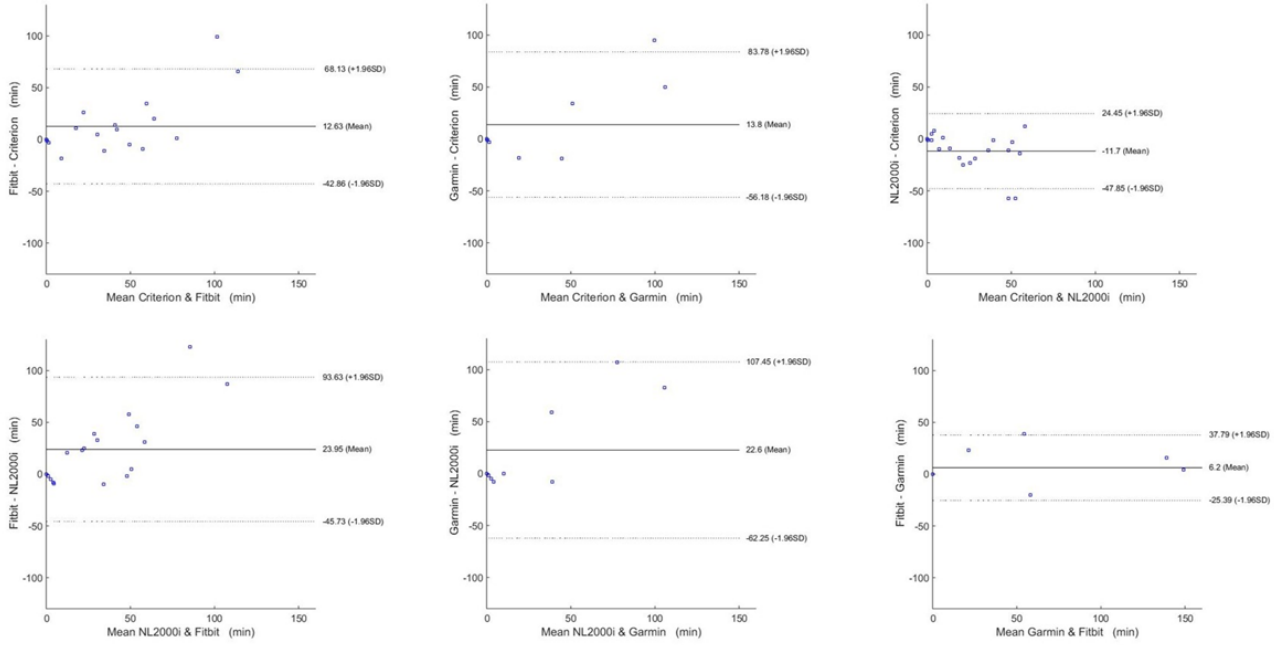


Figure 6. Bland-Altman plots for sleep parameters. Top row: total sleep time; bottom row: wake after sleep onset.

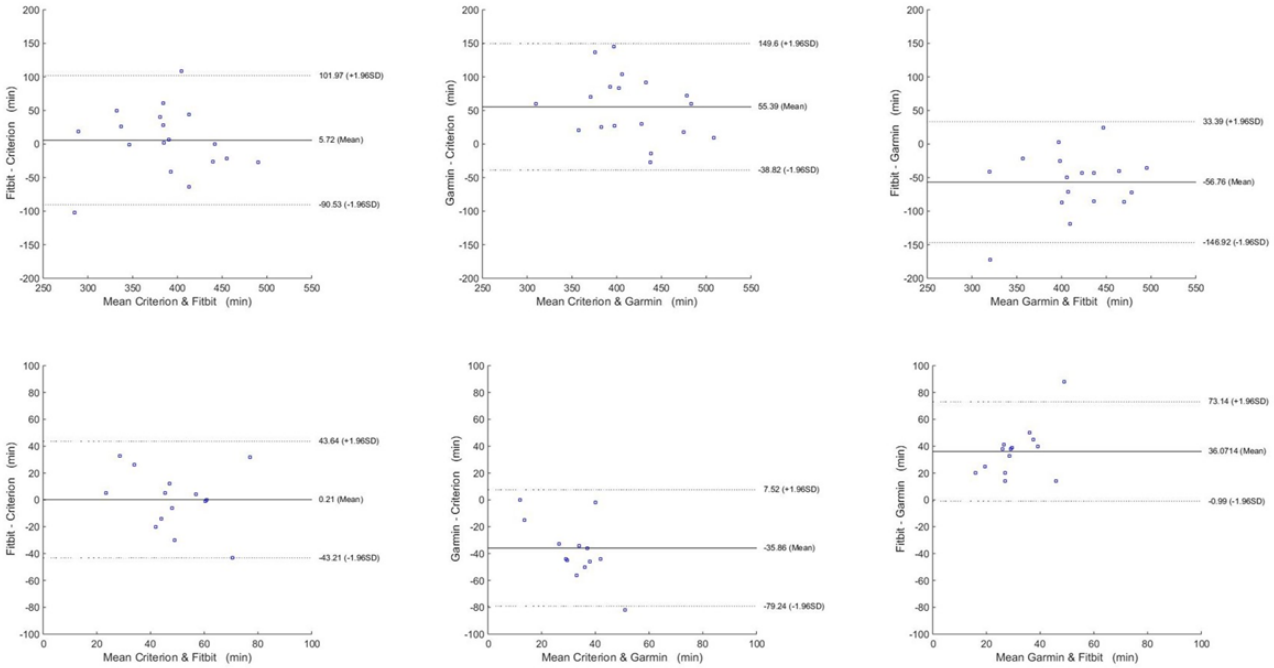


Table 4. Bland-Altman plots summary for each parameter and tracker (N=20).

Parameter	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Step					
Fitbit					
Mean error (SD)	__[a]	1253.60 (1437.60)	1712.75 (1726.43)	—	698.60 (1491.27)
95% LoA[b]	—	–1564.09, 4071.29	–1671.05, 5096.55	—	–2224.29, 3621.49
Garmin					
Mean error (SD)	—	—	459.15 (969.78)	—	–555 (928.11)
95% LoA	—	—	–1441.62, 2359.92	—	–2374.10, 1264.10
ActiGraph (waist)					
Mean error (SD)	—	—	—	—	–1014.15 (846.81)
95% LoA	—	—	—	—	–2673.89, 645.59
Energy expenditure (cal)					
Fitbit					
Mean error (SD)	—	–110.40 (586.13)	—	221.95 (309.00)	–175.7 (293.18)
95% LoA	—	–1259.21, 1038.41	—	–383.69, 827.59	–750.34, 398.94
Garmin					
Mean error (SD)	—	—	—	332.35 (671.37)	–65.3 (682.94)
95% LoA	—	—	—	–983.53, 1648.23	–1403.87, 1273.27
NL2000i					
Mean error (SD)	—	—	—	—	–397.65 (306.13)
95% LoA	—	—	—	—	–997.67, 202.37
MVPA[c] (min)					
Fitbit					
Mean error (SD)	—	6.20 (16.12)	—	23.95 (35.55)	12.63 (28.31)
95% LoA	—	–25.39, 37.79	—	–45.73, 93.63	–42.86, 68.13
Garmin					
Mean error (SD)	—	—	—	22.60 (43.29)	13.80 (35.70)
95% LoA	—	—		–62.25, 107.45	–56.18, 83.78
NL2000i					
Mean error (SD)	—	—	—	—	–11.70 (18.45)
95% LoA	—	—	—	—	–47.85, 24.45
TST[d] (min)					
Fitbit					
Mean error (SD)	—	–56.76 (45.99)	—	—	5.72 (49.11)
95% LoA	—	–146.92, 33.39	—	—	–90.53, 101.97
Garmin					
Mean error (SD)	—	—	—	—	55.39 (48.07)
95% LoA	—	—	—	—	–38.82, 149.60
WASO[e] (min)					
Fitbit					
Mean error (SD)	—	36.07 (18.91)	—	—	0.21 (22.15)
95% LoA	—	–0.99, 73.14	—	—	–43.21, 43.64

Parameter	Fitbit	Garmin	ActiGraph (waist)	NL2000i	Criterion
Garmin					
Mean error (SD)	—	—	—	—	–35.86 (22.13)
95% LoA	—	—	—	—	–79.24, 7.52

[a]Not applicable.

[b]LoA: limits of agreement.

[c]MVPA: moderate-to-vigorous physical activity.

[d]TST: total sleep time.

[e]WASO: wake after sleep onset.

Discussion

This investigation is one of the first studies to investigate the reliability and accuracy of two consumer-level, wrist-based activity trackers in the estimation of daily step count, total calories expenditure, MVPA, and sleep parameters in a home environment for a cohort of healthy older adults.

Results show that the mainstream monitors may be adopted to estimate steps, energy expenditure, and some sleep parameters (eg, TST) with a certain level of accuracy in a healthy older adult population in free-living settings, whereas other variables (MVPA and WASO) may show excessively large errors.

The measured mean values are largely consistent with data reported studying the same population of interest in other studies adopting non-consumer-level technologies for energy expenditure [26], steps and MVPA [35], and sleep analysis [36].

Regarding step-counting performance, all the trackers presented a good strength of agreement among one another and against the reference device. Also, this study confirms previous findings [17,18] indicating a slight overcounting by the Fitbit device and an undercounting by the ActiGraph. In absolute terms, as shown by the MAPE, there is no significant difference between the Fitbit and the Garmin monitors when monitoring steps.

Similar considerations could be drawn for energy expenditure; however, only the Fitbit shows moderate-to-good agreement with the other trackers. The Fitbit underestimated energy expenditure in our study, confirming findings illustrated in the review by Feehan et al [37], which cited a number of studies in which Fitbits worn on the wrist in free-living settings slightly underestimated METs by 7% (when compared against doubly labeled water), and showed a −10% measurement error (against the SenseWear), and provided MAPE values varying from 16% to 30% when compared with measurements from an ActiGraph or Actiheart accelerometer. However, most of the studies reviewed considered healthy adults and not older adults; thus, the lower MAPE values reported in our study (mean 12.25%, SD 7.33%) may be due to the generally limited amount of moderate-to-vigorous activity performed by older adults. Fitbit showed the narrowest limit of agreement among the trackers, which indicated the device could underestimate the amount of calories per day up to 750.34 kcal and overestimate up to 398.94 kcal.

Due to its many health benefits reported, MVPA may represent an important aspect in people's life and, with aging, this may become even more useful to guarantee independent living and prevention of noncommunicable diseases [35]. The current national MVPA recommendations consider a threshold of 30 minutes per day or 150 minutes per week. Therefore, a correct and reliable estimation of MVPA bouts helps support behavior change techniques applied to sedentary older adults. All the trackers considered (Fitbit, Garmin on the wrist, and New-Lifestyles NL-2000i on the waist) were moderately correlated with the reference. However, the Fitbit and Garmin showed an excellent strength of agreement between each other. The Fitbit and Garmin tended to overestimate MVPA with an average error of 12.63 minutes per day and 13.8 minutes per day, consistent with previously reported results [18], whereas the New-Lifestyles NL-2000i underestimated by 11.7 minutes per day. However, due to the limited moderate-to-vigorous activities performed, MAPE values are large, especially for the wrist-worn devices. The MAPE was mean 75.73% (SD 75.31%) and mean 91.98% (SD 47.13%) for the Fitbit and Garmin, respectively, confirming the large overestimation errors observed in Fitbit devices estimating MVPA in free-living settings compared with an ActiGraph accelerometer in healthy young adults and older adults living with a variety of chronic diseases (MAPEs >30%) [37]. The waist-worn device showed slightly better results both in terms of MAPE and limits of agreement (–47.85 to 24.45).

Finally, aging also impacts sleep, and changes occur in sleep patterns with aging (for example, decrease in the amount of slow wave sleep, increases in non-rapid eye movement sleep, increase in the number of spontaneous arousals, changes in the normal circadian sleep cycle) [38]. Moreover, older adults are more prone to develop sleep-related respiratory disorders, which are associated with cardiovascular disease, metabolic disorders, and impaired neurocognition [38]. Thus, low-cost, unobtrusive, and effective sleep monitoring devices such as consumer-level activity trackers are ideal for providing insightful details on the normal changes in sleeping patterns with advancing age. Between the Fitbit and the Garmin, the Fitbit was moderately correlated with the ActiGraph worn on the wrist, and only for the estimation of the sleeping time. TST was overestimated by a mean 5.72 minutes per day with a MAPE equal to 10.13% (SD 9.12%), which are largely consistent with findings reported in other studies adopting Fitbit devices to investigate sleep measurement accuracy in healthy young adults in free-living settings (MAPE approximately 10%) [37]. The Garmin showed larger errors with a MAPE of 16.8% (SD 13.3%). In contrast, WASO measurements were poorly correlated against the ActiGraph for both devices. Although the lowest mean error

was 0.21 minutes per day for the Fitbit, MAPE was large (mean 49.7%, SD 72%) due to the generally limited amount of time spent awake overnight. Conversely, the Garmin significantly underestimated the measurements. Limits of agreement were similar for both trackers for both parameters. These performances may not be suitable for clinical-grade investigations because they require accurate measurements for supporting the decision-making process. For example, WASO is typically adopted as a criterion for discriminating insomnia and normal-sleeper groups (the general threshold is WASO ≥31 minutes per day occurring at least three times per week for at least 6 months) [39]; thus, the WASO estimation inaccuracy may hinder the adoption of mainstream wristband devices for clinical assessments in populations expected to have abnormal sleep patterns [40].

It is worth clarifying that there is no universally accepted definition of an acceptable degree of error for physical activity wearable devices. Some studies recommend that an acceptable measurement error under controlled conditions or for research purposes is within ±3% [41,42] and under free-living conditions is within ±10% [41,42]. Other studies recommend that mean errors of less than 20% have acceptable validity for clinical purposes [43]. This investigation considers the validity criteria between the tested and criterion physical activity measures for clinical purposes when the mean error is less than 20%. Results suggest that the tested devices could be adopted to estimate steps, energy expenditure, and sleep duration with an acceptable level of accuracy in the population of interest, whereas clinicians should be cautious in considering other parameters (eg, MVPA, awakenings) for clinical and research purposes. Although performance estimation is modest in some variables, it may still be adequate for guidance purposes. For instance, the ever-growing acceptance of wearable technologies by older people may push the adoption of wrist-worn trackers in behavior change investigations [11,44].

This study was limited to healthy older adults. As a consequence, it is difficult to indicate if these findings are generalizable to less active older adults or impaired or hospitalized older adults. Indeed, as shown in the literature, step-counting accuracy in people using a walking aid in a laboratory-structured protocol represents a challenge for all consumer-level trackers as evidenced by large MAPE values. Moreover, the small number of studied participants and the reduced intervention duration may also limit the generalizability of these findings. Thus, further studies would be needed to investigate activity trackers' performance in a large cohort and also in nonhealthy populations.

Although the most common commercial trackers were considered in this study, it is difficult to indicate if results may translate to other consumer monitors on the market, due to the different algorithms they may employ.

This analysis was limited to some health parameters, whereas other variables, which may be of interest in older adults, could not be taken into account due to the lack of a gold-standard for nonlaboratory settings. Some examples are sedentary bouts, light activity bouts, the amount of time spent in different postures, distance traveled, speed, additional sleep measures (eg, sleep efficiency, sleep latency), and physiological measures, such as continuous heart rate measurements, blood oxygen saturation levels, galvanic skin response, blood pressure, or photoplethysmography, and these should be further investigated in future studies.

This study explored the performance of two wrist-worn trackers (Fitbit Charge2 and Garmin vivosmart HR+) estimating steps, energy expenditure, MVPA levels, and sleep parameters against gold-standard technologies in a free-living environment in a cohort of healthy participants aged 65 years and older.

This study confirmed that the wrist-worn devices are effective in estimating steps, energy expenditure, and some sleep parameters with a certain level of accuracy in healthy older adults (lower MAPE values: 12.89% for step counting with the Garmin, 12.25% for energy expenditure with the Fitbit, and 10.13% for TST estimation with the Fitbit). The results were coherent with previous studies, and the observed accuracy was acceptable for monitoring everyday activities. However, clinicians should be cautious in considering other parameters (eg, MVPA levels and WASO) for clinical and research purposes.

Acknowledgments

The authors would like to thank all the participants who voluntarily took part in the study. This publication was from research supported by EU H2020-funded project ProACT under grant agreement No 689996. Aspects of this work have been supported in part by a research grant from Science Foundation Ireland (SFI) and it is cofunded under the European Regional Development Fund under Grant Number 13/RC/2077. Aspects of this work have been supported in part by INTERREG NPA funded project SenDOC.

Conflicts of Interest

None declared.

References

1. Kamišalić A, Fister I, Turkanović M, Karakatič S. Sensors and functionalities of non-invasive wrist-wearable devices: a review. Sensors (Basel) 2018 May 25;18(6):1714 [FREE Full text] [doi: 10.3390/s18061714] [Medline: 29799504]

2. Alley S, Schoeppe S, Guertler D, Jennings C, Duncan MJ, Vandelanotte C. Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. BMJ Open 2016 Dec 07;6(7):e011243 [FREE Full text] [doi: 10.1136/bmjopen-2016-011243] [Medline: 27388359]
3. Tao W, Liu T, Zheng R, Feng H. Gait analysis using wearable sensors. Sensors (Basel) 2012 Feb;12(2):2255-2283 [FREE Full text] [doi: 10.3390/s120202255] [Medline: 22438763]
4. Tedesco S, Urru A, Clifford A, O'Flynn B. Experimental validation of the Tyndall Portable Lower-limb Analysis System with wearable inertial sensors. Procedia Engineer 2016;147:208-213. [doi: 10.1016/j.proeng.2016.06.215]
5. Fortune E, Lugade VA, Kaufman KR. Posture and movement classification: the comparison of tri-axial accelerometer numbers and anatomical placement. J Biomech Eng 2014 May;136(5):051003 [FREE Full text] [doi: 10.1115/1.4026230] [Medline: 24337255]
6. Ancillao A, Tedesco S, Barton J, O'Flynn B. Indirect measurement of ground reaction forces and moments by means of wearable inertial sensors: a systematic review. Sensors (Basel) 2018 Aug 05;18(8):E2564 [FREE Full text] [doi: 10.3390/s18082564] [Medline: 30081607]
7. Xie J, Wen D, Liang L, Jia Y, Gao L, Lei J. Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: comparative study. JMIR Mhealth Uhealth 2018 Apr 12;6(4):e94 [FREE Full text] [doi: 10.2196/mhealth.9754] [Medline: 29650506]
8. Rosenberger ME, Buman MP, Haskell WL, McConnell MV, Carstensen LL. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc 2016 Mar;48(3):457-465 [FREE Full text] [doi: 10.1249/MSS.0000000000000778] [Medline: 26484953]
9. Tedesco S, Sica M, Garbay T, Barton J, O'Flynn B. A comprehensive comparison of commercial wrist-worn trackers in a young cohort in a lab-environment. 2018 Mar 19 Presented at: IEEE PERCOM Workshops 2018; March 19-23, 2018; Athens, Greece. [doi: 10.1109/PERCOMW.2018.8480386]
10. Malwade S, Abdul S, Uddin M, Nursetyo A, Fernandez-Luque L, Zhu X, et al. Mobile and wearable technologies in healthcare for the ageing population. Comput Methods Programs Biomed 2018 Jul;161:233-237 [FREE Full text] [doi: 10.1016/j.cmpb.2018.04.026] [Medline: 29852964]
11. Ehn M, Eriksson LC, Åkerberg N, Johansson A. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. JMIR Mhealth Uhealth 2018 Feb 01;6(2):e34 [FREE Full text] [doi: 10.2196/mhealth.8345] [Medline: 29391342]
12. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. JMIR Mhealth Uhealth 2017 Nov 15;5(11):e173 [FREE Full text] [doi: 10.2196/mhealth.8211] [Medline: 29141837]
13. Tedesco S, Barton J, O'Flynn B. A review of activity trackers for senior citizens: research perspectives, commercial landscape and the role of the insurance industry. Sensors (Basel) 2017 Jun 03;17(6):1277 [FREE Full text] [doi: 10.3390/s17061277] [Medline: 28587188]
14. Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. Res Gerontol Nurs 2015 May;8(5):213-219. [doi: 10.3928/19404921-20150429-03] [Medline: 25942386]
15. Floegel TA, Florez-Pregonero A, Hekler EB, Buman MP. Validation of consumer-based hip and wrist activity monitors in older adults with varied ambulatory abilities. J Gerontol A Biol Sci Med Sci 2017 Feb;72(2):229-236. [doi: 10.1093/gerona/glw098] [Medline: 27257217]
16. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
17. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med 2015 Jul;1(1):e000013 [FREE Full text] [doi: 10.1136/bmjsem-2015-000013] [Medline: 27900119]
18. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol 2016 Sep;23(14):1476-1485. [doi: 10.1177/2047487316634883] [Medline: 26907794]
19. Boeselt T, Spielmanns M, Nell C, Storre JH, Windisch W, Magerhans L, et al. Validity and usability of physical activity monitoring in patients with chronic obstructive pulmonary disease (COPD). PLoS One 2016 Jun;11(6):e0157229 [FREE Full text] [doi: 10.1371/journal.pone.0157229] [Medline: 27305105]
20. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. J Aging Phys Act 2018 Jan 01;26(1):128-135. [doi: 10.1123/japa.2016-0344] [Medline: 28595019]
21. Burton E, Hill KD, Lautenschlager NT, Thøgersen-Ntoumani C, Lewin G, Boyle E, et al. Reliability and validity of two fitness tracker devices in the laboratory and home environment for older community-dwelling people. BMC Geriatr 2018 Dec 03;18(1):103 [FREE Full text] [doi: 10.1186/s12877-018-0793-4] [Medline: 29724191]
22. Grant PM, Dall PM, Mitchell SL, Granat MH. Activity-monitor accuracy in measuring step number and cadence in community-dwelling older adults. J Aging Phys Act 2008 Apr;16(2):201-214. [Medline: 18483442]

23. O'Connell S, ÓLaighin G, Quinlan LR. When a step is not a step! Specificity analysis of five physical activity monitors. PLoS One 2017;12(1):e0169616 [FREE Full text] [doi: 10.1371/journal.pone.0169616] [Medline: 28085918]
24. An HS, Jones GC, Kang SK, Welk GJ, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017 Apr;17(3):360-368. [doi: 10.1080/17461391.2016.1255261] [Medline: 27912681]
25. Imboden MT, Nelson NB, Kaminsky LA, Montoye AH. Comparison of four Fitbit and Jawbone activity monitors with a research-grade ActiGraph accelerometer for estimating physical activity and energy expenditure. Br J Sports Med 2018 Jul;52(13):844-850. [doi: 10.1136/bjsports-2016-096990] [Medline: 28483930]
26. Patterson M, Wang W, Ortiz A. Dietary and physical activity outcomes determine energy balance in US adults aged 50-74 years. J Aging Phys Act 2018 Oct 01;26(4):561-569. [doi: 10.1123/japa.2017-0304] [Medline: 29283747]
27. Kim M, Yoshida H, Sasai H, Kojima N, Kim H. Association between objectively measured sleep quality and physical function among community-dwelling oldest old Japanese: A cross-sectional study. Geriatr Gerontol Int 2015 Aug;15(8):1040-1048. [doi: 10.1111/ggi.12396] [Medline: 25312049]
28. Johnson ST, Thiel D, Al Sayah F, Mundt C, Qiu W, Buman MP, et al. Objectively measured sleep and health-related quality of life in older adults with type 2 diabetes: a cross-sectional study from the Alberta's Caring for Diabetes Study. Sleep Health 2017 Dec;3(2):102-106. [doi: 10.1016/j.sleh.2016.12.002] [Medline: 28346155]
29. Crouter SE, Kuffel E, Haas JD, Frongillo EA, Bassett DRJ. Refined two-regression model for the ActiGraph accelerometer. Med Sci Sports Exerc 2010 May;42(5):1029-1037 [FREE Full text] [doi: 10.1249/MSS.0b013e3181c37458] [Medline: 20400882]
30. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
31. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. Sleep 1992 Oct;15(5):461-469. [Medline: 1455130]
32. Qin J, Barbour KE, Nevitt MC, Helmick CG, Hootman JM, Murphy LB, et al. Objectively measured physical activity and risk of knee osteoarthritis. Med Sci Sports Exerc 2018 Dec;50(2):277-283. [doi: 10.1249/MSS.0000000000001433] [Medline: 28976494]
33. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, et al. Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010 Dec;62(12):1724-1732 [FREE Full text] [doi: 10.1002/acr.20305] [Medline: 20806273]
34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
35. Aspvik NP, Viken H, Zisko N, Ingebrigtsen JE, Wisløff U, Stensvold D. Are older adults physically active enough - a matter of assessment method? The Generation 100 Study. PLoS One 2016;11(11):e0167012 [FREE Full text] [doi: 10.1371/journal.pone.0167012] [Medline: 27893785]
36. Marino M, Li Y, Rueschman MN, Winkelman JW, Ellenbogen JM, Solet JM, et al. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. Sleep 2013 Nov 01;36(11):1747-1755 [FREE Full text] [doi: 10.5665/sleep.3142] [Medline: 24179309]
37. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
38. Edwards BA, O'Driscoll DM, Ali A, Jordan AS, Trinder J, Malhotra A. Aging and sleep: physiology and pathophysiology. Semin Respir Crit Care Med 2010 Oct;31(5):618-633 [FREE Full text] [doi: 10.1055/s-0030-1265902] [Medline: 20941662]
39. Lineberger MD, Carney CE, Edinger JD, Means MK. Defining insomnia: quantitative criteria for insomnia severity and frequency. Sleep 2006 Apr;29(4):479-485. [Medline: 16676781]
40. Kang SG, Kang JM, Ko KP, Park SC, Mariani S, Weng J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. J Psychosom Res 2017 Dec;97:38-44. [doi: 10.1016/j.jpsychores.2017.03.009] [Medline: 28606497]
41. Schneider PL, Crouter SE, Bassett DR. Pedometer measures of free-living physical activity: comparison of 13 models. Med Sci Sports Exerc 2004 Feb;36(2):331-335. [doi: 10.1249/01.MSS.0000113486.60548.E9] [Medline: 14767259]
42. Tudor-Locke C, Sisson SB, Lee SM, Craig CL, Plotnikoff RC, Bauman A. Evaluation of quality of commercial pedometers. Can J Public Health 2006;97(1):S10-S15. [doi: 10.17269/cjph.97.1544]
43. Schneider PL, Crouter SE, Lukajic O, Bassett DR. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. Med Sci Sports Exerc 2003 Oct;35(10):1779-1784. [doi: 10.1249/01.MSS.0000089342.96098.C4] [Medline: 14523320]
44. McMahon SK, Lewis B, Oakes JM, Wyman JF, Guan W, Rothman AJ. Assessing the effects of interpersonal and intrapersonal behavior change strategies on physical activity in older adults: a factorial experiment. Ann Behav Med 2017 Jun;51(3):376-390. [doi: 10.1007/s12160-016-9863-z] [Medline: 28188585]

Abbreviations

MAPE: mean absolute percentage error
MVPA: moderate-to-vigorous physical activity
TST: total sleep time
WASO: wake after sleep onset
ICC: intraclass correlation
MET: metabolic equivalent

Edited by G Eysenbach; submitted 11.12.18; peer-reviewed by B Kim, M Ehn, M Alharbi; comments to author 30.01.19; revised version received 08.02.19; accepted 09.02.19; published 19.06.19.

Please cite as:
Tedesco S, Sica M, Ancillao A, Timmons S, Barton J, O'Flynn B
Validity Evaluation of the Fitbit Charge2 and the Garmin vivosmart HR+ in Free-Living Environments in an Older Adult Cohort
JMIR Mhealth Uhealth 2019;7(6):e13084
URL: https://mhealth.jmir.org/2019/6/e13084/
doi:10.2196/13084
PMID:31219048

Original Paper

Consumer Wearable Deployments in Actigraphy Research: Evaluation of an Observational Study

Ciara Duignan[1,2], BSc (Hons); Patrick Slevin[1,2], MSc; Niladri Sett[1], PhD; Brian Caulfield[1,2], PhD

[1]Insight Centre for Data Analytics, University College Dublin, Dublin, Ireland

[2]School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

Corresponding Author:
Ciara Duignan, BSc (Hons)
Insight Centre for Data Analytics
University College Dublin
Belfield
Dublin,
Ireland
Phone: 353 17166500
Email: ciara.duignan@insight-centre.org

Abstract

Background: Consumer wearables can provide a practical and accessible method of data collection in actigraphy research. However, as this area continues to grow, it is becoming increasingly important for researchers to be aware of the many challenges facing the capture of quality data using consumer wearables.

Objective: This study aimed to (1) present the challenges encountered by a research team in actigraphy data collection using a consumer wearable and (2) present considerations for researchers to apply in the pursuit of robust data using this approach.

Methods: The Nokia Go was deployed to 33 elite Gaelic footballers from a single team for a planned period of 14 weeks. A bring-your-own-device model was employed for this study where the Health Mate app was downloaded on participants' personal mobile phones and connected to the Nokia Go via Bluetooth. Retrospective evaluation of the researcher and participant experience was conducted through transactional data such as study logs and email correspondence. The participant experience of the data collection process was further explored through the design of a 34-question survey utilizing aspects of the Technology Acceptance Model.

Results: Researcher challenges included device disconnection, logistics and monitoring, and rectifying of technical issues. Participant challenges included device syncing, loss of the device, and wear issues, particularly during contact sport. Following disconnection issues, the data collection period was defined as 87 days for which there were 18 remaining participants. Average wear time was 79 out of 87 days (90%) and 20.8 hours per day. The participant survey found mainly positive results regarding device comfort, perceived ease of use, and perceived usefulness.

Conclusions: Although this study did not encounter some of the common published barriers to wearable data collection, our experience was impacted by technical issues such as disconnection and syncing challenges, practical considerations such as loss of the device, issues with personal mobile phones in the bring-your-own-device model, and the logistics and resources required to ensure a smooth data collection with an active cohort. Recommendations for achieving high-quality data are made for readers to consider in the deployment of consumer wearables in research.

(JMIR Mhealth Uhealth 2019;7(6):e12190) doi:10.2196/12190

KEYWORDS

wearable electronic device; digital divide; activity trackers; technology; wearable challenges

Introduction

The advancement of wearable technology has brought with it the promise of expanding the capabilities of health care. Wearable sensors have the potential to augment and transform the diagnosis and ongoing management of both physical and physiological conditions. Due to this potential, wearables have been used increasingly for research in many areas of health and performance science. Recently, we note the progression of this work to incorporate activity tracking and its association with

behavioral support and biological markers, respectively [1-3]. Although fitness trackers have been shown to increase physical activity levels, current evidence does not support their use in improving health outcomes beyond clinical interventions [4,5]. There has been a rapid growth of registered studies using consumer wearables, thus the body of published work in this area is expected to continue to mature [6].

Consumer wearables are a familiar, accessible, and cost-effective solution for remote actigraphy measurement, making them an attractive option for research studies [7,8]. However, ensuring the capture of high-quality data with consumer wearables presents a growing challenge. Frequently cited challenges include usability problems, unreliable technology, activity detection, limited device battery power, and user adherence [9,10]. Adherence is a crucial barrier to data capture and can be influenced by factors such as device convenience and comfort, and its charging, complexity, and interaction requirements [10]. User interest and motivation can be central to adherence and can be further determined by aspects such as the perceived usefulness of the device [11].

Knowledge regarding the challenges of data capture using wearables represents a pressing need, especially as research and industry deployments continue to grow in rapidly advancing areas such as digital health and connected health. If fresh understandings are created regarding how to implement, deploy, organize, and execute a wearable-led study, there is an opportunity to enhance deployments and avoid the creation of an unsatisfactory experience and unsuccessful outcomes for researcher and participant.

The aim of this study was to present and discuss the challenges experienced in a longitudinal data collection using a consumer wearable device with a healthy, active cohort where initial recruitment had been established. Second, the study will aim to identify how we might mitigate these issues, with the desired outcome being recommendations to promote high-quality data collection using consumer wearables.

This study was not originally designed to investigate the challenges of deploying a consumer wearable, rather, we set out to conduct a larger observational study where actigraphy was one of the measures included in our data collection. The content presented here has emerged from the experience of conducting this study, where we outline our observations and evaluation of the data capture process.

Methods

Part 1: Original Study

Study Design

The original study conducted was a longitudinal observation of activity and wellness measures in elite Gaelic footballers over a single season. Objective activity and sleep measures were obtained via a consumer wearable device, and self-reported wellness measures were recorded via a commercial mobile application daily. A total of 34 male elite Gaelic footballers (age: mean 23.4, SD 2.8 years) from a single team were recruited as a purposeful sample for this study and were eligible to take part if they were older than 18 years, part of the senior football team, and had the ability to provide informed consent. There were no exclusion criteria. Ethical approval for this study was granted by University College Dublin Human Research Ethics Committee. Participants were required to provide informed written consent before participation and were advised of their right to withdraw from the study at any time.

Device

Participants were an active group of sportspeople; a diverse cohort including students and working professionals in both sedentary and active jobs. Due to the duration of the data collection and the nature of the cohort, the device requirements were intended to be passive and involve as little disruption as possible for the participants. The user requirements for the device were defined as (1) comfortable and unobtrusive; (2) durable and water resistant, with a secure clasp; (3) long battery life (noncharging); and (4) clock function. The authors searched for suitably neat wrist-worn wearables and researched the features as required (Table 1).

Table 1. Device comparison.

Device name	Replaceable battery	Water resistant	Clock function
Fitbit Flex	No	No	Yes
Fitbit Flex 2	No	Yes	No
Polar Loop	No	Yes	Yes
Jawbone UP	No	No	No
Jawbone UP 2	No	No	No
Misfit Ray	Yes	Yes	No
Misfit Shine	Yes	Yes	Yes
Archon Touch	No	No	Yes
Striiv	No	No	Yes
Garmin Vivofit 2	Yes	Yes	Yes
Nokia Go	Yes	Yes	Yes

Figure 1. The Nokia Go wearable activity and sleep tracker.

The study device was not limited to any consumer or medical grade device; however, it was necessary to access raw intraday data from the device, requiring an open application programming interface (API). Access to the Garmin API carries a substantial fee; therefore, the Garmin Vivofit 2 was not considered for selection. The authors trialed the 2 devices that fit the defined criteria (Misfit Shine and Nokia Go) before choosing to provide participants with the Nokia Go because of its perceived durability and comfort during exercise. The Nokia Go (Nokia Technologies, Finland) is a wrist-worn activity and sleep tracker. Form factor design is a watch shape, with an E Ink display of your activity progression, which can also display an analog clock when pressed. The body of the device sits tightly into the strap, and the clasp is a thread-through with a button press. The watch is water resistant up to 50 m with automatic activity and sleep recognition and a replaceable battery lasting up to 8 months (Figure 1).

Data Collection

The study design required participants to wear the device for 24 hours a day over 2 specified periods: Preseason, which was 6 weeks spread across November and December and in-season, which was 8 weeks in February and March. The Health Mate app (Nokia Technologies, Finland) was downloaded and installed on each participant's personal mobile phone. A confidential research account email address and password was set up for each participant, and their device was connected to the Health Mate app on their personal mobile phone via Bluetooth. Outcomes of interest in this study were step count and sleep duration; therefore, no additional app settings were applied (such as body mass index). Deployment was conducted during the preseason phase of competition by the lead researcher. Access to the account data was set up via Nokia's API.

Daily individual step count was recorded through the API to monitor device wearing and syncing by participants. The Nokia Go device purported to sync with the Health Mate app automatically, provided Bluetooth was enabled on the mobile device. Participants were advised to sync the watch daily if they did not have Bluetooth always enabled on their mobile phone. If a participant had no data recorded on the Health Mate app for that day, the lead researcher would contact them with reminders to wear and sync the device. The device was to be worn at all times with the exception of during competitive Gaelic games, where jewelry wear is not permitted because of health and safety reasons. Participants were required to have data collected on at least 75% (65) of the study days to be included in the original study analysis.

Part 2: Retrospective Study Evaluation

During the data collection for the original study, considerable challenges were encountered by the study team. We subsequently decided to review and interrogate these issues through retrospective researcher and participant analysis, which were not part of the original study design. The results have, therefore, been structured into 4 sections: sections 1 and 2 detail the researcher and participant challenges, which were collated from transactional data such as research team correspondence, study logs, and messages from participants, whereas the third section outlines the actual data collected as part of the original study design, and the fourth was derived from the participant survey, which was designed as follows: A 34-question survey was designed to further understand the user experience of study participation and gain insight into data collection challenges from the participant's perspective. Survey design leveraged the Technology Acceptance Model [12], that is, perceived usefulness, perceived ease of use, attitude toward use, behavioral intention to use, and actual use, in addition to some practical questions. In 12 questions, a positive statement was made, and participants were asked to rate their agreement on a 5-point Likert scale. A total of 9 questions were multiple choice, 3 questions were open ended, and a further 10 questions were follow-up questions. The anonymous survey was deployed to participants online via Google Forms in the month following conclusion of the study (see Multimedia Appendix 1).

Results

Researcher Challenges

Disconnection and Product Support

The devices connected to each of the participants' phones without issue, with the exception of 1 Huawei PRA-LX1, which was listed on the Nokia website as incompatible. Therefore, 33 participants were included in the device deployment in November 2017. In a 9-day period, 20 participants experienced a disconnection. Disconnection happened between 3 to 12 days after deployment and had no pattern to phone type. This was generally reported by participants as "my watch won't sync" or "sleep data isn't working." On investigation, it appeared that the devices had disassociated from the user account. On the devices' page of the Health Mate app, where the connected device normally appeared, there was no device connected. The Nokia Go appeared to be recording data as shown on the watch display, but this data could not be collected as we were unable to sync it to a phone without performing a factory reset.

Where participants reported that their "sleep data isn't working," it transpired that the Nokia Go had disconnected from the Health Mate app, but the app had switched to collecting step data from the phone. This meant that many participants did not recognize there was an issue and subsequently, did not report the problem immediately. It was also difficult for the researchers to identify these anomalies; as to us, the participant may just have not been wearing the device at night.

At one point during data capture, the Nokia API requirements for accessing summary activity data changed. There had been no prior announcement from Nokia regarding this change, and the cause of the problem was discovered by the authors through developer chat rooms. In addition, it took multiple requests from the research team to Nokia product support to gain authorization for accessing intraday activity through the API.

Logistics

The research plan was to deploy the devices for 6 weeks in preseason (November-December 2017) as a trial period whereby we would assess feasibility and iron out potential issues and thus, redeploying the devices for an 8-week data collection during in-season (February-March 2018). After the initial disconnection issue, we reset and resynced the devices with the hope that this would be a once off issue. This took up to 3 weeks by the time the issue was identified and rectified in each case. One disassociated device would not reconnect; again, this was a Huawei phone but not a model listed as incompatible on the Nokia website.

When issues arose in the first 9 days of the study, we were very aware of the burden on participants in reporting and rectifying these issues. Potentially, the initial novelty of participation was dissipated during that time, and the logistics of collecting and redeploying the devices seemed like a process that would have been burdening on participants and challenging to maintain compliance. At this point we made the decision to continue the data collection uninterrupted from initial deployment (November 2017) to the end of the study (March 2018).

When researchers contacted participants to remind them to wear the device, it became apparent that although participants were wearing the device, they had not manually synced it or did not have Bluetooth enabled on their phone for automatic syncing. Participants were already receiving a daily reminder for another action as part of the wider study participation, and we became cognizant of the burden and potential annoyance of this contact, particularly as the data collection period had been extended. We made the decision to reduce syncing reminders to once a week, but often, this meant we could lose data if in fact the participant was not wearing the device.

The lead researcher (CD) worked clinically with the participant group 3 times per week on average, and when problems or technical issues arose, she planned to address these during her clinical visits. This was often impeded, however, by simple barriers such as the participant forgetting their phone, their phone battery being depleted, their phone not having internet access because of data restrictions, and frequently, they would forget to bring the watch, having taken it off when it appeared to or did stop working. Many participants also lost or had issues with their mobile phone over the course of the study, which resulted in lost data. Some participants changed their mobile phone, which required the action of another setup process, and again, may have resulted in lost data if they did not inform the lead researcher before changing.

Participant Challenges

Syncing

As many of the participants did not normally leave Bluetooth enabled on their mobile phone, they were required to manually sync the device daily. This method was suboptimal in that it required an action from the participant, which they would often forget to do. Participants reported that syncing could be very slow, and if the phone screen timed out, it would lead them to believe the sync had not worked; a belief which was reinforced for many by the initial disassociation. In addition, for those who did have their Bluetooth enabled, Nokia support informed us that "If you occasionally turn off Bluetooth on your device, it is possible that the Nokia Go will stop syncing even after Bluetooth is turned back on. If you encounter this issue, you will need to Force Stop the Health Mate app, turn on Bluetooth, and then launch the Health Mate app again," meaning that automatic syncing did not occur for the entire study. When the device stopped syncing automatically for those participants, their study participation routine changed from a passive process to one that required daily action.

Wearing

Anecdotally, the device appeared to be acceptable for most participants. There was 1 dropout from the study. This participant was a teacher and cited his decision in the comment section of the survey as occupational impracticality: "As a practical teacher one of the key things in each lesson is having the room tidied before the bell. The watch was not functional in this sense. In particular, the clock would twist in the socket and I would be reading the time wrong by a few minutes."

Loss of the device was a common issue with 7 permanent losses reported over the study duration, 3 of which were during the

Christmas period. Feedback from the participants on how this happened is presented in the survey results section below. A total of 3 participants opted not to wear the device during sport, citing glove wear and tackling as a reason. Early in the study, it appeared that the devices were being mislaid during sport because of the watch falling off during a tackle or similar contact. Some of them were retrieved and handed to a staff member, and others were not found. At this point it was decided that participants should no longer wear the device during sport. Although this was going to disrupt data collection, it appeared to be more desirable than losing further devices. As expected, many participants would either forget to remove the device for sport or forget to put it back on afterwards. The disruption this caused to data collection also meant that overall step counts were no longer reliable, as some could reflect training activity, and some may not. Furthermore, regulating wear during training was not feasible under the study conditions.

There was 1 reported instance of 2 participants confusing each other's device for their own. Although devices were marked with a study code by both waterproof stickers and permanent marker, neither method stood the test of time with this cohort.

Actual Data Collected

After rectifying the initial dissociation issues, the data collection period was defined as a total of 87 days from January to March 2018. A total of 9 participants were lost from the study because of cuts to the team panel, injury, and personal reasons, which were beyond our control. As previously mentioned, 1 device would not reconnect to a participants' mobile phone after the initial dissociation, and another had technical difficulties with syncing the device. A total of 4 participants lost their device during this data collection period, which totals to 6 participants who were removed from the cohort as they did not reach the 75% threshold of data collection.

Therefore, 18 of the remaining 24 participants were included in the analysis of the full data collection period. Nonwear time was calculated by considering 90 min of inactivity as the cut-off for nonwear [13]. The average overall wear was 79 out of 87 days (90%). On those days when the device was worn, the average wear time was 20.8 hours/day. Average wear time was calculated using a weighted average (by wear days) of the individual players.

Participant Feedback (Survey)

The 34-question feedback survey was deployed online via Google forms to participants in the month following completion of the study with a response rate of 91% (30/33). Results from the multiple-choice questions are presented in Multimedia Appendix 2.

Practical Considerations: Follow-Up Questions

Of those who did not leave their Bluetooth on, all 70% (21/30) indicated phone battery concerns as their reasoning.

A total of 47% (14/30) reported losing the watch, a higher number than what we had recorded, but most of those outliers explained that the watch had fallen off during sport and been found afterwards. A total of 3 of those participants mentioned that they felt that the strap opened too easily. Of the 9 responders who reported forgetting their watch somewhere, 6 (67%) reported that they would often leave it in their sport bag after training.

Attitude Toward Use (Open-Ended 1 and 2)

In describing the best thing or things about the watch, 50% (n=15) mentioned sleep tracking, and 57% (n=17) mentioned step and activity tracking. A total of 2 responders reported that the best aspect was achieving their daily step goals, whereas 2 enjoyed having the clock function, and other participants reported the app itself and the style of the device.

Whereas, when describing the worst thing(s) about the watch, 40% (n=12) reported that there was none, 17% (n=5) reported syncing the device, 17% (n=5) reported poor strap security, 10% (n=3) reported discomfort at night, with a further 6 mentioning aspects such as aesthetics, difficulty reading the time, inaccuracy, and forgetting to wear the watch.

Behavioral Intention to Use (Open-Ended 3)

In stating what might motivate participants to wear the watch, 37% (n=11) reported tracking their steps and sleep, 13% (n=4) said that nothing would motivate them to wear it, 10% (n=3) reported that an activity goal or competition would be motivating, 2 responders wanted it to be more comfortable, 2 wanted information on calories and energy expenditure, and further responses concerned the need for a better strap for training, a digital clock, a heart rate monitor, more information related to performance, and the watch being of better quality.

Actual Use: Follow-Up Questions

Of the 43% (n=13) who reported that they have continued to wear the watch, 9 reported that they find it useful for tracking steps and sleep, 2 reported that they use it primarily for telling the time, and 1 reported that they have become used to using it.

Whereas of the 57% (n=17) who have not continued wearing the watch, 6 reported that they lost the watch, 3 reported that they normally do not wear watches, 3 reported that they wear another watch, 2 reported that it was uncomfortable, 2 reported that they have no use for it, and 1 responder commented that they had forgotten about it.

Discussion

Disconnection and Field Test

In this study, the disassociation event had a sizeable impact on how the remainder of the data collection was conducted. Initial deployment of the wearable was designed to act as a field test, whereby we could identify and tackle issues as they arose. This field test consequently developed into the entire data collection phase, molded by the gravity of the initial disassociation and our understanding of how this might impact the participants. Considering the survey results reported only 23% (n=7) having an issue with their watch, we feel that we may have overestimated the potential burden of the disassociation event.

It is not clear whether a field test with a sample of the participants would have prevented the disassociation (which appeared to be a once-off event). However, it is possible that it

may have provided insights to modify the study design or preempt other challenges. Certainly, we feel that the issue with device comfort and strap security experienced during sport could have been identified through a field test, which may have enabled us to implement a stronger contingency plan to mitigate wear disruption. However, this prior knowledge may not have affected our ability to regulate device wear during and after training because of the personnel and time it would have required. Under the study conditions, the lead researcher was the sole point of contact for participants while also acting as the team physiotherapist during her regular interaction with participants.

Participant contact from the lead researcher when no data were recorded changed from daily to weekly, corresponding to strategies previously employed [14]. The burden of this monitoring and subsequent contact was evidently time-consuming but also a necessary component when dealing with wearable technology. Weekly checks were less burdensome than daily ones but also allowed for the increased possibility of data loss. In hindsight, as our survey results suggest, the daily reminders were not as burdensome for the participant (19/30 disagree or strongly disagree) as we may have perceived, suggesting this is a more favorable option when feasible.

Regarding our technical challenges with disconnection, API access, and product support, of significance is that Nokia were in the process of evaluating their digital health business at the time, announcing a strategic review and ultimately selling their digital health sector back to Withings from whom they had purchased it in 2016. The availability of product support and customer confidence in the maintenance of data access is of utmost importance when investing financial, time, and human resources in a research project but will not always be predictable for researchers.

Logistics and Syncing

As the lead researcher worked clinically with this cohort an average of 3 times per week, it was envisaged at the beginning of the study that identifying and rectifying technical or other issues would be relatively easy with this level of access. However, when issues arose, such as a disconnection, a syncing problem, or the loss of a device, participants were generally quite slow to report them or were less likely to make outside contact immediately when an issue was identified. In some cases, participants were not aware of the issue because the Health Mate app had now connected to display their phone's step count. In other cases, such as a loss, participants were familiar to the lead researcher and may have felt guilty that they had lost the device. In these scenarios, the fact that participants did not sync the watch daily meant that issues were not always readily identifiable on our part.

The survey identified the variety of syncing strategies employed by the participants, with only 23% having a daily syncing routine and no participants reporting that the device synced automatically, as originally purported by Nokia. This resulted in greater burden for the participants while pairing the device, a practice that has been previously shown to negatively influence the sustained use of wearable devices [15]. To avoid scenarios where user maintenance negatively effects data quality, device syncing should be seamless and passive for the participant. A defined and robust syncing strategy is central to success.

When such issues were eventually recognized, simple barriers could make them slow to fix. For example, the assumptions we made that the participants would always carry their phones and that they would be charged and have on-going access to the internet proved to be untrue. In reality, the cohort were busy, used many social media apps, and had to travel long distances for training both of which meant drained batteries were common, whereas internet access was sometimes a problem because of phone data packages. These results challenge our understanding of the preconceived ease of using technologies with young adults, who in fact appear to display their own obstacles in the digital divide [16]. The practical learnings from these issues include the researcher being armed with multiple phone chargers and having a portable Wi-Fi or hotspot setup. In addition, sending extra reminders to participants to bring their device, and if feasible fully charged, is advised when their attendance at a venue is not for the sole purpose of troubleshooting device issues, for instance, they may be attending training (as in this case) or a medical appointment.

Moreover, a number of participants changed their mobile phone over the course of the study—a basic issue that we did not foresee and were generally not preinformed about. If the participant had changed their phone and not synced the device for several days or weeks, those data were lost. This particularly happened around the Christmas period, and these are important challenges to be aware of when using the bring-your-own-device strategy in longitudinal data collection. The alternative being obviously higher cost and also the disruption of employing a device that the participant doesn't already use daily for other purposes [17].

Wearing

The results of this study demonstrate that device wear during contact sport was simply not suitable. Although the authors believed the strap of the device to be sturdy and secure, this belief was not shared by participants with regards to contact sport. Results of general comfort were quite favorable but dropped considerably when asked about comfort during sport. These findings are consistent with previously published reasons for taking off smart watches: discomfort during sport and concerns about breakage of the device [18]. Perhaps no watch-like activity tracker is suitable for wear during a physical contact sport. and this is probably reflected in the illegality of same during competitive games. However, contact sports are a common part of life, and when dealing with an active cohort, this is an eventuality that may need to be planned for. In addition, for an active cohort such as this, permanent marking of devices would be necessary, for example, engraving a participant code into the strap.

As mentioned, the disruption of nonwear during training meant that many participants would either forget to take off or put back on the device, and unfortunately, we did not have the resources to regulate this under the study conditions. If this eventuality could have been planned for, we might have taken the initiative of 1 participant who was observed putting the device in his shoe rather than in his sports bag, as a reminder

to put it back on his wrist afterwards. A basic gesture but an effective one nonetheless.

Perceptions and Attitudes

Although there were many challenges involved in this data collection, if we were to evaluate adherence to the device based on the wear characteristics of the remaining 18 participants, we could conclude that this study was successful. However, the issues experienced meant that our cohort size was depleted, in addition to the original 14-week split observation being reduced to 1 12-week observation. When asked if they needed more information about the watch or the app, only 1 responder reported that they did, which appears to reflect that the cohort felt well-informed and confident in their use of the device and negates many of the current adoption challenges on the other side of the digital divide [19]. The open-ended responses regarding the best thing about the watch and what might motivate them to continue wearing the watch had a common theme relating to the participants desire for more information related to their health and performance. These answers appear to reflect the interests of a group of sportspeople who want to better their sporting performance and understand how their activity and sleep can influence this, which follows previous research underpinning motivation and feedback as leading facets of engagement [20,21]. However, the participant characteristics should be considered in these interpretations: as an already active group of males, they may be more likely to use a wearable activity tracker than other cohorts [22]. There was poor agreement as to whether the team had a competitive attitude to step counting or not, which probably reflects that only some individuals treated it as a competition. This is important to note with regards to its effect, or lack thereof, on user motivation and compliance.

Limitations

A limitation of the presented study has been mentioned throughout; the fact that we did not set out to investigate and report on a specific research question; it is rather, an unstructured, retrospective evaluation of another study. Therefore, this report does not follow the standard layout and presentation of results that one might expect.

Although this study presents common issues that can be experienced in a wearable deployment, it is important to consider that this case study represents a single regional cohort using a specific activity tracker and mobile app. The cohort were a small, purposeful sample chosen for the requirements of the original study and cannot be assumed to represent a general population.

The authors acknowledge the limitations of the survey, although anonymous, the participants' familiarity with the lead researcher may have affected their honesty in reporting on their experience.

Conclusions

Our findings contrast many of the published barriers to wearable use such as discomfort [23] and technology familiarity [24] and support many referenced facilitators such as user acceptance and perceived usefulness [23]. Nevertheless, our data collection was highly impacted by challenges, namely logistics, technical issues, and loss of the device. These factors are not always predictable and will need to be strongly considered to ensure satisfactory data quality is achieved. It is reasonable to suggest that some of the issues we encountered might have been mitigated with careful planning and pilot studies. Nonetheless, the frequency and range of issues encountered did not reflect the level of preparatory work that was done. Previously published potential solutions to the user adherence challenge lists many of the strategies employed in this study, such as providing clear instructions, reminders, and checking in with participants regularly to identify issues [25]. On the basis of our experience, we suggest readers to consider their research context in the selection of a consumer wearable [26], including the availability of an extensive support team to undertake these requirements. Issues will arise despite extensive preparation and exacting due diligence, and they will need to be identified and rectified promptly to maintain data quality. Although we do not expect exact replication of these challenges in other contexts, our experience can be a useful example for learnings in a wider readership for wearables research.

Learning Outcomes and Recommendations

The Field Test

Although our field testing did not work out as planned, we would recommend field testing your device with a subset of the cohort to identify practical and technical issues that may arise. This will enable the refinement of implementation and adoption strategies that align with the needs of your cohort.

The Syncing Strategy

Passive (automatic) and seamless device syncing is optimal for user compliance and maintained use as it facilitates immediate identification of issues such as nonwear or disconnection. If automated syncing and upload is not possible, the strategy should include clear instructions for users to perform manual sync on a daily basis.

Bring-Your-Own-Device

When using a participant-owned mobile phone, the deployment team should take time to adjust the phone settings to ensure that no partner apps can share and extract data, such as access to the phone's step counter. This unsolicited data sharing could conceal technical issues and may be indiscernible from activity tracker data when the raw dataset is downloaded. It is essential to plan for and be informed about personal mobile phone changes from participants.

Support Team and Logistics

With such a sizeable time-burden involved in monitoring daily syncing, it is imperative to invest in the appropriate resources to monitor this wear and resolve issues. When dealing with a working population, it may be difficult to contact participants during the day, and home visits may also be unavailable to you. Although we had regular access to the participants, resolving issues did not work as smoothly as planned, partly because the access was not for the sole purpose of partaking in the study.

Acknowledgments

This research has been funded by the Science Foundation of Ireland under grant number SFI/12/RC/2289.

Authors' Contributions

CD, PS, and BC developed the study concept. CD conducted the recruitment, data collection, and analysis. NS was involved in data processing and technical assistance. CD and PS were involved in drafting the manuscript. All authors reviewed, edited, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Participant survey.

[PDF File (Adobe PDF File), 38KB - mhealth_v7i6e12190_app1.pdf]

Multimedia Appendix 2

Survey results.

[PDF File (Adobe PDF File), 52KB - mhealth_v7i6e12190_app2.pdf]

References

1. Lim WK, Davila S, Teo JX, Yang C, Pua CJ, Blöcker C, et al. Beyond fitness tracking: the use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. PLoS Biol 2018;16(2):e2004285 [FREE Full text] [doi: 10.1371/journal.pbio.2004285] [Medline: 29485983]
2. Li X, Dunn J, Salins D, Zhou G, Zhou W, Schüssler-Fiorenza Rose SM, et al. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol 2017 Jan;15(1):e2001402 [FREE Full text] [doi: 10.1371/journal.pbio.2001402] [Medline: 28081144]
3. 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] [Medline: 28652228]
4. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EE, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Dec;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]
5. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
6. Strath SJ, Rowley TW. Wearables for promoting physical activity. Clin Chem 2018 Jan;64(1):53-63. [doi: 10.1373/clinchem.2017.272369] [Medline: 29118062]
7. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. Clin Pharmacol Ther 2018 Jul;104(1):42-52 [FREE Full text] [doi: 10.1002/cpt.966] [Medline: 29205294]
8. Lee HA, Lee HJ, Moon JH, Lee T, Kim MG, In H, et al. Comparison of wearable activity tracker with actigraphy for sleep evaluation and circadian rest-activity rhythm measurement in healthy young adults. Psychiatry Investig 2017 Mar;14(2):179-185 [FREE Full text] [doi: 10.4306/pi.2017.14.2.179] [Medline: 28326116]
9. Reeder B, David A. Health at hand: a systematic review of smart watch uses for health and wellness. J Biomed Inform 2016;63:269-276 [FREE Full text] [doi: 10.1016/j.jbi.2016.09.001] [Medline: 27612974]
10. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. Am J Physiol Regul Integr Comp Physiol 2017 Mar 01;312(3):R358-R367. [doi: 10.1152/ajpregu.00349.2016] [Medline: 28052867]
11. Kim KJ, Shin DH. An acceptance model for smart watches: Implications for the adoption of future wearable technology. Internet Research 2015 Aug 03;25(4):527-541. [doi: 10.1108/IntR-05-2014-0126]
12. Davis F. MIT Libraries. 1986. Original Technology Acceptance Model URL: https://dspace.mit.edu/handle/1721.1/15192
13. Choi L, Ward SC, Schnelle JF, Buchowski MS. Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. Med Sci Sports Exerc 2012 Oct;44(10):2009-2016 [FREE Full text] [doi: 10.1249/MSS.0b013e318258cb36] [Medline: 22525772]

14. Silva de Lima AL, Hahn T, Evers LJ, de Vries NM, Cohen E, Afek M, et al. Feasibility of large-scale deployment of multiple wearable sensors in Parkinson's disease. PLoS One 2017;12(12):e0189161 [FREE Full text] [doi: 10.1371/journal.pone.0189161] [Medline: 29261709]
15. Lazar A, Koehler C, Tanenbaum J, Nguyen DH. Why We Use and Abandon Smart Devices. : ACM; 2015 Presented at: ACM Int Jt Conf Pervasive Ubiquitous Comput (UbiComp); September 7-11, 2015; Osaka, Japan p. 635-646 URL: https://dl.acm.org/citation.cfm?id=2804288
16. Dijkstra A, Heida A, van Rheenen PF. A web-based telemonitoring strategy for teenagers with IBD to improve efficiency of care: abandon or ready for sustained adoption? J Med Internet Res 2018;21(3):e11761 [FREE Full text] [doi: 10.2196/11761] [Medline: 30924785]
17. Pugliese L, Woodriff M, Crowley O, Lam V, Sohn J, Bradley S. Feasibility of the "Bring Your Own Device" Model in Clinical Research: Results from a Randomized Controlled Pilot Study of a Mobile Patient Engagement Tool. Cureus 2016 Mar 16;8(3):e535 [FREE Full text] [doi: 10.7759/cureus.535] [Medline: 27096135]
18. Jeong H, Kim H, Kim R, Lee U, Jeong Y. Smartwatch wearing behavior analysis: a longitudinal study. Proc ACM Interact Mob Wearable Ubiquitous Technol 2017 Sep;1(3):1-31 Article 60. [doi: 10.1145/3131892]
19. Hastall MR, Dockweiler C, Mühlhaus J. Achieving End User Acceptance: Building Blocks for an Evidence-Based User-Centered Framework for Health Technology Development and Assessment. 2017 Presented at: 11th International Conference on Universal Access in Human-Computer Interaction UAHCI; July 9-14, 2017; Vancouver, Canada p. 13-25. [doi: 10.1007/978-3-319-58700-4_2]
20. Standage M, Sebire SJ, Loney T. Does exercise motivation predict engagement in objectively assessed bouts of moderate-intensity exercise? A self-determination theory perspective. J Sport Exerc Psychol 2008 Aug;30(4):337-352. [Medline: 18723896]
21. Brown E, Cairns P. A Grounded Investigation of Game Immersion. In: CHI'04 Extended Abstracts on Human Factors in Computing Systems. 2004 Presented at: CHI'04; April 24-29, 2004; Vienna, Austria p. 1297-1300. [doi: 10.1145/985921.986048]
22. Alley S, Schoeppe S, Guertler D, Jennings C, Duncan MJ, Vandelanotte C. Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. BMJ Open 2016;6(7):e011243 [FREE Full text] [doi: 10.1136/bmjopen-2016-011243] [Medline: 27388359]
23. Chan M, Estève D, Fourniols JY, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
24. Tana J, Forss M, Hellstén T. Theseus. 2017. The Use of Wearables in Healthcare - Challenges and Opportunities URL: https://www.theseus.fi/bitstream/handle/10024/140584/AWP_6-2017_Tana_Forss_Hellsten.pdf?sequence=2&isAllowed=y
25. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable technology and physical activity in chronic disease: opportunities and challenges. Am J Prev Med 2018 Jan;54(1):144-150. [doi: 10.1016/j.amepre.2017.08.015] [Medline: 29122356]
26. Caulfield B, Reginatto B, Slevin P. Not all sensors are created equal: a framework for evaluating human performance measurement technologies. NPJ Digit Med 2019 Feb 14;2(1). [doi: 10.1038/s41746-019-0082-4]

Abbreviations

API: application programming interface

Edited by G Eysenbach; submitted 12.09.18; peer-reviewed by B Arnoldussen, A Henriksen, K Ng, E Lattie; comments to author 23.03.19; revised version received 08.04.19; accepted 20.05.19; published 24.06.19.

Please cite as:
Duignan C, Slevin P, Sett N, Caulfield B
Consumer Wearable Deployments in Actigraphy Research: Evaluation of an Observational Study
JMIR Mhealth Uhealth 2019;7(6):e12190
URL: http://mhealth.jmir.org/2019/6/e12190/
doi:10.2196/12190
PMID:31237237

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org