
JMIR Formative Research

Impact Factor (2023): 2.0
Volume 6 (2022), Issue 12 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Learning the Treatment Process in Radiotherapy Using an Artificial Intelligence–Assisted Chatbot: Development Study (e39443) Nathanael Rebelo, Leslie Sanders, Kay Li, James Chow.	7
A Web-Based Program About Sustainable Development Goals Focusing on Digital Learning, Digital Health Literacy, and Nutrition for Professional Development in Ethiopia and Rwanda: Development of a Pedagogical Method (e36585) Katarina Bälter, Feben Javan Abraham, Chantal Mutimukwe, Reuben Mugisha, Christine Persson Osowski, Olle Bälter.	20
Perceptions and Attitudes Toward an Interactive Voice Response Tool (Call for Life Uganda) Providing Adherence Support and Health Information to HIV-Positive Ugandans: Qualitative Study (e36829) Phoebe Kajubi, Rosalind Parkes-Ratanshi, Adelline Twimukye, Agnes Bwanika Naggirinya, Maria Nabaggala, Agnes Kiragga, Barbara Castelnovo, Rachel King.	33
Design of a Patient Voice App Experience for Heart Failure Management: Usability Study (e41628) Antonia Barbaric, Cosmin Munteanu, Heather Ross, Joseph Cafazzo.	49
Certified Peer Support Specialists Training in Technology and Delivery of Digital Peer Support Services: Cross-sectional Study (e40065) Karen Fortuna, Julia Hill, Samantha Chalker, Joelle Ferron.	62
Changes in Glycemic Control Following Use of a Spanish-Language, Culturally Adapted Diabetes Program: Retrospective Study (e40278) Caitlyn Edwards, Elisa Orellana, Kelly Rawlings, Mirta Rodriguez-Pla, Aarathi Venkatesan.	71
Predicting Overweight and Obesity Status Among Malaysian Working Adults With Machine Learning or Logistic Regression: Retrospective Comparison Study (e40404) Jyh Wong, Miwa Yamaguchi, Nobuo Nishi, Michihiro Araki, Lei Wee.	81
An Emotional Bias Modification for Children With Attention-Deficit/Hyperactivity Disorder: Co-design Study (e36390) Melvyn Zhang, Vallabhajosyula Ranganath.	92
Co-design and Development of EndoSMS, a Supportive Text Message Intervention for Individuals Living With Endometriosis: Mixed Methods Study (e40837) Kerry Sherman, Melissa Pehlivan, Anna Singleton, Alexandra Hawkey, Julie Redfern, Mike Armour, Blake Dear, Tanya Duckworth, Donna Ciccia, Michael Cooper, Kelly Parry, Esther Gandhi, Sara Imani.	104

Use of an Interactive Obesity Treatment Approach in Individuals With Severe Mental Illness: Feasibility, Acceptability, and Proposed Engagement Criteria ([e38496](#))
 Ginger Nicol, Madeline Jansen, Rita Haddad, Amanda Ricchio, Michael Yingling, Julia Schweiger, Katie Keenoy, Bradley Evanoff, John Newcomer. 120

Detailed Versus Simplified Dietary Self-monitoring in a Digital Weight Loss Intervention Among Racial and Ethnic Minority Adults: Fully Remote, Randomized Pilot Study ([e42191](#))
 Michele Patel, Angel Cleare, Carly Smith, Lisa Rosas, Abby King. 134

Understanding the Role of Support in Digital Mental Health Programs With Older Adults: Users' Perspective and Mixed Methods Study ([e43192](#))
 Judith Borghouts, Elizabeth Eikey, Cinthia De Leon, Stephen Schueller, Margaret Schneider, Nicole Stadnick, Kai Zheng, Lorraine Wilson, Damaris Caro, Dana Mukamel, Dara Sorkin. 159

Efficacy of the Mental Health App Intellect to Reduce Stress: Randomized Controlled Trial With a 1-Month Follow-up ([e40723](#))
 Sean Toh, Jessalin Tan, Feodora Kosasih, Oliver Sündermann. 174

End User Participation in the Development of an Ecological Momentary Intervention to Improve Coping With Cannabis Cravings: Formative Study ([e40139](#))
 Molly Anderson, Alan Budney, Nicholas Jacobson, Inbal Nahum-Shani, Catherine Stanger. 191

Electronic Source Data Transcription for Electronic Case Report Forms in China: Validation of the Electronic Source Record Tool in a Real-world Ophthalmology Study ([e43229](#))
 Bin Wang, Junkai Lai, Mimi Liu, Feifei Jin, Yifei Peng, Chen Yao. 202

Assessing and Promoting Cardiovascular Health for Adolescent Women: User-Centered Design Approach ([e42051](#))
 Kolbi Bradley, Santiago Arconada Alvarez, Amanda Gilmore, Morgan Greenleaf, Aayahna Herbert, Melissa Kottke, Maren Parsell, Sierra Patterson, Tymirra Smith, Mercedes Sotos-Prieto, Elizabeth Zeichner, Holly Gooding. 215

Using Vocal Characteristics To Classify Psychological Distress in Adult Helpline Callers: Retrospective Observational Study ([e42249](#))
 Ravi Iyer, Maja Nedeljkovic, Denny Meyer. 227

Factors Associated With Intention and Use of e-Mental Health by Mental Health Counselors in General Practices: Web-Based Survey ([e34754](#))
 Ann De Veirman, Viviane Thewissen, Matthijs Spruijt, Catherine Bolman. 238

Analysis of Patient Cues in Asynchronous Health Interactions: Pilot Study Combining Empathy Appraisal and Systemic Functional Linguistics ([e40058](#))
 Elena Rey Velasco, Hanne Pedersen, Timothy Skinner, Impact Diabetes B2B Collaboration Group. 256

Navigating the Systemic Conditions of a Digital Health Ecosystem in Alberta, Canada: Embedded Case Study ([e36265](#))
 Chad Saunders, Devon Currie, Shane Virani, Jill De Groot. 274

Comparing a Fitbit Wearable to an Electrocardiogram Gold Standard as a Measure of Heart Rate Under Psychological Stress: A Validation Study ([e37885](#))
 Joel Gagnon, Michelle Khau, Léandre Lavoie-Hudon, François Vachon, Vicky Drapeau, Sébastien Tremblay. 290

A Voice App Design for Heart Failure Self-management: Proof-of-Concept Implementation Study ([e40021](#))
 Antonia Barbaric, Cosmin Munteanu, Heather Ross, Joseph Cafazzo. 301

A Mobile Phone Text Messaging Intervention to Manage Fatigue for People With Multiple Sclerosis, Spinal Cord Injury, and Stroke: Development and Usability Testing (e40166)	
Kerri Morgan, Alex Wong, Kim Walker, Rachel Desai, Tina Knepper, Pamela Newland.	313
Two-way Automated Text Messaging Support From Community Pharmacies for Medication Taking in Multiple Long-term Conditions: Human-Centered Design With Nominal Group Technique Development Study (e41735)	
Gemma Donovan, Nicola Hall, Felicity Smith, Jonathan Ling, Scott Wilkes.	330
Digital Medicine System in Veterans With Severe Mental Illness: Feasibility and Acceptability Study (e34893)	
Sarah Gonzales, Olaoluwa Okusaga, J Reuteman-Fowler, Megan Oakes, Jamie Brown, Scott Moore, Allison Lewinski, Cristin Rodriguez, Norma Moncayo, Valerie Smith, Shauna Malone, Justine List, Raymond Cho, Amy Jeffreys, Hayden Bosworth.	349
Evaluating User Preferences, Comprehension, and Trust in Apps for Environmental Health Hazards: Qualitative Case Study (e38471)	
Annabelle Workman, Fay Johnston, Sharon Campbell, Grant Williamson, Chris Lucani, David Bowman, Nick Cooling, Penelope Jones.	361
Understanding Emergency Room Visits for Nontraumatic Oral Health Conditions in a Hospital Serving Rural Appalachia: Dental Informatics Study (e31433)	
Raj Khanna, Alfred Cecchetti, Niharika Bhardwaj, Bobbi Muto, Usha Murughiyan.	376
Perceptions of a Secure Cloud-Based Solution for Data Sharing During Acute Stroke Care: Qualitative Interview Study (e40061)	
Marcela Tuler de Oliveira, Lúcio Amorim Reis, Henk Marquering, Aeilko Zwinderman, Sílvia Delgado Olabarriga.	383
State-Level COVID-19 Symptom Searches and Case Data: Quantitative Analysis of Political Affiliation as a Predictor for Lag Time Using Google Trends and Centers for Disease Control and Prevention Data (e40825)	
Alex Turvy.	400
Ethical Issues in the Use of Smartphone Apps for HIV Prevention in Malaysia: Focus Group Study With Men Who Have Sex With Men (e42939)	
Antoine Khati, Jeffrey Wickersham, Aviana Rosen, Jeffrey Lucas, Nicholas Copenhaver, Alma Jeri-Wahrhaftig, Mohd Ab Halim, Iskandar Azwa, Kamal Gautam, Kai Ooi, Roman Shrestha.	411
Barriers and Facilitators of Obtaining Social Determinants of Health of Patients With Cancer Through the Electronic Health Record Using Natural Language Processing Technology: Qualitative Feasibility Study With Stakeholder Interviews (e43059)	
Jordan Alpert, Hyehyun Kim, Cara McDonnell, Yi Guo, Thomas George, Jiang Bian, Yonghui Wu.	425
Effectiveness of Secondary Risk-Reducing Strategies in Patients With Unilateral Breast Cancer With Pathogenic Variants of BRCA1 and BRCA2 Subjected to Breast-Conserving Surgery: Evidence-Based Simulation Study (e37144)	
Jelena Maksimenko, Pedro Rodrigues, Miki Nakazawa-Miklašević, David Pinto, Edvins Miklašević, Genadijs Trofimovičs, Jānis Gardovskis, Fatima Cardoso, Maria Cardoso.	434
The Acceptance, Usability, and Utility of a Web Portal for Back Pain as Recommended by Primary Care Physicians: Qualitative Interview Study With Patients (e38748)	
Christian Schlett, Nicole Röttele, Piet van der Keylen, Andrea Schöpf-Lazzarino, Miriam Klimmek, Mirjam Körner, Kathrin Schnitzius, Sebastian Voigt-Radloff, Andy Maun, Mario Sofroniou, Erik Farin-Glattacker.	448
Theoretical Approach and Scale Construction of Patient Privacy Protection Behavior of Doctors in Public Medical Institutions in China: Pilot Development Study (e39947)	
Jie Xu, Lu Lu, Kaichen Xing, Huwei Shi, Ruiyao Chen, Yujun Yao, Sichen Liu, Zhongzhou Xiao, Xinwei Peng, Shuqing Luo, Yun Zhong.	466

<p>Job Disengagement Among Physical Education Teachers: Insights From a Cross-sectional Web-Based Survey With Path Modeling Analysis (e29130) Nasr Chalghaf, Wen Chen, Amayra Tannoubi, Noomen Guelmami, Luca Puce, Noureddine Ben Said, Maher Ben Khalifa, Fairouz Azaiez, Nicola Bragazzi.</p>	491
<p>Perceptions and Aspirations Toward Peer Mentoring in Social Media–Based Electronic Cigarette Cessation Interventions for Adolescents and Young Adults: Focus Group Study (e42538) Joanne Lyu, Aliyyat Afolabi, Justin White, Pamela Ling.</p>	501
<p>The Persian Version of the Mobile Application Rating Scale (MARS-Fa): Translation and Validation Study (e42225) Saeed Barzegari, Ali Sharifi Kia, Marco Bardus, Stoyan Stoyanov, Marjan GhaziSaeedi, Mouna Rafizadeh.</p>	511
<p>Social Determinants of Digital Health Adoption: Pilot Cross-sectional Survey (e39647) Sharvil Patel, Elizabeth Sun, Alec Reinhardt, Sanjaly Geevarghese, Simon He, Julie Gazmararian.</p>	523
<p>Data Quality and Study Compliance Among College Students Across 2 Recruitment Sources: Two Study Investigation (e39488) Abby Braitman, Megan Strowger, Jennifer Shipley, Jordan Ortman, Rachel MacIntyre, Elizabeth Bauer.</p>	534
<p>Development of an mHealth App–Based Intervention for Depressive Rumination (RuminAid): Mixed Methods Focus Group Evaluation (e40045) Eve Rosenfeld, Cassandra Lyman, John Roberts.</p>	554
<p>Detecting Elevated Air Pollution Levels by Monitoring Web Search Queries: Algorithm Development and Validation (e23422) Chen Lin, Safoora Yousefi, Elvis Kahoro, Payam Karisani, Donghai Liang, Jeremy Sarnat, Eugene Agichtein.</p>	572
<p>Using Social Media to Engage Justice-Involved Young Adults in Digital Health Interventions for Substance Use: Pilot Feasibility Survey Study (e37609) Anna Harrison, Johanna Folk, Christopher Rodriguez, Amanda Wallace, Marina Tolou-Shams.</p>	593
<p>Using Continuous Glucose Monitoring to Detect and Intervene on Dietary Restriction in Individuals With Binge Eating: The SenseSupport Withdrawal Design Study (e38479) Adrienne Juarascio, Paakhi Srivastava, Emily Presseller, Mandy Lin, Anna Patarinski, Stephanie Manasse, Evan Forman.</p>	607
<p>Preferences of Older Adult Veterans With Heart Failure for Engaging With Mobile Health Technology to Support Self-care: Qualitative Interview Study Among Patients With Heart Failure and Content Analysis (e41317) Marva Foster, Wei Xiong, Lisa Quintiliani, Christine Hartmann, Stephan Gaehde.</p>	622
<p>Feasibility, Usability and Acceptability of a mHealth Intervention to Reduce Cardiovascular Risk in Rural Hispanic Adults: Descriptive Study (e40379) Sheri Rowland, Athena Ramos, Natalia Trinidad, Sophia Quintero, Rebecca Johnson Beller, Leeza Struwe, Bunny Pozehl.</p>	634
<p>Influences, Barriers, and Facilitators to COVID-19 Vaccination: Cross-sectional Survey on Vaccine Hesitancy in 2 Rural States (e39109) Elaine Nguyen, Melanie Wright, John Holmes, Kevin Cleveland, Catherine Oliphant, Mary Nies, Renee Robinson.</p>	655
<p>Impact of Telehealth on the Delivery of Prenatal Care During the COVID-19 Pandemic: Mixed Methods Study of the Barriers and Opportunities to Improve Health Care Communication in Discussions About Pregnancy and Prenatal Genetic Testing (e38821) Caitlin Craighead, Christina Collart, Richard Frankel, Susannah Rose, Anita Misra-Hebert, Brownsyne Tucker Edmonds, Marsha Michie, Edward Chien, Marissa Coleridge, Oluwatosin Goje, Angela Ranzini, Ruth Farrell.</p>	667

Assessing Associations Between COVID-19 Symptomology and Adverse Outcomes After Piloting Crowdsourced Data Collection: Cross-sectional Survey Study (e37507) Natalie Flaks-Manov, Jiawei Bai, Cindy Zhang, Anand Malpani, Stuart Ray, Casey Taylor.	682
Engagement, Use, and Impact of Digital Mental Health Resources for Diverse Populations in COVID-19: Community-Partnered Evaluation (e42031) Kenneth Wells, April Thames, Alexander Young, Lily Zhang, MarySue Heilemann, Daniela Romero, Adrian Oliva, Felica Jones, Lingqi Tang, Melissa Brymer, Thomas Elliott, Armen Arevian, Together for Wellness/Juntos Collaborators and Writing Group.	697
Continued Use of Contact-Tracing Apps in the United States and the United Kingdom: Insights From a Comparative Study Through the Lens of the Health Belief Model (e40302) Zhan Zhang, Isaac Vaghefi.	712
Using a Proximity-Detection Technology to Nudge for Physical Distancing in a Swedish Workplace During the COVID-19 Pandemic: Retrospective Case Study (e39570) My Villius Zetterholm, Lina Nilsson, Päivi Jokela.	726
Perceptions of, and Obstacles to, SARS-CoV-2 Vaccination Among Adults in Lebanon: Cross-sectional Online Survey (e36827) Nadeem Abou-Arraj, Diana Maddah, Vanessa Buhamdan, Roua Abbas, Nadine Jawad, Fatima Karaki, Nael Alami, Pascal Geldsetzer.	744
Using Twitter Data to Estimate the Prevalence of Symptoms of Mental Disorders in the United States During the COVID-19 Pandemic: Ecological Cohort Study (e37582) Ruilie Cai, Jiajia Zhang, Zhenlong Li, Chengbo Zeng, Shan Qiao, Xiaoming Li.	759
IT and the Quality and Efficiency of Mental Health Care in a Time of COVID-19: Case Study of Mental Health Providers in England (e37533) Frederick Konteh, Russell Mannion, Rowena Jacobs.	770
Pilot Evaluation of Possible Airborne Transmission in a Geriatric Care Facility Using Carbon Dioxide Tracer Gas: Case Study (e37587) Yo Ishigaki, Shinji Yokogawa, Yuki Minamoto, Akira Saito, Hiroko Kitamura, Yuto Kawauchi.	782
Monitoring and Identifying Emerging e-Cigarette Brands and Flavors on Twitter: Observational Study (e42241) Qihang Tang, Runtao Zhou, Zidian Xie, Dongmei Li.	798
Sample Bias in Web-Based Patient-Generated Health Data of Dutch Patients With Gastrointestinal Stromal Tumor: Survey Study (e36755) Anne Dirkson, Dide den Hollander, Suzan Verberne, Ingrid Desar, Olga Husson, Winette van der Graaf, Astrid Oosten, Anna Reyners, Neeltje Steeghs, Wouter van Loon, Gerard van Oortmerssen, Hans Gelderblom, Wessel Kraaij.	810

Viewpoint

Spotlighting Disability in a Major Electronic Health Record: Michigan Medicine's Disability and Accommodations Tab (e38003) Heather Halkides, Tyler James, Michael McKee, Michelle Meade, Christa Moran, Sophia Park.	647
--	-----



Corrigenda and Addenda

Addendum: The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees ([e44497](#))

Joris van Gend, Jan van 't Klooster, Catherine Bolman, Julia van Gemert-Pijnen. 796

Original Paper

Learning the Treatment Process in Radiotherapy Using an Artificial Intelligence–Assisted Chatbot: Development Study

Nathanael Rebelo¹; Leslie Sanders², PhD; Kay Li³, PhD; James C L Chow^{4,5}, PhD

¹Department of Physics, Toronto Metropolitan University, Toronto, ON, Canada

²Department of Humanities, York University, Toronto, ON, Canada

³Department of English, University of Toronto, Toronto, ON, Canada

⁴Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

⁵Department of Radiation Oncology, University of Toronto, Toronto, ON, Canada

Corresponding Author:

James C L Chow, PhD

Radiation Medicine Program

Princess Margaret Cancer Centre

University Health Network

700 University Ave

7/F, Room 7-606

Toronto, ON, M5G 1X6

Canada

Phone: 1 416 946 4501 ext 5089

Email: james.chow@rmp.uhn.ca

Abstract

Background: In knowledge transfer for educational purposes, most cancer hospital or center websites have existing information on cancer health. However, such information is usually a list of topics that are neither interactive nor customized to offer any personal touches to people facing dire health crisis and to attempt to understand the concerns of the users. Patients with cancer, their families, and the general public accessing the information are often in challenging, stressful situations, wanting to access accurate information as efficiently as possible. In addition, there is seldom any comprehensive information specifically on radiotherapy, despite the large number of older patients with cancer, to go through the treatment process. Therefore, having someone with professional knowledge who can listen to them and provide the medical information with good will and encouragement would help patients and families struggling with critical illness, particularly during the lingering pandemic.

Objective: This study created a novel virtual assistant, a chatbot that can explain the radiation treatment process to stakeholders comprehensively and accurately, in the absence of any similar software. This chatbot was created using the IBM Watson Assistant with artificial intelligence and machine learning features. The chatbot or bot was incorporated into a resource that can be easily accessed by the general public.

Methods: The radiation treatment process in a cancer hospital or center was described by the radiotherapy process: patient diagnosis, consultation, and prescription; patient positioning, immobilization, and simulation; 3D-imaging for treatment planning; target and organ contouring; radiation treatment planning; patient setup and plan verification; and treatment delivery. The bot was created using IBM Watson (IBM Corp) assistant. The natural language processing feature in the Watson platform allowed the bot to flow through a given conversation structure and recognize how the user responds based on recognition of similar given examples, referred to as intents during development. Therefore, the bot can be trained using the responses received, by recognizing similar responses from the user and analyzing using Watson natural language processing.

Results: The bot is hosted on a website by the Watson application programming interface. It is capable of guiding the user through the conversation structure and can respond to simple questions and provide resources for requests for information that was not directly programmed into the bot. The bot was tested by potential users, and the overall averages of the identified metrics are excellent. The bot can also acquire users' feedback for further improvements in the routine update.

Conclusions: An artificial intelligence–assisted chatbot was created for knowledge transfer regarding radiation treatment process to the patients with cancer, their families, and the general public. The bot that is supported by machine learning was tested, and it was found that the bot can provide information about radiotherapy effectively.

KEYWORDS

chatbot; artificial intelligence; machine learning; radiotherapy chain; radiation treatment process; communication; diagnosis; cancer therapy; internet of things; radiation oncology; medical physics; health care

Introduction

Rationale

To educate patients with cancer, their families, and the general public about the radiation treatment process in a cancer hospital or center, understanding the radiotherapy chain, including all the steps that a patient visiting the clinic has to go through, is necessary. For those people who are unfamiliar with radiotherapy, the radiotherapy chain is complex and intimidating [1,2]. By developing a simple and comprehensive internet-based virtual assistant such as chatbot or bot through which the radiation treatment process can be explained to the user at their own pace, the user can be guided through each step in the radiotherapy chain with simple and relevant information. This was the basis for creating a chatbot capable of explaining this process and reducing the risk of misinformation for the user, while also benefiting from availability [3,4]. Patients must schedule an appointment with the radiation oncologist, who has a fixed period during which they can ask questions, but a bot will be permanently available to them.

Although the bot is functional for purposes of explaining radiotherapy chain to the users, it serves as a proof of concept for more sophisticated and open-ended health care chatbot. The IBM Watson application programming interface (API) allows the programmer to easily add to the program and update any connected applications automatically [5]. Alternatively, additional dialogue chains can be added into a single assistant, which could be used to combine a general assistant such as this one with more specialized assistants. For example, one could focus on a specific part of the radiotherapy chain and can respond to queries specific to a single portion.

Background

The Radiotherapy Chain

The radiotherapy chain refers to the various steps during a patient's treatment, from the initial diagnosis until their eventual radiation treatment in the cancer hospital or center. Generally, the steps in the radiotherapy chain consist of the major checkpoints in the radiation treatment. In this bot, the radiotherapy chain includes seven key stages in the following order: (1) patient diagnosis, consultation, and prescription; (2) patient positioning, immobilization, and simulation; (3) 3D imaging for radiation treatment planning; (4) target volume and organ delineation; (5) radiation treatment planning; (6) patient setup and plan verification; and (7) treatment delivery. First, the bot will provide a big picture of the radiotherapy chain to the user and state these 7 steps before explaining each of them in detail.

The first step in the chain is diagnosis, consultation, and provision of prescription. This is when the radiation oncologist meets the patient with cancer and conducts a medical

examination. After reviewing the patient's medical record, laboratory test result, and radiology report, the radiation oncologist decides whether radiotherapy is suitable for the patient. If this is the case, the radiation oncologist will prescribe radiation dose for the tumor and prepare the patient for the next step of patient positioning, immobilization, and simulation in the chain. This step, along with the next step, 3D imaging for radiation treatment planning, are sometimes combined and referred to as a single *patient simulation* step. The purpose of these steps is to acquire the patient's 3D image set of anatomy including the tumor through computed tomography (CT) for treatment planning. This involves positioning the patient on the treatment couch such that they can be seated comfortably, while at the same time, the radiotherapists are able to place the patient in the same position in daily treatment. As even a minor positional or orientation deviation would change the radiation beam targeting during irradiation, a variety of immobilization devices such as headrests and body molds are used to ensure minimal movement [6]. After the patient has been immobilized, a CT scan is performed to acquire the information about the patient's anatomy using a CT simulator. This allows the radiation staff to identify the tumor and critical organs. In some cases, a contrast agent will be injected into the patient to improve the quality of the scan [7]. Once the tumor has been identified and its size and shape have been determined, the patient is usually marked on the treatment couch, to indicate the targeted radiation dose delivery in the final step of the radiotherapy chain. For the acquired CT image set of the patient, the radiation oncologist contours the 3D shape and location of the tumor. This process is becoming more automated with the increasing accuracy of software image recognition and deep learning [8,9]. The tumor is referred to as the gross tumor volume. Once this has been obtained, the clinical target volume and planning target volume are identified [10]. The planning target volume includes portions of healthy tissue, which is simply an inevitable consequence of current radiation treatment. The target volume is the focus of the next step—treatment planning. The main goal is to determine a treatment plan, which delivers a high dose to the target volume, while at the same time, minimizes radiation exposure to the surrounding healthy tissues. Radiation oncologist, radiotherapist, and medical physicist work together to create a plan that is tailored to the treatment based on plan variables such as positions and sizes of the patient, target, and critical organs. After the treatment plan is created, patient setup verification is initiated. This is referred to as verification simulation. In this step, the patient's treatment is conducted as a *dry run* before the patient's actual treatment. This process is performed with the help of computers to simulate the patient's anatomy and cancer location, for both practice and quality assurance purposes [11]. A study conducted in 2018 confirmed that this process resulted in reduction in the rate of incidents during treatment and anxiety levels of both patients and staff during treatment [12]. After patient setup and

plan verification, the patient is finally ready to receive radiation treatment. In most cases, the course of treatment includes one to thirty fractions. In the patient's first fraction, the patient is placed in the position established during their initial immobilization. Radiotherapists deliver the dosage as prescribed by the radiation oncologist. Patient images from cone-beam CT or portal imager are taken at this point to guide the patient setup. This informs the radiation staff about whether the treatment dosage is effective [13,14].

Health Care Chatbots

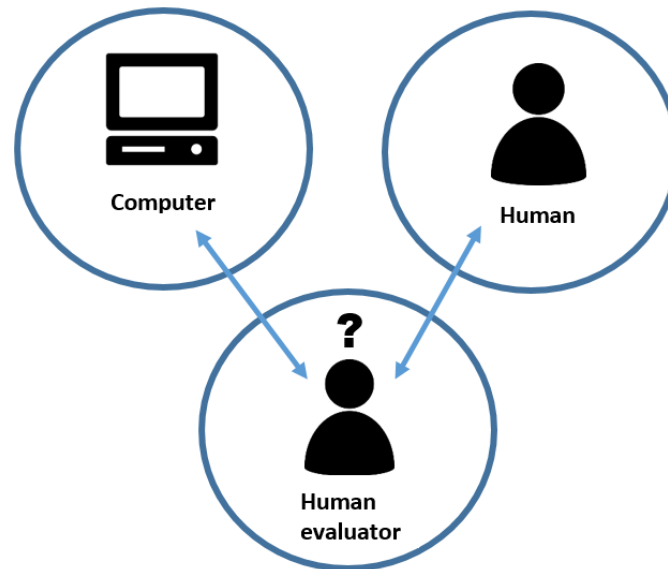
The radiotherapy chain is a lot of information for a patient or family member to process in a single session. It is important for them to understand the basic treatment process when visiting the cancer center, with a similar level of detail as a patient received during briefing from a radiation staff in an in-person education section. One of the most prominent advantages that health care chatbots have over physicians is their accessibility. A chatbot can be accessed by patients at any time through internet. An autonomous health care chatbot can learn from all available information on any given topic in minutes, as opposed to years. All physicians and medical specialists are limited by information gathered through their own experiences or readings. A software bot such as a patient-facing bot can feasibly learn from the experiences of every health care bot and process information on a given topic at a rate incomparable with that of humans [15]. A sufficiently advanced software bot could make connections that no human would be capable of simply owing to the incomparable processing capabilities of modern computers. Health care chatbots are an inevitable consequence of these advancements. They provide a means by which a practically infinite information source can be simplified and communicated to suit the needs of a human [16]. This can be achieved through scripting and computer programming on the chatbot virtual assistant platform. Chatbots for the purposes of health care have existed since the 1960s, but the computers of the time simply lacked both the information and optimization to have any practical use [17]. This was much before computers were capable of natural language processing (NLP) or machine

learning (ML). In recent years, many health care chatbots have been adopted for a variety of purposes, ranging from chatbots encouraging physical fitness to diagnostic chatbots [18]. Although health care chatbots have not yet been fully adopted by the medical field, receptiveness toward the idea of increased use of health care chatbots is moderately high and shows signs of continuing to increase simultaneously with improvements in NLP and ML [19,20]. This is especially true for informational health care chatbots, which show high levels of receptiveness among medical professionals. A study in the United Kingdom found high demand for increased chatbot use in health care, especially among young populations [21].

Turing Test

Turing test was suggested by Alan Turing in 1950 to justify a machine's ability to exhibit intelligent behavior equal to that of a human [22,23]. In the test, an evaluator would judge the conversations between a human and a machine (computer) designed to mimic human-like responses (Figure 1). The human evaluator, human, and computer are isolated from one another, and the conversation would be limited to the text-only channel. To pass the test, the evaluator must not distinguish the computer from the human. Since the Turing test was introduced, it has become an important concept in the philosophy of artificial intelligence to justify whether the intelligence of a computer is similar to that of a human. The Turing test affects the development of digital communication program such as NLP [24]. ELIZA is an NLP computer program created at the Artificial Intelligence laboratory in Massachusetts Institute of Technology between 1964 and 1966 [25]. ELIZA simulated the conversation with users using the pattern matching approach, which demonstrated superficiality of communication between humans and machines. As people attributed human-like feelings to ELIZA, it is believed that the program could positively influence the lives of many people. The example of ELIZA inspired the idea to create chatbot with artificial intelligence to help users who need human-like touch and communication [15,16].

Figure 1. Schematic diagram showing the Turing test. In the test with the human evaluator, in which the human and computer are separate from one another, the human evaluator asks a series of questions to both the human and computer. Then, the evaluator tries to identify which side is operated by the human and which side is operated by the computer.



Objectives

The goal of this study was to develop a bot capable of explaining the radiation treatment process in terms of radiotherapy chains to the user in a simple and informational manner. Maintaining the bot as less conversational would also benefit from having a reduced risk of uncanny valley effects. Using IBM Watson's API, a basic conversation structure was created to explain the purpose of a radiotherapy chain and teach the user about each step in the chain. During bot development, it was important to program sufficient number of expected intents for the program to understand what the user is trying to communicate and generate a response. Another goal was to enable the bot to respond to requests for additional information with more detail and, if this was insufficient, direct the user to a source for more in-depth information. Although this bot was designed with various restrictions in mind, an important intention was to allow the bot to be directly implemented into a variety of applications. Many rudimentary chatbots are restricted to a single conversation flow, which makes it time consuming for a user to make requests for information. This bot was created with user efficiency as a priority.

Methods

IBM Watson Assistant

ML and NLP

The bot was created using IBM Watson Assistant with ML feature. The current state of NLPs allows for Watson to flow through a given conversation structure and recognize how the user responds based on recognition of similar given examples, which are referred to as *intents* during development [26]. Although Watson cannot answer any question posed by the user, it can recognize a given information based on expected input. This is done by receiving a variety of example responses and recognizing similar users' responses using Watson's NLP. Once completed, Watson Assistant can be integrated into a

variety of different applications or websites depending on its purpose. One of the main advantages of Watson is its API [27], which is simple and user-friendly, allowing for a variety of very simple implementations into commonly used apps such as SMS text messaging and WhatsApp.

Watson API

Intent

Watson recognizes user input as *intents*. These are values chosen by the programmer based on expected user input. The values are separate from the dialogue chain and are effectively used by Watson as variables. If the chatbot is expected to request information from the user, the programmer provides as many predictable responses as possible to represent a specific user request. Some of these intents will be referred to frequently, whereas others will be referred to only once.

Nodes

Basic conversation structure is created using user-created nodes. Every unique node represents a specific section of the dialogue. In Watson, these nodes can be connected either in parallel or sequentially according to the dialogue tree. For example, the dialogue has the potential to split based on user input; then, the node can be created in parallel to another node. If the dialogue is only designed to flow from one point to another, the dialogue will be written in series. Nodes that lead to a series are referred to as parent nodes, and the nodes that follow are referred to as child nodes. Nodes that have no child nodes run in parallel with the first node and are referred to as root nodes. Each dialogue node can be activated by a specific condition—almost always in the form of intents. Once a node is completed, the chatbot has 2 options. It can either wait for a reply if user input is required for the next child node or jump to another node at any point in the dialogue.

Digressions

Dialogue nodes have many different options depending on their function in relation to the chatbot. Digression settings allow the

programmer to allow or disallow nodes from being activated, when the node exists outside the current dialogue chain. The programmer can use *digression* to set responses that can be activated at any point in time. This is especially effective for developing methods to allow the user to skip certain portions.

Multiple Conditioned Responses

Dialogue nodes can also be set to respond to multiple conditions and have varying responses depending on the user input. This is useful for forming responses to conversation-specific questions. A dialogue node can contain responses to multiple questions and then immediately return to the dialogue chain. Without this functionality, questions would need to be represented by individual nodes. Then, the nodes would require *jump to* commands to return to the original dialogue chain or duplicate sets of child nodes to continue the conversation.

Testing

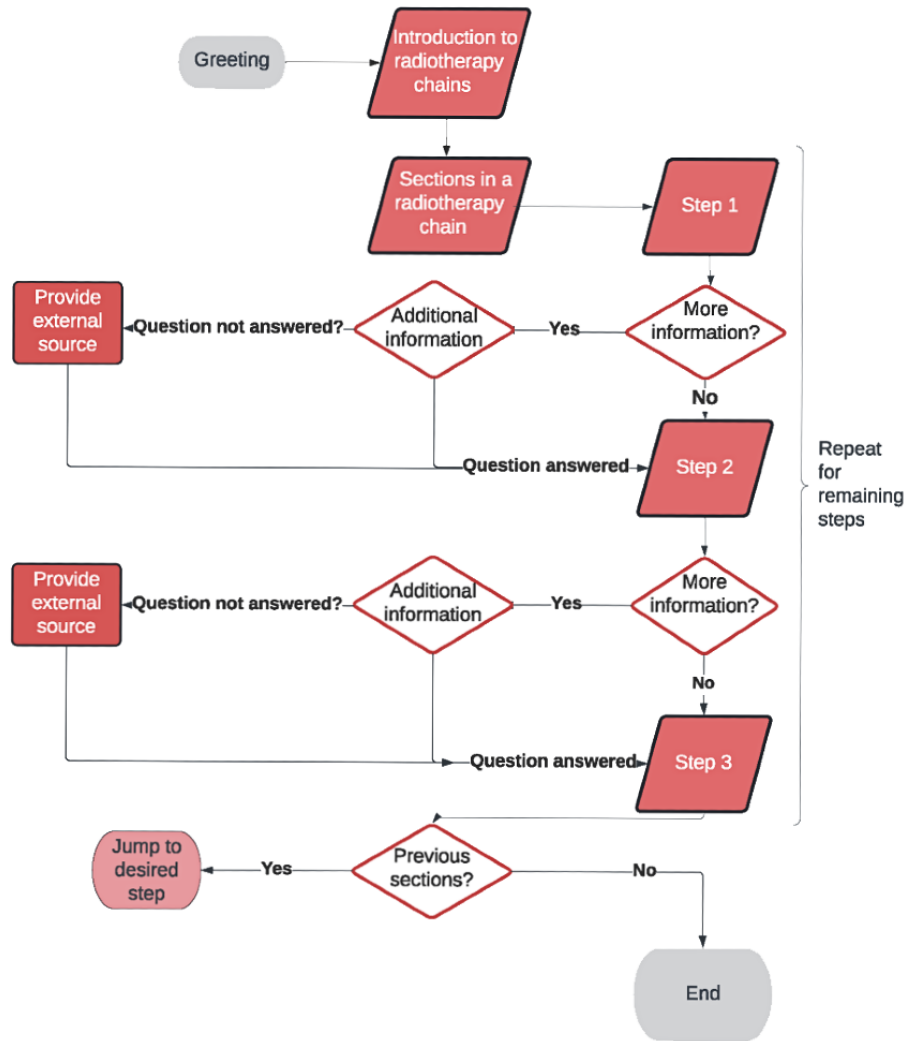
The testing feature in the API allows the programmer to examine the program without requiring an integration of the bot. The tester will display the user's intent when the user makes a statement. This can be used to solve disambiguation issues when the program has >1 similar intent.

The Radiotherapy Chain Dialogue Tree

The bot was designed by establishing a dialogue tree based on a series of predefined rules to guide the users by offering them a conditional *if or then* at each step. This conversation flow is

based on the workflow of the bot, as shown in [Figure 2](#). The tree was created by the programmer by considering the feedback from the users, and the complexity of the tree depends on the amount of content in the bot. The intent was to create an informational tool capable of being accessed and used by the user in the most efficient manner possible. From this, the basic dialogue tree was formed. This was represented in the Watson platform by a single parent node indicating the start of the conversation, with each step of the radiotherapy chain represented as a child node of the previous step. Once this was finalized, a variety of global variables were created in parallel to the initial parent chain. These were programmed using the *jump to* commands, because they are not part of the conversation structure. They can be accessed at any time if the user asks the bot about a specific section of the radiotherapy chain or if the user simply wants to start from the beginning. At this point, the bot was already functional, but it did not yet satisfy the goal of responding to queries for information. Child nodes were added to each link in the dialogue chain to respond to requests for more information and another child node was added to indicate whether the information was unsatisfactory to the user. [Figure 2](#) shows a flowchart of the dialogue chain. As indicated, the user can either go through the radiotherapy chain by selecting *continue* at each dialogue node. In [Figure 2](#), the nodes in the form of a red rhombus represent nodes along the parent chain, which require user input to continue. The white diamonds represent potential forks in the chain. After each of these forks, the bot will return to the parent chain.

Figure 2. Flowchart of the dialogue chain.



Integration

For integration, a website was chosen to host the bot for the sake of simplicity. As the topic is relatively specific, integrating it into a website allows the bot to be found by web crawlers. This ensured that if someone looks for a radiotherapy chatbot on a search engine, this bot can easily be found. The website was designed and hosted through Weebly (Block Inc)—a simple website builder that allows for easy-to-use customization options and hosting. The website was designed with simplicity in mind and to serve purely as a site for the bot.

Bot Testing

The bot was reviewed by the radiation staff and researchers regarding the idea, scope, content, and infrastructure in the cancer center, to fine-tune the bot for the nonexpert users among the general public, including patients and their family members. Specifically, a sample population of 50 people from the general public was asked to use and evaluate the bot and score their experience on a scale ranging from 1 to 5. This was measured based on 3 different metrics: information quality, user

experience, and navigability. Upon completion, they were asked to copy and paste their chat logs for review. This was conducted to establish trends in use. These trends could be used to guide further development of the bot.

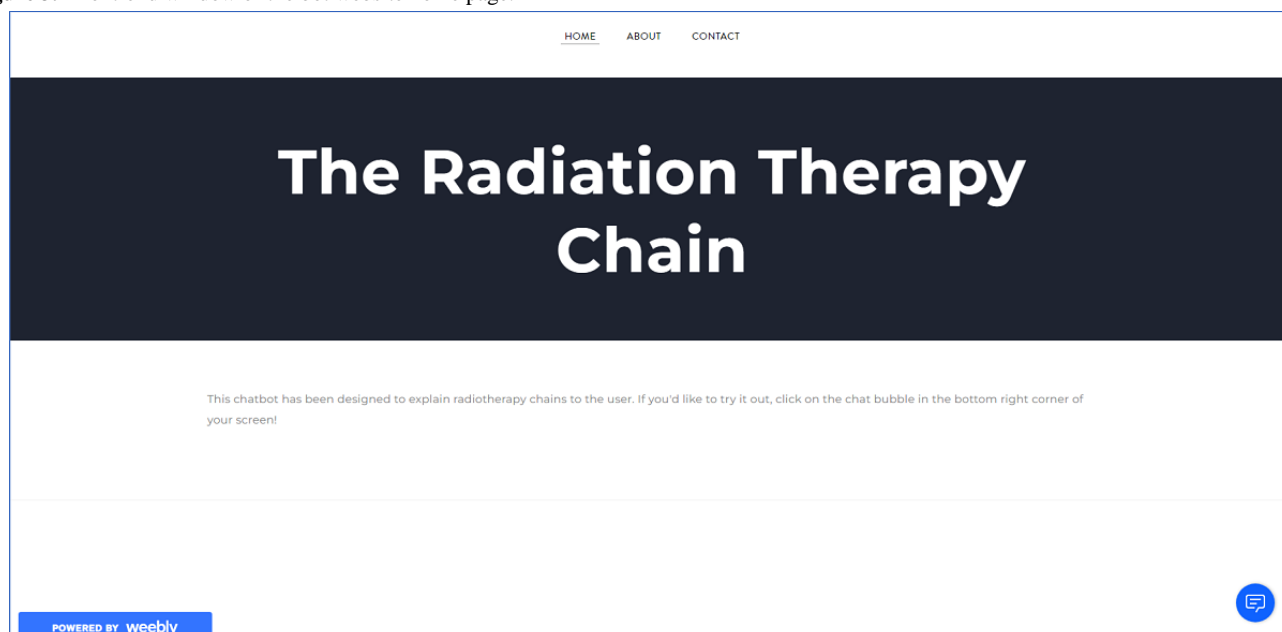
Ethics Approval

This study did not require ethics approval because it is only related to the creation of a software of an educational chatbot for information transfer with a performance evaluation and quality assurance. The data in this study are anonymous and deidentified.

Results

Overview of the Bot

The bot can be accessed via the web [28]. Currently, the bot is fully functional. As anticipated, the bot is fully capable of guiding the user through the conversation structure and can respond to simple questions and provide resources for requests of information that were not directly programmed into the bot. The front-end window of the bot is shown in Figure 3.

Figure 3. Front-end window of the bot website home page.

Communication Between the Bot and User

In [Figure 3](#), the user can start the bot by clicking the chat bubble in the bottom right-hand corner. An introductory message to say hello to the user and explain the functionalities of the bot is shown in [Figure 4A](#). The introductory message informs the user that the bot will guide them to understand the radiotherapy chain step by step. In addition, the user is welcome to ask question at any time and jump to different steps in the chain. When the user is ready, they can type “ready” in the textbox or click the “ready” icon to continue. In this case, the bot will send the next message to explain what a radiotherapy chain in the radiation treatment process is ([Figure 4B](#)). The user can answer “yes” to continue, and the bot will inform the user generally about the procedures the patient with cancer will need to follow one by one when they visit a cancer hospital or center ([Figure 4C](#)). When the user understands a step, they can continue to the next step. However, the user can answer “no” to the bot. In this case, the bot will request the user to inform when they will be ready to continue. This is shown in [Figure 4D](#).

When the bot has finished explaining a procedure in the radiotherapy chain such as consultation and diagnosis of a patient with cancer, as shown in [Figure 5A](#), the user will have

the option to select continuing to proceed to the next step in the treatment process or to have more information. If the user selects the latter, the bot will provide more information about the topic and ask again if the user still wants to know more. When the user answers that they would still want to know more, the bot will provide a link, so that the user can use it to access more related information on the internet, as shown in [Figure 5B](#). A connection to other websites for more information allows the bot to maintain simplicity and update from the user’s response.

When the bot receives unpredictable response from the user, for example, a random string of letters (eg, “aaaaabbbbcccccc”) instead of expected answers such as “ready,” “yes,” and so on, the bot will provide another window of suggestions to guide the user to answer the question ([Figure 6](#)). In that window of suggestions, the user only needs to select what they want to know from a list of items. Alternatively, the user can remove that window and answer the question again by rephrasing the answer. This can avoid incorrect communication between the bot and user when the latter cannot answer well (eg, typographical error in the answer). The guidance can also help the user to communicate with the bot when they cannot understand the question in the conversation.

Figure 4. Communications between the bot and the user regarding (A) an introductory message about the bot requiring the user to respond whether they are ready to start, (B) a brief explanation about the radiotherapy chain, and (C) more detailed description of the radiotherapy chain with each step in the process. The user can say “yes” to continue or “no” to discontinue, and (D) when the user answers “no,” the bot will stop until the user is ready to continue.

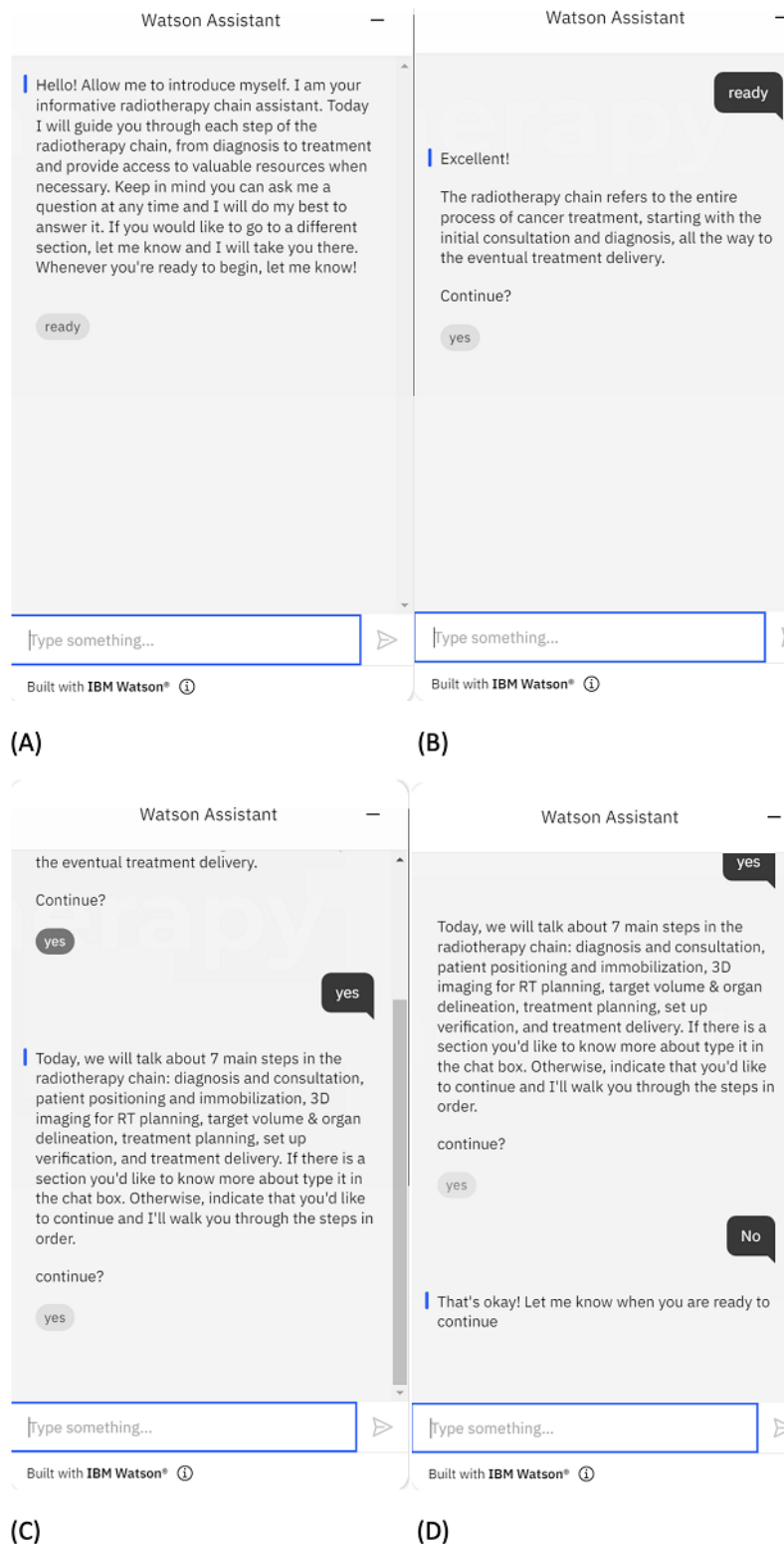


Figure 5. Responses of the bot when the user wants to have more information. (A) After the bot explains a procedure in the radiotherapy chain, the user is asked if they want to proceed to the next step or have more information. If the answer is “I’d like to know more!” the bot will provide more information. (B) After the information is provided, the bot will ask again if the user still wants to know more or continue; if the answer is “I still want to know more,” the bot will provide a link to the user so that they can assess more detailed information according to the topic from an external source.

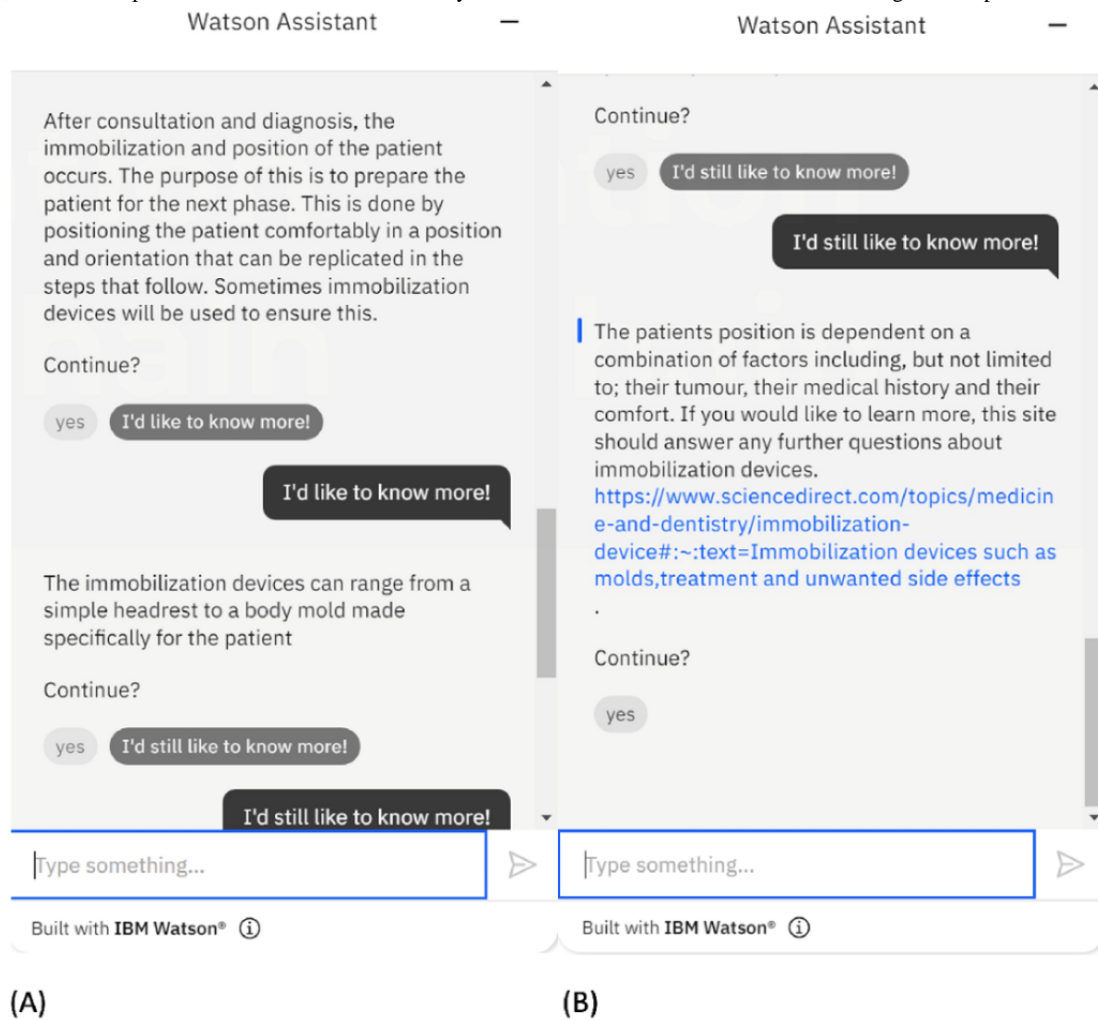
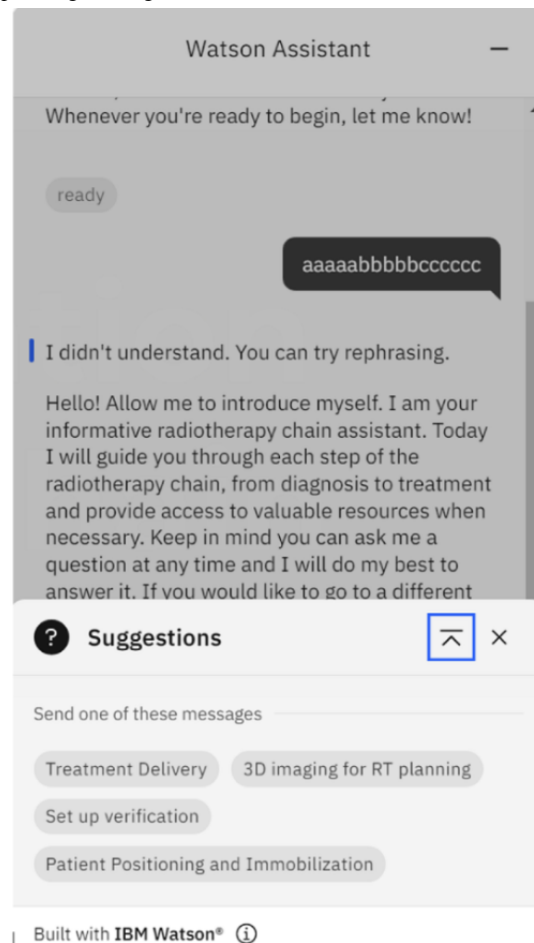


Figure 6. When the bot cannot receive a predictable response from the user, guidance will be provided with a list of items to select from. This can help the user to proceed to the next step even if they cannot answer well in the communication. Alternatively, the user can remove the window of suggestions and provide an answer to the bot again by rephrasing the original answer.



Performance Evaluation

The performance of the bot was evaluated randomly by a group of 50 people from the public. The participants were anonymous and unrelated to the development of the bot. The evaluation was based on three criteria including (1) information quality, (2) user experience, and (3) navigability. The score ranges from 1 to 5, corresponding to *poor*, *unsatisfactory*, *satisfactory*, *very satisfactory*, and *outstanding*, respectively. Here, 5 is the highest score, which indicates that the participant justified the bot as outstanding, whereas 1 is the lowest score, which indicates that the participant considered the bot as poor. The averages of the identified metrics were as follows: information quality=4.3 (SD 7.6%), user experience=3.6 (SD 5.2%), and navigability=4.7 (SD 4.1%). All metrics scored >3 (between satisfactory and very satisfactory), and both information quality and navigability scored >4 (between very satisfactory and outstanding). The scores indicated that participants in this evaluation were at least satisfied with the performance of the bot.

Discussion

Principal Findings

The aim of this study was to create a virtual assistant chatbot that can explain the radiotherapy chains to the user. This bot was designed using the help of IBM's Watson Assistant. Once

built, it will be incorporated into a resource that can be easily accessed by the general public. This chatbot was successfully created and hosted on an independent website [28]. The basic chain communicates about radiotherapy chains and their purpose to the user, with the ability to answer a few programmed questions.

The bot encountered very few issues during testing and was rated highly by users regarding navigability and quality of information. Inspecting the chat logs revealed that most users simply went through the root dialogue chain without requesting more information. The average score of 4.3 for information quality indicates that the bot's primary purpose is adequately served, as this indicates an improved understanding of the radiotherapy chain. The lowest of the 3 scores was found to be 3.6 for user experience. This was potentially owing to the limited interactions that the user has with the bot and the limited variability in content. The score for navigability, despite being high, does not necessarily indicate the success of *jump* to commands or restart commands.

Future Directions

As mentioned in the *Introduction* section, the IBM Watson's API makes it easy for the programmer to modify the underlying structure or intents at any given time and add additional nodes to the conversation. As such, the successful functionality of the parent chain suggests that this bot can be improved with very

few changes to the existing nodes. Adding additional root nodes could increase the number of relevant questions. With the addition of a multitude of nodes and the use of nodes with multiple conditioned responses, the bot could also be used for patient-specific information. For example, if a patient was diagnosed with cancer and had already progressed through their initial consultation, the bot could give the patient a more specific description of their remaining treatment based on the patient's circumstances. Similarly, the first node could ask the user the reason for their interest in the radiation treatment process and use a tailored dialogue chain based on their response. This would provide a more user-specific experience. Alternatively, combining this bot with other chatbots or creating other dialogue chains that discuss different topics in radiotherapy could generalize the user experience.

Limitations

There are a variety of unavoidable limitations in a bot of this nature. Currently, it will not be able to answer any questions outside the scope of radiotherapy and will not recognize unpredictable responses because this is the next stage of the configuration. An example is shown in [Figure 6](#). By providing a list of items to select in case of communication impasse, the user can choose one of them to return to the scope of the bot and be back on track of communication. Another way to improve the limitation of the scope of the bot is to link it to an external website for further information. This can maintain the database of the bot as simple and comprehensive but can still satisfy the user if they want to have more information beyond the ability

of the bot. Regarding the miscommunication of the bot when users entering unpredictable responses, the user can send feedback about the issue to the programmer of the bot. Then, the programmer can modify the bot to solve the problem in the routine update. It should be noted that comments and feedback from users are significant sources to improve the bot and reduce its limitations continuously.

Conclusions

An education chatbot was created for patients, their family members, and the general public to understand the basic radiation treatment process by navigating the radiotherapy chain. Self-reported scores for information quality indicate that it also successfully conveys the information to the user in a manner that can be understood by the general population. In addition, the ease with which this bot can be improved and the potential uses of a chatbot of this nature establish it successfully as an effective proof of concept for informative chatbots in health care. The chatbot can successfully navigate its users through a radiotherapy chain. Current studies on public opinion and continuing improvements in NLP, ML, and internet-based technologies [29] indicate that these chatbots will become more advanced. Health care chatbots such as the one described in this paper have great potential [30], as health care chatbots can surpass humans in both speed and accuracy. Future studies include continuously updating the bot based on the users' feedback. Thus, the performance of the bot can be improved from its preset process.

Acknowledgments

This study was supported by a Canadian Institutes of Health Research Planning and Dissemination Grant—Institute Community Support (CIHR PCS - 168296).

Conflicts of Interest

None declared.

References

1. Svensson GK. Quality assurance in radiation therapy: physics efforts. *Int J Radiat Oncol Biol Phys* 1984 Jun;10 Suppl 1:23-29. [doi: [10.1016/0360-3016\(84\)90441-3](https://doi.org/10.1016/0360-3016(84)90441-3)] [Medline: [6735791](https://pubmed.ncbi.nlm.nih.gov/6735791/)]
2. Webb S. The physics of radiation treatment. *Phys World* 2015 Sep 21;11(11):39-44. [doi: [10.1088/2058-7058/11/11/29](https://doi.org/10.1088/2058-7058/11/11/29)]
3. Sabater S, Roviroso À, Arenas M. In response to Korreman s. et al. Radiation oncologists are, above all, medical doctors. *Clin Transl Radiat Oncol* 2021 May;28:116-117 [FREE Full text] [doi: [10.1016/j.ctro.2021.03.005](https://doi.org/10.1016/j.ctro.2021.03.005)] [Medline: [33937531](https://pubmed.ncbi.nlm.nih.gov/33937531/)]
4. Chow JC. Computer method and modeling: medical biophysics applications in cancer therapy, medical imaging and drug delivery. *AIMS Biophy* 2021;8(3):233-235. [doi: [10.3934/biophy.2021017](https://doi.org/10.3934/biophy.2021017)]
5. da Silva Oliveira J, Espíndola DB, Barwaldt R, Ribeiro LM, Pias M. IBM WATSON application as FAQ assistant about moodle. In: *Proceedings of the 2019 IEEE Frontiers in Education Conference*. 2019 Presented at: FIE '19; October 16-19, 2019; Covington, KY, USA p. 1-8. [doi: [10.1109/fie43999.2019.9028667](https://doi.org/10.1109/fie43999.2019.9028667)]
6. Saw CB, Yakoob R, Enke CA, Lau TP, Ayyangar KM. Immobilization devices for intensity-modulated radiation therapy (IMRT). *Med Dosim* 2001;26(1):71-77. [doi: [10.1016/s0958-3947\(00\)00059-5](https://doi.org/10.1016/s0958-3947(00)00059-5)] [Medline: [11417510](https://pubmed.ncbi.nlm.nih.gov/11417510/)]
7. Nowak T, Hupfer M, Brauweiler R, Eisa F, Kalender WA. Potential of high-Z contrast agents in clinical contrast-enhanced computed tomography. *Med Phys* 2011 Dec;38(12):6469-6482. [doi: [10.1118/1.3658738](https://doi.org/10.1118/1.3658738)] [Medline: [22149830](https://pubmed.ncbi.nlm.nih.gov/22149830/)]
8. Zabel WJ, Conway JL, Gladwish A, Skliarenko J, Diodato G, Goorts-Matthews L, et al. Clinical evaluation of deep learning and atlas-based auto-contouring of bladder and rectum for prostate radiation therapy. *Pract Radiat Oncol* 2021;11(1):e80-e89 [FREE Full text] [doi: [10.1016/j.prro.2020.05.013](https://doi.org/10.1016/j.prro.2020.05.013)] [Medline: [32599279](https://pubmed.ncbi.nlm.nih.gov/32599279/)]
9. Chow JC. Artificial intelligence in radiotherapy and patient care. In: Lidströmer N, Ashrafian H, editors. *Artificial Intelligence in Medicine*. Cham, Switzerland: Springer; 2022:1275-1286.

10. Weiss E, Hess CF. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlenther Onkol* 2003 Jan;179(1):21-30. [doi: [10.1007/s00066-003-0976-5](https://doi.org/10.1007/s00066-003-0976-5)] [Medline: [12540981](https://pubmed.ncbi.nlm.nih.gov/12540981/)]
11. Kron T, Fox C, Ebert MA, Thwaites D. Quality management in radiotherapy treatment delivery. *J Med Imaging Radiat Oncol* 2022 Mar;66(2):279-290. [doi: [10.1111/1754-9485.13348](https://doi.org/10.1111/1754-9485.13348)] [Medline: [35243785](https://pubmed.ncbi.nlm.nih.gov/35243785/)]
12. Gresswell S, Renz P, Hasan S, Werts M, Fortunato M, Werts D. Determining the impact of pre-radiation treatment verification simulation/dry run by analyzing intradepartmental reported incidents and surveying staff and patients. *Pract Radiat Oncol* 2018;8(6):468-474 [FREE Full text] [doi: [10.1016/j.prro.2018.05.007](https://doi.org/10.1016/j.prro.2018.05.007)] [Medline: [30195926](https://pubmed.ncbi.nlm.nih.gov/30195926/)]
13. Gillan C, Giuliani M, Harnett N, Li W, Dawson LA, Gospodarowicz M, et al. Image guided radiation therapy: unlocking the future through knowledge translation. *Int J Radiat Oncol Biol Phys* 2016 Oct 01;96(2):248-250. [doi: [10.1016/j.ijrobp.2016.05.028](https://doi.org/10.1016/j.ijrobp.2016.05.028)] [Medline: [27598801](https://pubmed.ncbi.nlm.nih.gov/27598801/)]
14. Chow JC. Cone-beam CT dosimetry for the positional variation in isocenter: a Monte Carlo study. *Med Phys* 2009 Aug;36(8):3512-3520. [doi: [10.1118/1.3166934](https://doi.org/10.1118/1.3166934)] [Medline: [19746785](https://pubmed.ncbi.nlm.nih.gov/19746785/)]
15. Siddique S, Chow JC. Machine learning in healthcare communication. *Encyclopedia* 2021 Feb 14;1(1):220-239. [doi: [10.3390/encyclopedia1010021](https://doi.org/10.3390/encyclopedia1010021)]
16. Xu L, Sanders L, Li K, Chow JC. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR Cancer* 2021 Nov 29;7(4):e27850 [FREE Full text] [doi: [10.2196/27850](https://doi.org/10.2196/27850)] [Medline: [34847056](https://pubmed.ncbi.nlm.nih.gov/34847056/)]
17. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J* 2019 Jun;6(2):94-98 [FREE Full text] [doi: [10.7861/futurehosp.6-2-94](https://doi.org/10.7861/futurehosp.6-2-94)] [Medline: [31363513](https://pubmed.ncbi.nlm.nih.gov/31363513/)]
18. Kovacek D, Chow JC. An AI-assisted chatbot for radiation safety education in radiotherapy. *IOPSciNotes* 2021 Aug 30;2(3):034002. [doi: [10.1088/2633-1357/ac1f88](https://doi.org/10.1088/2633-1357/ac1f88)]
19. Siddique S, Chow JC. Artificial intelligence in radiotherapy. *Rep Pract Oncol Radiother* 2020;25(4):656-666 [FREE Full text] [doi: [10.1016/j.rpor.2020.03.015](https://doi.org/10.1016/j.rpor.2020.03.015)] [Medline: [32617080](https://pubmed.ncbi.nlm.nih.gov/32617080/)]
20. Adamopoulou E, Moussiades L. Chatbots: history, technology, and applications. *Mach Learn Appl* 2020 Dec;2:100006. [doi: [10.1016/j.mlwa.2020.100006](https://doi.org/10.1016/j.mlwa.2020.100006)]
21. Palanica A, Flaschner P, Thommandram A, Li M, Fossat Y. Physicians' perceptions of chatbots in health care: cross-sectional Web-based survey. *J Med Internet Res* 2019 Apr 05;21(4):e12887 [FREE Full text] [doi: [10.2196/12887](https://doi.org/10.2196/12887)] [Medline: [30950796](https://pubmed.ncbi.nlm.nih.gov/30950796/)]
22. Turing AM. Computing machinery and intelligence. In: Epstein R, Roberts G, Beber G, editors. *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Dordrecht, The Netherlands: Springer; 2009:23-65.
23. Turing AM. I.—Computing machinery and intelligence. *Mind* 1950 Oct;LIX(236):433-460. [doi: [10.1093/mind/lix.236.433](https://doi.org/10.1093/mind/lix.236.433)]
24. Chow JC. Applications of artificial intelligence, mathematical modeling and simulation in medical biophysics. *AIMS Biophys* 2021;8(1):121-122. [doi: [10.3934/biophy.2021009](https://doi.org/10.3934/biophy.2021009)]
25. Weizenbaum J. *Computer Power and Human Reason: From Judgment to Calculation*. New York, NY, USA: W. H. Freeman and Co; 1976.
26. Packowski S, Lakhana A. Using IBM Watson cloud services to build natural language processing solutions to leverage chat tools. In: *Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering*. 2017 Nov Presented at: CASCON '17; November 6-8, 2017; Markham, Canada p. 211-218.
27. Tan W, Fan Y, Ghoneim A, Hossain MA, Dustdar S. From the service-oriented architecture to the web API economy. *IEEE Internet Comput* 2016 Aug 2;20(4):64-68. [doi: [10.1109/mic.2016.74](https://doi.org/10.1109/mic.2016.74)]
28. The Radiation Therapy Chain. URL: <https://radiotherapy-chatbot.weebly.com> [accessed 2022-11-16]
29. Chow JC. Internet-based computer technology on radiotherapy. *Rep Pract Oncol Radiother* 2017 Nov;22(6):455-462 [FREE Full text] [doi: [10.1016/j.rpor.2017.08.005](https://doi.org/10.1016/j.rpor.2017.08.005)] [Medline: [28932174](https://pubmed.ncbi.nlm.nih.gov/28932174/)]
30. Milne-Ives M, de Cock C, Lim E, Shehadeh MH, de Pennington N, Mole G, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *J Med Internet Res* 2020 Oct 22;22(10):e20346 [FREE Full text] [doi: [10.2196/20346](https://doi.org/10.2196/20346)] [Medline: [33090118](https://pubmed.ncbi.nlm.nih.gov/33090118/)]

Abbreviations

- API:** application programming interface
 - CT:** computed tomography
 - ML:** machine learning
 - NLP:** natural language processing
-

Edited by A Mavragani; submitted 10.05.22; peer-reviewed by E Stroulia, W Zhang; comments to author 27.09.22; revised version received 29.09.22; accepted 11.10.22; published 02.12.22.

Please cite as:

Rebelo N, Sanders L, Li K, Chow JCL

Learning the Treatment Process in Radiotherapy Using an Artificial Intelligence-Assisted Chatbot: Development Study

JMIR Form Res 2022;6(12):e39443

URL: <https://formative.jmir.org/2022/12/e39443>

doi: [10.2196/39443](https://doi.org/10.2196/39443)

PMID: [36327383](https://pubmed.ncbi.nlm.nih.gov/36327383/)

©Nathanael Rebelo, Leslie Sanders, Kay Li, James C L Chow. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Web-Based Program About Sustainable Development Goals Focusing on Digital Learning, Digital Health Literacy, and Nutrition for Professional Development in Ethiopia and Rwanda: Development of a Pedagogical Method

Katarina Bälter^{1,2}, MSc, PhD; Feben Javan Abraham¹, MD, MMSc; Chantal Mutimukwe³, PhD; Reuben Mugisha³, MSc; Christine Persson Osowski¹, MSc, PhD; Olle Bälter³, PhD

¹Department of Public Health, School of Health, Care and Social Welfare, Mälardalen University, Västerås, Sweden

²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

³Department of Media Technology and Interaction Design, Kungliga Tekniska Högskolan, Royal Institute of Technology, Stockholm, Sweden

Corresponding Author:

Katarina Bälter, MSc, PhD

Department of Public Health

School of Health, Care and Social Welfare

Mälardalen University

Universitetsplan 1

Västerås, 722 20

Sweden

Phone: 46 021151753

Email: katarina.balter@mdu.se

Abstract

Background: East African countries face significant societal challenges related to sustainable development goals but have limited resources to address these problems, including a shortage of nutrition experts and health care workers, limited access to physical and digital infrastructure, and a shortage of advanced educational programs and continuing professional development.

Objective: This study aimed to develop a web-based program for sustainable development with a focus on digital learning, digital health literacy, and child nutrition, targeting government officials and decision-makers at nongovernmental organizations (NGOs) in Ethiopia and Rwanda.

Methods: A web-based program—OneLearns (Online Education for Leaders in Nutrition and Sustainability)—uses a question-based learning methodology. This is a research-based pedagogical method developed within the open learning initiative at Carnegie Mellon University, United States. Participants were recruited during the fall of 2020 from ministries of health, education, and agriculture and NGOs that have public health, nutrition, and education in their missions. The program was conducted during the spring of 2021.

Results: Of the 70 applicants, 25 (36%) were selected and remained active throughout the entire program and filled out a pre- and postassessment questionnaire. After the program, of the 25 applicants, 20 (80%, 95% CI 64%-96%) participants reported that their capacity to drive change related to the sustainable development goals as well as child nutrition in their organizations had increased *to large extent* or *to a very large extent*. Furthermore, 17 (68%, 95% CI 50%-86%) and 18 (72%, 95% CI 54%-90%) participants reported that their capacity to drive change related to digital health literacy and digital learning had increased *to a large extent* and *to a very large extent*, respectively.

Conclusions: Digital learning based on a question-based learning methodology was perceived as a useful method for increasing the capacity to drive change regarding sustainable development among government officials and decision-makers at NGOs in Ethiopia and Rwanda.

(JMIR Form Res 2022;6(12):e36585) doi:[10.2196/36585](https://doi.org/10.2196/36585)

KEYWORDS

digital learning; digital health literacy; sustainable development goals; public health; nutrition; question-based learning; open learning initiative; Rwanda; Ethiopia

Introduction

Background

East African countries face significant societal challenges covered by the global sustainable development goals (SDG) [1]. A total of 7 million people are at risk of starvation and >33 million people face acute food insecurity [1]. Furthermore, the pooled prevalence of chronic undernutrition among children aged <five years was 33% (95% CI 33-36%) in East Africa, ranging from 22% in Kenya to 53% in Burundi [2]. At the same time, East African countries have limited resources to address these problems [3], including a shortage of nutrition experts and health care workers [4,5], limited access to infrastructure, and a shortage of advanced educational programs and continuing professional development [4]. The SDGs were introduced to “achieve a better and sustainable future for all” [6]. To achieve the SDGs in general, good health and well-being (goal 3) and quality education (goal 4) in particular, effective education initiatives are needed [7]. Digital learning has the potential to overcome the lack of professional education and development [8], especially during the COVID-19 pandemic.

Although information and communication technology (ICT) is central to meeting new skills and training demands in most low- and middle-income countries [9], internet connectivity in most African institutions, especially those in rural and semiurban areas, is limited, expensive, unstable, or poorly managed [10,11]. Even though these countries are lagging in the adoption and implementation of effective digital learning tools [8], the advantages of digital learning in achieving the SDGs in East Africa are clear. In this context, digital learning, or e-learning, is based on the use of electronic media and devices as tools to improve access to training, communication, and interaction as well as to facilitate the adoption of new ways of understanding and learning [12]. However, research on digital learning issues in East Africa is scarce [4]. Pedagogical research from high-income countries may potentially be inapplicable to a low-resource context because of differences in teachers' educational backgrounds and experiences. In addition, local cultural contexts and perceptions must be considered [13].

Similar to digital learning, digital health literacy is a fundamental aspect of any country's economic development and a major contributor to the realization of SDGs [14]. Digital literacy is the ability to understand and use information in multiple formats from a wide variety of sources when it is presented via electronic sources [15]. Digital health literacy is the ability to seek, understand, and appraise health information from electronic sources and apply the knowledge gained to address or solve health problems [16]. However, most individuals holding decision-making positions in East Africa lack basic knowledge of ICT and related training [17]. Training in information technology literacy is therefore essential for the successful integration of digital health initiatives into existing health care services in low- and middle-income countries [18].

Rwanda encourages web-based learning and the digitization of several other service sectors, such as health care, commerce, and governance [19]. However, academics and students have been reluctant to fully engage in web-based education, mainly because most web-based platforms and applications fail to reach the core of the learning crisis in Rwanda, because some learning materials are difficult to comprehend, and there is no documented evidence of their effectiveness compared with physical classrooms. In Ethiopia, digitalization of education in schools and universities is still in its infancy. This is primarily because of the lack of infrastructure development in ICT and insufficient human resources with knowledge and training, creating barriers to move forward [20]. Therefore, we developed the OneLearns (Online Education for Leaders in Nutrition and Sustainability) program aimed at increasing human resource capacity within the digitalization of the health and education fields. Specifically, OneLearns capitalizes on the documented evidence of question-based learning methodology as an effective and time-saving method for both students and teachers [21].

Objectives

The OneLearns program targeted government officials and decision-makers in nongovernmental organizations (NGOs) in Ethiopia and Rwanda and covered the SDGs with a focus on digital learning, digital health literacy, and nutrition. We described the development of the OneLearns program, learning objectives, question-based learning methodology, process of recruiting participants, and the results of the program evaluation. Finally, we highlighted some lessons learned and ways to improve the program in the future.

Methods

The Teaching Team

The OneLearns program was designed and developed by a team of 6 experts at the KTH Royal Institute of Technology and Mälardalen University, Sweden, between September 2020 and December 2020. The team included a professor, 2 associate professors, a postdoctoral researcher, a PhD student, and a research assistant with a master's degree. The team had expertise in different domains including digital learning technologies, digital health literacy, sustainable development, public health, and nutrition. The program was based on the team's research fields, but to ensure that the program would be relevant for the target participants from Rwanda and Ethiopia, the content was discussed with research colleagues from Rwanda and Ethiopia, including members of the teaching team from Ethiopia and Rwanda.

Question-Based Learning Methodology

The program used a question-based learning methodology in which the learning material was organized around formative questions linked to one or more skills, which in turn were linked to the learning objectives [21]. More concrete examples from the program are shown in Figures 1 and 2, but in general, this

question-based learning methodology is a research-based pedagogical method developed within the open learning initiative at Carnegie Mellon University, United States [22]. The foundation of the methodology is that questions with constructive feedback are scattered over the learning material to stimulate students to engage with it. The methodology is used both in fully web-based courses and on campus, often in a flipped classroom setting [23], meaning that students worked with the web-based question-based material before coming to class, and the teacher used the learning data accumulated from the students' activities in the learning material to plan the lecture. This type of active learning is 6 times more efficient than reading and watching videos [24]. The methodology used basic Bayesian hierarchical models to predict student mastery

[25], and when used repeatedly and refined with the data from each course iteration, this method became very effective for learners. In a randomized case-control study on campus at Carnegie Mellon University, Pittsburgh, United States with this type of web-based learning material, learning time was reduced by 50%, while maintaining the learning outcomes compared with a parallel traditional course [22]. For newly developed material (without iteratively improved material, such as in this project), the reduction in learning time was estimated to be 25%, with maintained learning outcomes [21]. This method was therefore considered suitable for the target participants, as they are professionals with full-time work and, hence, do not have much time.

Figure 1. An example of a formative question from the module on Child nutrition in the web-based material. The green rectangle shows feedback that the participant received after selecting the correct answer. It serves the purpose of reinforcing the correct answer and teaches the student something new about the topic.

Learn by doing

Which of the following statements about energy is true?

- Children need to eat more energy than adults since they are growing.
- Sweet foods such as cookies and candy should be provided in children's diets as they provide energy needed for growth.
- Fruit and berries should be avoided as they contain sugar.
- If the child eats a vegetarian diet, it may be necessary to include, for instance, extra vegetable oil in order to provide enough energy.

Correct. The content of carbohydrates is generally high, and the content of fat is generally low in plant-based food items. Vegetable oil contains healthy fatty acids and may be added to the diet in order to provide enough energy.

Figure 2. An example of a formative question from the module on Child nutrition in the web-based material. The red rectangle shows feedback that the participant received after selecting an incorrect answer and targets common misconceptions associated with the wrong answer, thus contributing to the learning process.

Learn by doing

Which of the following statements about energy is true?

- Children need to eat more energy than adults since they are growing.
- Sweet foods such as cookies and candy should be provided in children's diets as they provide energy needed for growth.
- Fruit and berries should be avoided as they contain sugar.
- If the child eats a vegetarian diet, it may be necessary to include, for instance, extra vegetable oil in order to provide enough energy.

Incorrect. Children's energy needs are lower than adults', but it is important that the food they eat is nutrient dense. For instance, children have high iron needs and therefore have to eat foods rich in iron.

Program Design and Development

The program was designed for participants' self-paced studies of digital material followed by web-based video seminars at the end of each unit, in line with the principle of the flipped classroom. The program comprised 3 units (Multimedia Appendix 1). The first unit was titled "Web-based learning" and subdivided into 2 modules: "Becoming a web-based learner" and "Effective digital learning." The second unit focused on

"Digital health literacy" and included the 2 modules "Introduction to digital health literacy" and "Fundamentals of digital literacy in contemporary health care." Finally, the third unit, "Nutrition and the SDGs" consisted of a module "Nutrition and sustainability."

To ensure that the material was relevant to the target group and countries, half (3/6, 50%) of the OneLearns teachers were from Ethiopia or Rwanda. Moreover, policy documents and other

teaching materials were reviewed to ensure that the topics included in the program were relevant for and adjusted to their corresponding countries. The first step in the development of the program was to determine the learning objectives and related skills for each module. In total, 17 learning objectives and 47 skills were identified; an example of a skill can be as follows: “Explain the pros and cons of web-based learning.” Thereafter, program designers formulated a series of formative questions for each skill. The formative questions formed the backbone of the digital learning material [22] and were designed as multiple-choice questions, ordering questions, select-all-that-apply questions, and drag-and-drop questions. The idea behind the formative questions was to engage students in collaborative tasks that supported authentic practice with the concepts and skills they were learning. The questions aimed for, at minimum, the *understanding* level of the Bloom taxonomy of learning, which is a hierarchical model that categorizes learning objectives into varying levels of complexity, and the level of “understanding” is the second lowest level [26]. Each formative question generated automatic feedback to the participant regarding whether the answer was correct or incorrect. Feedback reinforced the correct answer and targeted common misconceptions associated with incorrect answers (Figures 1 and 2). Construction of the formative questions with

relevant answering alternatives and constructive feedback was based on the principles developed by Glassey and Bälter [27] (Textbox 1).

Next, the program designers compiled the digital learning material in the form of short texts, images, simulations, short videos, and working examples to support the learning process, ensuring that the participants could answer the formative questions. Part of the learning material was reused from open sources (websites, video clips, and so on, especially from Athabasca University, Stanford University, and Carnegie Mellon University), and the rest was based on the teaching team’s own research. Finally, questions for the final assessment tests were formulated. The final assessment tests took place at the end of modules B, C, D, and E (Multimedia Appendix 1) and were referred to as module tests. No feedback besides right or wrong answers was offered, but the participants had the opportunity to retake each module test up to three times for revision and refinement. The final score was either pass or fail, and 100% of correct answers were required to pass the tests in modules B, C, and D, whereas it was 80% for module E because that module had more questions than the others. The module tests were either autocorrected or graded manually, as in the case of essay or short-answer questions.

Textbox 1. The construction of formative questions with relevant answering alternatives and constructive feedback was based on the principles developed by Glassey and Bälter [27].

Principles of good questions

- Question is from the program domain (ie, related to the learning objectives of the program).
- Question is targeted toward a misconception.
- Question is not based on a reference lookup.
- Question is reasonable to solve without external systems.

Principles of good answering alternatives

- Three or more answer alternatives are provided.
- Answer alternatives are plausible and linked to the misconception.
- Answer alternatives are formulated to maximize readability.

Principles of good feedback

- Feedback is constructive.
- Feedback is unique and provided for each answer alternative.
- Feedback for answer alternatives does not reveal the answer.

Recruitment Process

The recruitment process continued between September 2020 and December 2020, targeting potential participants from various ministries with a focus on the ministries of health, education, and agriculture and NGOs with public health, nutrition, and education in their missions. We collaborated with the embassies of Rwanda and Ethiopia in Stockholm, who distributed calls for applications to ministries and NGOs in their respective countries. To reach the different ministries, the information had to go via the ministry of foreign affairs, whereas the NGOs were reached via less formal networks. In addition, we sent information to our own network of NGO contact persons

and to the Swedish embassies in Ethiopia and Rwanda, which they then shared further.

We received 70 applications, including 42 from Ethiopia and 28 from Rwanda. Applicants were screened and scored according to the professional field, leadership experience, educational background, current job position, and their potential for decision-making, as well as the ability to express oneself in English. Furthermore, we aimed to achieve a gender balance of close to 1:1 ratio for men and women. On the basis of applicant scoring, 27 applicants were selected for the second phase of recruitment, where web-based interviews were conducted primarily to verify sufficient communication and verbal skills

in English. Although 3 (11%) participants were recognized as having less experience communicating orally in English, none of them were excluded. Of the 27 participants, 13 (48%) were men, 14 (52%) women, 15 (56%) from Ethiopia, and 12 (44%) from Rwanda. Of these, 26 (96%) successfully registered to the program, but 1 (4%) requested withdrawal from the program after the registration, leaving 25 active participants. For the 2 (7%) participants who did not start the program, one reported a change in work circumstances that conflicted with the program's schedule, while the other did not communicate a reason and failed to register completely.

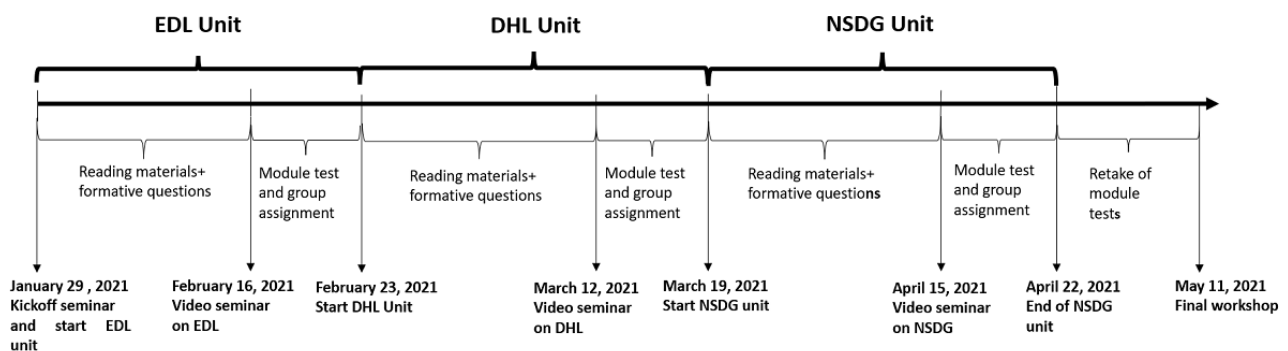
Program Delivery and Timeline

The program continued between January 29, 2021, and May 11, 2021 (Figure 3), and it was conducted entirely on the web owing to the ongoing COVID-19 pandemic. The original plan was to invite a representative subgroup of participants to Sweden at the beginning of the program and conduct the final workshop on-site in Ethiopia and Rwanda, but this part of the program was canceled as the pandemic did not allow travel. Each participant was expected to spend approximately 40 hours of

study time during the program, and the program design allowed a self-paced learning approach to a large extent.

Each unit was supposed to be completed within approximately 25 days. The participants began by working with the digital material and answering formative questions, followed by a web-based video seminar with teachers from Sweden and a final group assignment and module test. In the absence of physical interaction, it was important to ensure that the participants interacted and networked with each other. Therefore, 1 group assignment per program unit was included. Reminders were sent weekly with deadlines, including a general group progress report. The program ended on May 11, 2021, with a full-day web-based workshop and included invited speakers from Sweden, Rwanda, and Ethiopia and "learning by doing" activities. Unlimited data packages were provided to all participants in Ethiopia and Rwanda to ensure that they could be involved in the workshop. In addition, the participants in Rwanda were invited to meet in a large conference room at a hotel in Kigali and take part in the workshop from there, whereas the COVID-19 pandemic restrictions in Ethiopia did not allow for physical gatherings at that point in time.

Figure 3. Timeline for the web-based program OneLearns (Online Education for Leaders in Nutrition and Sustainability). The program ran between January 29, 2021, and May 11, 2021, and comprised 3 modules. DHL: digital health literacy; EDL: Effective Digital Learning; NSDG: nutrition and the sustainable development goal.



Means of Digital Communications and Interaction

The predominant means of communicating program information, progress updates, and deadline reminders to the participants was email. For less formal communication, that is, sharing scientific articles or pertinent seminars, a WhatsApp group was set up so that the participants could network as well. Google Docs was used for group assignment submissions, and teachers used the comment function to provide feedback on assignments. The initial kickoff, video seminars at the end of each module, and final workshop were conducted via Zoom (Zoom Video Communications, Inc). Moreover, during the final workshop, a Miro board, a web-based creative collaboration platform [28], was used for brainstorming, and the sticky note function was used to share ideas on a web-based and collaborative board (Multimedia Appendix 2).

Data Collection and Analysis

Pre- and postassessment surveys were administered to assess the participants' knowledge, capacity, and experience both

before and after the program. These web-based surveys included questions provided by the grant provider and the national foreign aid agency (Swedish Institute). To the best of our knowledge, these questions have not been validated but are used in all programs that they support. The surveys were distributed using the survey software Survey Generator and included automatic reminders to nonresponders. The first survey had 17 questions, whereas the latter had 9 (Multimedia Appendices 3 and 4).

The overall aim of the preassessment survey was to capture the participants' demographic and professional profiles and their own perception of their capacity in the various topics, that is, the current level of knowledge and confidence to bring about change. Questions such as "To what extent do you have sufficient knowledge and skills to drive change regarding digital learning" were used, with a nominal scale with options "not at all," "to a small extent," "to some extent," "to a large extent," or "to a very large extent" or "not applicable." Owing to the small number of participants, categories were merged into three main categories (Figures 4 and 5), (1) *not at all and to a small extent*, (2) *to some extent*, and (3) *to a large or to a very large*

extent. Descriptive data were generated by Survey Generator from Altra or analyzed using IBM SPSS 26 Statistics or Microsoft Excel Analysis ToolPak.

To evaluate the program material, an end-of-unit evaluation with open-ended questions was sent to the 25 participants after the *Efficient digital learning* and the *Nutrition and sustainability* units were completed, to which 12 (48%) and 18 (72%) participants responded, respectively. Questions such as “Which section did you learn the most from?” and ‘What did you think was missing from the unit?’ were used to identify areas of improvement for future programs; however, an analysis of the results was beyond the scope of this study.

Data on the performance of the participants, that is, their engagement with the learning material and learning objectives, were automatically collected by the open learning initiative system. Response rates and answering patterns of the formative questions were assessed; for example, how many times different answer alternatives were selected for the multiple-choice questions. In addition, based on the data that the open learning initiative system collected from participant engagement with the program material, it automatically generated learning curves for each of the 47 skills in the program.

Figure 4. Participants self-reported knowledge and skills to drive change before the program for various topics.

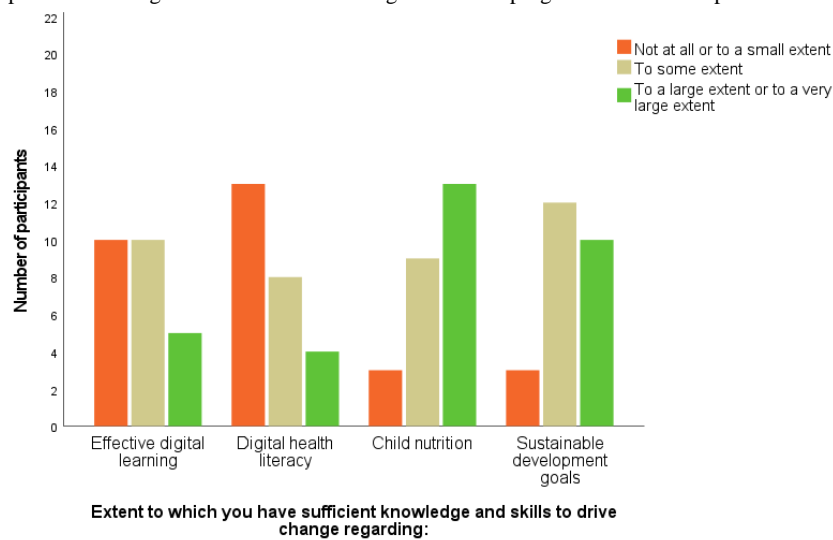
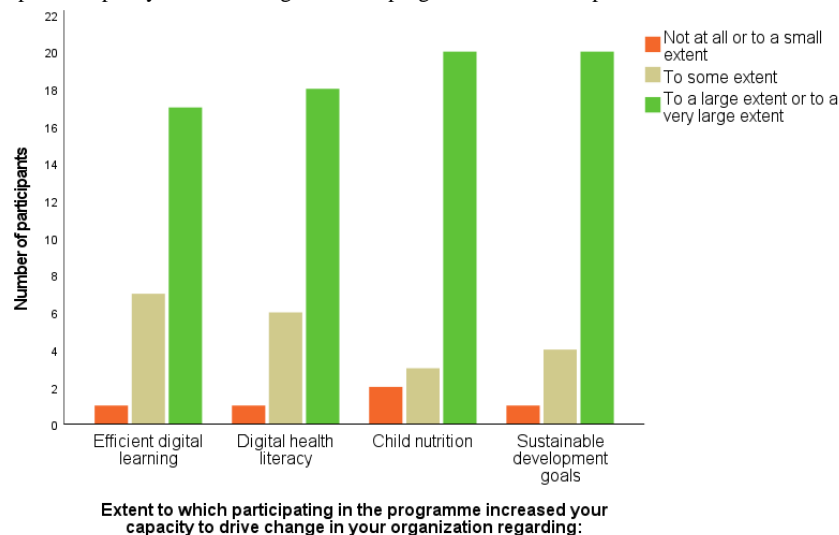


Figure 5. Participants self-reported capacity to drive change after the program for various topics.



Ethical Considerations

Paragraph 2 of the Swedish Ethical Review Authority states that ethics approval is only necessary if a project collects sensitive personal data from the participants, involves physical procedures, uses a method that will impact the participants physically or physiologically or if there is a risk of harm, or involves biological material [29]. None of this was the case in the OneLearns program; therefore, we did not apply for ethics

approval. Participation in the web-based program was voluntary, and it was possible to discontinue the program at any point without informing why. All participants received written and oral information about the program before it started, and filling out the pre- and postassessment questionnaires was considered as consent to help evaluate and improve the educational program.

Results

The Participants and Program Completion Rate

Table 1 shows demographic characteristics of the 25 participants enrolled in the program. Most participants (22/25, 88%) were aged <40 years, the gender distribution was almost equal with 13 (52%) women and 12 (48%) men, and 13 (52%) were from Ethiopia and 12 (48%) from Rwanda. Almost half (12/25, 48%) of the participants were working in governmental institutions, followed by 10 (40%) working in NGOs, and finally 3 (12%) in private organizations and other organizations that engaged with civil society; for example, project management or consultancy or other sectors. Most of the participants had educational backgrounds related to health, such as public health

(8/25, 32%), nutrition (5/25, 20%), and the medical field (3/25, 12%). The remaining participants had backgrounds in agriculture (2/25, 8%) or other various backgrounds (7/25, 28%), including software, IT, project management and business administration.

All 25 participants remained active and participated throughout the entire program, but 3 (12%) did not complete all the tasks required by the program, such as formative questions and module tests. Divided by units, of the 25 participants, 24 (96%) successfully completed all the requirements of units 1 and 2, and 23 (92%) completed those of unit 3. Participants who completed all the tasks were given a certificate of completion, and the 3 (%) participants who remained active but did not complete all the tasks required by the program received a certificate of participation.

Table 1. Characteristics of the participants in the program (N=25).

Characteristics	Participants, n (%)
Gender	
Men	12 (48)
Women	13 (52)
Country	
Ethiopia	13 (52)
Rwanda	12 (48)
Age (years)	
<30	7 (28)
30-39	15 (60)
40-49	1 (4)
50-65	2 (8)
Occupation	
Government staff and public sector	12 (48)
NGOs ^a	10 (40)
Private sector	2 (8)
Other	1 (4)
Education level	
Master's level	19 (76)
Bachelor's degree	4 (16)
Others	2 (8)
Educational background	
Public health	8 (32)
Nutrition	5 (20)
Project management	3 (12)
Medical field	3 (12)
Agriculture	2 (8)
Others	4 (16)

^aNGO: nongovernmental organization.

Participants' Knowledge Before and After the Program

On the basis of the results from the preassessment survey, the topic that the participants self-reported having the least previous knowledge about was digital health literacy, and 17 of the 25 (68%, 95% CI 50%-86%) participants expressed having *none* or *very little* knowledge before the program started. In addition, 12 (48%, 95% CI 28%-68%) participants reported having *none* or *very little* knowledge about digital learning before the program started. Corresponding numbers for knowledge level rated as *none* or *very little* for child nutrition and SDGs were 2 (8%, 95% CI 0%-19%) participants and 1 (4%, 95% CI 0%-12%) participant, respectively.

When asked about to what extent they possessed sufficient knowledge and skills to drive change in their organizations before the program started, only 4 (16%, 95% CI 2%-30%) participants replied *to a large extent* or *to a very large extent* for digital health literacy and 5 (20%, 95% CI 4%-36%) participants regarding digital learning. The corresponding numbers for SDGs and child nutrition were 10 (40%, 95% CI 21%-59%) and 13 (52%, 95% CI 32%-73%) participants, respectively (Figure 4).

Least previous knowledge was noted for digital health literacy and digital learning before the program started, but after the program, of the 25 participants, 24 (96%, 95% CI 88%-100%) reported that their knowledge regarding the digital health literacy had increased *to a large extent* or *to a very large extent* after the program, and 23 (92%, 95% CI 81%-100%) reported that their knowledge regarding digital learning increased *to a large extent* or *to a very large extent*. Moreover, despite having extensive self-reported knowledge about the SDGs before the program, the postassessment survey showed that all 25 participants reported that their knowledge regarding the SDGs had increased *to a large extent* or *to a very large extent* after the program, whereas 20 (80%, 95% CI 64%-96%) participants reported that their knowledge regarding child nutrition increased *to a large extent* or *to a very large extent*. Regarding the participants' capacity to drive change in their organizations after completing the program, 20 (80%, 95% CI 64%-96%) participants reported that their capacity related to the SDGs had increased *to a large extent* or *to a very large extent*, and 20 (80%, 95% CI 64%-96%) reported that their capacity to drive change related to child nutrition increased *to a large extent* or *to a very large extent*. Furthermore, 18 (72%, 95% CI 54%-90%) participants reported that their capacity to drive change related to digital health literacy had increased *to a large extent* or *to a very large extent*, and 17 (68%, 95% CI 50%-86%) reported that their capacity to drive change regarding digital learning had increased "to a large extent or to a very large extent" (Figure 5).

Participants' Feedback About the Program

Different types of digital learning materials were provided and participants were asked in the postassessment survey about which materials had been the most helpful. Multiple answers were allowed, and almost all participants reported that they found the formative questions helpful (24/25, 96%), followed by module test questions (22/25, 88%), reading the program text and web pages (19/25, 76%), watching videos (18/25, 72%),

and web-based workshops (16/25, 64%), and more than half (14/25, 56%) of the participants reported that group assignments were helpful. Overall, approximately three-fourths (19/25, 76%) of the participants felt that their participation in the program was worth the time and effort they had put in.

The postassessment survey included 2 open-ended questions: (1) "Did you miss something in the program?" and (2) "Did the program meet your expectations?" For the first question, more than half (15/25, 60%) of the participants reported that they did not feel anything was missing in the program. The second most common response was that the participants felt that they would have benefited from physical interactions. The second open-ended question yielded more uniform results, in which all but 1 (4%) respondent felt that the program met their expectations. For example, one of the participants wrote the following:

...The content and topics were beyond my expectations. For some modules, it was my first time to hear about them especially in Unit 1 (digital learning) and unit 2 (module 3 and 4) digital health literacy in digital ecosystem was an eye opener. I am looking forward to explore it in my routine work...I am now confidently organizing a Facebook campaign on Covid-19 in Haiti...I do not feel intimidated anymore. Even if it's something I do not understand, I search online, YouTube and get hands-on..."

Regarding web-based networking and interaction, more than half (14/25, 56%) of the participants reported having been in contact with other program participants on a monthly or weekly basis. In addition, more than half (14/25, 56%) of the participants reported having been in contact on a monthly or weekly basis with teachers participating in the program. No participant had been completely out of touch with the other participants, whereas 1 (4%) participant reported that they had never been in contact with teachers outside of the class.

Discussion

Principal Findings

The OneLearns program showed that a digital program is a useful approach to enhance knowledge about issues related to sustainable development among government officials and leaders of NGOs in Ethiopia and Rwanda. All participants (25/25, 100%) reported that their knowledge regarding the SDGs had increased to a large extent or more after the program, and 80% (20/25) reported that their capacity to drive change in their organizations related to the SDGs had increased to a large extent or more. In this program, midlevel to high-ranking government officials and leaders of NGOs were targeted because they influence policy making in their respective countries and often collaborate to bring about changes in communities and policies [30]. Hence, both sectors were included to ensure maximum outreach.

Knowledge, skills, and abilities after completion of a web-based learning experience are common ways to define the effectiveness of a web-based education program [31]; the results from OneLearns show the potential of web-based learning in general

and question-based methodology in particular, in Rwanda and Ethiopia. This outcome is in line with previous studies highlighting the benefits of international digital learning to promote good health and well-being as well as advanced education [8,32] and achieving SDGs [33].

Participants also reported that they gained substantial knowledge on topics in which they had initially believed they were already well-versed, such as in the case of the SDGs. At the same time, for topics that the participants reported as having low previous knowledge, such as digital learning and digital health literacy, 92% (23/25) and 96% (24/25), respectively, of the participants reported that their knowledge regarding these issues increased to a large extent or to a very large extent. This indicated the importance of offering advanced professional development to experts in the ministries and NGOs in these countries.

The participants were expected to spend 40 hours on the program over a 3-month period at the same time as they were working in their regular jobs. An advantage of self-study of digital material is that it was possible for the participants to adjust their reading to their own schedules. However, the disadvantage was time management issues, such as interference of daily life routines in studies and some personal problems, which have previously been reported from similar programs [34] and were echoed in the OneLearns program. Moreover, a limitation when conducting web-based learning was poor internet access and connectivity, a well-known challenge in Rwanda and Ethiopia [5] and a frequently reported reason for missed deadlines and poor engagement. For this reason, we offered a physical conference room in Kigali for the final workshop (the bans on meetings in Ethiopia owing to the pandemic prevented the same setup in Addis Ababa) to assure a stable internet connection, but all participants, regardless of physical location, were asked to complete the same tasks during the workshop. Another challenge with professional education in this target group was that the participants had full-time jobs that may have required them to “go out in the field” (as they expressed it) with short notice, which affected their possibilities to partake in program events. We tried to accommodate this by providing extended deadlines, but in the end, it was up to the participants and their employer to prioritize between the program and their ordinary work tasks. This is in line with the study by Heller et al [35], who also noted that those who required more time or had difficulty completing their tasks often reported conflicts with their work schedules.

An example of a digital barrier was that we provided unlimited data packages for all participants for the final workshop, but the poor infrastructure from the local service provider impeded a few participants from having reliable connectivity during this full-day collaborative event. In addition, although weekly reminders were sent containing deadline dates and a general group progress report, deadlines for all 4 module tests had to be extended for some of the participants based on their personal needs. This was done to accommodate the needs of the participants, as the program was a professional development program and not a formal academic course. The foremost goal of the program was to ensure that the participants learned the course material.

An incentive to fulfill all the requirements, announced at the first kickoff seminar, was that the participants in OneLearns were promised a printed and digital certificate at the end of the program. Moreover, a few ambitious and high-performing participants were appointed as course ambassadors. This initiative was implemented after the program started to recognize their efforts and form long-lasting relations between participants and teachers. For example, 2 ambassadors were invited to speak at the kickoff of the second edition of the OneLearns program, which started in September 2021. In addition, all participants were invited to become alumni members of the international network run by the Swedish Institute.

Regarding communication with the participants, the WhatsApp group proved useful and easy to use. All participants had previous experience with the app, and it was their preferred method of communication with the teachers as compared with email. Sending reminders via WhatsApp ensured faster responses or follow-ups than expected from the participants; thus, we increased the use of the app for smoother and faster communication during the course of the program. WhatsApp was used for informal communication among the participants, contributing to networking among professionals in different areas and countries and was available for future communication after the end of the program.

In the open learning initiative platform, data were collected when students answered questions, that is, click-data, and these are used in a machine learning model to predict students' mastery of the learning objectives [21,22]. Further, these mastery prediction data were used when the teachers prepared the video seminars at the end of each module to assist the teachers; therefore, the focus was on the learning objectives that most students had difficulty understanding. Moreover, 96% (24/25) of the participants indicated that formative questions were helpful in the learning process, and the corresponding number of module test questions was 88% (22/25). This result highlighted the pedagogical method of question-based learning as a successful approach in this context. The click-data were also used after the program to identify which parts of the program needed improvement when preparing for the second edition of OneLearns. In other words, data-driven methodology can be used to improve learning, learning materials, and teaching. For example, for multiple-choice questions, answering alternatives that were not chosen at all or had been answered correctly 100% of the time were interpreted as being too easy or not addressing any sort of misconception, thus not allowing the participants to learn anything new and were therefore changed and improved in the second edition of OneLearns.

Comparison With Prior Work

Open learning initiative courses using question-based learning methodology are being used at more than a hundred colleges and universities in the United States, but their spread outside North America is limited [36]. We know from previous research that this question-based learning methodology can be very efficient for learners, with the possibility of reducing learning time by 50% for a well-iterated course [22] and 25% for a newly developed course [21], as in our case. Thus, the methodology offered a time-saving alternative for the professional

development of experts working full-time in countries with limited resources, as in Rwanda and Ethiopia. However, this depends on internet access, which requires infrastructure. We hope that by letting government officials experience first-hand what state-of-the-art digital learning can offer, they in turn will influence the progression of internet infrastructure in a wise way in their countries.

Limitations

A potential limitation of this program was selection bias in the recruitment process because participants need to be able to communicate in English, although this is not the official language in Ethiopia and has just recently become one in Rwanda [37]. Although English is widely used in academic and professional settings in both countries, potential key players in decision-making and policy making within the government and NGOs could potentially have been overlooked. Another potential limitation is that we did not assess the participants' knowledge in the form of formal tests at the beginning and end and therefore, were not able to assess changes in knowledge over time in test results. Now, we rely on their self-reported assessment of their knowledge; thus, future programs should consider formal testing. Another way to assess the capacity for driving change in future programs is to plan for follow-up questionnaires or interviews after the program ended to document the initiatives that the participants had started. Moreover, there was a range of measures of digital health literacy to be used at the research, clinical, organizational, and societal level [38-41] that were not covered in this program. These tools include questions about skills for searching, finding, evaluating, determining relevance, and applying electronic health information to health problems, as well as about engaging

with and navigating the health care system. In the next generation of OneLearns, the participants will learn these methods and how to use them. In addition, future pedagogical research may include experiments on the effect of sending reminders and program process reports.

The next step for participants is to implement what they learned and build capacity for change in their organizations. If done on a large scale, it will take time, but may include, for example, increased use of digital learning in schools, universities, and other educational organizations. Digital health literacy may lead to increased ability of the general population and professionals to seek information on self-care, access to care, increased use of internet-based medical consultations to reduce emergency department visits, and electronic receipts of drugs [37]. In addition, a broad knowledge of child nutrition will allow for future programs with intersectoral collaborations to address the multifactorial nature of malnutrition. Finally, developing the infrastructure for digitalization is of high priority for both Rwanda and Ethiopia, but it is costly and requires trained human resources. Rwanda has a national information and communication infrastructure development plan aimed at developing and facilitating the establishment and growth of its ICT sector [42], and Ethiopia has a national ICT sector policy and strategy [43].

Conclusions

Digital learning based on question-based learning methodology is a useful method to increase knowledge and the capacity to drive change related to sustainable development among government officials and decision-makers at NGOs in Rwanda and Ethiopia.

Acknowledgments

The OneLearns (Online Education for Leaders in Nutrition and Sustainability) program was funded by a grant from the Swedish Institute a government agency responsible for foreign aid with a focus on culture, education, science, and business to strengthen international relations and development. The program was a part of the public sector innovation program, aiming to strengthen the capacity of government professionals and contribute to innovation in the public sector abroad. Much emphasis is placed on supporting the implementation of the 2030 agenda for sustainable development worldwide.

The authors are grateful to Benti Geleta Buli at Mälardalen University for the important input in the planning of the program, Nisha Singh at Mälardalen University for new statistical analyses during the revision process, Susan Stetson Clarke for careful language editing, and the embassies of Rwanda and Ethiopia in Stockholm for their valuable collaboration.

Authors' Contributions

All authors contributed to this work. OB, KB, and RM designed the program; OB, FJA, CM, RM, and CPO created the learning material; all authors participated in the teaching process; KB collected quantitative data; KB and FJA analyzed the data; FJA, CM, and KB drafted the manuscript; and all authors contributed to the revision of the manuscript and have read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A layout of web-based program.

[[PDF File \(Adobe PDF File\), 115 KB - formative_v6i12e36585_app1.pdf](#)]

Multimedia Appendix 2

Sticky note function in the Miro program.

[[PDF File \(Adobe PDF File\), 247 KB - formative_v6i12e36585_app2.pdf](#)]

Multimedia Appendix 3

The preassessment survey questionnaire.

[[PDF File \(Adobe PDF File\), 129 KB - formative_v6i12e36585_app3.pdf](#)]

Multimedia Appendix 4

The postassessment survey questionnaire.

[[PDF File \(Adobe PDF File\), 103 KB - formative_v6i12e36585_app4.pdf](#)]

References

1. Sibanda LM, Mwamakamba SN. Policy considerations for African food systems: towards the United Nations 2021 Food Systems Summit. *Sustainability* 2021 Aug 12;13(16):9018. [doi: [10.3390/su13169018](#)]
2. Tesema GA, Yeshaw Y, Worku MG, Tessema ZT, Teshale AB. Pooled prevalence and associated factors of chronic undernutrition among under-five children in East Africa: a multilevel analysis. *PLoS One* 2021 Mar 25;16(3):e0248637 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0248637](#)] [Medline: [33765094](#)]
3. Lopes C. Financing Africa's development agenda. *Africa Renewal*. 2015 Dec. URL: <https://www.un.org/africarenewal/magazine/december-2015/financing-africa%E2%80%99s-development-agenda> [accessed 2021-11-23]
4. Barteit S, Guzek D, Jahn A, Bärnighausen T, Jorge MM, Neuhann F. Evaluation of e-learning for medical education in low- and middle-income countries: a systematic review. *Comput Educ* 2020 Feb;145:103726 [[FREE Full text](#)] [doi: [10.1016/j.compedu.2019.103726](#)] [Medline: [32565611](#)]
5. Frehywot S, Vovides Y, Talib Z, Mikhaïl N, Ross H, Wohltjen H, et al. E-learning in medical education in resource constrained low- and middle-income countries. *Hum Resour Health* 2013 Feb 04;11:4 [[FREE Full text](#)] [doi: [10.1186/1478-4491-11-4](#)] [Medline: [23379467](#)]
6. Issever Grochová L, Litzman M. The efficiency in meeting measurable sustainable development goals. *Int J Sustainable Dev World Ecol* 2021 Feb 09;28(8):709-719. [doi: [10.1080/13504509.2021.1882606](#)]
7. Transforming our world: the 2030 Agenda for Sustainable Development. United Nations Department of Economic and Social Affairs. URL: <https://sdgs.un.org/2030agenda> [accessed 2021-11-23]
8. Krönke M. Africa's digital divide and the promise of e-learning. *Africa Portal*. 2020 Jun 30. URL: <https://www.africaportal.org/publications/africas-digital-divide-and-promise-e-learning/> [accessed 2021-11-23]
9. Joyne C, Rossignoli S. 21st century skills: evidence of issues in definition, demand and delivery for development contexts. Department for International Development, United Kingdom. 2019 Aug 1. URL: <https://www.gov.uk/research-for-development-outputs/21st-century-skills-evidence-of-issues-in-definition-demand-and-delivery-for-development-contexts> [accessed 2022-11-04]
10. Gakio K. African Tertiary Institutions Connectivity Survey (ATICS) - 2006 Report. International Development Research Centre. 2006. URL: <https://idl-bnc-idrc.dspace.direct.org/handle/10625/50710> [accessed 2022-05-30]
11. Mbunge E, Akinuwasu B, Fashoto SG, Metfula AS, Mashwama P. A critical review of emerging technologies for tackling COVID-19 pandemic. *Hum Behav Emerg Technol* 2021 Jan;3(1):25-39 [[FREE Full text](#)] [doi: [10.1002/hbe2.237](#)] [Medline: [33363278](#)]
12. Sangrà A, Vlachopoulos D, Cabrera N. Building an inclusive definition of e-learning: an approach to the conceptual framework. *Int Rev Res Open Distributed Learn* 2012 Apr 13;13(2):145. [doi: [10.19173/irrodl.v13i2.1161](#)]
13. Belaïd A. Étude empirique de l'impact de la question de la confiance valeur du capital social sur la gouvernance territoriale en Kabylie. *Unegmu tadamsa magazine* 2021 Mar 10;1(1):15-39 [[FREE Full text](#)]
14. Azzopardi-Muscat N, Sørensen K. Towards an equitable digital public health era: promoting equity through a health literacy perspective. *Eur J Public Health* 2019 Oct 01;29(Supplement_3):13-17 [[FREE Full text](#)] [doi: [10.1093/eurpub/ckz166](#)] [Medline: [31738443](#)]
15. Gilster P. *Digital Literacy*. Hoboken, NJ, USA: Wiley; 1997.
16. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun 16;8(2):e9 [[FREE Full text](#)] [doi: [10.2196/jmir.8.2.e9](#)] [Medline: [16867972](#)]
17. Omotosho A, Ayegba P, Emuoyibofarhe J, Meinel C. Current state of ICT in healthcare delivery in developing countries. *Int J Onl Eng* 2019 May 14;15(08):91. [doi: [10.3991/ijoe.v15i08.10294](#)]
18. Machin-Mastromatteo JD. Information and digital literacy initiatives. *Inf Dev* 2021 Jul 19;37(3):329-333. [doi: [10.1177/02666669211031695](#)]
19. Chakravorti B, Chaturvedi R. Research: how technology could promote growth in 6 African countries. *Harvard Business Review*. 2019 Dec 4. URL: <https://hbr.org/2019/12/research-how-technology-could-promote-growth-in-6-african-countries> [accessed 2022-05-30]

20. Yigezu M. Digitalization in teaching and education in Ethiopia. International Labour Organization. 2021 Apr 27. URL: https://www.ilo.org/sector/Resources/publications/WCMS_783664/lang--en/index.htm [accessed 2022-11-04]
21. Bälter O, Glassey R, Wiggberg M. Reduced learning time with maintained learning outcomes. In: Proceedings of the 52nd ACM Technical Symposium on Computer Science Education. 2021 Presented at: SIGCSE '21; March 13-20, 2021; Virtual. [doi: [10.1145/3408877.3432382](https://doi.org/10.1145/3408877.3432382)]
22. Lovett M, Meyer O, Thille C. The open learning initiative: measuring the effectiveness of the OLI statistics course in accelerating student learning. *J Interact Media Educ* 2008 May 20;2008(1):13. [doi: [10.5334/2008-14](https://doi.org/10.5334/2008-14)]
23. Bishop J, Verleger MA. The flipped classroom: a survey of the research. In: Proceedings of the 2013 ASEE Annual Conference & Exposition. 2013 Presented at: ASEE '13; June 23-26, 2013; Atlanta, GA, USA. [doi: [10.18260/1-2--22585](https://doi.org/10.18260/1-2--22585)]
24. Koedinger K, McLaughlin E, Jia J, Bier N. Is the doer effect a causal relationship?: how can we tell and why it's important. In: Proceedings of the 6th International Conference on Learning Analytics & Knowledge. 2016 Presented at: LAK '16; April 25-29, 2016; Edinburgh, UK. [doi: [10.1145/2883851.2883957](https://doi.org/10.1145/2883851.2883957)]
25. Bier N, Lip S, Strader R, Thille C, Zimmaro D. An approach to knowledge component/skill modeling in online courses. Open Learning Initiative (Carnegie Mellon University, Stanford University) and Google Inc. 2014 Apr 30. URL: <https://research.google/pubs/pub48380/> [accessed 2022-11-04]
26. Anderson L, Krathwohl D, Bloom B. A taxonomy for learning, teaching, and assessing: a revision of bloom's taxonomy of educational objectives. Semantic Scholar. 2000 Dec 29. URL: <https://www.semanticscholar.org/paper/A-Taxonomy-for-Learning%2C-Teaching%2C-and-Assessing%3A-A-Anderson-Krathwohl/23eb5e20e7985fca5625548d2ee6d781a2861d41> [accessed 2021-11-23]
27. Glassey R, Balter O. Put the students to work: generating questions with constructive feedback. In: Proceedings of the 2020 IEEE Frontiers in Education Conference. 2020 Presented at: FIE '20; October 21-24, 2020; Uppsala, Sweden. [doi: [10.1109/fie44824.2020.9274110](https://doi.org/10.1109/fie44824.2020.9274110)]
28. Visual Collaboration Platform for Teamwork. Miro. URL: <https://miro.com/> [accessed 2021-07-22]
29. Vad säger lagen. Etikprövningsmyndigheten. URL: <https://etikprovningmyndigheten.se/for-forskare/vad-sager-lagen/> [accessed 2022-06-30]
30. Obiyan AS. A critical examination of the state versus non-governmental organizations (NGOs) in the policy sphere in the global south: will the state die as the NGOs thrive in Sub-Saharan Africa and Asia? *Afr Asian Stud* 2005;4(3):301-326. [doi: [10.1163/156920905774270475](https://doi.org/10.1163/156920905774270475)]
31. Noesgaard SS, Ørngreen R. The effectiveness of e-learning: an explorative and integrative review of the definitions, methodologies and factors that promote e-Learning effectiveness. *Electron J e-Learn* 2015 Apr 1;13(4):278-290.
32. Heller RF. Experience with a "social model" of capacity building: the Peoples-uni. *Hum Resour Health* 2009 May 29;7(1):43 [FREE Full text] [doi: [10.1186/1478-4491-7-43](https://doi.org/10.1186/1478-4491-7-43)] [Medline: [19476652](https://pubmed.ncbi.nlm.nih.gov/19476652/)]
33. Tjulin Å, MacEachen E, Vinberg S, Selander J, Bigelow P, Larsson R. Virtual internationalization – we did it our way. *Högre Utbildning* 2021;11(2):1-13. [doi: [10.23865/hu.v11.2344](https://doi.org/10.23865/hu.v11.2344)]
34. Aydin S, Öztürk A, Büyükköse GT, Er F, Sönmez H. An investigation of drop-out in open and distance education. *Educ Sci Theory Pract* 2019;19(2):40-57. [doi: [10.12738/estp.2019.2.003](https://doi.org/10.12738/estp.2019.2.003)]
35. Heller RF, Strobl J, Madhok R. Online education for public health capacity building in low- to middle-income countries. *Int Rev Res Open Distributed Learn* 2019 Feb 28;20(1):79-93. [doi: [10.19173/irrodl.v20i1.3927](https://doi.org/10.19173/irrodl.v20i1.3927)]
36. Glassey R, Bälter O. Sustainable approaches for accelerated learning. *Sustainability* 2021 Oct 29;13(21):11994. [doi: [10.3390/su132111994](https://doi.org/10.3390/su132111994)]
37. Cecilio-Fernandes D, Parisi MC, Santos TM, Sandars J. The COVID-19 pandemic and the challenge of using technology for medical education in low and middle income countries. *MedEdPublish* 2020 Apr 23;9:74. [doi: [10.15694/mep.2020.000074.1](https://doi.org/10.15694/mep.2020.000074.1)]
38. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006 Nov 14;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
39. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of health 1.0 and health 2.0 skills. *J Med Internet Res* 2017 Jan 24;19(1):e27 [FREE Full text] [doi: [10.2196/jmir.6709](https://doi.org/10.2196/jmir.6709)] [Medline: [28119275](https://pubmed.ncbi.nlm.nih.gov/28119275/)]
40. Epstein HB. The health literacy questionnaire. *J Consum Health Internet* 2021 Sep 09;25(3):303-309. [doi: [10.1080/15398285.2021.1954399](https://doi.org/10.1080/15398285.2021.1954399)]
41. Yoon J, Lee M, Ahn JS, Oh D, Shin S, Chang YJ, et al. Development and validation of digital health technology literacy assessment questionnaire. *J Med Syst* 2022 Jan 24;46(2):13 [FREE Full text] [doi: [10.1007/s10916-022-01800-8](https://doi.org/10.1007/s10916-022-01800-8)] [Medline: [35072816](https://pubmed.ncbi.nlm.nih.gov/35072816/)]
42. National ICT Strategy and Plan NICI. Government of Rwanda. 2015. URL: https://www.itu.int/en/ITU-D/Cybersecurity/Documents/National_Strategies_Repository/Rwanda%20NCSS%20NICI_III.pdf [accessed 2022-05-30]
43. Ababa A. The National Information and Communication Technology (ICT) Policy and Strategy. Government of Ethiopia. 2016 Sep. URL: <https://comesabusinesscouncil.org/wp-content/uploads/2020/04/6-ICT-Policy-and-Strategy.pdf> [accessed 2022-05-24]

Abbreviations

ICT: information communication technology

NGO: nongovernmental organization

OneLearns: Online Education for Leaders in Nutrition and Sustainability

SDG: sustainable development goal

Edited by A Mavragani; submitted 20.01.22; peer-reviewed by S Andersson, T Zan; comments to author 10.04.22; revised version received 31.05.22; accepted 01.07.22; published 05.12.22.

Please cite as:

Bälter K, Javan Abraham F, Mutimukwe C, Mugisha R, Persson Osowski C, Bälter O

A Web-Based Program About Sustainable Development Goals Focusing on Digital Learning, Digital Health Literacy, and Nutrition for Professional Development in Ethiopia and Rwanda: Development of a Pedagogical Method

JMIR Form Res 2022;6(12):e36585

URL: <https://formative.jmir.org/2022/12/e36585>

doi: [10.2196/36585](https://doi.org/10.2196/36585)

PMID: [36469416](https://pubmed.ncbi.nlm.nih.gov/36469416/)

©Katarina Bälter, Feben Javan Abraham, Chantal Mutimukwe, Reuben Mugisha, Christine Persson Osowski, Olle Bälter. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Perceptions and Attitudes Toward an Interactive Voice Response Tool (Call for Life Uganda) Providing Adherence Support and Health Information to HIV-Positive Ugandans: Qualitative Study

Phoebe Kajubi¹, BEd, MEd, PhD; Rosalind Parkes-Ratanshi^{1,2}, MBBS, MA, PhD; Adelline Twimukye¹, BASS, MAPAM; Agnes Bwanika Naggirinya^{1,3}, MSc, MMed, MD; Maria Sarah Nabaggala¹, BSTAT, BCOM, MSc; Agnes Kiragga¹, BSTAT, BCOM, MSc, PhD; Barbara Castelnuovo¹, MMed, MD, PhD; Rachel King^{1,4}, DiP, BA, MPH, PhD.

¹Infectious Diseases Institute, College of Health Sciences, Makerere University, Kampala, Uganda

²Department of Public Health & Primary Care, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom

³Department of Internal Medicine, College of Health Sciences, School of Medicine, Makerere University, Kampala, Uganda

⁴Institute for Global Health Sciences, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:

Phoebe Kajubi, BEd, MEd, PhD
Infectious Diseases Institute
College of Health Sciences
Makerere University
PO Box 22418
Kampala
Uganda
Phone: 256 312 307000
Email: phoebekajubi@yahoo.com

Abstract

Background: The continuing decline in AIDS-related deaths in the African region is largely driven by the steady scale-up of antiretroviral therapy. However, there are challenges to retaining people living with HIV on treatment. Call for Life Uganda (CFLU) is an interactive voice response tool using simple analogue phones. CFLU supports patients with daily pill reminders, preappointment reminders, symptom reporting and management, and weekly health promotion tips. Mobile health tools are being increasingly used in resource-limited settings but are often adopted without rigorous evaluation.

Objective: This qualitative study conducted at 12 months after enrollment assessed patients' experiences, perceptions, and attitudes regarding CLFU.

Methods: We conducted a qualitative substudy within an open-label randomized controlled trial titled "Improving outcomes in HIV patients using mobile phone based interactive software support." Data were collected through 6 focus group discussions with participants sampled based on proportion of calls responded to—<25%, between 25% and 50%, and >50%—conducted at the Infectious Diseases Institute, Mulago, and the Kasangati Health Centre IV. NVivo (version 11; QSR International) was used in the management of the data and in the coding of the emerging themes. The data were then analyzed using content thematic analysis.

Results: There was consensus across all groups that they had more positive than negative experiences with the CFLU system. Participants who responded to >50% of the calls reported more frequent use of the specific elements of the CFLU tool and, consequently, experienced more benefits from the system than those who responded to calls less frequently. Irrespective of calls responded to, participants identified pill reminders as the most important aspect in improved quality of life, followed by health promotion tips. The most common challenge faced was difficulty with the secret personal identification number.

Conclusions: Findings showed participants' appreciation, high willingness, and interest in the intervention, CFLU, that demonstrated great perceived potential to improve their access to health care; adherence to treatment; health awareness; and, consequently, quality of life.

Trial Registration: ClinicalTrials.gov NCT02953080; <https://clinicaltrials.gov/ct2/show/NCT02953080>

KEYWORDS

mobile health; mHealth; mobile communication technologies; people living with HIV; antiretroviral therapy; quality of life; Uganda

Introduction

Background

The decline in AIDS-related deaths in the African region is largely driven by the steady scale-up of antiretroviral therapy (ART) [1]. Of the 16.3 million (64%) people living with HIV accessing treatment in 2018, a total of 52% had a suppressed viral load [1]. Similarly, in Uganda, a total of 1.16 million of the estimated 1.38 million people living with HIV were on ART by December 2018, giving a coverage of >90% on ART [2]. In addition, evidence shows that there has been remarkable progress toward ensuring that people initiated on ART have their viral load suppressed [2,3]. For example, Uganda reached a viral load suppression rate of 90.4% in 2017 to 2018 [4].

Despite these notable achievements, retention on treatment of people living with HIV is an increasing challenge [5], coupled with challenges of maintaining adherence [5-8]. In 2010, the average rate of early mortality and loss to follow-up in resource-limited countries was estimated at 23% across the sub-Saharan region [9]. Finitsis et al [3] conducted a meta-analysis of 84 observational studies where nearly 40% of participants had <90% adherence.

Amid these challenges, adherence to ART is reportedly influenced by certain factors that differ by region of the world, which include socioeconomic and disease- and ART-related factors as well as factors related to people living with HIV and their families [10,11]. A study on adherence to ART and its determinants conducted among patients infected with HIV in Nigeria revealed that forgetfulness, busy daily tasks, occurrence of adverse effects, and *too many pills to take* constituted the major reasons for missing doses [10]. Consequently, these factors cause poor adherence to therapy, resulting in treatment failure and the development of viral resistance to antiretroviral medications [6,11,12].

In light of this, it is important to develop additional innovative, practical, targeted, and feasible interventions to improve retention and increase and maintain levels of adherence among patients with HIV on ART if treatment failure and resistance is to be avoided to achieve maximal viral load suppression [8,11,12]. Odili et al [10] suggest that interventions to improve adherence to ART should address challenges such as forgetfulness among the patients and frequent occurrence of adverse effects and consider specific patient-related factors such as daily tasks.

In response to this urgent need, the potential of mobile health (mHealth) communication technologies in closing the gaps in the HIV treatment continuum and their use have grown significantly over the years [13-15]. The World Health Organization recommends the use of mobile phone-based technologies for management of chronic diseases and ART adherence [16,17]. The benefits of mHealth technologies in

health care have been reported worldwide and recommended as an opportunity to increase the quality and cost-effectiveness of health care, particularly in resource-constrained settings amid the growing number of patients [18-21]. Studies have shown that mobile phones are used throughout lower-income countries more than any other modern technology; have the potential to revolutionize health care, particularly in low-resource settings where health care infrastructure and services are often insufficient; and offer great promise for improving the quality of life [6,15,21-25]. Important benefits of using mHealth technologies highlighted in systematic reviews and other studies include adherence to treatment (for people living with HIV and other categories of patients) being the most significant [6,14,22,26-34], high effectiveness for the dissemination of health promotion messages and lifestyle tips [19,20,22,25,35,36], reminders for physician appointments or improvement in clinic attendance [6,23,25,28,33,37,38], remote diagnosis [39,40], emergency medical response [22,41,42], improvement in communication and information delivery and retrieval processes over vast distances between health care service providers and patients [22], privacy and convenience allowing the user to be in charge of the process [13], and improvement in viral load suppression [28]. However, a major concern raised by some studies regarding SMS text messaging-based interventions is the variability in the magnitude of study outcomes [43,44]. Chib et al [15] observed that the mHealth literature in low- and middle-income countries (LMICs) reveals a growing body of knowledge, yet existing reviews suggest that projects yield mixed results.

Objectives

We undertook a randomized controlled trial (RCT) in Uganda (trial registration: ClinicalTrials.gov NCT02953080) using an intervention entitled Call for Life Uganda (CFLU), an mHealth tool or software that is based on the open-source Mobile Technology for Community Health (Grameen Foundation), described elsewhere in a separate publication [45]. The RCT hypothesis was that CFLU would increase medication adherence, virological outcomes, and HIV knowledge to give an overall increased quality of life in vulnerable populations starting or established on ART or switching to second-line ART, including special populations such as pregnant women, discordant couples, and young people in Uganda.

This paper describes a qualitative substudy conducted at 12 months after enrollment aimed at assessing *similarities* and *differences* in the experiences, perceptions, and attitudes of people living with HIV regarding the CFLU tool by response rate category.

Methods

RCT Study Design

This qualitative study was part of an open-label RCT titled “Improving outcomes in HIV patients using mobile phone based interactive software support” (NCT02953080). The intervention was Call for Life, a software adapted from Connect for Life, and was first piloted in India and Ghana [46,47]. It was developed by the Grameen Foundation and Janssen, the pharmaceutical companies of Johnson and Johnson [47]. CFLU was adapted to support people living with HIV in Uganda in a variety of ways. It interacts with people living with HIV using SMS text messaging or interactive voice response (IVR) functionalities (a technology that allows a computer to interact with humans through the use of voice and tones input via keypad). Participants in the RCT received either the usual standard of care plus the CFLU mHealth tool (intervention arm) or standard of care with clinic support and provider-initiated counseling (control arm). Intervention participants received the usual standard of care plus daily adherence IVR reminders (or SMS text messaging) delivered just before the usual pill-taking time. The system sends a phone call to the participant and, when they answer it, plays music until they enter a personal identification number (PIN) to access further services. They also received preappointment reminders; had access to educational health tips once weekly to increase knowledge of HIV and comorbid conditions; and had an option to call a toll-free line and report symptoms or medication side effects, which would prompt health care workers to call back within 24 hours. The primary objective of the RCT was to determine the effect of the CFLU tool on the quality of life of people living with HIV in urban and semiurban Ugandan health facilities at 6, 12, and 24 months or at closeout of the project [45]. The study was conducted at 2 sites. The first was the Infectious Diseases Institute (IDI), Mulago, a specialist private not-for-profit urban HIV clinic established within Makerere University, located near the National Mulago Hospital Complex serving >8000 people living with HIV. The second study site was Kasangati Health Centre IV (KSG HCIV), a semiurban government-owned clinic that provides comprehensive HIV care and treatment and serves approximately 5000 patients actively undergoing HIV care. It is located in the Wakiso District, which comprises both rural and urban areas and has a population of >2 million people. Eligible patients were randomized into either the control (standard of care) or intervention arm (1:1 ratio) in the RCT. At baseline, participants in the intervention arm were trained on how to initiate and receive calls, and the messages were provided in four languages: English, Luganda, Kiswahili, and Runyankore. A description of how the tool was introduced to the participants is detailed in 2 separate study publications that include the RCT study and the qualitative exploratory study findings [45,48]. At baseline, all patients underwent a quality-of-life assessment, which was repeated at months 6, 12, and 24.

Qualitative Substudy

Patients enrolled in the intervention arm of the CFLU RCT were purposively sampled to participate in the qualitative substudy. Two lists (one for KSG HCIV and another for IDI Mulago)

were generated, categorizing the proportion of calls answered by the patients (ie, patients who had a proportion of <25% of calls answered; those who had a proportion of 25%-50% of calls answered; and, finally, those who had a proportion of >50% of calls answered). Patients from these 3 categories based on the proportion of all calls answered were contacted by telephone. The purpose and study procedures were explained to them in a call, and they were recruited based on their availability to participate in the study.

We used a descriptive qualitative design following a phenomenological research approach to explore lived experiences and perspectives of participants who used the CFLU tool to gain a deeper insight to address the research questions. Specifically, the qualitative study conducted at 12 months after enrollment assessed *similarities* and *differences* in the experiences, perceptions, and attitudes of people living with HIV regarding the CFLU tool by response rate category. Participants were asked to describe their experiences with the CFLU tool, specific elements of the tool that they had used (daily pill reminders, health tips, appointment reminders, symptom reporting, and adherence checks), what they particularly liked about the CFLU tool and why they liked those specific elements, frequency of use, the secret PIN code, what they disliked about the tool, comfortability of use, lessons learned, suggestions on how the tool could help other patients, how the tool could be improved, and patients' willingness to pay for the CFLU system.

Data Collection

The qualitative study used focus group discussions (FGDs) as the method of data collection to stimulate a rich discussion. The FGDs were conducted by 2 social scientists, which included a graduate counselor (AT) and a senior social scientist (PK) experienced in qualitative methods of data collection and conversant in Luganda, a language widely spoken at the 2 study sites. They worked as moderator and note-taker for every FGD. The FGDs were conducted following a topic guide. Each FGD had between 5 and 14 participants, lasted between 1 and 1 and a half hours, and was audio recorded with participant permission. The FGDs were scheduled at a time convenient to participants and were conducted at the 2 study sites in venues that allowed for privacy. Confidentiality was maintained by use of a coding system during the FGDs. Participants were told of this anonymity and the measures to be taken. Data were collected using handwritten notes plus audio recordings to capture details of the FGD responses.

Data Management and Analysis

The recorded FGDs were transcribed and translated from Luganda into English verbatim. The two social scientists (PK and AT) cross-checked and verified the transcripts for consistency and accuracy and carried out preliminary thematic coding through multiple readings of the transcripts. The next step involved PK and AT understanding the data and developing a list of thematic categories for constructing a codebook. The research objectives, FGD topic guide, and transcripts guided the generation of a codebook. There were 7 major codes, namely, positive experiences with the CFLU tool, negative experiences with the CFLU tool, specific elements of the tool

that patients used, frequency of calls, comfort of use, lessons learned, and suggestions for improvement. The data set was imported into NVivo (version 11; QSR International) [49] for coding, managing, and retrieving. Once all the transcripts were brought together, systematic coding was undertaken. Codes were assigned to relevant segments of the text; similar codes were aggregated to form themes that were then used to address the research questions and develop coherent narratives [50]. This involved the 2 social scientists, who independently read through the relevant section of a transcript within the NVivo project and pooled together the relevant segments into a node (which is the NVivo equivalent of a theme). Independent codes were compared and discussed to ensure rigor and trustworthiness [51]. Coding helped identify all segments of data that related to the particular node listed in the codebook. In some cases, there were more complex relationships where data were coded into more than one node.

Ethical Considerations

Ethical clearance was obtained from Makerere University School of Public Health Higher Degrees Research and Ethics Committee (378) and the Uganda National Council for Science and Technology (HS 3005), and the study was registered at ClinicalTrials.gov (NCT02953080). Written informed consent was obtained from the participants for involvement in the main RCT; participation in the study was voluntary based on informed

consent. The substudy team were trained to administer informed consent in the language best understood by the participants. Objectives of the study and procedures to be followed during the FGDs were explained to all participants. Written consent for participation in the qualitative study was sought after the participants received the study information, objectives, and procedures. The research assistants (PK and AT) read the consent form out loud, and all participants provided written informed consent. Results were disseminated to participants and stakeholders, which served as a member-check activity.

Results

Demographics and Sample Characteristics

A total of 300 participants were enrolled in the intervention arm of the parent RCT, and 256 (85.3%) of them completed the 12-month follow-up [45]. Findings from the parent RCT showed that at 12 months, 9.8% (25/256) of the active participants were low users (0%-24%), 41.8% (107/256) of the participants were moderate users (25%-50%), and 48.4% (124/256) of the participants were high users (>50%).

For the qualitative study, a total of 52 participants (n=19, 37% male and n=33, 63% female) took part in 6 FGDs: 2 (33%) from the KSG HCIV study site and 4 (67%) from the IDI Mulago study site (Table 1). The mean age was 43 years and the overall age range was 22 to 64 years.

Table 1. Characteristics of the focus groups.

FGD ^a category	Study site	Participants, n (%)		Age range (years)
		Male participants	Female participants	
FGD 1—responded to <25% of calls (n=8)	KSG HCIV ^b	0 (0)	8 (100)	23-32
FGD 2—responded to <25% of calls (n=5)	Infectious Diseases Institute, Mulago	2 (40)	3 (60)	32-38
FGD 3—responded to 25% to 50% of calls (n=14)	KSG HCIV	4 (29)	10 (71)	22-47
FGD 4—responded to 25% to 50% of calls (n=9)	Infectious Diseases Institute, Mulago	4 (44)	5 (56)	25-64
FGD 5—responded to >50% of calls (n=9)	Infectious Diseases Institute, Mulago	9 (100)	0 (0)	34-50
FGD 6—responded to >50% of calls (n=7)	Infectious Diseases Institute, Mulago	0 (0)	7 (100)	29-56

^aFGD: focus group discussion.

^bKSG HCIV: Kasangati Health Centre IV.

Participants' Use and Experiences With the CFLU Tool

Overview

Participants were asked what they knew about the CFLU tool and to describe the specific elements of the tool that they had used, their experiences with the tool, and what they particularly liked and disliked about the CFLU system and why they liked those specific elements. Irrespective of proportion of calls responded to, there were more similarities than differences among the 3 categories (ie, those who had a proportion of <25%, 25%-50%, and >50% of calls answered) regarding what participants knew about CFLU, their use, their experiences with the tool, and what they particularly liked about it. Findings

revealed that *Pill reminders* were reportedly the most popular and were perceived to be most beneficial among the specific elements of the CFLU tool, followed by health tips, symptom reporting, and appointment reminders.

Pill Reminders

Overview

Throughout the responses, it was evident that participants across all FGDs largely attributed the positive experiences with the CFLU system first to pill reminders that assisted them in taking their medicines on time, resulting in improved adherence and better health and quality life compared with the health challenges before their involvement with the CFLU system:

Before enrolment, I was very sick; smelling bark cloth (cloth made out of the bark of a tree used to wrap dead bodies ready for burial) i.e. I was so close to death and I was about to be buried in my home village. Then they put me on CFLU. But ever since I joined the system, they remind me to take my medication, educate me about healthy living. From that time, my health has been okay. The virus is even asleep. I feel much better than before so please continue with it. [KSG HCIV, responded to 25%-50% of calls, FGD 3]

Across the different categories, there was concurrence in participants' description of their *medicine-taking practices* before and after enrollment in the CFLU system. Most participants from the 3 categories reported that, before their involvement with the CFLU system, they used to forget to take their medicines on time because of busy work schedules, reluctance to take the pills, pill burden, absence of a person to remind them, and lack of proper food to take with the medicines. However, with the pill reminders, participants reportedly took their medicines on time even amid busy schedules and other challenges:

Before, I would forget, but ever since I started getting pill reminders, I don't forget however busy I may be; I still remember the time for taking. I work in a hotel so sometimes I may be so busy, but when I see the call, I remember that it is time for taking pills. I had poor adherence and it used to affect me; but now it doesn't. [KSG HCIV, responded to <25% of calls, FGD 1]

Findings further revealed that even participants who responded to <25% of calls used and benefitted from pill reminders. Low responders often had challenges with their PINs or secret PIN codes, which prevented them from having access to other CFLU components. Consequently, most of them felt that CFLU system was just about pill reminders. They explained that, whenever the phones rang, they were unable to access other services but took the call only as a reminder to take their pills:

Yes, reminding me of the time to take my pills. Because there's nothing much to it; as soon as the phone rings, I know straight away that it is a pill reminder...that very number is for pill reminders. They told us to put in the PIN and talk, but my PIN refused to work, in fact I have never used it. [KSG HCIV, responded <25% of calls, FGD 1]

Benefits Associated With Pill Reminders

In addition to improved adherence, experiences of improved self-esteem and hope to live positively were topics within all FGDs attributed to the CFLU phone calls that enhanced clients' psychosocial support. Participants passionately used terms such as "*parent, counselor, friend*" to refer to CFLU indicating what the system meant to them. Across all FGDs, participants explained that the answering machine spoke very calmly, kindly, and politely to them, making them feel loved, cared for, and accepted:

They don't ignore you; they show you love and care. For instance, some of us who are HIV positive, we reach a point and cut off relations with people, we get fed up of people. But now, that gentleman who speaks (the voice), he speaks so gently and calmly to me; there's a way he speaks to you, he puts you in the mood, he makes you feel loved...they have not abandoned us. They have helped us a lot. [KSG HCIV, responded 25%-50% of calls, FGD 3]

Health Tips and Associated Benefits

Overview

In addition to pill reminders, health tips were particularly appreciated and found very useful in raising awareness about health-related information and were linked to improvement in health and quality of life and promoting a healthy lifestyle. Again, this was a topic across all FGDS, especially among those who responded to >50% of calls, followed by those in the 25% to 50% category and, finally, the few among the <25% category who did not have challenges with their secret PIN codes and could access the option for health tips on their phones. Participants mentioned that the educative health tips provided a range of very useful information that they had not been exposed to during previous counseling and educational sessions at health facilities and other forums, which included information on chronic illnesses (such as breast and cervical cancer and diabetes), encouraging them to attend medical checkups for early detection and treatment, and on other diseases, including tuberculosis and sexually transmitted infections such as candidiasis. Other health tips mentioned included information on behavior change such as abstinence, faithfulness to one sexual partner and condom use, positive living, nutrition, family planning, pregnancy, and breastfeeding. Improved self-esteem and boldness to teach others were also attributed to the educative health-related information from the health tips. This was mostly recounted by participants who responded to calls >50% of the time as they listened more to the health tips than the other 2 categories of participants. The former purportedly disclosed their status readily to family members and nonrelatives and practiced positive living:

It has educated us, and has also made us educate others. Because in our villages people still point fingers that so and so is HIV+. But it has helped us be doctors to others; you explain to someone that if they have these signs they need to see a doctor because it may be such and such a disease. In other words, they keep us informed about various diseases that attack people. [IDI Mulago, responded to >50% of calls, FGD 5]

They boldly shared the learned health tips with their families and community members; encouraged members of the community to attend HIV counseling and testing and start medication if found HIV-positive; and, for those found HIV-negative, they encouraged them to abstain, be faithful to their partners if married, or practice safe sex using condoms:

What I know about it is, it teaches us about health, diseases which are dangerous and all other ordinary

diseases. And they continue to educate you even on what you don't know because they reach a time and ask you, "do you still want to listen to more health tips?" The benefits are many, I cannot exhaust them all. That is what I like most about it. [IDI Mulago, responded to >50% of calls, FGD 6]

Most participants, especially those who responded to calls >50% of the time, requested to be called on weekends particularly to listen to the health tips, some with their family members because this was when they had quality time. Participants pointed out that the CFLU system also gave them the option to listen to the same health tips several times:

This time when I listen to health tips about HIV/AIDS, the next time I listen to cancer health tips, the following time I listen to health tips concerning sexual behaviour. I keep changing because there may be health tips about another disease that they may have brought. I press different health tips one at a time so that I can memorize. [IDI Mulago, responded to >50% of calls, FGD 6]

Positive and Healthy Living

Participants attributed positive (healthy) living to messages from the health tips, which encouraged them to discard fear about being HIV-positive and imminent death, accept their status, and live positively. Consequently, this helped them have hope in life and lead productive lives:

Another thing, it has helped me at least to come out boldly, and I can confront the fear which I had in me that I am going to die. I know that yes I have HIV but I won't die, I'll live longer. The more I take my medication, the more I will live longer by being healthy. Positive living in everything I do. [IDI Mulago, responded to >50% of calls, FGD 6]

Nutritional advice encouraged consuming more fruits and vegetables, drinking a lot of water, eating a balanced diet, and having meals on time.

Behavioral Change

Participants described aspects of behavior change that they were practicing from the knowledge acquired from the health tips, as illustrated as follows:

What I have put in practice, is condom use. I and my husband are HIV-positive but I don't want his virus (Laughs) and I would not want my virus to mix with his. Yes, we have put that in practice. In addition to that, sleeping under treated mosquito nets, I boil my water, I don't mix local herbs with the medicine. Even these condoms, sometimes we sit in youth groups and I educate them that using a condom does not mean you are HIV+, it means you are simply protecting yourselves. [IDI Mulago, responded to >50% of calls, FGD 6]

Symptom Reporting and Associated Benefits

Patients were given a toll-free number to report any danger signs, request information, and receive medical advice and emergency assistance. In addition, health workers would call

them, inquire about any health-related problem they had, and advise accordingly. Again, this was a topic across all FGDs; participants appreciated the "symptom reporting" component.

Many participants passionately stated that the symptom reporting component demonstrated to them that the *basawo* (health workers) really cared for them and went the extra mile of calling them to find out if they had any health issues. They appreciated the fact that they had someone to talk to anytime they had health problems and that, most of the time, the *basawo* endeavored to call them or follow up with them advising them on what medicines to take or where to buy the medicines or instructed them to go to the health facilities for treatment even when it was not their appointment day. They found this beneficial as sometimes they did not have money for transport to hospitals; other times, such health issues happened when it was not their appointment day. They had trusted and informed people to confide in about their problems and, in most cases, they were helped:

Another thing about CFLU, it has helped me so much in a way that you are like our parents. You call us and ask, "do you have any illness?"...you take the responsibility to call us and give us advice on what to do, or how to get help...I am so happy about it and my health, because if it was not for CFLU, my health would have been really bad. [IDI Mulago, responded to >50% of calls, FGD 6]

Consequently, this improved the relationship between the participants and health workers, resulting in increased trust and confidence of the participants in them, which encouraged them to call any time and report health-related problems:

It makes you feel proud that someone cares for you. We used to think you "basawo" (health workers-doctors) do not care for us because people always complain about that. But here we are more than certain that we are cared for and that makes us happy. [IDI Mulago, responded to >50% of calls, FGD 6]

By contrast, a few participants, particularly among those in the <25% category, reported challenges that hindered full use of the symptom-reporting component. Some reportedly were not aware of the toll-free number, and some did not know how to use it, as stated as follows:

There is a toll free number for CFLU, I also tried to call it but they said it does not exist on the network. And then I thought to myself, did the doctor give me a wrong number? [IDI Mulago, responded to <25% of calls, FGD 2]

Across all categories, many participants complained about the long waiting hours to receive a response from the physician after reporting a health problem.

Appointment Reminders and Associated Benefits

Among the CFLU system elements, this was the least used. Most participants did not receive appointment reminders. Of the few who received this service, most reported having received appointment reminders once or twice, and this was mostly

among those who responded to >50% of calls, followed by those in the 25% to 50% category and, finally, the few among the <25% category who did not have challenges with their secret PIN codes. Among participants who received appointment reminders, this component was most appreciated by those employed in jobs that involved frequent traveling across long distances and also those in mobile businesses or trade. Participants in this category explained that such jobs kept them so busy that they depended on these reminders to keep their medical appointments:

CFLU has helped me to keep my appointment. They always give me a message that you have an appointment this Tuesday. [IDI Mulago, responded to >50% of calls, FGD 5]

Frequency of Calls and Comfort of Use

At enrollment into the CFLU system, all participants reportedly asked for an appropriate time and day to receive calls. Most reportedly received calls daily, some of them twice daily:

If you take your pills twice a day, in the morning and evening, they call you at those times. But if you take once a day, they'll also call once. [KSG HCIV, responded to 25%-50% of calls, FGD 3]

Furthermore, participants were asked to rate their comfort level with using the CFLU system on a scale of 1 to 5 (5 being most comfortable) and give reasons for the rating. Most rated it 4-5, pointing out that they were comfortable using it and could clearly explain the different options. Other reasons included the training given to use the system, easy-to-follow prompts, expeditious rectification of technical issues, and the confidentiality of the PIN code.

However, participants who had challenges with the secret PIN code could not explain clearly whether they were comfortable using the system.

Negative Experiences With the CFLU Tool

The study further explored participants' challenges regarding using the CFLU tool. The number and proportions of mentions of the challenges are summarized in [Table 2](#).

Table 2. Number and percentages of responses of negative experiences with the Call for Life Uganda (CFLU) tool (N=52).

Themes and subthemes	Responses, n (%)
PIN^a challenges	
Blocked PIN	11 (21)
PIN refusing to work	8 (15)
Incorrect PIN code	7 (13)
Technical issues	
Phone ringing endlessly	9 (17)
Wrong timing of the pill reminders (calling before or after agreed time)	8 (15)
Inconsistent and irregular calls	4 (8)
Calling participants from different CFLU system numbers	4 (8)
Nonresponse of CFLU system to participants' calls	3 (6)
Long waits for physician's response after reporting a symptom	4 (8)
Challenges with toll-free number	3 (6)

^aPIN: personal identification number.

Several challenges were reported by participants, most centered on the secret PIN code. Participants narrated challenges they had with the PIN code, most explaining that, even when they reported and received another number, it could not work:

You have to put in the PIN code because that's the only way they will know whether you took the pills on time or not. I tried it the first time and it refused, I tried it again and they gave me another PIN code which also refused to work. Then they called me and I told the doctor that they refused; then I was told to read the other one, which I did, and I tried it as well but it also didn't work. [KSG HCIV, responded to <25% of calls, FGD 1]

Most participants in this category reportedly resorted to listening only to pill reminders when the attempt with the second PIN code failed. They explained that, whenever the phone rang, they

would just know that it was time to take their pills and would not press for other options. These complaints were largely expressed by participants who responded to <25% of calls:

As for me doctor mine refused from the very beginning; even when I came here they told me the system is still being rectified. So I have never spoken to anyone on that call or listening to anything. Most times when I pick up I only hear the kadodi but most times I don't even pick up the call, I just end it because I already know the essence of the call. [KSG HCIV, responded to <25% of calls, FGD 1]

The most reported challenge in this category were blocked PINs when responding to a phone call:

My view about the PIN is that, my PIN was blocked; I was on 2255 and I thought that maybe they found that 2255 is commonly used by many people and they

blocked some so that they could get other PINs. But I do not think one can forget their pin code; because I did it for a month and they would keep telling me it's not the one, I should try again. Until I called the CFL contacts and they told me to get another PIN because it seems the previous one has some issues. But I think they should do it for us; they should get for us the pin codes that we shall use. That's it. [IDI Mulago, responded to <25% of calls, FGD 2]

These were followed by participants whose PIN codes reportedly refused to work, as explained as follows:

My PIN code also refused to work; the one they gave us. So I came back here and was given a new PIN which also refused to work. However much I keep pressing it says the PIN is incorrect. So it just keeps ringing. [KSG HCIV, responded to <25% of calls, FGD 1]

Similar to the aforementioned issues were challenges reported by participants who were told repeatedly that their PIN codes were wrong or incorrect even when they received new PINs:

About the pin code, there should be a change, I am not sure if it is us who forget our PINs but I am sure they are the ones; but then they tell me it is incorrect. So I do not know if I am the wrong one or is it the callers that are wrong. So I would like to be helped in that area. [IDI Mulago, responded to <25% of calls, FGD 2]

Challenges further included technical issues with the CFLU tool, which comprised the phone ringing endlessly, wrong timing of the pill reminders, inconsistent and irregular pill reminders, calling participants from different CFLU tool numbers, and nonresponse of the CFLU tool to participants' calls.

Participants who complained about the phone ringing endlessly found this challenge wearisome. They stated that, even when they entered the PIN, the phone would continue ringing without communicating any message:

Doctor sometimes it rings so endlessly. It gets to a point beyond what I can handle; and get fed up and just switch off the phone. [KSG HCIV, responded to <25% of calls, FGD 1]

Participants further complained about wrong timing of the pill reminders, that is, being called before or after the agreed time. Participants who faced this challenge strongly believed that it was a CFLU system error:

What I don't like about it is, I am supposed to get my pill reminder at night, but then at around 11am and I get the call. At first I did not know and when I come I am told to press some things and it would not respond. Then the doctor called me to ask about my health, then I also told her about that issue. I was told, when it calls at a wrong time I just press option 5. That was my only issue. I would wonder why they called at that time yet I am supposed to get the call at night. [IDI Mulago, responded to >50% of calls, FGD 6]

Related to the aforementioned challenge, participants further complained about inconsistent and irregular calls whereby, unlike the aforementioned challenge where the timing was inconsistent, participants facing this challenge were called on diverse days, which they found perplexing:

Doctor one more thing; sometimes they take a week without calling. I don't really know why. You wait for the pill reminders in vain; then after a week they resume. [IDI Mulago, responded to 25%-50% of calls, FGD 4]

Participants further complained about the CFLU tool calling them from different numbers, which they found disturbing and confusing. The numbers were reportedly from different countries that included Kenya, Burundi, Sudan, and South Africa. Some recounted that, when the "kadodi" played and they entered the password, no information was communicated until the call ended:

I think what they are complaining about is faced by more than one person. Because I personally get calls from 3 different numbers; when you receive all of them; they play the "kadodi" but when you enter the password, it does not give you the information you want; until the call switches itself off. It has now taken 2 weeks when I enter the password and it fails. [IDI Mulago, responded to >50% of calls, FGD 5]

The last technical issue participants pointed out was nonresponse of the CFLU tool to participants' calls. The few participants who experienced this challenge explained that, sometimes, when they called back after receiving a call from the CFLU tool, either the latter would go silent or would not pick up the call:

In addition to what she said, sometimes they may call and you wonder what is happening; then you try to call them back but they don't answer you. Airtime is spent, and yet you haven't gotten any response. [IDI Mulago, responded to 25%-50% of calls, FGD 4]

Another type of challenge concerned two aspects of CFLU namely, symptom reporting and the toll-free number. A few participants complained about long waits for the physician's response or never talking to a physician at all after reporting a symptom:

I may have a headache or a fever; on the phone; and they tell me they are going to give me a professional doctor to speak to me; but I have never talked to them or gotten any feedback from the doctor, or telling me to go to hospital. That's it. [KSG HCIV, responded to <25% of calls, FGD 1]

The last negative experience stated by participants was the nonfunctional toll-free number. The few who experienced this challenge complained about not receiving a response when they called the toll-free number, being told that they should not call that number, or receiving a response that the number did not exist on the network:

As for me, I would like to know, if I want to talk to the doctor directly, the toll free number that we were given directly, whenever I call it, no one answers. Unless if it is the counsellor that calls me directly.

But if it is me calling directly, it is hard for them to pick up. [IDI Mulago, responded to >50% of calls, FGD 6]

Patients' Ideas and Suggestions to Improve the CFLU Tool

The study captured patients' suggestions on how to improve the CFLU tool during the FGDs. The following suggestions emerged from the FGDs on how to improve the CFLU tool.

Expanding CFLU to Other Areas, Social Places, and Media and Including More People

This was reportedly the most common suggestion made to improve the CFLU tool. Participants suggested this idea based

on the benefits from the tool, most importantly, improvement in quality of life attributed to pill reminders and health tips. Participants considered reminders necessary for all patients, especially those residing in rural areas. They recommended that the tool should be expanded to include other people with a similar condition and not be limited to only urban health facilities but also extended to rural facilities, schools, and churches and extended further to media such as radios and televisions such that, on hearing these messages, affected or infected people would be inspired to get tested for HIV and be started on treatment. Participants were aware of the importance of adherence to medication and felt that CFLU would motivate others to adhere if enrolled. Illustrative quotes are provided in [Textbox 1](#).

Textbox 1. Illustrative quotes from participants regarding expanding Call for Life Uganda.

More people should be allowed to join Call for Life Uganda

- *Doctor, I am suggesting, even those who have not yet joined it, they should put them all on the system because all of us came here to work on our health. Whoever is seeking for good health should join CFL...because it helps to remind them; what brought them was their quest for good health. So when it helps them, it is not bad on their side. They will just be appreciative just like we are here who have already joined it. [IDI Mulago, responded to 25%-50% of calls, FGD 4]*

Call for Life Uganda should be scaled to rural health facilities

- *As for me, what I am suggesting is, they should not only have the system in the hospitals within the city; but they should also take the system to rural hospitals and health centers so that they can also benefit...so that they also enjoy a good life; because some people fear coming to these big public hospitals due to self-stigma. So if you also go to those places and educate them as well, one may come to hospital without any problem about it. [KSG HCIV, responded to 25%-50% of calls, FGD 3]*

The Call for Life Uganda tool should be extended to the media

- *Doctor; even if it is not rural areas alone; the doctors can come up with something; just like the herbalists; they are always on radios and TVs talking about what they do. Many people watch TVs and listen to radios; someone may be changing stations on the radio and randomly land on those herbalists speaking. So even if the person feels hopeless, they may hear this herbalist teaching about their medicine; and also see doctors teaching the same on TV. I think it would be of good help. Even those that had stopped taking their medicine will start; those that hadn't gone for testing will go and test themselves. [KSG HCIV, responded to <25% of calls, FGD 1]*

The Call for Life Uganda tool should also be extended to schools

- *One more thing, they should also go to schools, talk to the youth about how to prevent themselves from catching HIV/AIDS; in secondary schools, they could do free HIV testing; basically extend their health services to these schools. [KSG HCIV, responded to <25% of calls, FGD 1]*

Including Additional Health Tips (Topics) and Providing Detailed Information

This was the second most common suggestion proposed by participants to improve the CFLU tool. Interestingly, even participants who had challenges with the secret PIN code and were not receiving the health tips suggested health tips to include in the CFLU tool. Regarding additional health tips, the following were suggested: nutrition (ie, improved diet that comprises healthy eating and increased consumption of fruits for their nutritional benefits), behavior change (including topics on domestic violence and its effects on health), safe conception for discordant couples (eg, prevention of mother-to-child transmission and sexually transmitted diseases), and sensitization about medications and their effects.

The following is an illustrative statement:

According to me, I think what should be added, is first of all about the fruits. You could say maybe in a week or month, decide that this week we are talking

about this fruit. You outline its benefits and all other information concerning it, and this helps to urge more people to consume that fruit. For example, you may say this week you are talking about beetroot, and you go in depth in information concerning this fruit, and then the following week you talk about another fruit. You could also add some information which may not be medical as such. For instance, domestic violence; if it is in a family, even taking medicine changes; the health takes a negative turn, as well as other things. So if you decide that this week you are talking about domestic violence, and then you get good words to use for that. It could even be a case study, and you get a true story which is educative. It will make a person more anxious and anticipate for when they will get the next health tips. [IDI Mulago, responded to 25%-50% of calls, FGD 4]

A participant expressed the need for frequent provision of health tips:

About the health tips, what you can add on is maybe call on a daily basis to increase in knowledge; because learning never stops. I am sure that it will really help us to understand more about our health, and what we are following. [IDI Mulago, responded to <25% of calls, FGD 2]

Conversely, a participant was of the view that health tips may also be sent as SMS text messages to be read later during free time by patients with busy work schedules, as expressed as follows:

According to me; as a transporter, that is why sometimes I have very little time to look into those tips because I'm always having constant phone calls. Is there anyway those tips can be sent to us who are part of CFL, and we get some of these tips as messages on the phone. So that in your free time you can read through them; because us transporters we are always traveling; from here to Rwanda, Burundi, Congo. We are always on the move so there is hardly any time for us to listen to those health tips. Just like we get messages, WhatsApp messages, depending on

your PIN. [IDI Mulago, responded to 25%-50% of calls, FGD 4]

Providing Appointment Reminders to All Participants

Several participants suggested sending appointment reminders to all those involved with the CFLU tool as it was only a few who reportedly received them. Some suggested that reminders to get their medicine refills should be sent a few days before the appointment date:

In my opinion, they should send the appointment reminders 2 or 3 days to the date of appointment. They can call on Friday and say, "do you know that you have a medical checkup on Monday" "do not forget to go to the hospital." It would be very good but I have never received it. They should really do it like CFL. [IDI Mulago, responded to >50% of calls, FGD 5]

Resolving CFLU Tool Technical Issues Faced by Participants

Regarding specific technical issues revolving around the CFLU tool, participants made the suggestions outlined in [Textbox 2](#).

Textbox 2. Illustrative quotes from participants regarding solving technical issues.

<p>The Call for Life Uganda tool or system should have the same number (one common number) used to call patients to avoid confusion</p> <ul style="list-style-type: none"> What I am saying is; please have a permanent number on which you call us. There are about 3 different numbers that call me; there are numbers that call, you put in the password but the “kadodi” continues ringing. Yet the main number I know, when I put in the password it works very well. But some of those other numbers that call, we do not know them. I saved the CFL main number as “dawa” and as soon as I see it ringing I just put in the password and listen. But when the other numbers call, they reject my password. [IDI Mulago, responded to >50% of calls, FGD 5] <p>Resolve issues of irregular or inconsistent calling times</p> <ul style="list-style-type: none"> What I don't like about CFL, let them make sure if it is 10 o'clock that I chose to get my pill reminders, let them not call me at 1am or 2am. I am kindly requesting; that it should strictly be at 10pm. [IDI Mulago, responded to <25% of calls, FGD 2] <p>Rectify issues of blocked personal identification numbers</p> <ul style="list-style-type: none"> In most cases they call me; But I stopped getting the health tips because my pin code was blocked; so that is one of the issues I wanted to address to my doctor so that it can be rectified and I continue getting those health tips; because they help me a lot; when it comes to positive living. [IDI Mulago, responded to <25% of calls, FGD 2] <p>Address the challenges encountered with the toll-free number</p> <ul style="list-style-type: none"> One more thing; there is a toll free number for CFL I also tried to call it but they said it does not exist on the network. And then I think to myself; did the doctor give me a wrong number? I also request that they work on that. [IDI Mulago, responded to <25% of calls, FGD 2] <p>Continuation of the Call for Life Uganda intervention even after completion of the study</p> <ul style="list-style-type: none"> I request that even when we complete the study you continue giving us calls. [IDI Mulago, responded to >50% of calls, FGD 6] <p>Prompt response to symptom reporting</p> <ul style="list-style-type: none"> I personally think they should respond on that very same day that you report the symptom; so that you get help...sometimes you may be in great pain and you need urgent help...Doctor I would want to get my response as soon as I report the symptom. [IDI Mulago, responded to 25%-50% of calls, FGD 4] <p>Preference for two-way communication</p> <ul style="list-style-type: none"> I would want to speak directly to the doctor...The doctor will speak to you, and you can inquire about anything. But as for the answering system, it will just do the talking but you won't be able to talk back; all you do is just listen. [KSG HCIV, responded to <25% of calls, FGD 1] <p>Call for Life Uganda should have a separate space for its patients (clients) at health facilities</p> <ul style="list-style-type: none"> As for me what I am saying; us the ones part of CFL, we should have our own space where we get our medicine, and also bring others to join which means we will be spreading CFL to our peers...so that means you can slowly by slowly bring other patients to join the system. If I bring some, and someone else brings in another, the CFL circle will get bigger. [KSG HCIV, responded to <25% of calls, FGD 1]
--

Discussion

Principal Findings

This study complements the existing body of literature related to use and outcomes of mHealth communication technologies in health care delivery [30,42]. This substudy also complements our parent study that was conducted to determine the impact of IVR technology on Medical Outcomes HIV Quality of Life scores and viral suppression at 12 months, which showed better study outcomes and higher quality-of-life scores for participants with higher use of the CFLU tool than for low users [45]. It further complements another study (qualitative) by the team that explored the acceptability of a mobile phone support tool (CLFU) for promoting adherence to ART among young adults in a substudy of the RCT [48]. However, our study explored similarities and differences in the experiences, perceptions, and attitudes of people living with HIV regarding the CFLU tool among 3 categories based on the proportion of calls responded to. The study revealed that there were more similarities than differences in patients' experiences with the CFLU tool. In

particular, there was consensus across all groups that they had more positive experiences than negative ones with the CFLU system. However, participants who responded to >50% of calls reported more frequent use of the specific elements of the CFLU tool; hardly complained about the timing of the phone calls; and, consequently, experienced more benefits from the system than those who responded to calls <25% and between 25% and 50% of the time. This finding concurs with the results of the parent study, which revealed that, among participants in the intervention arm who were active at 12 months and had higher use of the CFLU tool than low users, there was a trend toward better study outcomes [45]. The challenges faced by the 2 categories who responded to <50% of the calls (mainly repeated challenges with the secret PIN code) are reflected in findings from the parent study and of another study that examined the feasibility of using IVR and SMS for automated collection of weekly individual-level ART adherence data in (rural) Southwestern Uganda [52]. In that study (which involved caregivers of children on ART who owned phones that were used to collect adherence data), Haberer et al [25,52] reported

similar challenges, especially with the PIN, which reportedly caused most confusion, where >66% of the Ugandan patients with HIV studied were unable to use their mobile phone PINs to feed back information to the health care provider. Despite these challenges, the study recommended that the use of IVR and SMS in resource-limited settings is technically feasible. Suggestions made to improve response rates to address the aforementioned challenges were found applicable to our study as well, which included repeated trainings over time, training in groups so shy participants can learn from each other, and testing knowledge from the trainings.

Generally, participants across the 3 categories valued the confidentiality associated with the secret PIN code, privacy (place of choice), convenience (individual time), and comfort of using the CFLU tool provided. Given that HIV-related stigma is still prevalent, CFLU was generally acceptable as it provided participants with privacy. A similar observation was reported by Adeagbo et al [13,53], who found that mobile phone-connected HIV testing and web-based clinical care and prevention have the potential to support access to these services, particularly for young people and men whose uptake of these services remains low, especially in Africa.

A key finding is that our study unveiled that the 2 most popular application elements were pill reminders and health promotion tips. Irrespective of calls responded to, all participants largely attributed improved quality of life to pill reminders, followed by health promotion tips. Self-reported compliance to treatment resulting from pill reminders was a topic that ran across the 3 categories and was even more pronounced among participants who responded to <25% of the calls as most of them used the system solely as a pill reminder. Our findings are broadly consistent with those from a study conducted in 2 rural Ugandan districts, where adherence levels were significantly higher during mobile phone intervention [6]. Similar findings were reported in Kenya, where 2 RCTs sent ART reminders to people living with HIV, whose adherence reportedly improved significantly [28,29]. In addition, our results correspond to the findings from reviews and other studies that include both LMICs and the global north, where mHealth communication technologies improved adherence through SMS text messages and voice pill reminders [3,14,30].

Furthermore, participants who responded to >50% of calls listened more to health tips than their counterparts in the <25% and 25% to 50% FGDs and ably reported detailed health information that they had put into practice. Our findings are in agreement with other studies, which revealed that mHealth technologies are highly effective for the dissemination of health promotion messages such as nutrition advice and other lifestyle tips [19,33,35] and directly increased disease awareness with regard to transmission, prevention, treatment, and care among study participants (people living with HIV), who, in turn, reported that they sensitized their families and community members. Mobile phones are being increasingly used worldwide in health promotion and health care [19], where improvement in communication and information delivery was observed through symptom reporting and appointment reminders for clinic attendance [37]. However, contrary to this finding, a pilot study that explored the efficacy of an mHealth campaign using

SMS as a platform to disseminate and measure HIV and AIDS knowledge and promote HIV and AIDS testing at clinics in rural Northwest Uganda had proportionately limited success in increasing knowledge levels on a mass scale [54]. Despite this challenge, the study recognized the potential of mHealth tools when extended to large numbers of mobile phone users as part of an integrated health campaign approach and suggests that mHealth campaigns need to be combined with other forms of dissemination in low-income countries where mobile phone access and literacy disparities exist.

Another discussion point is that most of the participants had good therapeutic relationships with the health care providers (eg, physicians, nurses, and counselors), which boosted their self-esteem, confidence, and trust in them and was a source of psychosocial support. Odili et al [10] believe that patient-care provider relationships and trust in the provider could be motivating factors for adherence.

An attempt has been made to provide concrete ideas on how to improve aspects of mHealth interventions in LMIC settings. As observed from the study findings, participants attributed their improved adherence, better health, and improved quality of life to CFLU and recommended that the tool be scaled to rural health facilities, nonparticipants, social places, and the media. Benefits of mHealth tools have been reported by other studies; for example, a systematic review of sociotechnical factors affecting patients' adoption of mHealth tools revealed that mHealth adoption may improve health outcomes [55]. This review further explained that patients who perceive potential benefits such as better health effects and enhanced health behaviors from the use of the tools are more likely to use them.

Among the recommendations to improve aspects of mHealth tools is addressing technical challenges that reportedly affect mHealth tool performance, such as those mentioned in our study. This is because technical difficulties have been frequently cited as barriers, creating frustration and discouraging the embracing of mHealth tools [56]. Ease of use has been quoted as one of the leading factors affecting mHealth acceptance [57] and, if made possible, can greatly improve the usability of the tools in LMICs.

Training involving a more participatory approach is recommended given the different sociodemographic characteristics of patients, especially those in the older age groups and those with low levels of education that are common in LMICs [58]. In addition, provision of timely and adequate technical support may help users overcome technical challenges that were reportedly common among the study participants [59]. Coupled with the aforementioned issues, follow-up or continuous monitoring of the users by the health care providers may help provide feedback on use and challenges that can improve the performance of the different aspects of the mHealth tools [60].

Furthermore, studies have pointed out the importance of combining web-based and traditional health care provider communication to enable quicker and easier exchange of information between health care providers and patients [55,61,62]. Preference for two-way communication was also suggested by study participants.

Finally, provision of relevant and up-to-date and appropriate health education tailored to patients' needs may greatly improve mHealth tools in LMICs where access to health education is constrained [56]. This may address knowledge gaps, raise disease awareness, and encourage healthier behaviors, consequently helping patients better understand their medications and possible side effects and symptoms and achieve better health results [55,56].

Study Limitations

A potential limitation of this study is the fact that it was conducted in a setting where poor mobile phone coverage and frequent power and mobile phone network outages exist and loss of mobile phones is a common phenomenon. These findings may not be applicable to higher-income settings as such challenges would have implications in terms of sustainability

of such an intervention. However, the findings of our study are not intended to be generalized to other settings. The resonance of our findings with other studies suggests that the findings may be applicable beyond the study area. Finally, the data in our study were based on informant responses, which might be subject to social desirability.

Conclusions

In conclusion, our findings showed participants' appreciation, high willingness, and interest in CFLU that demonstrated great potential to improve access to health care; adherence to treatment; health awareness; and, consequently, quality of life. The technology was well received, but the use of PIN codes for confidentiality was a challenge, and other confidentiality checks should be considered in our environment.

Acknowledgments

The authors are very grateful to the Grameen Foundation and Janssen, the pharmaceutical companies of Johnson and Johnson, who developed Call for Life Uganda (CFLU), a mobile health tool adapted from Connect For Life designed to help patients in a number of ways. Special thanks go to the study participants, who provided valuable information during the focus group discussions; the health workers in the study sites for their endless support; and the team that participated in data collection and data analysis.

Authors' Contributions

PK and AT led the data collection and analysis. PK drafted the first manuscript. AT contributed to manuscript writing and reviewing until the final version approval. RPR led the study design and implementation and contributed to manuscript reviewing. ABN led the study implementation and data collection and contributed to manuscript writing and reviewing. RK, AK, BC, and MSN contributed to the study design and manuscript reviewing until the final version approval.

Conflicts of Interest

The Infectious Diseases Institute, RP-R, and ABN have received research funding from Janssen, the pharmaceutical companies of Johnson and Johnson and Pfizer.

References

1. HIV/AIDS. World Health Organization. URL: <https://www.afro.who.int/health-topics/hiv/aids> [accessed 2022-11-22]
2. WHO Uganda Annual Report for 2018. World Health Organization. URL: <https://www.office.com/?auth=2> [accessed 2022-11-22]
3. Finitis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: [10.1371/journal.pone.0088166](https://doi.org/10.1371/journal.pone.0088166)] [Medline: [24505411](https://pubmed.ncbi.nlm.nih.gov/24505411/)]
4. Uganda AIDS country progress report July 2017-June 2018. Government of Uganda. 2018 Sep. URL: https://www.unaids.org/sites/default/files/country/documents/UGA_2019_countryreport.pdf [accessed 2022-11-22]
5. Joint United Nations Programme on HIV/AIDS (UNAIDS) Special Analysis 2019. unaids.org 2022.
6. Kunutsor S, Walley J, Katabira E, Muchuro S, Balidawa H, Namagala E, et al. Using mobile phones to improve clinic attendance amongst an antiretroviral treatment cohort in rural Uganda: a cross-sectional and prospective study. AIDS Behav 2010 Dec 11;14(6):1347-1352. [doi: [10.1007/s10461-010-9780-2](https://doi.org/10.1007/s10461-010-9780-2)] [Medline: [20700644](https://pubmed.ncbi.nlm.nih.gov/20700644/)]
7. Venter W, Coleman J, Chan VL, 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](https://doi.org/10.2196/mhealth.8376)] [Medline: [30021706](https://pubmed.ncbi.nlm.nih.gov/30021706/)]
8. Rosen S, Fox MP. Retention in HIV care between testing and treatment in sub-Saharan Africa: a systematic review. PLoS Med 2011 Jul 19;8(7):e1001056 [FREE Full text] [doi: [10.1371/journal.pmed.1001056](https://doi.org/10.1371/journal.pmed.1001056)] [Medline: [21811403](https://pubmed.ncbi.nlm.nih.gov/21811403/)]
9. Fox M, Rosen S. Patient retention in antiretroviral therapy programs up to three years on treatment in sub-Saharan Africa, 2007-2009: systematic review. Trop Med Int Health 2010 Jun;15 Suppl 1:1-15 [FREE Full text] [doi: [10.1111/j.1365-3156.2010.02508.x](https://doi.org/10.1111/j.1365-3156.2010.02508.x)] [Medline: [20586956](https://pubmed.ncbi.nlm.nih.gov/20586956/)]
10. Odili VU, Obieche AO, Amibor KC. Adherence to antiretroviral therapy and its determinants among HIV-infected patients in Nigeria. J Pharm Pract 2017 Jun;30(3):291-295. [doi: [10.1177/0897190016633978](https://doi.org/10.1177/0897190016633978)] [Medline: [26921233](https://pubmed.ncbi.nlm.nih.gov/26921233/)]

11. Heestermans T, Browne JL, Aitken SC, Vervoort SC, Klipstein-Grobusch K. Determinants of adherence to antiretroviral therapy among HIV-positive adults in sub-Saharan Africa: a systematic review. *BMJ Glob Health* 2016 Dec 30;1(4):e000125 [[FREE Full text](#)] [doi: [10.1136/bmjgh-2016-000125](https://doi.org/10.1136/bmjgh-2016-000125)] [Medline: [28588979](#)]
12. Reda AA, Biadgilign S. Determinants of adherence to antiretroviral therapy among HIV-infected patients in Africa. *AIDS Res Treat* 2012;2012:574656-574658 [[FREE Full text](#)] [doi: [10.1155/2012/574656](https://doi.org/10.1155/2012/574656)] [Medline: [22461980](#)]
13. Adeagbo O, Herbst C, Blandford A, McKendry R, Estcourt C, Seeley J, et al. Exploring people's candidacy for mobile health-supported HIV testing and care services in rural KwaZulu-Natal, South Africa: qualitative study. *J Med Internet Res* 2019 Nov 18;21(11):e15681 [[FREE Full text](#)] [doi: [10.2196/15681](https://doi.org/10.2196/15681)] [Medline: [31738174](#)]
14. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. *J Med Internet Res* 2014 Oct 17;16(10):e222 [[FREE Full text](#)] [doi: [10.2196/jmir.3442](https://doi.org/10.2196/jmir.3442)] [Medline: [25326646](#)]
15. 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 Mar 27;20(1):4-34. [doi: [10.1080/10810730.2013.864735](https://doi.org/10.1080/10810730.2013.864735)] [Medline: [24673171](#)]
16. *MHealth New Horizons for Health Through Mobile Technologies*. Geneva: World Health Organization; 2011.
17. *Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection Recommendations for a Public Health Approach*. Geneva: World Health Organization; 2016.
18. Marent B, Henwood F, Darking M, EmERGE Consortium. Development of an mHealth platform for HIV care: gathering user perspectives through co-design workshops and interviews. *JMIR Mhealth Uhealth* 2018 Oct 19;6(10):e184 [[FREE Full text](#)] [doi: [10.2196/mhealth.9856](https://doi.org/10.2196/mhealth.9856)] [Medline: [30339132](#)]
19. Fischer H, Moore SL, Ginosar D, Davidson AJ, Rice-Peterson CM, Durfee MJ, et al. Care by cell phone: text messaging for chronic disease management. *Am J Manag Care* 2012 Feb 01;18(2):e42-e47 [[FREE Full text](#)] [Medline: [22435883](#)]
20. Kruse C, Betancourt J, Ortiz S, Valdes Luna SM, Bamrah IK, Segovia N. Barriers to the use of mobile health in improving health outcomes in developing countries: systematic review. *J Med Internet Res* 2019 Oct 09;21(10):e13263 [[FREE Full text](#)] [doi: [10.2196/13263](https://doi.org/10.2196/13263)] [Medline: [31593543](#)]
21. Kahn JG, Yang JS, Kahn JS. 'Mobile' health needs and opportunities in developing countries. *Health Aff (Millwood)* 2010 Feb;29(2):252-258. [doi: [10.1377/hlthaff.2009.0965](https://doi.org/10.1377/hlthaff.2009.0965)] [Medline: [20348069](#)]
22. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. *Matern Child Health J* 2012 Jul 19;16(5):1092-1101. [doi: [10.1007/s10995-011-0836-3](https://doi.org/10.1007/s10995-011-0836-3)] [Medline: [21688111](#)]
23. Chan CV, Kaufman DR. A technology selection framework for supporting delivery of patient-oriented health interventions in developing countries. *J Biomed Inform* 2010 Apr;43(2):300-306 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2009.09.006](https://doi.org/10.1016/j.jbi.2009.09.006)] [Medline: [19796709](#)]
24. Déglise C, Suggs LS, Odermatt P. SMS for disease control in developing countries: a systematic review of mobile health applications. *J Telemed Telecare* 2012 Jul 23;18(5):273-281. [doi: [10.1258/jtt.2012.110810](https://doi.org/10.1258/jtt.2012.110810)] [Medline: [22826375](#)]
25. Bastawrous A, Armstrong MJ. Mobile health use in low- and high-income countries: an overview of the peer-reviewed literature. *J R Soc Med* 2013 Apr 01;106(4):130-142 [[FREE Full text](#)] [doi: [10.1177/0141076812472620](https://doi.org/10.1177/0141076812472620)] [Medline: [23564897](#)]
26. Vervloet M, Linn AJ, van Weert JC, de Bakker DH, Bouvy ML, van Dijk L. The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. *J Am Med Inform Assoc* 2012 Sep 01;19(5):696-704 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2011-000748](https://doi.org/10.1136/amiajnl-2011-000748)] [Medline: [22534082](#)]
27. Georgette N, Siedner MJ, Petty CR, Zandoni BC, Carpenter S, Haberer JE. Impact of a clinical program using weekly Short Message Service (SMS) on antiretroviral therapy adherence support in South Africa: a retrospective cohort study. *BMC Med Inform Decis Mak* 2017 Feb 20;17(1):18 [[FREE Full text](#)] [doi: [10.1186/s12911-017-0413-9](https://doi.org/10.1186/s12911-017-0413-9)] [Medline: [28219372](#)]
28. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. *Lancet* 2010 Nov;376(9755):1838-1845. [doi: [10.1016/s0140-6736\(10\)61997-6](https://doi.org/10.1016/s0140-6736(10)61997-6)]
29. Pop-Eleches C, Thirumurthy H, Habyarimana J, Zivin J, Goldstein M, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS* 2011 Mar 27;25(6):825-834 [[FREE Full text](#)] [doi: [10.1097/QAD.0b013e32834380c1](https://doi.org/10.1097/QAD.0b013e32834380c1)] [Medline: [21252632](#)]
30. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. *BMC Health Serv Res* 2015 Jan 22;15(1):33 [[FREE Full text](#)] [doi: [10.1186/s12913-014-0654-6](https://doi.org/10.1186/s12913-014-0654-6)] [Medline: [25609559](#)]
31. Mbuagbaw L, van der Kop ML, Lester RT, Thirumurthy H, Pop-Eleches C, Ye C, et al. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): an individual patient data meta-analysis of randomised trials. *BMJ Open* 2013 Dec 17;3(12):e003950 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2013-003950](https://doi.org/10.1136/bmjopen-2013-003950)] [Medline: [24345901](#)]
32. Crankshaw T, Corless IB, Giddy J, Nicholas PK, Eichbaum Q, Butler LM. Exploring the patterns of use and the feasibility of using cellular phones for clinic appointment reminders and adherence messages in an antiretroviral treatment clinic,

- Durban, South Africa. *AIDS Patient Care STDS* 2010 Nov;24(11):729-734. [doi: [10.1089/apc.2010.0146](https://doi.org/10.1089/apc.2010.0146)] [Medline: [21039181](https://pubmed.ncbi.nlm.nih.gov/21039181/)]
33. Littman-Quinn R, Chandra A, Schwartz A, Chang A, Fadlelmola F, Ghose S, et al. mHealth applications for clinical education, decision making, and patient adherence in Botswana. In: *Proceedings of the 2011 IST-Africa Conference Proceedings*. 2011 Presented at: 2011 IST-Africa Conference Proceedings; May 11-13, 2011; Gaborone, Botswana.
 34. Rana Y, Haberer J, Huang H, Kambugu A, Mukasa B, Thirumurthy H, et al. Short message service (SMS)-based intervention to improve treatment adherence among HIV-positive youth in Uganda: focus group findings. *PLoS One* 2015 Apr 16;10(4):e0125187 [FREE Full text] [doi: [10.1371/journal.pone.0125187](https://doi.org/10.1371/journal.pone.0125187)] [Medline: [25881059](https://pubmed.ncbi.nlm.nih.gov/25881059/)]
 35. John M, Samson-Akpan P, Etowa J, Akpabio I, John E. Enhancing self-care, adjustment and engagement through mobile phones in youth with HIV. *Int Nurs Rev* 2016 Dec 21;63(4):555-561. [doi: [10.1111/inr.12313](https://doi.org/10.1111/inr.12313)] [Medline: [27653013](https://pubmed.ncbi.nlm.nih.gov/27653013/)]
 36. Head KJ, Noar SM, Iannarino NT, Grant Harrington N. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. *Soc Sci Med* 2013 Nov;97:41-48. [doi: [10.1016/j.socscimed.2013.08.003](https://doi.org/10.1016/j.socscimed.2013.08.003)] [Medline: [24161087](https://pubmed.ncbi.nlm.nih.gov/24161087/)]
 37. Car J, Gurol - Urganici I, de Jongh T, Vodopivec - Jamsek V, Atun R. Mobile phone messaging reminders for attendance at healthcare appointments. *Cochrane Database of Systematic Reviews*. 2012 Jul 11. URL: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD007458.pub2/full> [accessed 2022-11-22]
 38. Guy R, Hocking J, Wand H, Stott S, Ali H, Kaldor J. How effective are short message service reminders at increasing clinic attendance? A meta-analysis and systematic review. *Health Serv Res* 2012 Apr 08;47(2):614-632 [FREE Full text] [doi: [10.1111/j.1475-6773.2011.01342.x](https://doi.org/10.1111/j.1475-6773.2011.01342.x)] [Medline: [22091980](https://pubmed.ncbi.nlm.nih.gov/22091980/)]
 39. Berger E. Telemedicine: has its time come? *Ann Emerg Med* 2010 Nov;56(5):A15-A17. [doi: [10.1016/j.annemergmed.2010.09.008](https://doi.org/10.1016/j.annemergmed.2010.09.008)] [Medline: [21064235](https://pubmed.ncbi.nlm.nih.gov/21064235/)]
 40. Blake H. Innovation in practice: mobile phone technology in patient care. *Br J Community Nurs* 2008 Apr;13(4):160, 162-160, 165. [doi: [10.12968/bjcn.2008.13.4.29024](https://doi.org/10.12968/bjcn.2008.13.4.29024)] [Medline: [18595303](https://pubmed.ncbi.nlm.nih.gov/18595303/)]
 41. Mechael P. Barriers and gaps affecting mHealth in low and middle income countries : policy white paper. Columbia university. 2010. URL: <http://bibalex.org/baifa/en/resources/document/452419> [accessed 2022-11-22]
 42. Abaza H, Marschollek M. mHealth application areas and technology combinations. *Methods Inf Med* 2018 Jan 31;56(S 01):e105-e122. [doi: [10.3414/me17-05-0003](https://doi.org/10.3414/me17-05-0003)]
 43. Qiao S, Zhang Y, Li X, Menon JA. Facilitators and barriers for HIV-testing in Zambia: a systematic review of multi-level factors. *PLoS One* 2018 Feb 7;13(2):e0192327 [FREE Full text] [doi: [10.1371/journal.pone.0192327](https://doi.org/10.1371/journal.pone.0192327)] [Medline: [29415004](https://pubmed.ncbi.nlm.nih.gov/29415004/)]
 44. de Tolly K, Skinner D, Nembaware V, Benjamin P. Investigation into the use of short message services to expand uptake of human immunodeficiency virus testing, and whether content and dosage have impact. *Telemed J E Health* 2012 Jan;18(1):18-23. [doi: [10.1089/tmj.2011.0058](https://doi.org/10.1089/tmj.2011.0058)] [Medline: [22150712](https://pubmed.ncbi.nlm.nih.gov/22150712/)]
 45. Byonanebye DM, Nabaggala MS, Naggirinya AB, Lamorde M, Oseku E, King R, et al. An interactive voice response software to improve the quality of life of people living with HIV in Uganda: randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Feb 11;9(2):e22229 [FREE Full text] [doi: [10.2196/22229](https://doi.org/10.2196/22229)] [Medline: [33570497](https://pubmed.ncbi.nlm.nih.gov/33570497/)]
 46. Macleod B, Phillips J, Stone A, Walji A, Awoonor-Williams JK. The architecture of a software system for supporting community-based primary health care with mobile technology: the mobile technology for community health (MoTech) initiative in Ghana. *Online J Public Health Inform* 2012 May 17;4(1):ojphi.v4i1.3910 [FREE Full text] [doi: [10.5210/ojphi.v4i1.3910](https://doi.org/10.5210/ojphi.v4i1.3910)] [Medline: [23569631](https://pubmed.ncbi.nlm.nih.gov/23569631/)]
 47. Swendeman D, Jana S, Ray P, Mindry D, Das M, Bhakta B. Development and pilot testing of daily interactive voice response (IVR) calls to support antiretroviral adherence in India: a mixed-methods pilot study. *AIDS Behav* 2015 Jun;19 Suppl 2:142-155 [FREE Full text] [doi: [10.1007/s10461-014-0983-9](https://doi.org/10.1007/s10461-014-0983-9)] [Medline: [25638037](https://pubmed.ncbi.nlm.nih.gov/25638037/)]
 48. Twimukye A, Bwanika Naggirinya A, Parkes-Ratanshi R, Kasirye R, Kiragga A, Castelnuovo B, et al. Acceptability of a mobile phone support tool (call for life Uganda) for promoting adherence to antiretroviral therapy among young adults in a randomized controlled trial: exploratory qualitative study. *JMIR Mhealth Uhealth* 2021 Jun 14;9(6):e17418 [FREE Full text] [doi: [10.2196/17418](https://doi.org/10.2196/17418)] [Medline: [34121665](https://pubmed.ncbi.nlm.nih.gov/34121665/)]
 49. Download NVivo. NVivo. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/support-services/nvivo-downloads> [accessed 2022-11-22]
 50. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
 51. Baxter P, Jack S. Qualitative case study methodology: study design and implementation for novice researchers. *Qual Report* 2015 Jan 14;13(4):544-559. [doi: [10.46743/2160-3715/2008.1573](https://doi.org/10.46743/2160-3715/2008.1573)]
 52. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Challenges in using mobile phones for collection of antiretroviral therapy adherence data in a resource-limited setting. *AIDS Behav* 2010 Dec 8;14(6):1294-1301 [FREE Full text] [doi: [10.1007/s10461-010-9720-1](https://doi.org/10.1007/s10461-010-9720-1)] [Medline: [20532605](https://pubmed.ncbi.nlm.nih.gov/20532605/)]
 53. Siu GE, Wight D, Seeley JA. Masculinity, social context and HIV testing: an ethnographic study of men in Busia district, rural eastern Uganda. *BMC Public Health* 2014 Jan 13;14(1):33 [FREE Full text] [doi: [10.1186/1471-2458-14-33](https://doi.org/10.1186/1471-2458-14-33)] [Medline: [24417763](https://pubmed.ncbi.nlm.nih.gov/24417763/)]

54. Chib A, Wilkin H, Ling LX, Hoefman B, Van Biejma H. You have an important message! Evaluating the effectiveness of a text message HIV/AIDS campaign in Northwest Uganda. *J Health Commun* 2012 May 02;17 Suppl 1(sup1):146-157. [doi: [10.1080/10810730.2011.649104](https://doi.org/10.1080/10810730.2011.649104)] [Medline: [22548607](https://pubmed.ncbi.nlm.nih.gov/22548607/)]
55. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. *JMIR Mhealth Uhealth* 2022 May 05;10(5):e36284 [FREE Full text] [doi: [10.2196/36284](https://doi.org/10.2196/36284)] [Medline: [35318189](https://pubmed.ncbi.nlm.nih.gov/35318189/)]
56. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
57. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. *JMIR Mhealth Uhealth* 2020 Jul 09;8(7):e17588 [FREE Full text] [doi: [10.2196/17588](https://doi.org/10.2196/17588)] [Medline: [32673249](https://pubmed.ncbi.nlm.nih.gov/32673249/)]
58. De La Cruz Monroy MF, Mosahebi A. The use of smartphone applications (apps) for enhancing communication with surgical patients: a systematic review of the literature. *Surg Innov* 2019 Apr 02;26(2):244-259. [doi: [10.1177/1553350618819517](https://doi.org/10.1177/1553350618819517)] [Medline: [30602332](https://pubmed.ncbi.nlm.nih.gov/30602332/)]
59. Abelson JS, Kaufman E, Symer M, Peters A, Charlson M, Yeo H. Barriers and benefits to using mobile health technology after operation: a qualitative study. *Surgery* 2017 Sep;162(3):605-611. [doi: [10.1016/j.surg.2017.05.007](https://doi.org/10.1016/j.surg.2017.05.007)] [Medline: [28651777](https://pubmed.ncbi.nlm.nih.gov/28651777/)]
60. Li J, Varnfield M, Jayasena R, Celler B. Home telemonitoring for chronic disease management: perceptions of users and factors influencing adoption. *Health Informatics J* 2021 Mar 08;27(1):1460458221997893 [FREE Full text] [doi: [10.1177/1460458221997893](https://doi.org/10.1177/1460458221997893)] [Medline: [33685279](https://pubmed.ncbi.nlm.nih.gov/33685279/)]
61. Runz-Jørgensen SM, Schiøtz ML, Christensen U. Perceived value of eHealth among people living with multimorbidity: a qualitative study. *J Comorb* 2017 Aug 24;7(1):96-111 [FREE Full text] [doi: [10.15256/joc.2017.7.98](https://doi.org/10.15256/joc.2017.7.98)] [Medline: [29359124](https://pubmed.ncbi.nlm.nih.gov/29359124/)]
62. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study. *J Med Internet Res* 2017 May 30;19(5):e187 [FREE Full text] [doi: [10.2196/jmir.7479](https://doi.org/10.2196/jmir.7479)] [Medline: [28559223](https://pubmed.ncbi.nlm.nih.gov/28559223/)]

Abbreviations

ART: antiretroviral therapy
CFLU: Call for Life Uganda
FGD: focus group discussion
IDI: Infectious Diseases Institute
IVR: interactive voice response
KSG HCIV: Kasangati Health Centre IV
LMIC: low- and middle-income country
mHealth: mobile health
PIN: personal identification number
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 27.01.22; peer-reviewed by M Tummalacherla, J Mitchell; comments to author 30.03.22; revised version received 18.07.22; accepted 28.07.22; published 06.12.22.

Please cite as:

Kajubi P, Parkes-Ratanshi R, Twimukye A, Bwanika Naggirinya A, Nabaggala MS, Kiragga A, Castelnuovo B, King R
Perceptions and Attitudes Toward an Interactive Voice Response Tool (Call for Life Uganda) Providing Adherence Support and Health Information to HIV-Positive Ugandans: Qualitative Study
JMIR Form Res 2022;6(12):e36829
URL: <https://formative.jmir.org/2022/12/e36829>
doi: [10.2196/36829](https://doi.org/10.2196/36829)
PMID: [36472904](https://pubmed.ncbi.nlm.nih.gov/36472904/)

©Phoebe Kajubi, Rosalind Parkes-Ratanshi, Adelline Twimukye, Agnes Bwanika Naggirinya, Maria Sarah Nabaggala, Agnes Kiragga, Barbara Castelnuovo, Rachel King. Originally published in *JMIR Formative Research* (<https://formative.jmir.org/>), 06.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Design of a Patient Voice App Experience for Heart Failure Management: Usability Study

Antonia Barbaric^{1,2,3}, MASc; Cosmin Munteanu^{4,5}, PhD; Heather Ross^{6,7,8}, MD, MHSc; Joseph A Cafazzo^{1,2,3,9,10}, PEng, PhD

¹Centre for Digital Therapeutics, Techna Institute, University Health Network, Toronto, ON, Canada

²Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

³Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

⁴Institute for Communication, Culture, Information, and Technology, University of Toronto, Mississauga, ON, Canada

⁵Technologies for Aging Gracefully Lab, University of Toronto, Toronto, ON, Canada

⁶Ted Rogers Centre for Heart Research, University Health Network, Toronto, ON, Canada

⁷Department of Medicine, University of Toronto, Toronto, ON, Canada

⁸Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada

⁹Department of Computer Science, University of Toronto, Toronto, ON, Canada

¹⁰Healthcare Human Factors, Techna Institute, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Antonia Barbaric, MASc

Centre for Digital Therapeutics

Techna Institute

University Health Network

Toronto General Hospital - RFE Building, 4th floor

190 Elizabeth St

Toronto, ON, M5G2C4

Canada

Phone: 1 416 340 4800 ext 4765

Email: antonia.barbaric@mail.utoronto.ca

Abstract

Background: The use of digital therapeutics (DTx) in the prevention and management of medical conditions has increased through the years, with an estimated 44 million people using one as part of their treatment plan in 2021, nearly double the number from the previous year. DTx are commonly accessed through smartphone apps, but offering these treatments through additional platforms can improve the accessibility of these interventions. Voice apps are an emerging technology in the digital health field; not only do they have the potential to improve DTx adherence, but they can also create a better user experience for some user groups.

Objective: This research aimed to identify the acceptability and feasibility of offering a voice app for a chronic disease self-management program. The objective of this project was to design, develop, and evaluate a voice app of an already-existing smartphone-based heart failure self-management program, *Medly*, to be used as a case study.

Methods: A voice app version of *Medly* was designed and developed through a user-centered design process. We conducted a usability study and semistructured interviews with patients with heart failure (N=8) at the Peter Munk Cardiac Clinic in Toronto General Hospital to better understand the user experience. A *Medly* voice app prototype was built using a software development kit in tandem with a cloud computing platform and was verified and validated before the usability study. Data collection and analysis were guided by a mixed methods triangulation convergence design.

Results: Common themes were identified in the results of the usability study, which involved 8 participants with heart failure. Almost all participants (7/8, 88%) were satisfied with the voice app and felt confident using it, although half of the participants (4/8, 50%) were unsure about using it in the future. Six main themes were identified: changes in physical behavior, preference between voice app and smartphone, importance of music during voice app interaction, lack of privacy concerns, desired reassurances during voice app interaction, and helpful aids during voice app interaction. These findings were triangulated with the quantitative

data, and it concluded that the main area for improvement was related to the ease of use; design changes were then implemented to better improve the user experience.

Conclusions: This work offered preliminary insight into the acceptability and feasibility of a *Medly* voice app. Given the recent emergence of voice apps in health care, we believe that this research offered invaluable insight into successfully deploying DTx for chronic disease self-management using this technology.

(*JMIR Form Res* 2022;6(12):e41628) doi:[10.2196/41628](https://doi.org/10.2196/41628)

KEYWORDS

heart failure; self-management; digital therapeutics; voice-activated technology; smart speaker; usability study; formative evaluation; mobile phone; smartphone

Introduction

Background

The prevalence of heart failure (HF) continues to be on the rise as more people are surviving cardiovascular disease [1]. Cardiovascular diseases can cause the heart muscle to become damaged and weak, leading to the development of HF. HF occurs when the pumping action of the heart muscle is not strong enough to meet the needs of the body or when the heart muscle does not relax properly to accommodate blood flow back into the heart. When this occurs, fluid can build up in the lungs and other parts of the body, such as the ankles, creating congestion in the lungs, which results in a lack of oxygen being delivered to the rest of the body [2]. As of 2017, it is estimated that 64.3 million people are living with HF globally [3], and many countries are reporting a steady increase in this condition's prevalence [1].

HF creates a burden not only on health care resources and expenditures [1] but also on the patient's well-being if not cared for properly [4]. HF limits a patient's capacity to live well through physical, psychological, or social means [5]. Patient self-management plays an integral role in the treatment of HF, with studies reporting improved health outcomes, decreased clinic visits, and decreased health costs [6]. Mobile health (mHealth) is a type of digital health technology that involves the use of mobile devices (smartphone, patient monitoring device, wireless devices, etc) for medical and public health practices [7] and enables the integration of self-care support into a patient's daily routine [8]. Smartphone apps are one of the most popular tools for helping patients diagnosed with chronic conditions manage their health at home [9].

Medly is an evidence-based, HF self-management program that has been developed by the University Health Network (UHN) and is implemented as part of the standard of care at UHN's Peter Munk Cardiac Centre (PMCC) [10]. This program is available to patients as a prescribed digital therapeutic, cleared by Health Canada in 2020 [11]. *Medly* consists of a smartphone app that enables patients to log clinically relevant physiological measurements and symptoms daily, which is then used in the *Medly* algorithm to generate an automated self-care message. The care team is able to review the patient's data daily and can also view current trends and historic data. If the algorithm detects that the patient's health is deteriorating, the care team receives alerts via email and on the dashboard and is able to contact the patient to provide advice. Previous studies evaluating

Medly have proven that this program can reduce health service use and improve clinical, quality of life, and patient self-care outcomes [10].

Given these encouraging results, a longitudinal study measuring adherence was also performed on the *Medly* program. The data gathered over 12 months concluded that adherence to the program was inconsistent among all age groups and kept declining as time went on [12]. Furthermore, other studies have shown that various patient demographics struggle to use mobile technologies, such as touch screen devices and smartphones, specifically older adults and those with cognitive and physical impairments [13-19]. Although mHealth remains to be one of the most popular platforms for chronic disease self-management programs [9], there are still some patient demographics who struggle to interact with this technology. Voice apps have the potential to improve the user interaction of programs such as *Medly* based on emerging evidence of this technology being used in the self-management of chronic diseases. This study was performed and designed as the first step in determining whether it would be feasible and beneficial to create, deploy, and use a voice app of *Medly*. If results show promise, future studies can go beyond feasibility to then investigate implementation and whether certain demographics benefit more from this technology than others.

Voice apps are an emerging technology in the health care field, ranging from integration with in-clinic registration processes [20] to helping people manage their chronic illness and live independently in their homes [21-25]. There is growing interest in smart speakers such as Google Home (Google LLC) and Amazon Alexa (Amazon Inc) because of their simple setup, ease of use, and low cost. Voice apps are deployed through voice user interfaces (VUIs) and enable ubiquitous connectivity by allowing the user to access services using their speech, making for a more convenient experience. VUI is the underlying software that processes and handles speech inputs to allow the user to interact with various apps. A wide range of devices now includes VUIs, such as smartphones and smart speakers. Research relating to voice apps for chronic illness is still in the development and piloting phases and has limited efficacy in testing to support final outcomes and conclusions [26]. As described by Sezgin et al [24,26], most voice apps currently being developed provide information and assistance services, which include general educational content and guidance, reminders, and tracking. There is limited research showcasing voice apps as a tool to provide more personalized, user-specific support, creating an opportunity to further investigate whether

this technology can be used to provide health care services and support.

Objectives

Guided by the findings from the literature review and coupled with a user-centered design (UCD) process, we sought to design, develop, and evaluate a voice app for the smartphone-based HF self-management program *Medly* to determine whether a voice app version adds benefit to the program's current model of interaction and care. The *Medly* voice app was evaluated in a usability study, and the findings from this evaluation helped inform the final design and development of the *Medly* voice app, as presented in this paper.

Methods

Overview

The *Medly* voice app was created in two main phases: (1) design and development and (2) usability evaluation of the preliminary design. A prototype version of the *Medly* voice app was created and used in a usability study. Findings from the usability evaluation helped inform the redesign and redevelopment work needed to create a voice app that better met the needs of patients with HF. This section begins with an outline of the design and development specifications created for the *Medly* voice app, followed by the verification and validation protocols. The methods pertaining to the usability study design and data collection and analysis will then be explained.

The *Medly* voice app was designed and developed for deployment on VUIs and was built using an external cloud computing service. A mixed methods, triangulation convergence model was used to interpret the results. Both quantitative and qualitative data regarding the user's experience when interacting with the *Medly* voice app prototype were collected and analyzed separately. Following the analysis of each data set, the results were converged for interpretation to identify the main conclusions. The findings from this usability study influenced the redesign and redevelopment work that occurred after this evaluation.

Ethical Considerations

The usability study was performed in Toronto General Hospital, part of UHN, and was approved by UHN Research Ethics Board (19-5051.2). Interested participants were asked to sign a written consent form before being enrolled in the study. They were

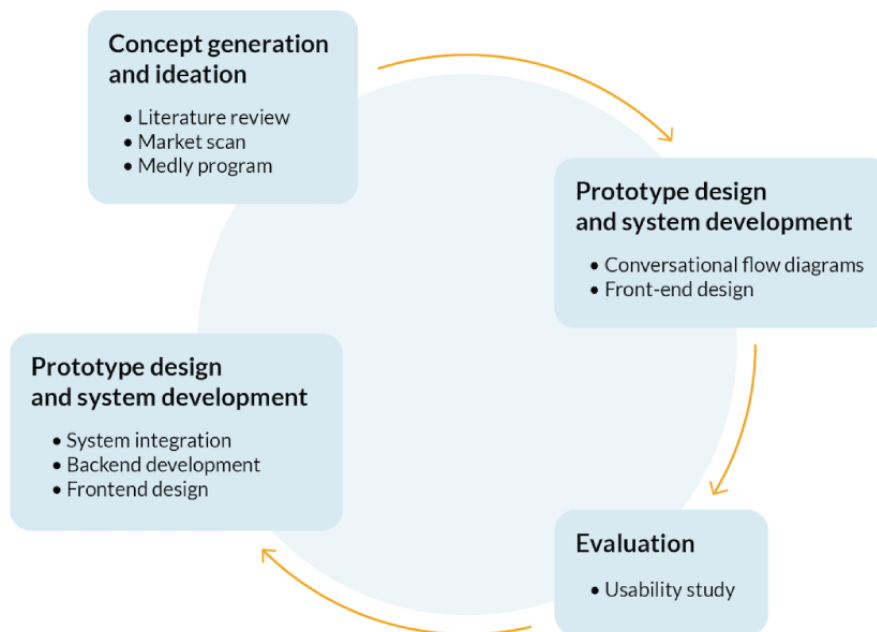
reminded that (1) they can contact the study coordinator at any point should they have questions, (2) their decision to participate or not has no impact on their clinical care in any way, and (3) they are free to withdraw from the study at any point with no impact on their clinical care. All information collected pertaining to observations, semistructured interviews, and questionnaire responses was identified by study number only. Neither participants' name nor any identifying information will be used in any publication or presentation. The participants received a CAD \$25 (US \$18.34) Visa gift card for participating in the study.

Design Process

The first phase of the UCD process, identifying user needs and ideation of the *Medly* voice app, was guided by a literature review, a market scan, and the existing *Medly* program requirements. Findings from the literature review and market scan revealed that many chronic disease self-management voice apps (not limited to just HF) used Amazon Alexa or Google Home devices for deployment. All the research studies shared the same purpose of identifying the feasibility and acceptability of using a voice app for the self-management of a specific chronic disease. The findings were preliminary but encouraging with high user satisfaction and usability, indicating that devices such as Amazon Alexa and Google Home are reliable to be used for self-management purposes. The design of the voice apps described in both the literature review and market scan revealed two useful findings: (1) the voice app conversation is designed to be short and succinct, and (2) some of the voice apps are deployed on devices that have large screens, meant to complement the interaction. These results encouraged the *Medly* voice app's conversational design to be concise and to explore the possibility of using a screen to complement the conversational experience.

The second phase of the UCD process involved using the information gathered in the first phase to develop and deploy a prototype of the *Medly* voice app on a VUI. User feedback about the *Medly* voice app was received through a usability study, performed with patients with HF in The Ted Rogers and Family Centre of Excellence in Heart Function clinic at the PMCC in Toronto General Hospital. On the basis of the usability study results, the prototype was further developed and redesigned to better meet the identified user needs and is expected to be evaluated in a future clinical study. [Figure 1](#) illustrates the process that was followed for this project.

Figure 1. The design process was guided by a user-centered design process; details describing what was done during each phase are included.



Design Specifications of the Medly Voice App

Design requirements were created before developing the *Medly* voice app to help guide the development process and define the voice app's functionalities. The main requirements implemented for the voice app are presented in Table S1 in [Multimedia Appendix 1](#), which summarize the main objectives of the program. The voice app was designed to have the same functionalities as those of the mobile app, which include the following: (1) asking the user to measure their weight, blood pressure, and heart rate; (2) saving and storing those values in the *Medly* clinical dashboard; (3) asking the user a series of yes or no questions relating to HF symptoms; (4) processing the data using the *Medly* algorithm; and (5) outputting a message to the user based on the algorithm result. These requirements helped create an app that was appropriate for voice interaction and were based on research as well as guidelines related to VUI design [27,28]. Conversational flow diagrams were also created using VUI design guidelines, and each scenario was tested on a VUI following its implementation [29,30].

Development Specifications of the Medly Voice App

The architecture of the *Medly* voice program consists of various components required to receive, store, and send data. A visual of the architectural design along with a description of how it operates is presented in Figure S1 in [Multimedia Appendix 1](#).

Voice apps consist of both a VUI and app logic. When the user speaks, the VUI invokes the voice app, and a request is created and processed in the context of the voice app's interaction model using tools such as machine learning, natural language understanding, and automatic speech recognition. Once the interaction model has processed the speech, a request is sent to the app logic, which provides a response back to the user. For the purposes of this project, an Alexa-hosted Skill built using the Amazon Alexa Skills Kit [31] processed and handled user requests and implemented the *Medly* voice app logic, but other frameworks could have been used as well.

Finally, the verification and validation of the *Medly* voice app were ongoing processes that were completed at various steps in the design process. Verification was done on the *Medly* voice app using a demonstration method to ensure that it met the design requirements and performed as expected. Different scenarios with various inputs were created to ensure that the produced results were as expected. Different conversational flows were brainstormed and used to help validate the voice app. These conversations were then tested using a web-based voice app simulator, which responds with voice and text and also includes the JavaScript Object Notation request to help fix errors.

Usability Study

Overview

The purpose of the usability study was to investigate the potential of VUI devices as an option for patients to access the *Medly* program. Patients with HF were recruited and asked to explore the intuitiveness of the *Medly* voice app. The usability study consisted of 3 main components, which each patient completed: a usability session, the standardized System Usability Scale (SUS) questionnaire ([Multimedia Appendix 2](#)), and a qualitative semistructured interview session. The SUS questionnaire, developed in 1986, is a validated and reliable tool created to measure usability. The survey consists of 10 items and evaluates various products and services; it is commonly used in health care when implementing newer technologies [32]. The findings from this study helped identify the acceptability and ease of use of the voice app and also informed the service design that would be most appropriate to be implemented for a voice app offering a self-management program.

Participant Recruitment

A total of 8 participants were recruited for this study, and all of them used *Medly* as part of their standard of care. All the participants recruited were adults living with chronic HF and

could read, write, and understand English. Patients using *Medley* were asked to participate so that unique insights and themes could be identified, as they already had experience with using the program. The participants were recruited through the *Medley* nurse coordinator, who provided a brief overview of the study to patients and then asked whether they were interested in participating before introducing them to the study coordinator. Most usability studies require 5 to 8 participants to receive response saturation [33,34]; therefore, we chose to recruit 8 participants to help identify the potential this platform may have as a health care delivery technology in the future, by identifying major themes emerging from the study.

Usability Study Design

Participants who volunteered attended a single, 1-hour session with the study coordinator after their scheduled appointment and were asked to complete four main tasks: (1) inputting measurements on the *Medley* smartphone app, (2) inputting measurements on the *Medley* voice app, (3) completing the SUS questionnaire, and (4) engaging in a semistructured interview with the study coordinator. Following this, the participants were then asked to verbally interact with the *Medley* voice app, with the help of an instruction card, using Amazon Alexa, which was also provided by the study coordinator. Observational notes were recorded throughout the session. The participants then completed a standardized, poststudy SUS questionnaire regarding their experience with the voice app. Once the participants completed the questionnaire, the study coordinator led individual semistructured interviews, which included questions pertaining to their perceptions of the usability of the *Medley* voice app and specific issues or areas for improvement. The results of this research highlighted key findings regarding the acceptability of integrating VUI platforms into patient self-care tools and informed decisions for future design iterations.

Data Collection and Analysis

Both quantitative and qualitative data were gathered during the usability study and triangulated when identifying the main conclusions from the results. The SUS questionnaire was analyzed as per standard protocol [32], and the results are graphically represented in the “Usability Study Findings” section below. The data were interpreted in conjunction with the qualitative interview findings and observational notes. An inductive, qualitative descriptive approach was used to analyze the interview responses and field notes; all transcripts were analyzed and coded by the study coordinator (AB), using Microsoft Word (Microsoft Corp). Because the purpose of the usability study was to collect user insights, it was not designed to be statistically powered; therefore, specific patient demographic characteristics such as age, ethnicity, etc were not collected or used in the analysis.

The mixed methods triangulation design followed a convergence model, with a larger emphasis being placed on the qualitative results. The findings from the SUS questionnaire were used to validate the main conclusions gathered from the interviews and

field notes. The rationale for this approach was due to the usability study design; because a small sample size was needed to reach saturation, the interpretations from the quantitative data were used to support the qualitative findings.

Results

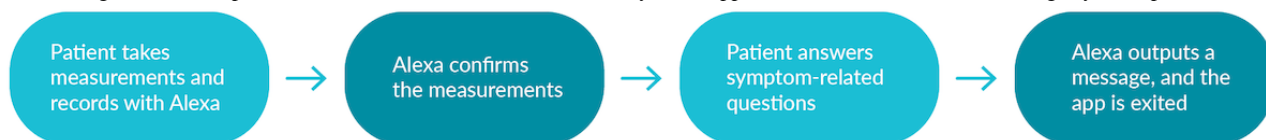
Preliminary Design of the Medley Voice App

A voice app prototype of *Medley* was created for the purpose of being used during a usability study to identify problems in the design, uncover any opportunities for improvement, and learn more about the users’ behaviors and preferences when interacting with the app. An instruction card was created and used to help guide the user through their interaction with the *Medley* voice app (Figure S2 in [Multimedia Appendix 1](#)). A preliminary design of the card was used by the participants in the study so that feedback could be collected about the design to make changes, if required.

A high-level overview of the conversation was first mapped out and created (Figure 2). Using this, the full conversation was designed and is described in more detail in Figure S3 in [Multimedia Appendix 1](#). The first half of the voice app interaction consisted of the user measuring and recording their weight, blood pressure, and heart rate with nonlyrical music playing in the background in between prompts. Playing music in the background not only created a more pleasant user experience but was also a strategy used to ensure that the app did not time out [35]. After the user inputted their heart rate, the voice app reiterated the measurements it captured to the user and gave them an opportunity to correct any wrong measurements. Once the user recorded their measurements, the app began to ask them a series of yes or no questions regarding HF-related symptoms. Once this was completed, the app outputted a message and exited. The *Medley* algorithm was not implemented for the prototype because it was not needed to accomplish the purpose of the usability study.

An interaction model was then created to implement the conversational flow described earlier, and it consisted of intents, training phrases, and slot types. An intent was created for each piece of required information: weight, blood pressure, heart rate, and symptom-related questions. Within each intent, training phrases were added and consisted of predictable utterances the user may say to record their responses. Within each training phrase, the most important pieces of data were identified and labeled as a slot type (weight, blood pressure, and heart rate were categorized as *numbers*, and symptom responses were labeled as “yes” or “no”). The slot types were used to help extract the measurements needed to be sent to the *Medley* voice server. The measurements were extracted from the interaction model, and the app logic was used to help direct the flow of conversation. Once the required measurements were captured, a POST request was made to the *Medley* server, and a self-care message was sent back to the voice app to be relayed to the user.

Figure 2. A high-level description of the conversational flow for the Medly voice app. “Milestones” are outlined and signify an important event.



Usability Study Findings

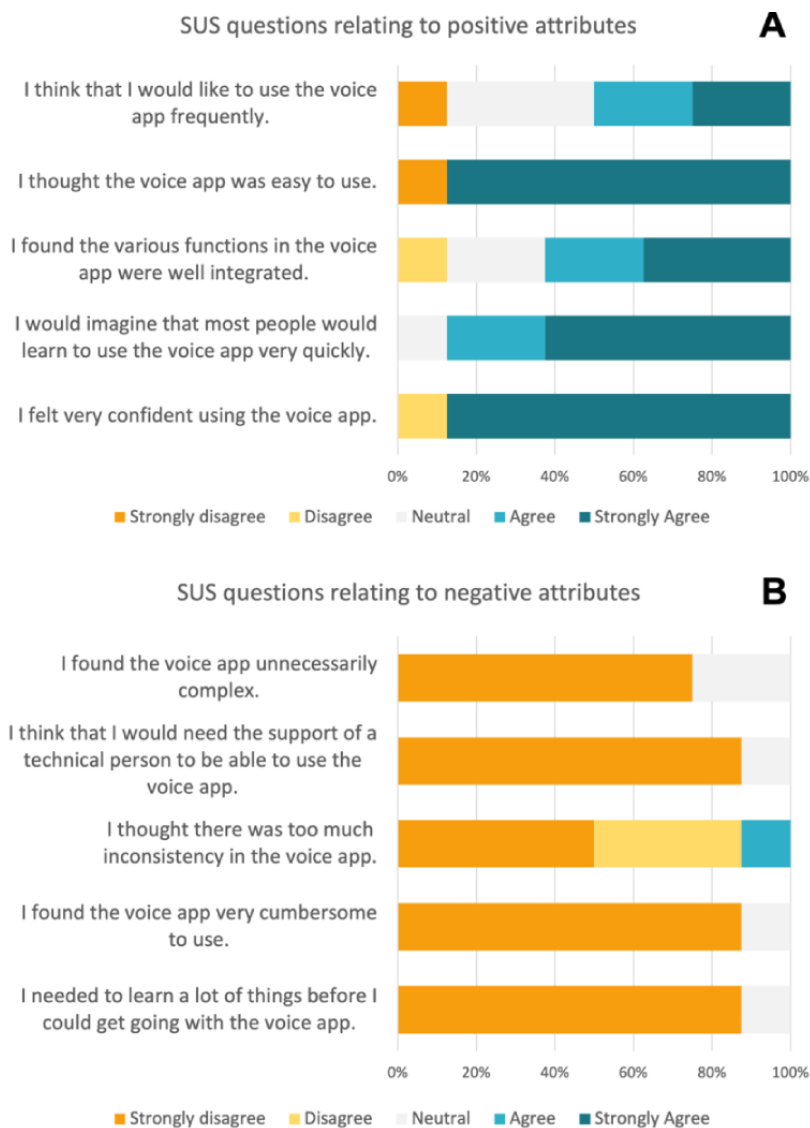
SUS Results

A usability study was performed with 8 participants with HF from UHN’s Ted Rogers Center of Excellence for Heart Failure clinic. Saturation was successfully achieved with the sample size because of the common themes that emerged across the 8 sessions. After interacting with the voice app, the participants filled out the SUS questionnaire to describe their experience. The average SUS score was 92, out of 100 (SD 5.2), ranking the voice app in the 98th percentile based on previous studies. According to these data, most participants were satisfied with

the design and development of the voice app. Results for the positive and negative statements of the SUS questionnaire are shown in Figures 3A and 3B, respectively.

Looking at the positive attribute statements, almost all participants (7/8, 88%) felt confident using the app and thought it was easy to use. This was further supported by the responses to the negative attribute statements, with the most popular response being “strongly disagree.” However, when asked whether they would use the voice app frequently, there was a greater divide in opinion, with the most popular response (3/8, 38%) being neutral.

Figure 3. (A) Data showcasing the participant results for the questions in the System Usability Scale (SUS) relating to positive attributes of the Medly voice app system. (B) Data showcasing the participant results for the questions in the SUS relating to the negative attributes of the Medly voice app system.



SUS Triangulation With Qualitative Data

Overall, these findings showed promise when determining user acceptability because of the overall positive responses from the SUS questionnaire. Triangulating these results with the field notes and interview findings indicated areas for improvement regarding the ease of use. Although most of the responses were “strongly agree” or “strongly disagree,” 3 of the 10 statements had more variation: “I think that I would like to use the voice app frequently,” “I found the various functions in the voice app were well integrated,” and “I thought there was too much inconsistency in the voice app.” This can potentially be explained using 2 observations that were commonly seen during the user sessions. First, the voice app had difficulties understanding measurements when users said each digit separately for their weight, blood pressure, and heart rate (Table S1 in [Multimedia Appendix 3](#)). Second, the voice app would often time out whenever there was a small pause in between the prompt provided and the user’s response. In this scenario, it became difficult for the user to converse with the voice app. Given these observations (and their validation through questionnaire responses), some changes have been made to improve the patient experience.

Qualitative Findings

The observational notes and semistructured interview responses were analyzed to (1) identify and describe common themes and (2) create a list of potential errors that could occur while the user is interacting with the app. Both these deliverables provided insights and helped shape the redesign and redevelopment of the voice app. Six main themes were identified: (1) changes in physical behavior, (2) preference between voice app and smartphone, (3) importance of music during voice app interaction, (4) lack of privacy concerns, (5) desired reassurances during voice app interaction, and (6) helpful aids during voice app interaction.

One of the more notable themes was the observed change in behavior when the participants interacted with the voice app; most participants showed signs of nervousness and looked tense. However, as the conversation progressed, there was a noticeable difference in their demeanor, as they became more relaxed once they better understood the conversational flow. Some voice app features helped the users navigate through their session, making for a more pleasant experience, such as having music play in the background so that they did not feel rushed through the interaction and having the voice app reiterate measurements back to them to ensure that the device correctly recorded the values. Because the users were accustomed to seeing visual feedback of their measurements on their smartphones, they preferred having some type of feedback (verbal or visual) to ensure that the voice app heard them correctly. The participants heavily relied on the instruction card as an aid, and those who followed it closely were the most successful in completing the interaction. The use of music, reiteration feature, and instruction card was observed to be helpful during the interactions, and this was confirmed with the participants during the interview sessions. Insights were also provided as to why some patients found the smartphone device difficult to use, and the responses typically revolved around the idea of convenience. Some

participants (4/8, 50%) found that inputting their data through a conversation (voice app) was easier for them than using their smartphone device. One of the participants said that they use their smartphone only for the Medly app and otherwise have it turned off. Finally, concerns about privacy were brought up only by 25% (2/8) of participants, indicating it may not be a prominent barrier to use as originally predicted. Further details describing these themes are outlined in Table S2 in [Multimedia Appendix 3](#).

As mentioned previously, a list of potential risks, related to the interaction of the user with the voice app, was identified. The severity of each risk was assessed based on its likelihood and consequence. Mitigation strategies for each risk were developed and implemented in the voice app design during the redevelopment work. These data can be seen in Table S1 in [Multimedia Appendix 3](#).

Redevelopment of the Medly Voice App

Overview

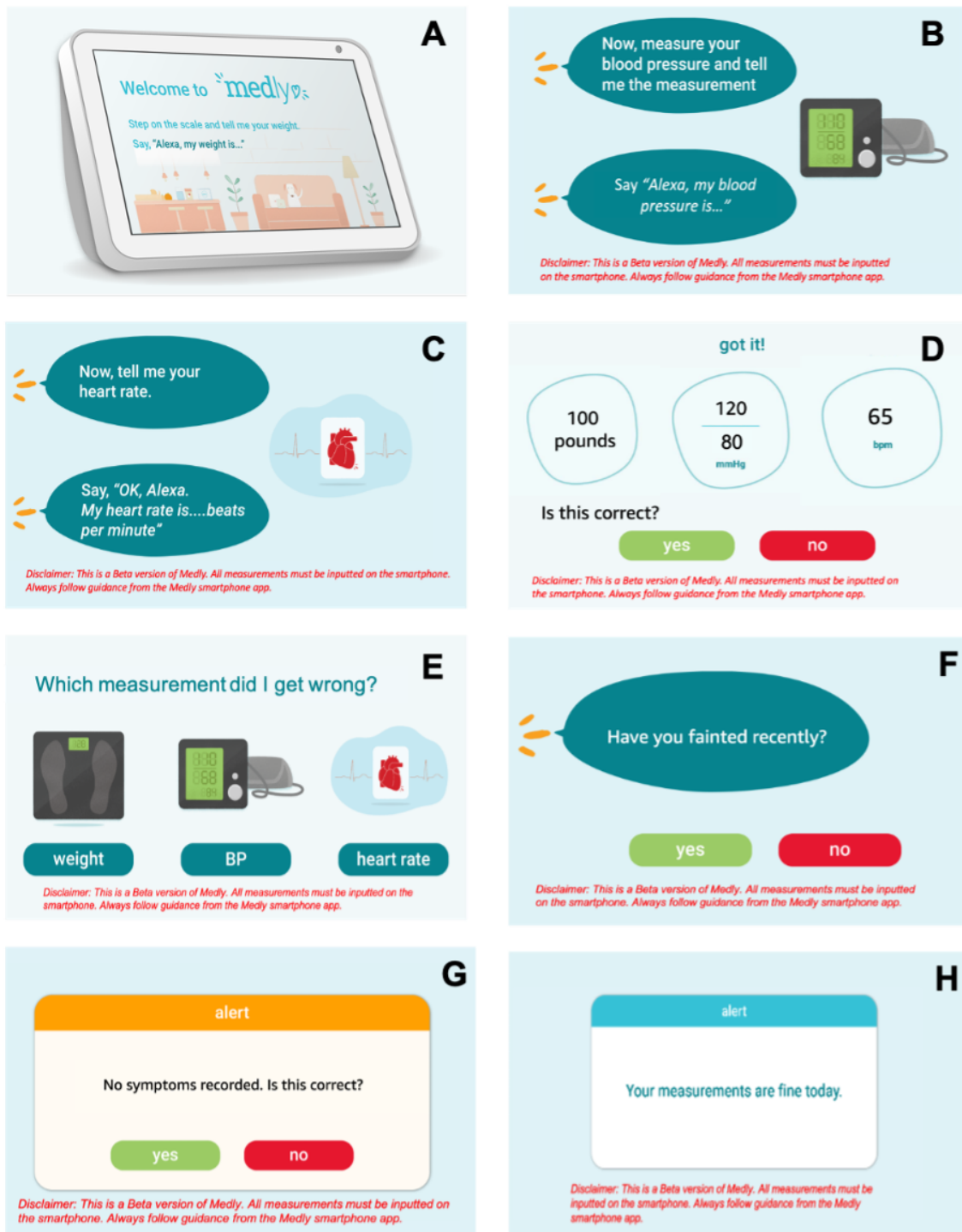
The prototype built for the usability study was further developed to meet all the design requirements of the *Medly* voice app. The VUI device was also changed to incorporate screen and touch screen capabilities; this decision was based on the findings from the usability study. We expected that the touch screen will not be suitable for all patients; however, we were also led to believe (based on user feedback) that some may prefer it over the smartphone and screenless VUI device, especially for visual feedback.

Screen Design

Having a screen display on the device meant that the users would be able to visually see instructions and measurements and also have the option of responding to yes or no questions using the touch screen. Although some demographics struggle to use touch screen, these interaction features are optional for the user. Although we predicted that the visual feedback will be helpful for most users and that having a bigger screen will be easier than the smartphone, the users will still be able to interact with the app using only their voice if they find it to be the easiest and simplest mode of interaction for them. Each screen was designed to only include the information the user needed to avoid any confusion. A consistent design theme, similar to the one used in the *Medly* smartphone app, was used for each screen display using the same color scheme, layout, and font size and style. The written cues on the screen aligned with the verbal prompts to avoid user confusion. Design guidelines were used for font size and layout to ensure that the prompts are presented at an appropriate reading size [35]. The screen design at various points of the conversation is shown in [Figure 4](#).

The incorporation of a screen display for the *Medly* voice app also allows users to see their previous measurements and data trends, similar to what they were used to seeing when using the smartphone app. Although the implementation of seeing data history was not in scope for this project, it is a feature that is common in most mHealth apps, as it helps users understand their health status better.

Figure 4. Screen designs showcased through images A-H, representing different events that occur in the conversation.

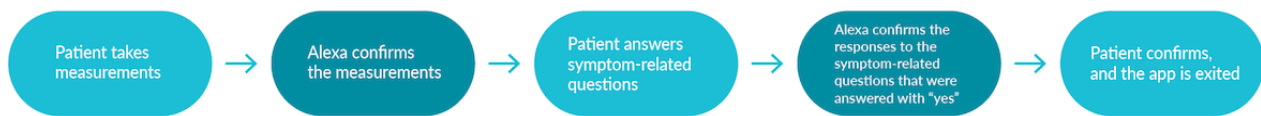


Conversational Design

The conversational script used in the preliminary design stayed the same during the redevelopment work. Findings from the usability study helped inform which parts of the conversation were most prone to error and needed to be changed. More training phrases and symptom confirmations were new additions to the conversation, and the summary of the interaction is shown in Figure 5. A more detailed description of the new additions to the conversation can be seen in Figure S4 in Multimedia

Appendix 1. With these new changes, the app only confirms the symptoms the user answered yes to for a more efficient process by giving them an opportunity to correct any wrong responses recorded. Once all the responses have been corrected, the data are sent to the *Medly* voice server. The voice server retrieves patient thresholds, sends the daily data to the *Medly* research server, and uses the *Medly* algorithm to send an output message back to the voice app, informing the patient on the status of their health.

Figure 5. A high-level description of the conversational flow for the final version of the Medly voice app. A new milestone has been added to describe the Medly voice app, which confirms the responses to the symptom-related questions that were answered with “yes.”.



Backend Development

The software architecture and *Medly* algorithm were implemented during the redevelopment phase of the project. Each output message generated to the user was a personalized response based on their measurements and preexisting thresholds. To receive data from the voice app and pull relevant patient thresholds, account linking was required to be set up to connect the patient’s identity from 2 different services. Personalized *Medly* accounts were created for each study participant and connected to their respective Amazon account. The *Medly* algorithm was built and deployed on the *Medly* voice server; it pulls data from the *Medly* app server as well as data sent from the VUI to generate an output message.

Discussion

Principal Findings

This manuscript describes the UCD approach that was used to develop a voice app experience for HF management. Accessing chronic disease management programs through voice apps has the potential to increase the uptake of these programs by making them more accessible, thereby reaching more patients. The evaluation of the voice app demonstrated user acceptability through the SUS questionnaire (92/100 average score), and the qualitative findings provided assurance of deployment feasibility, as the users were able to successfully interact with the voice app. Common themes from the usability sessions (preference for music playing during interaction, providing oral or visual feedback, and the instruction card being a helpful aid) served as feedback for the redevelopment phase of the project.

The impact of the findings from this study is promising, especially given the near-ubiquitous nature of smart speakers, such as Amazon Alexa and Google Home. These devices can be easily integrated within households, and they make it easy and convenient to perform tasks through simple conversations and offer flexibility to do multiple tasks simultaneously. Patients with chronic illness experience a constant need to record and transmit data to better track their health, and voice apps, deployed through smart speakers, can help them accomplish this seamlessly, not only because it uses speech to collect these data but also because it can guide the user through their tasks more explicitly than the smartphone.

As this platform requires less technical knowledge than the smartphone (mHealth apps), those who struggle to use technology have the opportunity to participate in the program. According to the usability study observations, most participants felt that it was easy to interact with the voice app because it only required conversation and no technical background. The platform’s multilingual capability offers another opportunity, which may make some more individuals willing to use it to record measurements. VUIs also allow users to initiate a call

to anyone on their list through voice command; this can be useful in cases where they need assistance but cannot reach their phone to call for help. This scenario often occurs with older adults, who are also the most common group to be diagnosed with chronic illnesses and, as a result, are prescribed digital therapeutics.

It is also important to consider the implications of using voice apps for chronic disease self-management. When inputting measurements on an mHealth app, it is typical for the user to see their data on the screen before submitting. With a VUI that solely relies on verbal communication, visual feedback is no longer a possibility. The participants commented on the importance of knowing which measurements were recorded on the voice app and valued the ability to see historical data and trends on the smartphone app. With these findings, we incorporated verbal feedback as a necessary component in the conversational design to help with user confidence. Adding this step in the conversation comes at the cost of prolonging the time it takes to complete the interaction but is not necessarily a drawback because of the lesser focus required when compared with using a smartphone.

Although users only need to use speech when interacting with VUIs, the conversation style may vary significantly when compared with human-to-human interaction. During this study, the users who had never used these types of devices before felt the need to raise their voice in hopes that it heard and understood them better. Expelling this energy repeatedly may make the user feel tired, as one of the participants noted, especially because the demographic that uses these digital therapeutics are older adults. Without knowing how the voice app will respond, a difference in body language was also observed, namely users tensed up, felt the need to sit up straighter, and focused more when interacting with the device. This behavior is in stark contrast to how users interacted with the smartphone, as their physical body language while using the smartphone indicated a much more relaxed behavior. Despite these differences, it is expected that user body language will shift toward a more relaxed behavior as they become more comfortable with using voice apps.

Finally, it is also important to consider the potential barriers to entry that exist as a result of using this technology. For example, most VUI devices require the user to have an account to interact with the device, creating an accessibility issue. Another barrier to entry is the requirement to have a constant, reliable internet connection to use the device. This is in contrast to the *Medly* smartphone app which can be accessed offline and is mobile. Not only can this create an equity issue for those who do not have internet, but it can also be problematic when internet issues and power outages occur.

Comparison With Prior Work

To our knowledge, this study is one of the first to investigate the acceptability and feasibility of a personalized chronic disease self-management voice app, designed to be used by patients with HF in their homes. Our work is similar to other recent studies that have used either an Amazon Alexa or a Google Home device to deploy a self-management voice app for various chronic diseases, including type 2 diabetes [23]. The usability study provided by Cheng et al [23] collected quantitative data through a survey, asking participants whether they preferred a smartphone or voice app; similar to the findings in our study, 8 out of 10 participants preferred the voice app. In our study, qualitative data were used to further explore the high acceptability response. Voice apps have also been developed for heart disease but for purposes other than self-management [21,22]; Jadczyk et al [20] developed one for patients to interact with during in-clinic registration to streamline the process. The usability study performed was in a controlled laboratory environment but revealed 97.5% accuracy when converting speech to text. In summary, our findings are consistent with the existing literature, which provide preliminary assessments of voice apps, an emerging technology in health care. Using a systematic approach, we were able to gather more insights about the feasibility and acceptability of voice apps through quantitative and qualitative data collection.

Limitations

Although the results from this research project showed promise for future use, it is important to acknowledge the limitations associated with this work.

The main limitation of usability studies is the short interaction time. Although the observations and feedback from the participants were helpful in identifying the potential that voice apps may have for chronic disease management, they only interacted with the technology for a short period. Future work should include a study that requires the users to interact with the device for a longer period to give them more experience with the voice app to reflect more grounded feedback.

A sample size of 8 participants was chosen because the usability study was only focused on gathering user insights and not statistically relevant data. According to the literature, the total participants needed for a usability study is dependent on when

saturation is achieved in results. In most cases, this occurs when the sample size is typically between 5 and 8 participants [36]. Therefore, it is recommended that multiple usability tests occur with fewer users (between 5 and 8 participants) and that changes be made between usability tests to mitigate challenges observed in previous sessions. The research showcased in this paper underwent 1.5 cycles of the UCD process and requires further testing to identify whether the challenges uncovered in the usability study have been resolved with the latest iteration of design and development work. Owing to the small sample size, the usability study was also not statistically powered. In this case, the evaluation of the results is mostly limited to qualitative data analysis. Although a questionnaire was completed by each user after the session, these results were only used to support the conclusions drawn from the qualitative data.

Finally, the participants may have had social desirability bias when answering interview questions and filling out the questionnaire, especially because they knew that the study coordinator also had involvement in the design and development of the voice app [37]. When things went wrong, the participants most often felt as though it was their fault for not understanding how to use the technology and, as a result, would not bring up comments regarding what they disliked about the experience.

Conclusions

This project involved the design, development, and usability evaluation of a voice app for HF management. A UCD process was followed to systematically create a voice app that would meet user needs and be easy to use. The usability study performed at the PMCC at UHN provided insightful user feedback about the voice app's design, with the overall response being positive with high user satisfaction. The findings from this usability study impacted the redevelopment of the voice app, which will be used in a future clinical evaluation. As advancements in VUIs progress, we believe they will play an integral role in providing access to chronic disease management programs by (1) helping more patients complete their tasks independently, (2) offering a more convenient experience to record relevant data, and (3) allowing those with limited technical and English skills access to these programs. The findings from this research show promise in using VUIs to help with chronic disease management.

Acknowledgments

The authors wish to thank Jacqueline Simpson for designing the *Medly* voice app screensavers and figures presented in this paper. The authors are also grateful to all the patients who gave their time and participated in this study.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to an agreement with the Research Ethics Board of the University Health Network, which describes very specific circumstances that prohibits us from sharing the data on a public platform; however, the data are available from the corresponding author on reasonable request.

Conflicts of Interest

JAC and HR are part of the team that founded the *Medly* system under the intellectual property policies of the University Health Network and may benefit from future commercialization of this technology.

Multimedia Appendix 1

Design requirements (including both user and software) that were considered during the design stage (Table S1). Software architecture diagram of the Medly voice app system (Figure S1). Instruction card provided to users to help them navigate the conversation (Figure S2). An example of the “script” describing the conversation for the Medly voice app (Figure S3). A “script” describing the conversation that will occur if the Medly voice app did not correctly capture the user’s symptom responses (Figure S4).

[[DOCX File , 955 KB - formative_v6i12e41628_app1.docx](#)]

Multimedia Appendix 2

The System Usability Scale questionnaire that was handed out to the participants at the end of the usability session.

[[DOCX File , 16 KB - formative_v6i12e41628_app2.docx](#)]

Multimedia Appendix 3

List of potential errors identified from usability study results, along with risk levels and mitigation strategies (Table S1). Main themes derived from the usability sessions (Table S2).

[[DOCX File , 16 KB - formative_v6i12e41628_app3.docx](#)]

References

1. Savarese G, Lund LH. Global public health burden of heart failure. *Card Fail Rev* 2017 Apr;3(1):7-11 [[FREE Full text](#)] [doi: [10.15420/cfr.2016:25:2](#)] [Medline: [28785469](#)]
2. What is Heart Failure? American Heart Association. URL: <https://www.heart.org/en/health-topics/heart-failure/what-is-heart-failure> [accessed 2022-02-17]
3. Lippi G, Sanchis-Gomar F. Global epidemiology and future trends of heart failure. *AME Med J* 2020 Jun;5:15. [doi: [10.21037/amj.2020.03.03](#)]
4. McHorney CA, Mansukhani SG, Anatchkova M, Taylor N, Wirtz HS, Abbasi S, et al. The impact of heart failure on patients and caregivers: a qualitative study. *PLoS One* 2021 Mar 11;16(3):e0248240 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0248240](#)] [Medline: [33705486](#)]
5. Megari K. Quality of life in chronic disease patients. *Health Psychol Res* 2013 Sep 24;1(3):e27 [[FREE Full text](#)] [doi: [10.4081/hpr.2013.e27](#)] [Medline: [26973912](#)]
6. Seto E, Leonard KJ, Cafazzo JA, Masino C, Barnsley J, Ross HJ. Self-care and quality of life of heart failure patients at a multidisciplinary heart function clinic. *J Cardiovasc Nurs* 2011;26(5):377-385. [doi: [10.1097/JCN.0b013e31820612b8](#)] [Medline: [21263339](#)]
7. Chen X, Koskela M, Hyvääkä J. Image based information access for mobile phones. In: Proceedings of the 2010 International Workshop on Content Based Multimedia Indexing. 2010 Presented at: CBMI '10; June 23-25, 2010; Grenoble, France p. 130-135. [doi: [10.1109/cbmi.2010.5529886](#)]
8. Irfan Khan A, Gill A, Cott C, Hans PK, Steele Gray C. mHealth tools for the self-management of patients with multimorbidity in primary care settings: pilot study to explore user experience. *JMIR Mhealth Uhealth* 2018 Aug 28;6(8):e171 [[FREE Full text](#)] [doi: [10.2196/mhealth.8593](#)] [Medline: [30154073](#)]
9. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digit Health* 2019 Oct 10;5:2055207619882181 [[FREE Full text](#)] [doi: [10.1177/2055207619882181](#)] [Medline: [31656632](#)]
10. Ware P, Ross HJ, Cafazzo JA, Boodoo C, Munnery M, Seto E. Outcomes of a heart failure telemonitoring program implemented as the standard of care in an outpatient heart function clinic: pretest-posttest pragmatic study. *J Med Internet Res* 2020 Feb 08;22(2):e16538 [[FREE Full text](#)] [doi: [10.2196/16538](#)] [Medline: [32027309](#)]
11. Medly - Integrating heart failure care from hospital to home. 2018 Jun 13. URL: <https://medly.ca/> [accessed 2022-07-27]
12. Ware P, Dorai M, Ross HJ, Cafazzo JA, Laporte A, Boodoo C, et al. Patient adherence to a mobile phone-based heart failure telemonitoring program: a longitudinal mixed-methods study. *JMIR Mhealth Uhealth* 2019 Feb 26;7(2):e13259 [[FREE Full text](#)] [doi: [10.2196/13259](#)] [Medline: [30806625](#)]
13. Elboim-Gabyzon M, Weiss PL, Danial-Saad A. Effect of age on the touchscreen manipulation ability of community-dwelling adults. *Int J Environ Res Public Health* 2021 Feb 21;18(4):2094 [[FREE Full text](#)] [doi: [10.3390/ijerph18042094](#)] [Medline: [33669980](#)]
14. Naumann AB, Wechsung I, Hurtienne J. Multimodal interaction: a suitable strategy for including older users? *Interact Comput* 2010 Nov;22(6):465-474. [doi: [10.1016/j.intcom.2010.08.005](#)]
15. Nunes F, Silva PA, Cevada J, Correia Barros A, Teixeira L. User interface design guidelines for smartphone applications for people with Parkinson’s disease. *Univ Access Inf Soc* 2016;15(4):659-679. [doi: [10.1007/s10209-015-0440-1](#)]
16. -. Part 6: Rehabilitation and assistive robotics. *Technol Disabil* 2019 Jul 17;31(s1):147-154. [doi: [10.3233/tad-190009](#)]
17. Williams P, Shekhar S. People with learning disabilities and smartphones: testing the usability of a touch-screen interface. *Educ Sci* 2019 Oct 27;9(4):263. [doi: [10.3390/educsci9040263](#)]

18. Patrick PA, Obermeyer I, Xenakis J, Crocitto D, O'Hara DM. Technology and social media use by adult patients with intellectual and/or developmental disabilities. *Disabil Health J* 2020 Jan;13(1):100840. [doi: [10.1016/j.dhjo.2019.100840](https://doi.org/10.1016/j.dhjo.2019.100840)] [Medline: [31582357](https://pubmed.ncbi.nlm.nih.gov/31582357/)]
19. Kravitz N, Levanon Y, Cukierman-Yaffe T, Nota A, Kizony R, Rand D. Sensorimotor and cognitive abilities associated with touchscreen tablet app performance to support self-management of type 2 diabetes. *Am J Occup Ther* 2021;75(1):7501205080p1-7501205080p9. [doi: [10.5014/ajot.2021.040600](https://doi.org/10.5014/ajot.2021.040600)] [Medline: [33399056](https://pubmed.ncbi.nlm.nih.gov/33399056/)]
20. Jadczyk T, Kiwic O, Khandwalla RM, Grabowski K, Rudawski S, Magaczewski P, et al. Feasibility of a voice-enabled automated platform for medical data collection: CardioCube. *Int J Med Inform* 2019 Sep;129:388-393. [doi: [10.1016/j.ijmedinf.2019.07.001](https://doi.org/10.1016/j.ijmedinf.2019.07.001)] [Medline: [31445282](https://pubmed.ncbi.nlm.nih.gov/31445282/)]
21. Dojchinovski D, Ilievski A, Gusev M. Interactive home healthcare system with integrated voice assistant. In: Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics. 2019 Presented at: MIPRO '19; May 20-24, 2019; Opatija, Croatia p. 284-288. [doi: [10.23919/mipro.2019.8756983](https://doi.org/10.23919/mipro.2019.8756983)]
22. Apergi LA, Bjarnadottir MV, Baras JS, Golden BL, Anderson KM, Chou J, et al. Voice interface technology adoption by patients with heart failure: pilot comparison study. *JMIR Mhealth Uhealth* 2021 Apr 01;9(4):e24646 [FREE Full text] [doi: [10.2196/24646](https://doi.org/10.2196/24646)] [Medline: [33792556](https://pubmed.ncbi.nlm.nih.gov/33792556/)]
23. Cheng A, Raghavaraju V, Kanugo J, Handrianto YP, Shang Y. Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In: Proceedings of the 15th IEEE Annual Consumer Communications & Networking Conference. 2018 Presented at: CCNC '19; January 12-15, 2018; Las Vegas, NV, USA p. 1-5. [doi: [10.1109/ccnc.2018.8319283](https://doi.org/10.1109/ccnc.2018.8319283)]
24. Sezgin E, Noritz G, Elek A, Conkol K, Rust S, Bailey M, et al. Capturing at-home health and care information for children with medical complexity using voice interactive technologies: multi-stakeholder viewpoint. *J Med Internet Res* 2020 Feb 13;22(2):e14202 [FREE Full text] [doi: [10.2196/14202](https://doi.org/10.2196/14202)] [Medline: [32053114](https://pubmed.ncbi.nlm.nih.gov/32053114/)]
25. Maharjan B, Li J, Kong J, Tao C. Alexa, what should I eat? : A personalized virtual nutrition coach for native American diabetes patients using Amazon's smart speaker technology. 2019 IEEE International Conference on E-health Networking, Application & Services (HealthCom) 2019:1-6. [doi: [10.1109/HealthCom46333.2019.9009613](https://doi.org/10.1109/HealthCom46333.2019.9009613)]
26. Sezgin E, Militello LK, Huang Y, Lin S. A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors. *Transl Behav Med* 2020 Aug 07;10(3):606-628. [doi: [10.1093/tbm/ibz141](https://doi.org/10.1093/tbm/ibz141)] [Medline: [32766865](https://pubmed.ncbi.nlm.nih.gov/32766865/)]
27. What is the Alexa Skills Kit? alexa developer documentation. URL: <https://developer.amazon.com/en-US/docs/alexa/ask-overviews/what-is-the-alexa-skills-kit.html> [accessed 2021-06-28]
28. Murad C, Munteanu C, Clark L, Cowan BR. Design guidelines for hands-free speech interaction. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. 2018 Presented at: MobileHCI '18; September 3-6, 2018; Barcelona, Spain p. 260-276. [doi: [10.1145/3236112.3236149](https://doi.org/10.1145/3236112.3236149)]
29. Build Your Design Artifact. alexa developer documentation. URL: <https://developer.amazon.com/en-US/docs/alexa/alexa-design/build-artifact.html> [accessed 2021-06-28]
30. Chakraborty P. Creating voice interaction flows. UX Collective. 2018 Dec 31. URL: <https://uxdesign.cc/ux-design-for-voice-interfaces-part-ii-3b0056020cd3> [accessed 2021-06-28]
31. Alexa Skills Kit. Alexa. URL: <https://developer.amazon.com/en-US/alexa/alexa-skills-kit> [accessed 2022-11-24]
32. System Usability Scale (SUS). usability.gov. URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> [accessed 2021-06-28]
33. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 May Presented at: CHI '93; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213. [doi: [10.1145/169059.169166](https://doi.org/10.1145/169059.169166)]
34. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? *Hum Factors* 1992 Aug;34(4):457-468. [doi: [10.1177/001872089203400407](https://doi.org/10.1177/001872089203400407)]
35. Alexa Design System for Alexa Presentation Language (APL). alexa developer language. URL: <https://developer.amazon.com/en-US/docs/alexa/alexa-presentation-language/apl-alexa-packages-overview.html> [accessed 2021-06-28]
36. Nielsen J. Why You Only Need to Test with 5 Users. Nielsen Norman Group. 2000 Mar 18. URL: <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users> [accessed 2021-06-28]
37. Grimm P. Social desirability bias. In: Sheth J, Malhotra NK, editors. Wiley International Encyclopedia of Marketing. Hoboken, NJ, USA: Wiley Online Library; 2010.

Abbreviations

- HF:** heart failure
- mHealth:** mobile health
- PMCC:** Peter Munk Cardiac Centre
- SUS:** System Usability Scale
- UCD:** user-centered design

UHN: University Health Network
VUI: voice user interface

Edited by A Mavragani; submitted 02.08.22; peer-reviewed by E Sezgin, B Nørgaard; comments to author 09.09.22; revised version received 26.09.22; accepted 12.10.22; published 06.12.22.

Please cite as:

Barbaric A, Munteanu C, Ross H, Cafazzo JA

Design of a Patient Voice App Experience for Heart Failure Management: Usability Study

JMIR Form Res 2022;6(12):e41628

URL: <https://formative.jmir.org/2022/12/e41628>

doi: [10.2196/41628](https://doi.org/10.2196/41628)

PMID: [36472895](https://pubmed.ncbi.nlm.nih.gov/36472895/)

©Antonia Barbaric, Cosmin Munteanu, Heather Ross, Joseph A Cafazzo. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Certified Peer Support Specialists Training in Technology and Delivery of Digital Peer Support Services: Cross-sectional Study

Karen Fortuna¹, PhD, LICSW; Julia Hill²; Samantha Chalker³, PhD; Joelle Ferron¹, PhD

¹Geisel School of Medicine, Department of Psychiatry, Dartmouth College, Concord, NH, United States

²Dartmouth College, Hanover, NH, United States

³Veterans Affairs San Diego Healthcare System, San Diego, CA, United States

Corresponding Author:

Karen Fortuna, PhD, LICSW

Geisel School of Medicine

Department of Psychiatry

Dartmouth College

2 Pillsbury Street

Suite 401

Concord, NH, 03301

United States

Phone: 1 603 722 5727

Email: karen.l.fortuna@dartmouth.edu

Abstract

Background: When the COVID-19 pandemic lockdown measures were instituted, the wide-scale necessity for remote mental health care increased among professional clinicians, such as psychiatrists, psychologists, social workers, and certified peer support (CPS) specialists. Factors contributing to increased demand include concern for the safety of loved ones, the safety of oneself, overall well-being, unemployment, and loneliness for older individuals. While demand continues to increase and a shortage of mental health professionals persists, understanding the training, technology, media, and delivery of digital peer support services can facilitate community-based support services to assist patients in coping with mental health symptoms between clinical encounters with licensed professionals. Digital peer support consists of asynchronous and synchronous, live or automated, peer support services such as applications, social media, and phone calls.

Objective: The purpose of this cross-sectional study is to determine how digital peer support is delivered, by which technologies it is delivered, and how certified digital peer supporters are trained within the United States to inform future delivery of digital peer support.

Methods: We used an online cross-sectional self-report survey developed alongside certified peer specialists. The study included questions regarding the types of peer support training and the delivery methods used within their practices. We advertised the survey through a certified peer support specialist listserve, Facebook, and Twitter.

Results: Certified peer specialists provide mutual social emotional support to those with a similar mental health condition. Of certified peer specialists trained in CPS, the majority of CPS specialists were trained in peer support (418/426, 98.1%). Peer support specialists deliver services via telephone calls (182/293, 62.1%), via videoconference-based services (160/293, 54.6%), via SMS text messages (123/293, 42%), via smartphone apps (68/293, 23.2%), and via social media (65/293, 22.2%). Certified peer specialists deliver services through virtual reality (11/293, 3.8%) and through video games (6/293, 2%). Virtual reality and video games may represent emerging technologies to develop and deliver community-based support.

Conclusions: This study examined the modes of digital peer support intervention as well as the training and demographic background of peer supporters. Given the demand for mental health care, digital peer support emerges as one option to increase access. These results suggest that CPS specialists commonly use SMS text messaging, phone calls, and videoconferences to engage in peer support. Less frequently, they may use diverse modes such as apps, social media, and video games. It is important to consider the backgrounds of peer supporters and the mediums of communication to best accommodate areas where access to peer support is emerging. Larger longitudinal studies and a variety of experimental designs may be considered to understand the efficacy of digital interventions and digital peer support training to direct optimal care.

(JMIR Form Res 2022;6(12):e40065) doi:[10.2196/40065](https://doi.org/10.2196/40065)

KEYWORDS

digital peer support; mHealth; COVID-19; mental health; remote service; remote mental health; telehealth; peer support; psychological health

Introduction

Social distancing and lockdown measures due to the COVID-19 pandemic have transformed the delivery of mental health care from in-person to remote services supported through the use of technologies [1]. One study showed that, before the COVID-19 pandemic, 80% of mental health practitioners (ie, social workers, psychiatrists, and psychologists) did not offer remote services, while as of late March 2020, that number had declined dramatically to *only* 19% [1].

When the COVID-19 pandemic lockdown measures were instituted, a wide-scale necessity for remote mental health care emerged not only for professional clinicians, such as psychiatrists, psychologists, and social workers, but also for certified peer support (CPS) specialists [2]. CPS specialists (ie, individuals who provide mutual, social, and emotional support to those with a similar mental health condition and are trained and accredited to offer services like text-based advice or encouragement toward well-being in between clinical encounters with psychiatric treatment teams [3]) are mobilized to offer digital peer support services through technologies such as videoconferencing platforms and smartphone apps. Digital peer support has been shown to help patients in between clinical encounters and impact quality of life, functioning, and medical and mental health self-management skill development, and promote engagement with in-person and digital services [4]. Understanding the training and delivery of digital peer support services can facilitate community-based support services to assist patients in coping with mental health symptoms between clinical encounters with licensed professionals. In this cross-sectional study, we describe the training and delivery of digital peer support within the United States.

Methods

Development

A national online survey was developed with input from certified peer specialists: 2 CPS specialists, who were not authors of the study, and were selected from the Collaborative Design for Recovery and Health, a volunteer, virtual collaborative of service users, peer support specialists, caregivers, policy makers, and scientists with and without a lived experience engaging in community-based participatory research [5]. They reviewed and modified questions to ensure clarity and appropriateness for respondents. The requirements for authorship according to the New England Journal of Medicine are (1) substantial contributions to the conception and design or acquisition, analysis, or interpretation of data; (2) drafting of the article or critical revision for important intellectual content; (3) final approval of the version to be published; and (4) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved. The certified

peer specialists did not meet these requirements and were not interested in authorship.

Delivery

Qualtrics was used to deliver the online survey. We sampled participants through an online newsletter announcement through the last author's CPS specialist listserv. The newsletter was sent 3 times to individuals over the course of a month. The listserv includes 1500 CPS specialists. This link and scripted text announcing the study were also sent out via popular peer support Facebook media outlets and Twitter. The response rate is not known due to the unreported reach of social media. The survey was available online from September 2020 to November 2020. We included individuals who completed a state-accredited peer-support training program that resulted in certification, resided in the United States, and were older than 18 years. While there were 464 respondents to the survey, we included data from a total of 426 peer support specialists who had either fully responded to all survey questions or to the majority of questions; the 38 others with less than half to no responses were omitted from the analyses.

Ethical Considerations

This cross-sectional study was approved by the Dartmouth-Hitchcock Health institutional review board (#STUDY02000514). A consent form was provided to participants online. Typed consent was required to proceed to complete the survey. The study presents the reporting guidelines for cross-sectional studies according to the organization STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) [6].

Measures

The survey began with questions about sociodemographic information such as "What state in the United States do you represent? [drop down select]" and "What is your race/ethnicity? [select all that apply] White, Black or African American, Native American or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Hispanic." Given their role as CPS specialists, they were also asked, "If comfortable, what type of lived experience have you been impacted by [please select the primary]? Schizophrenia, Schizoaffective Disorder, Bipolar Disorder, Major Depression, Other Mental Health Concerns, Alcohol Use Disorder, Opioid Use Disorder, Other Substance Misuse Concerns, Physical Health Concerns, None of the above."

Next, they were asked about employment and training in peer support. For example, "Are you trained in offering peer support? [select all that apply] Yes, I am a peer support specialist (including certified and non-certified), Yes, I am a recovery coach, Yes, I am an older adult peer support specialist, Yes, I am a Veteran peer support specialist, No, Other." and "If you selected 'other,' please explain [open answer]." They also answered questions about the use of technology for the delivery of peer support, for example, "If you are offering digital peer support, how are you providing service to service users? Please

select all modes by which you are providing digital peer support: Smartphone application, Text messaging, Phone calls, Videoconference, Social media, Video game.” Last, we asked an open-ended question regarding how participants used technologies to deliver treatment, for example, “If you selected ‘other,’ please explain the other types of technology service users use to support their recovery: [open answer].” The online survey took approximately 10 minutes to complete.

Analysis

We used descriptive statistics (ie, frequencies and percentages) to assess the demographic characteristics of the sample. We used listwise deletion and description to address missing data. We conducted a chi-square analysis to examine the substantial association between the mode of therapy and the lived experience of mental illness. Analyses were conducted using SPSS 17 (IBM Corp).

Results

A total of 426 peer support specialists responded to the national online survey from 42 states as shown in [Table 1](#). Geographic location was distributed across the United States, with the largest proportions being from Tennessee (57/426, 12.3%) and from Virginia (64/426, 13.8%). Most offered digital peer support (295/426, 71.1%), and most were trained in peer support (418/426, 98.1%). Demographics are shown in [Table 2](#).

The technology used during digital peer support services is shown in [Table 3](#). Most CPS specialists’ services were delivered through phone calls (182/293, 62.1%), videoconference-based services like Zoom (160/293, 54.6%), or text (123/293, 42%). Peer support specialists also used smartphone apps (68/293, 23.2%), including apps such as Calm, 12 Step Tool Kit, NA

Meeting Guide, Affirmations, Celebrate Recovery, BoosterBuddy, Easyquit, Connections Companion app, DayCount, Daylio, Headspace, Mindfulness Coach, WRAP, and PeerTECH, and social media (65/293, 22.2%), including Facebook. Of the 120 smartphone apps mentioned, 4 (3.3%) targeted connection making, 5 (4.2%) targeted alcohol or cigarette use, and 6 (5%) targeted meditation and mindfulness. Some emerging digital peer support services were offered through virtual reality (11/293, 3.8%) and video games such as Words With Friends, Mah Jong, Sims, Fallout, and EVE Online (6/293, 2%).

For the section of [Table 2](#) labeled “type of organization as paid or volunteer PSS” some examples of other types of organizations include “I am both a provider of peer support and recovery coaching as well as Trauma-Informed Peer Support” or “Certified Community Health Worker with lived experience” or “NA.”

A chi-square test for goodness of fit showed a substantial relationship between CPS use of video games for their own recovery and primary lived experience of mental health ($\chi^2_{8}=24.4$; $P=.002$). The data indicate that the use of video games by CPS for peer support varies based on the primary type of lived experience. A chi-square test for goodness of fit showed a substantial relationship between CPS use of smartphone apps and primary lived experience of mental health ($\chi^2_{8}=16.4$; $P=.04$). Those with opioid use disorder and other substance misuse concerns used smartphone apps less than others with different primary lived experiences. Those with depression used smartphone apps more than others with different primary types of lived experience. The data indicate that the use of smartphone apps by CPS for peer support varies based on the primary type of lived experience.

Table 1. Proportion of respondents from each state of the United States.

State	Respondents, n (%)
Alabama	11 (2.4)
Alaska	1 (0.2)
Arizona	6 (1.3)
Arkansas	4 (0.9)
California	2 (0.4)
Colorado	3 (0.6)
Connecticut	2 (0.4)
Florida	5 (1.1)
Georgia	2 (0.4)
Hawaii	1 (0.2)
Idaho	1 (0.2)
Illinois	21 (4.5)
Indiana	38 (8.2)
Iowa	4 (0.9)
Kansas	1 (0.2)
Kentucky	3 (0.6)
Louisiana	4 (0.9)
Maryland	13 (2.8)
Massachusetts	32 (6.9)
Michigan	7 (1.5)
Minnesota	4 (0.9)
Missouri	2 (0.4)
Montana	1 (0.2)
Nebraska	14 (3.0)
Nevada	1 (0.2)
New Hampshire	10 (2.2)
New Jersey	2 (0.4)
New York	12 (2.6)
North Carolina	2 (0.4)
Ohio	9 (1.9)
Oregon	6 (1.3)
Pennsylvania	8 (1.7)
Rhode Island	1 (0.2)
South Carolina	7 (1.5)
Tennessee	57 (12.3)
Texas	7 (1.5)
Utah	43 (9.3)
Vermont	3 (0.6)
Virginia	64 (13.8)
Washington	2 (0.4)
West Virginia	8 (1.7)
Wisconsin	2 (0.4)

Table 2. Respondent characteristics.

Sociodemographic characteristics	Respondents, n (%)
Gender^a	
Male	83 (31.7)
Female	173 (66.0)
Nonbinary	5 (1.9)
Other	1 (0.4)
Age^b (years)	
20-26	9 (3.2)
27-49	144 (50.7)
50-64	108 (38.0)
≥65	23 (8.1)
Race or ethnicity	
White	344 (80.8)
Black or African American	53 (12.4)
American Indian or Alaska Native	3 (0.7)
Asian	4 (0.9)
More than one race and Hispanic	14 (3.3)
Hispanic only	8 (1.9)
Highest grade in school completed	
Some elementary schooling	2 (0.5)
Some high school	2 (0.5)
Completed high school or GED ^c	41 (9.6)
Some college	129 (30.3)
Completed college or technical school	29 (6.8)
Completed associate degree	46 (10.8)
Completed bachelor's degree	85 (20.0)
Some graduate school	35 (8.2)
Completed master's degree	50 (11.7)
Completed doctoral degree	7 (1.6)
Employment status	
Full-time	312 (73.2)
Part-time	61 (14.3)
Volunteer	20 (4.7)
Unemployed	15 (3.5)
Retired	12 (2.8)
Student	6 (1.4)
Employment as trained PSS^d	
No employment but trained in PSS (ie, unemployed, retired, student, volunteer)	32 (7.5)
Volunteer or paid PSS	386 (90.6)
Paid non-PSS	8 (1.9)
Type of organization as paid or volunteer PSS	
Peer-run organization	82 (19.2)

Sociodemographic characteristics	Respondents, n (%)
Hospital	20 (4.7)
Community mental health center	102 (23.9)
Research organization	4 (0.9)
Managed care organization	26 (5.4)
Veterans Administration	6 (1.4)
For-profit mental health center	9 (2.1)
Behavioral health home	9 (2.1)
Collaborative care model	14 (22.8)
Other	97 (22.8)
Trained in offering peer support^e	
PSS (including certified and non-certified)	347 (81.8)
Recovery coach	19 (4.5)
Older adult PSS	21 (5.0)
Veteran PSS	13 (3.1)
Other	14 (3.2)
No training	9 (2.1)
Primary type of lived experience impacted by^f	
Schizophrenia	11 (2.8)
Schizoaffective disorder	9 (2.3)
Bipolar disorder	81 (20.7)
Major depression	122 (31.1)
Other mental health concerns	53 (13.5)
Alcohol use disorder	44 (11.2)
Opioid use disorder	34 (8.7)
Other substance misuse concerns	35 (8.9)
Physical health concerns	3 (0.8)
Offers digital peer support^g	
Yes	295 (71.1)
No (still a peer specialist)	120 (28.9)

^aMissing 164 respondent answers.

^bMissing 142 respondent answers.

^cGED: General Educational Development test.

^dPSS: peer support specialist.

^eMissing 2 respondent answers.

^fMissing 34 respondent answers.

^gMissing 11 respondent answers.

Table 3. Digital peer support services technology modalities.

Modes used to provide trained digital peer support ^a	Respondents, n (%)
Smartphone apps	68 (23.2)
SMS text messaging	123 (42.0)
Phone calls	182 (62.1)
Videoconference	160 (54.6)
Social media	65 (22.2)
Video games	6 (2.0)
Virtual reality	11 (3.8)
Other	29 (9.9)

^aMissing 133 respondent answers.

Discussion

Principal Findings

This web-based, cross-sectional, self-reported survey examined the training, technology, media, and delivery of digital peer support services. We found the majority of CPS specialists were trained in digital peer support, and the majority of services were delivered via telephone calls, videoconference-based services, and SMS text messages. In addition, certified peer specialists offered peer support through smartphones and social media, as well as virtual reality and video games. Virtual reality and video games may represent emerging technologies to develop and deliver digital support. The majority of CPS specialists use popular commercially available technologies (eg, Calm app) rather than mental health technologies developed by scientists and available through academic or medical institutions. Clinicians' promotion of digital peer support services may support time between clinical encounters.

Comparison With Prior Work

Prior work in the field of digital peer support has indicated the preliminary effectiveness of the use of digital peer support via telephone, videoconferencing, and tablet or smartphone apps. The most recent systematic review found 4 peer-reviewed articles presented on apps, 2 of which had in-person features augmented by smartphone apps, 3 on websites, 1 with social media, 2 on in-person augmented with SMS text messaging and fitness trackers, and 1 with a noninteractive psychosocial online program and email messaging [7]. However, the actual use of digital peer support technologies was not reported on in this review.

Further, prior work has discussed training in peer support services. For example, the benefits of training on mental health during the COVID-19 pandemic and increased agency were reported by certified peer specialists [8,9]. In addition to certification and training, nontrained peer support occurs through YouTube with a sense of belonging and support. Yet, the health benefits remain unclear, and there are concerns about the perceived trustworthiness of online peers and dramatization being rewarded with attention [10]. Evidence indicates that trained CPS specialists produce more robust outcomes than

naturally occurring peer support. As such, there may be a need for training in this area.

Limitations

The results should be interpreted with caution. First, these findings are limited by the cross-sectional nature of our data. Second, the authors are unable to verify that the respondents were CPS specialists. However, since there was no incentive for participation, the likelihood of recruiting ineligible participants is unlikely. Additionally, while participants report they are trained in peer support (418/426, 98.1%), training is a requirement to participate in the listserve; therefore, this may be an underestimation of those trained in peer support. Third, since this was an online national survey advertised through the last author's listserve, Facebook, and Twitter, only people with internet access could complete the survey. This could potentially produce biased survey results (ie, those who have access to the internet or own and use smartphones might be more interested in a web-based survey on technology use, or those who cannot afford digital technology, but would be interested in using it, are unable to respond to the survey). We also did not clarify which CPS specialists were able to meet in person or did meet in person with peers, or whether or not they had access to the suggested digital technologies, such as videoconferencing. This may have caused the percent usage of digital peer support modalities to be lower as individuals with in-person opportunities or a lack of digital opportunities may use modes such as SMS text messaging, phone calls, or videoconferencing more infrequently.

Conclusions

This study examines the modes of digital peer support intervention as well as the training and demographic background of peer supporters. Given the demand for mental health care, digital peer support emerges as one option to increase access [11]. These results suggest that among CPS specialists who own or have access to smartphones, tablets, or computers with the internet, 90%, according to a recent survey, commonly use SMS text messaging, phone calls, and videoconferencing to engage in peer support [12]. Less frequently, they may use diverse modes such as apps, social media, and video games. It is important to consider the background of peer supporters and the modes of communication to best accommodate areas where access to peer support is emerging. Larger longitudinal studies

and a variety of experimental designs may be considered to support training to direct optimal care. understand the efficacy of digital interventions and digital peer

Acknowledgments

KF was supported in part by National Institutes of Mental Health (award K01MH117496).

Conflicts of Interest

KF is an employee of Social Wellness, LLC, and Emissary Health, Inc. All the other authors declare that they have no conflicts of interest.

References

1. Reilly SE, Zane KL, McCuddy WT, Soulliard ZA, Scarisbrick DM, Miller LE, et al. Mental health practitioners' immediate practical response during the COVID-19 pandemic: observational questionnaire study. *JMIR Ment Health* 2020 Oct 01;7(9):e21237 [FREE Full text] [doi: [10.2196/21237](https://doi.org/10.2196/21237)] [Medline: [32931440](https://pubmed.ncbi.nlm.nih.gov/32931440/)]
2. Taylor CB, Fitzsimmons-Craft EE, Graham AK. Digital technology can revolutionize mental health services delivery: the COVID-19 crisis as a catalyst for change. *Int J Eat Disord* 2020 Jul;53(7):1155-1157 [FREE Full text] [doi: [10.1002/eat.23300](https://doi.org/10.1002/eat.23300)] [Medline: [32449523](https://pubmed.ncbi.nlm.nih.gov/32449523/)]
3. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. *Psychiatr Rehabil J* 2004;27(4):392-401. [doi: [10.2975/27.2004.392.401](https://doi.org/10.2975/27.2004.392.401)] [Medline: [15222150](https://pubmed.ncbi.nlm.nih.gov/15222150/)]
4. Akin LB, De Neve J, Dunn EW, Fancourt DE, Goldberg E, Helliwell JF, et al. Mental health during the first year of the COVID-19 pandemic: a review and recommendations for moving forward. *Perspect Psychol Sci* 2022 Jul;17(4):915-936 [FREE Full text] [doi: [10.1177/17456916211029964](https://doi.org/10.1177/17456916211029964)] [Medline: [35044275](https://pubmed.ncbi.nlm.nih.gov/35044275/)]
5. Collaborative Design for Recovery and Health. URL: <http://collabr.org> [accessed 2022-11-29]
6. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Int J Surg* 2014 Dec;12(12):1500-1524 [FREE Full text] [doi: [10.1016/j.ijsu.2014.07.014](https://doi.org/10.1016/j.ijsu.2014.07.014)] [Medline: [25046751](https://pubmed.ncbi.nlm.nih.gov/25046751/)]
7. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. *JMIR Ment Health* 2020 Apr 03;7(4):e16460 [FREE Full text] [doi: [10.2196/16460](https://doi.org/10.2196/16460)] [Medline: [32243256](https://pubmed.ncbi.nlm.nih.gov/32243256/)]
8. Wright A, Kritikos K, Bhiku K, LeFeber L, Skiest H, Whitman A, et al. The impact of the COVID-19 pandemic on peer specialists. *Psychiatr Rehabil J* 2022 Sep;45(3):201-211 [FREE Full text] [doi: [10.1037/prj0000517](https://doi.org/10.1037/prj0000517)] [Medline: [35511511](https://pubmed.ncbi.nlm.nih.gov/35511511/)]
9. Salzer M, Darr N, Calhoun G, Boyer W, Loss RE, Goessel J, et al. Benefits of working as a certified peer specialist: results from a statewide survey. *Psychiatr Rehabil J* 2013 Sep;36(3):219-221 [FREE Full text] [doi: [10.1037/prj0000016](https://doi.org/10.1037/prj0000016)] [Medline: [24059632](https://pubmed.ncbi.nlm.nih.gov/24059632/)]
10. Naslund J, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. *PLoS One* 2014;9(10):e110171 [FREE Full text] [doi: [10.1371/journal.pone.0110171](https://doi.org/10.1371/journal.pone.0110171)] [Medline: [25333470](https://pubmed.ncbi.nlm.nih.gov/25333470/)]
11. Merwin EI. Psychiatric-mental health nursing workforce in 2018: implications for the future. *Arch Psychiatr Nurs* 2020 Oct;34(5):317-324. [doi: [10.1016/j.apnu.2020.08.007](https://doi.org/10.1016/j.apnu.2020.08.007)] [Medline: [33032752](https://pubmed.ncbi.nlm.nih.gov/33032752/)]
12. Fortuna K, Aschbrenner KA, Lohman MC, Brooks J, Salzer M, Walker R, et al. Smartphone ownership, use, and willingness to use smartphones to provide peer-delivered services: results from a national online survey. *Psychiatr Q* 2018 Dec;89(4):947-956 [FREE Full text] [doi: [10.1007/s11126-018-9592-5](https://doi.org/10.1007/s11126-018-9592-5)] [Medline: [30056476](https://pubmed.ncbi.nlm.nih.gov/30056476/)]

Abbreviations

CPS: certified peer support

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 03.06.22; peer-reviewed by B Chaudhry, E Da Silva, K Machin; comments to author 12.07.22; revised version received 01.09.22; accepted 26.09.22; published 07.12.22.

Please cite as:

Fortuna K, Hill J, Chalker S, Ferron J

Certified Peer Support Specialists Training in Technology and Delivery of Digital Peer Support Services: Cross-sectional Study
JMIR Form Res 2022;6(12):e40065

URL: <https://formative.jmir.org/2022/12/e40065>

doi: [10.2196/40065](https://doi.org/10.2196/40065)

PMID: [36476983](https://pubmed.ncbi.nlm.nih.gov/36476983/)

©Karen Fortuna, Julia Hill, Samantha Chalker, Joelle Ferron. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Changes in Glycemic Control Following Use of a Spanish-Language, Culturally Adapted Diabetes Program: Retrospective Study

Caitlyn Edwards¹, PhD, RD; Elisa Orellana¹, BA; Kelly Rawlings¹, MPH; Mirta Rodriguez-Pla¹, MS, RD, CDCES; Aarathi Venkatesan¹, PhD

Vida Health, San Francisco, CA, United States

Corresponding Author:

Caitlyn Edwards, PhD, RD

Vida Health

100 Montgomery St, Ste 750

San Francisco, CA, 94104

United States

Phone: 1 415 989 1017

Email: caitlyn.edwards@vida.com

Abstract

Background: Several barriers to diabetes treatment and care exist, particularly in underserved medical communities.

Objective: This study aimed to evaluate a novel, culturally adapted, Spanish-language mHealth diabetes program for glycemic control.

Methods: Professional Spanish translators, linguists, and providers localized the entirety of the Vida Health Diabetes Management Program into a culturally relevant Spanish-language version. The Spanish-language Vida Health Diabetes Management Program was used by 182 (n=119 women) Spanish-speaking adults with diabetes. This app-based program provided access to culturally adapted educational content on diabetes self-management, one-on-one remote counseling and coaching sessions, and on-demand in-app messaging with bilingual (Spanish and English) certified health coaches, registered dietitian nutritionists, and certified diabetes care and education specialists. Hemoglobin A_{1c} (HbA_{1c}) was the primary outcome measure, and a 2-tailed, paired *t* test was used to evaluate changes in HbA_{1c} before and after program use. To determine the relationship between program engagement and changes in glycemic control, a cluster-robust multiple regression analysis was employed.

Results: We observed a significant decrease in HbA_{1c} of -1.23 points between baseline (mean 9.65%, SD 1.56%) and follow-up (mean 8.42%, SD 1.44%; *P*<.001). Additionally, we observed a greater decrease in HbA_{1c} among participants with high program engagement (high engagement: -1.59%, SD 1.97%; low engagement: -0.84%, SD 1.64%; *P*<.001).

Conclusions: This work highlights improvements in glycemic control that were clinically as well as statistically significant among Spanish-preferring adults enrolled in the Vida Health Spanish Diabetes Management Program. Greater improvements in glycemic control were observed among participants with higher program engagement. These results provide needed support for the use of digital health interventions to promote meaningful improvements in glycemic control in a medically underserved community.

(*JMIR Form Res* 2022;6(12):e40278) doi:[10.2196/40278](https://doi.org/10.2196/40278)

KEYWORDS

type 2 diabetes; digital health; diabetes intervention; diabetes; mobile health; mhealth; app-based; health coaching; HbA_{1c}; glycemic improvements; localization; Spanish; health application; health education; patient education; nutrition; digital health intervention; health management

Introduction

Diabetes impacts more than 37.3 million, or 11.3%, of all people in the United States [1]. Attributable to a combination of genetic,

cultural, and socioeconomic factors, Hispanic/Latino Americans are more likely to have type 2 diabetes (17%) compared to non-Hispanic White Americans (8%), are more likely to develop

diabetes across their lifespan, and are more likely to develop diabetes at a younger age [2,3].

Clinical guidelines recommend that individuals diagnosed with diabetes achieve and subsequently maintain a hemoglobin A_{1c} (HbA_{1c}) level of $\leq 7\%$. Prospective studies, such as the United Kingdom Prospective Diabetes Study and the Diabetes and Aging Study, have identified associations between a higher level of HbA_{1c} in individuals with diabetes and an increased risk of both micro- and macrovascular health events [4,5]. Clinical guidelines for diabetes treatment therefore center around successful glycemic management achieved through behavior changes, including long-term self-monitoring of dietary intake and physical exertion, and in some cases, blood glucose, as well as other behaviors, such as taking medicine as directed, obtaining biometric-based feedback, and participating in high levels of provider and patient interaction. These behavior changes are encompassed and defined as diabetes self-management education and support (DSMES) [6]. Many structural and personal barriers to treatment exist, and effective diabetes management can be difficult for both patients and providers [7]. Lack of appointment availability, provider access, time, finances, and the influence of other co-occurring health conditions are just a few of the obstacles to meeting clinical guidelines and reducing individual disease burden.

Hispanic/Latino Americans may face additional barriers to seeking treatment that are directly related to communication and culture [7]. Accessing providers who speak their language and understand their cultural values, preferences, and approaches to health is an additional challenge for many communities, resulting in medically underserved communities in which individuals are not able to seek, implement, or sustain the recommendations of DSMES [8] and essential self-care behaviors [9]. Indeed, research suggests that many DSMES interventions in usual care settings are considerably less successful in patients from ethnic minority groups compared to their White counterparts, with worse outcomes, lower rates of participation, and higher attrition rates [10]. There have been robust findings of large gaps in the present diabetes treatment literature and the health care community as a whole, and there is a need for an evidence-based understanding of the unique needs of medically underserved communities. The development and subsequent evaluation of personalized and language-concordant interventions is needed.

Existing work examining personalized care addressing underserved populations confirms a prolonged decrease in HbA_{1c} when health care is used that is culturally adapted; in other words, when health care is used that is adapted to relevant culture-based food and eating patterns, familial hierarchies, or culture-based beliefs or knowledge about diabetes [11]. A recent Cochrane systematic review and meta-analysis found that culturally appropriate diabetes health education in ethnic minority groups can result in significant improvements in HbA_{1c}, triglycerides, and knowledge about diabetes and its management [8]. Additionally, improvements in HbA_{1c} were found to be higher with culturally appropriate health education interventions compared to “usual care” interventions, and these results were sustained for up to 2 years following the studied interventions

[8]. One potential mechanism of these improvements in outcomes is the reduction of language discordance between providers and patients with limited English proficiency (LEP). LEP has been associated with reports of suboptimal interactions between providers and patients with diabetes, and the prevalence of glycemic control has been shown to increase among Latinos with LEP who switched care from a language-discordant provider to a language-concordant provider [12,13]. This previous work highlights the achievability and potential effectiveness of the development and implementation of culturally appropriate interventions and suggests that health care organizations should make efforts toward implementing such interventions.

One proposed approach to increasing the implementation of such programs is the use of mobile health (mHealth) interventions. The concept of mHealth is broadly used to describe smartphone-, tablet-, or personal computer-based health care [14]. Interventions that use mHealth are highly customizable and have immense potential for localization, which is the “adaptation of a product, application, or document content to meet the language, cultural, and other requirements of a specific target market” [15]. Due to this, mHealth interventions for diabetes treatment are becoming more commonplace and have been shown to be effective in both improving glycemic control and providing cost savings for English-speaking participants [14,16]. Surprisingly, despite widespread usage of smartphones among the Hispanic/Latino community [17], as well as evidence of successful technology-based behavior change interventions, mHealth diabetes programs localized for use in the Hispanic/Latino community are lacking [18]. Of all diabetes apps offered on app stores, only 41% of Android apps and 21% of iOS apps are available in Spanish. Additionally, most apps offering Spanish-language content are written at a complex (ie, >fifth grade) reading level or do not provide their user interface in Spanish [19]. This is a concerning missed opportunity for providing equitable care, as properly designed mHealth provides a unique ability to increase access not only to cost-effective and easily accessible evidence-based diabetes care, but also to culturally appropriate, personalized, and language-concordant care and providers.

In response to this gap in the field of diabetes treatment and management, the Vida Health Diabetes Management Program was localized as the Vida Health Spanish Diabetes Management Program. The Vida Health Diabetes Management Program is an app-based digital health platform for the treatment and management of chronic diseases. The primary aim of this study was to evaluate the effectiveness of the Vida Health Spanish Diabetes Management Program on measures of glycemic control (ie, HbA_{1c}) among a cohort of Spanish-prefering adults from the United States. The secondary aim was to quantify and examine engagement metrics of the Vida Health Spanish Diabetes Management Program and its potential impact on glycemic control.

Methods

Ethical Considerations

This study was approved by an independent institutional review board (Western Institutional Review Board), which waived informed consent; the study was identified as having minimal risk because the data were fully anonymized before use in the analysis.

Recruitment and Enrollment

Adults from a single major-payer client of Vida Health were recruited for this study through brochures, calling campaigns, and email announcements. Potential participants were offered the app-based program as part of their medical insurance benefits. All participants enrolled in the study had smartphone- or web-based access. Participants accessed the program through the Apple App Store or the Google Play Store and entered an invitation code unique to their insurance carrier for complementary program usage. The program is offered in both English and Spanish; the language used is based on the default language preference of the participants' mobile phone. All subsequent engagement with the Vida Health app refers to the translated and adapted Spanish-language app interface and content.

After app installation, participants completed a number of internally designed intake forms, where they provided contact information for app notifications, basic demographic information including age and sex, and self-reported existing health conditions. Participants were excluded from study participation if they indicated the presence of type 1 diabetes, stage 4 or 5 chronic kidney disease, or class III or IV congestive heart failure; were pregnant; or were breastfeeding.

Laboratory data pertaining to glycemic control (ie, HbA_{1c} level) was provided by the insurance carrier at a monthly cadence using protected and standardized data-sharing arrangements.

Vida Health Spanish Diabetes Management Program

The Vida Health Diabetes Management Program is an app-based intervention designed to enable and empower individuals to manage their health through frequent provider interaction and rigorously designed multimedia content. Providers offered by the program include bilingual (Spanish and English) certified health coaches, registered dietitian nutritionists, and certified diabetes care and education specialists. Professional Spanish translators and linguists, as well as bilingual providers, localized the entirety of the Vida Health Diabetes Management Program into a culturally relevant Spanish-language version. Localization, as defined above, differs from translation. An example of localization, as opposed to mere translation, is how Vida approaches nutrition education content in the app. A translation approach entails term replacement, for example, swapping the term “1 rodaja pan” for “1 slice bread” in a lesson about using the Healthy Plate meal-planning approach at breakfast. The localization method looks at term use with a cultural and accessibility lens. An example of this includes providing a culturally equivalent recipe (eg, a favorite holiday dish traditionally made in Spanish-speaking countries instead of a more United States–influenced holiday special). Localization

also affects lists of ingredients on labels. Alongside the Spanish translation of a word, the English-language version is shown, to match what any person living in the United States would find at a grocery store (eg, “whole-grain flour”). If participants had any issue logging food or recipes in the app, they were encouraged to reach out to their provider for clarity or to send a picture of their meal to their provider, who could assist with logging.

After completing app intake, participants were paired with a bilingual (Spanish and English) certified health coach, a registered dietitian nutritionist, a certified diabetes care and education specialist, or a combination of these providers. To account for licensure requirements and cultural differences in geographic locations, providers were matched to participants by their state of residence. All providers received extensive evidence-based training on diabetes self-management and operated under a motivational interviewing framework to promote self-efficacy and autonomy for behavior change [20]. Motivational interviewing is an evidence-based skill set designed to aid individuals in identifying motivation for change and to facilitate personalized changes in lifestyle and behavior [21]. To provide support for bilingual providers, all training was available in both Spanish and English. Additional training was provided on the nutrition guidelines of various Spanish-speaking countries.

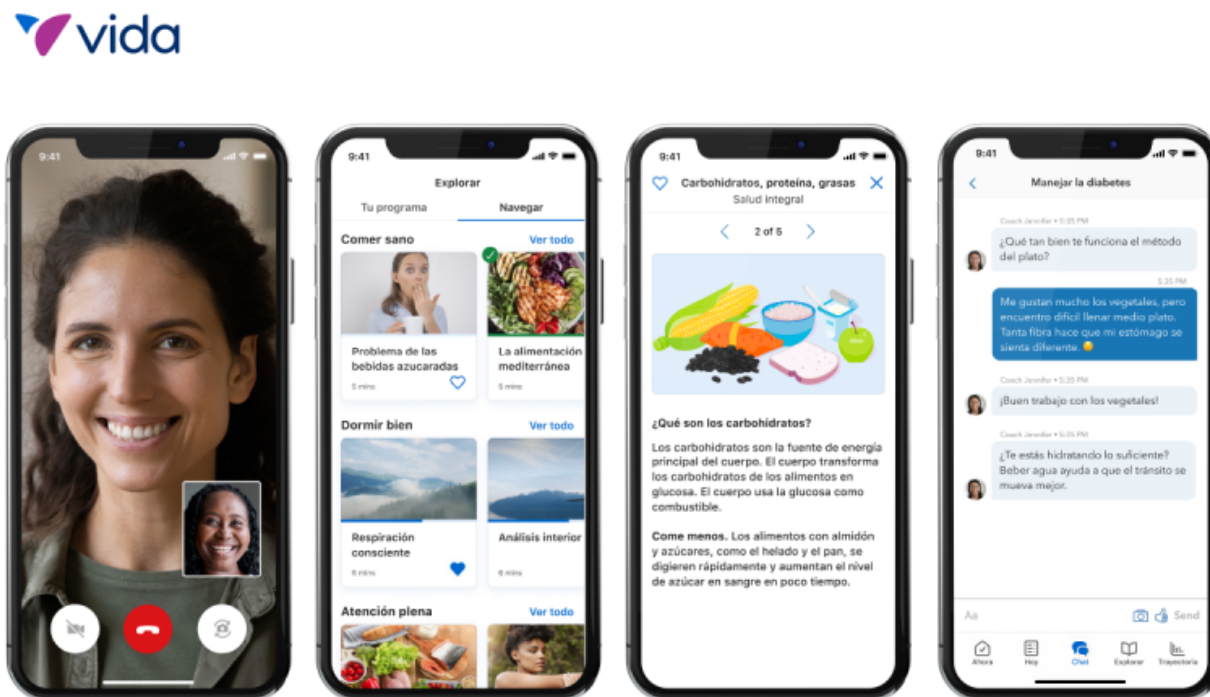
Full information on the program is described in previously published work [22]. Briefly, the program was designed from both the provider and participant perspectives to operationalize a complete feedback loop. The first interaction between provider and participant included a detailed assessment to identify personalized participant health goals and potential barriers to outcome success. All providers had the option to complete internal and external communications and notes in Spanish to prevent the need for switching between languages during consultations. Follow-up provider support was provided through on-demand in-app messaging and remote video sessions (the videos sessions were weekly for the first 12 weeks and monthly thereafter), in which providers and participants worked to follow up on goal progress and resolve ambivalence to change. Between provider sessions, participants were encouraged to message their providers as often as needed and received daily in-app content to provide education and activation opportunities and to support treatment goals (Figure 1). All app usage was tracked by the providers, and biometric data was collected at each video session to allow for up-to-date data monitoring, interpretation, and pertinent adjustment of treatment as needed. Participants were additionally encouraged to reach out to their providers for app support if any issues or questions arose regarding the app.

Daily app content included evidence-based structured lessons and multimedia content pertaining to nutrition, taking medicine, self-monitoring, and other behaviors outlined in the ADCES (Association of Diabetes Care and Education Specialists) 7 Self-Care Behaviors [9,23]. Alongside app-generated content, participants are encouraged to engage in a variety of data-logging activities within the program app: food logging, physical activity logging, and self-monitoring of blood glucose when appropriate. The Vida Health food logging system included a comprehensive list of culturally relevant food

selections for various Spanish-speaking countries, allowing for accessible food logging. Additionally, providers were trained in relevant cultural and regional clinical nutrition guidelines to ensure informed care was provided that considered an individual's unique traditions, customs, and values. This training and these potential resources included continuing education and

reference documentation on country- or region-specific staple foods and their corresponding nutrient breakdown, common traditions and celebrations, and customs surrounding meal timing. Screenshots of the Vida Health user interface are provided in [Figure 1](#).

Figure 1. Screenshots of the Vida Health Spanish Diabetes Management Program.



Statistical Analysis

All data preparation and analyses were conducted in Python (version 3.7.8) and R (version 4.1.2, R Studio Team). The primary outcome measure for this study was HbA_{1c}. HbA_{1c} was measured in clinical laboratories; the data were made available by the program's client. Baseline HbA_{1c} was defined according to the laboratory test results that were closest to the program start, measured between 1 year before (−365 days) to within 21 days after enrollment. Follow-up HbA_{1c} was defined as the most recent HbA_{1c} assessment within 12 months of enrollment and at least 90 days from the program start. A 2-tailed, paired *t* test was used to assess changes in HbA_{1c} from baseline.

Program engagement was a secondary focus of this study. First, we computed a cumulative sum for each of the following factors: (1) messages sent between a participant and their provider, (2)

the number of consultations with a provider completed by a participant, and (3) the amount of curriculum content completed within the first 12 weeks (the most intensive phase of the intervention) of the program start. Then, we created a binary engagement variable, in which “high engagement” was defined as engagement for each of our 3 defining engagement metrics that was greater than or equal to the median. A cluster-robust multiple regression analysis using the participants as a clustering factor was used to evaluate the association of level of engagement with HbA_{1c} change.

Results

Sample Demographics

Baseline HbA_{1c} data were available for 354 individuals. Follow-up HbA_{1c} data were available for 182 individuals. Full demographics for the final study cohort are shown in [Table 1](#).

Table 1. Demographic and engagement information for a cohort of Spanish-prefering individuals enrolled in a Spanish-language mHealth intervention (N=182).

Characteristics	Value
Sex^a, n	
Female	119
Male	59
Baseline HbA _{1c} ^b (%), mean (SD)	9.65 (1.56)
Follow-up HbA _{1c} (%), mean (SD)	8.42 (1.44)
Age (years), mean (SD)	55.67 (8.91)
Age category (years), n (%)	
20 to 40	11 (6)
40 to 60	119 (65.4)
Older than 60	52 (28.6)
Messages sent, mean (SD)	27.72 (44.26)
Consultations completed, mean (SD)	8.00 (7.63)
Curriculum completed, mean (SD)	8.95 (16.38)

^aFour individuals did not provide data for this category.

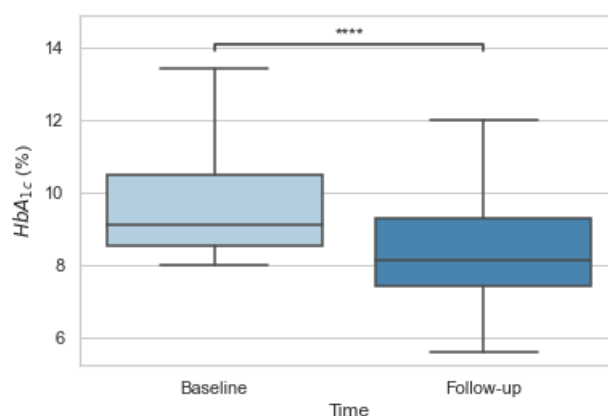
^bHbA_{1c}: hemoglobin A_{1c}.

Primary Outcome

Baseline HbA_{1c} measurements were completed on average –56.70 (SD 67.49) days prior to the program start, and follow-up HbA_{1c} measurements were completed on average 157.94 (SD

54.38) days from the program start. A paired *t* test revealed a significant reduction in HbA_{1c} of –1.23 percentage points between baseline (mean 9.65%, SD 1.56%) and follow-up (mean 8.42%, SD 1.44%; $t_{181}=-8.99$; $P<.001$; Figure 2).

Figure 2. Box plot showing significant differences between HbA_{1c} at baseline and at a minimum 90-day follow-up in a sample of Spanish-prefering (N=182) individuals enrolled in the Vida Health Spanish Diabetes Management Program. HbA_{1c}: hemoglobin A_{1c}.



Program Use Outcomes

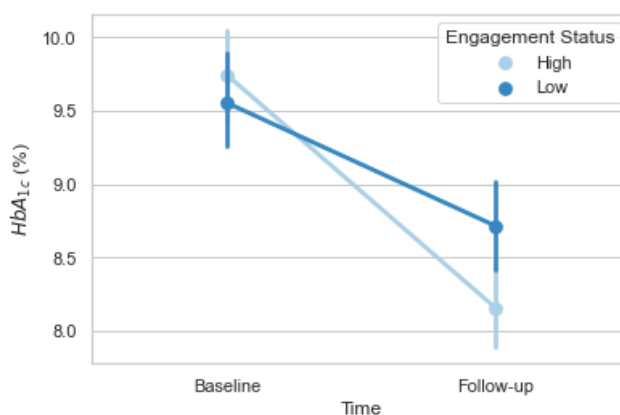
Program engagement data was available from all 182 individuals. Engagement metrics for up to 3 months (12 weeks) following program enrollment were calculated for (1) messages sent between a participant and their provider, (2) the number of consultations with their provider completed by a participant, and (3) the amount of the curriculum content completed. The engagement data were heavily leftward skewed, indicating the presence of a group of high program users. To account for this, the median of each engagement metric was used to dichotomize participants into “high” (95/182, 52.2%) or “low” (87/182,

47.8%) app users. No differences in age ($P=.16$) or baseline HbA_{1c} ($P=.41$) were observed between these engagement groups. To determine potential relationships between app engagement and glycemic control, a cluster-robust multiple regression analysis was conducted using participant as a clustering variable. This allowed the inherent standard error between providers to cluster on the participant. Age, sex, and the binary high or low engagement group status were included as fixed factors, and baseline HbA_{1c} was included as a covariate. Age ($P=.42$) and sex ($P=.89$) were not found to be predictive of a change in HbA_{1c} at follow-up, while baseline HbA_{1c} value was predictive of a change in HbA_{1c} ($\beta=-.77$, $P<.001$). A main

effect of engagement status was observed ($\beta=-.31$, $P=.003$), in which high program users exhibited a higher change in HbA_{1c}

over the course of the intervention than did low program users (high group mean change was -1.59% , SD 1.97%; low group mean change was -0.84% , SD 1.64; Figure 3).

Figure 3. Baseline and follow-up HbA_{1c} dichotomized by a median split of intervention engagement into “high” and “low” engagement categories (N=182). HbA_{1c}: hemoglobin A_{1c}.



Discussion

Principal Findings

This study aimed to evaluate the effectiveness of a localized Spanish-language diabetes management program for improving measures of glycemic control among a cohort of Spanish-prefering adults from the United States. We demonstrated that in a cohort of Spanish-prefering individuals with diabetes, enrollment in the Vida Health Spanish Diabetes Management Program was effective in significantly improving glycemic control. We also found that individuals with higher intervention engagement demonstrated higher improvements in glycemic control than individuals with lower intervention engagement.

Many barriers exist to diabetes treatment, and these barriers are only exacerbated in medically underserved communities [7]. The goal of developing the Vida Health Spanish Diabetes Management Program was to reduce barriers to diabetes education, diabetes self-management tools, and diabetes care provider access. The large effect on HbA_{1c} (-1.23 percentage points) observed in this study lends support to the belief that comprehensively designed, culturally adapted mHealth interventions can result in clinically significant health benefits. Additionally, this result is consistent in magnitude with findings from the non-Spanish Vida Health Diabetes Management Program, and the decrease we observed was larger than the 0.1% to 0.7% reduction found in a recent Cochrane review of the potential benefits for glycemic control of culturally appropriate diabetes interventions [8,22]. These findings indicate a significant clinical impact on the study participants, as an improvement in HbA_{1c} greater than 1% over 10 years has been associated with a 21% reduction in diabetes-related deaths and a 37% reduction in microvascular complications [5,24]. A reduction of over 1% was achieved as early as 90 days following program enrollment, potentially allowing for the “legacy effect” of the physiological benefits of HbA_{1c} reduction to begin earlier and be enjoyed for longer across the lifespan.

Comparison to Prior Work

Our secondary aim explored the hypothesis that participants with higher engagement with the program would see greater improvements in glycemic control, as has been demonstrated in both the non-Spanish Vida Health Diabetes Management Program [22] and other culturally adapted diabetes education programs [16]. The preferred type of engagement with programs designed to increase self-management of type 2 diabetes has been shown to vary, with participants expressing differing preferences for different design features [25]. Therefore, a composite measure of engagement was used to capture program engagement across potential individual-level differences in preferred methodology of engagement. While both high- and low-engagement participants demonstrated improvements in HbA_{1c}, we found that higher-engaging individuals, who completed more sessions with their providers, sent more messages to their provider team and completed more of the program curriculum; they also saw an improvement in glycemic control that was -0.75 percentage points greater than that seen by the lower-engaging participants. Achievement of this level of engagement was possible because this study is among the first, to the authors' knowledge, to use a smartphone app for a localized Spanish-language diabetes mHealth intervention. Recent advancements in smartphone technology have facilitated enhanced capabilities for mHealth interventions, allowing for app engagement opportunities not feasible in previous trials. While much work has been conducted in English-prefering populations to expand health-based interventions, the same level of expansion has not been applied to Spanish-prefering populations. As an example, the previous TExT-MED trial found trend-level improvements in HbA_{1c} in individuals with diabetes who participated in a unidirectional text-message intervention compared to those who completed a usual care intervention [26]. Had this, and other previous work, been able to expand upon text messaging with other smartphone-enabled technology to support localized and culturally adapted content, it is possible significant results would have been observed. The ability to mimic in-person “gold standard” diabetes treatments, alongside the enhanced ease of access due to the ubiquity of

technology and smartphone use, allows for comprehensive and bidirectional diabetes care in typically medically underserved communities. While future work is needed to explore potential factors associated with being a high- or low-engagement participant, what we can conclude is that the Vida Health Spanish Diabetes Management Program provides an effective and accessible way of providing diabetes care for those who engage with its content and experience.

It is also possible that participants who had higher engagement were more motivated to engage in and manage their health, as motivation to maintain a healthy lifestyle is key to diabetes management [27]. While motivation was not directly measured in the current work, the Vida Health Spanish Diabetes Management Program removed many obstacles to care that may have impeded participants' motivation to engage in diabetes treatment. Glazier et al [28] examined strategies to improve the response to cultural interventions in individuals with diabetes and found that most successful interventions used cultural tailoring of the intervention, provided more than 10 contacts between the individual and the intervention, provided a longer intervention duration than typical didactic education methods, and provided the opportunity for one-to-one individualized assessment. The Vida Health Spanish Diabetes Management Program provided all these opportunities. Particular care was taken in ensuring cultural congruency between the participants and the providers. Individuals in underserved medical communities are less likely to develop provider rapport and be offered evidence-based treatment interventions and achieve and sustain recommended outcomes when providers do not share the same language or do not have a thorough understanding of the cultural values of the patient [29]. In an ideal provider and patient setting, routine meetings with an educator, even at a distance and provided through an mHealth app, maintain active engagement and offer patients a high level of accountability for cocreated health goals [30]. The role of the provider has additionally been noted as a strong motivator in Hispanic/Latino culture, and a strong rapport with their providers may prompt participants to adopt new lifestyle habits and become better motivated to self-regulate their diabetes self-management behaviors [27,31,32].

Strengths and Limitations

This study is not without its limitations. First, it was retrospective in nature; therefore, limited assumptions can be

made about the causal relationship between the Vida Health Spanish Diabetes Management Program and the observed improvements in glycemic control. One component of diabetes management is the use of diabetes medications. While prescribing medications was outside the scope of the current program, the program content did provide education on the importance of medication adherence and management and encouraged regular visits with primary care physicians. Therefore, it is possible that diabetes medications contributed to the improvements in HbA_{1c} for some of the participants, particularly if they had been recently diagnosed with diabetes and initiated medications for the first time. Second, while great care was taken to ensure culturally appropriate matching of providers to patient needs, the Spanish-speaking community is complex and diverse, and cultural barriers may still have existed. Future work and future iterations of the Vida Health Spanish Diabetes Management Program will aim not only to expand Spanish-language cultural relevance but also to potentially develop other language-based diabetes health services. Third, this sample included participants from a United States-based payer organization, meaning all participants enrolled were employed and had some level of health insurance, which potentially limited the generalizability of our findings to individuals who may be unemployed or uninsured. Lastly, a number of individuals did not provide follow-up laboratory data, limiting the population sample size. Regardless, significant improvements in outcomes were observed among those who provided follow-up laboratory data.

Conclusion

This study demonstrated that a comprehensively designed and culturally relevant diabetes treatment program that provided access to localized, culturally adapted, Spanish-language diabetes education and self-management content alongside frequent access to bilingual certified health coaches, registered dietitian nutritionists, and certified diabetes care and education specialists led to clinically significant improvements in glycemic control in Spanish-speaking individuals with type 2 diabetes. Future work is needed to understand barriers to intervention engagement among potential participants to enable access to available culturally adapted interventions such as the Vida Health Spanish Diabetes Management Program.

Acknowledgments

We are grateful to the study participants and all members enrolled in Vida Health programs. We appreciate the health coaches, registered dietitians, and certified diabetes care and education specialists at Vida Health for their professional expertise, commitment, and ongoing support of members seeking to manage complex, polychronic conditions, including type 2 diabetes. We also wish to acknowledge and thank Terra Translations for their translation services.

Data Availability

The data sets analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

CE performed statistical analysis and wrote the first draft of the paper. EO, KR, and MR-P designed the study program. All authors (CE, EO, KR, MR-P, and AV) provided valuable feedback on the final submission.

Conflicts of Interest

All authors were employees of and held share options in Vida Health during the study.

References

1. National Diabetes Statistics Report. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/diabetes/data/statistics-report/index.html> [accessed 2022-11-17]
2. Candib LM. Obesity and diabetes in vulnerable populations: reflection on proximal and distal causes. *Ann Fam Med* 2007;5(6):547-556 [FREE Full text] [doi: [10.1370/afm.754](https://doi.org/10.1370/afm.754)] [Medline: [18025493](https://pubmed.ncbi.nlm.nih.gov/18025493/)]
3. Hispanic or Latino Americans and Type 2 Diabetes. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/diabetes/library/features/hispanic-diabetes.html> [accessed 2022-11-17]
4. Laiteerapong N, Ham SA, Gao Y, Moffet HH, Liu JY, Huang ES, et al. The Legacy Effect in Type 2 Diabetes: Impact of Early Glycemic Control on Future Complications (The Diabetes & Aging Study). *Diabetes Care* 2019 Mar;42(3):416-426 [FREE Full text] [doi: [10.2337/dc17-1144](https://doi.org/10.2337/dc17-1144)] [Medline: [30104301](https://pubmed.ncbi.nlm.nih.gov/30104301/)]
5. Srimanunthiphol J, Beddow R, Arakaki R. A review of the United Kingdom Prospective Diabetes Study (UKPDS) and a discussion of the implications for patient care. *Hawaii Med J* 2000 Jul;59(7):295-8, 313. [Medline: [10961040](https://pubmed.ncbi.nlm.nih.gov/10961040/)]
6. Powers MA, Bardsley J, Cypress M, Duker P, Funnell MM, Hess Fischl A, et al. Diabetes self-management education and support in type 2 diabetes: a joint position statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics. *Diabetes Care* 2015 Jul;38(7):1372-1382. [doi: [10.2337/dc15-0730](https://doi.org/10.2337/dc15-0730)] [Medline: [26048904](https://pubmed.ncbi.nlm.nih.gov/26048904/)]
7. Nam S, Chesla C, Stotts NA, Kroon L, Janson SL. Barriers to diabetes management: patient and provider factors. *Diabetes Res Clin Pract* 2011 Jul;93(1):1-9. [doi: [10.1016/j.diabres.2011.02.002](https://doi.org/10.1016/j.diabres.2011.02.002)] [Medline: [21382643](https://pubmed.ncbi.nlm.nih.gov/21382643/)]
8. Creamer J, Attridge M, Ramsden M, Cannings-John R, Hawthorne K. Culturally appropriate health education for Type 2 diabetes in ethnic minority groups: an updated Cochrane Review of randomized controlled trials. *Diabet Med* 2016 Feb 07;33(2):169-183. [doi: [10.1111/dme.12865](https://doi.org/10.1111/dme.12865)] [Medline: [26202820](https://pubmed.ncbi.nlm.nih.gov/26202820/)]
9. Association of Diabetes Care and Education Specialists, Kolb L. An effective model of diabetes care and education: The ADCES7 Self-Care Behaviors™. *Sci Diabetes Self Manag Care* 2021 Feb 11;47(1):30-53. [doi: [10.1177/0145721720978154](https://doi.org/10.1177/0145721720978154)] [Medline: [34078208](https://pubmed.ncbi.nlm.nih.gov/34078208/)]
10. Coonrod BA, Betschart J, Harris MI. Frequency and determinants of diabetes patient education among adults in the U.S. population. *Diabetes Care* 1994 Aug;17(8):852-858. [doi: [10.2337/diacare.17.8.852](https://doi.org/10.2337/diacare.17.8.852)] [Medline: [7956630](https://pubmed.ncbi.nlm.nih.gov/7956630/)]
11. McCurley JL, Gutierrez AP, Gallo LC. Diabetes prevention in U.S. Hispanic adults: a systematic review of culturally tailored interventions. *Am J Prev Med* 2017 Apr;52(4):519-529 [FREE Full text] [doi: [10.1016/j.amepre.2016.10.028](https://doi.org/10.1016/j.amepre.2016.10.028)] [Medline: [27989451](https://pubmed.ncbi.nlm.nih.gov/27989451/)]
12. Schenker Y, Karter AJ, Schillinger D, Warton EM, Adler NE, Moffet HH, et al. The impact of limited English proficiency and physician language concordance on reports of clinical interactions among patients with diabetes: the DISTANCE study. *Patient Educ Couns* 2010 Nov;81(2):222-228 [FREE Full text] [doi: [10.1016/j.pec.2010.02.005](https://doi.org/10.1016/j.pec.2010.02.005)] [Medline: [20223615](https://pubmed.ncbi.nlm.nih.gov/20223615/)]
13. Parker MM, Fernández A, Moffet HH, Grant RW, Torreblanca A, Karter AJ. Association of patient-physician language concordance and glycemic control for limited-English proficiency Latinos with type 2 diabetes. *JAMA Intern Med* 2017 Mar 01;177(3):380-387 [FREE Full text] [doi: [10.1001/jamainternmed.2016.8648](https://doi.org/10.1001/jamainternmed.2016.8648)] [Medline: [28114680](https://pubmed.ncbi.nlm.nih.gov/28114680/)]
14. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. *Adv Nutr* 2017 May;8(3):449-462 [FREE Full text] [doi: [10.3945/an.116.014100](https://doi.org/10.3945/an.116.014100)] [Medline: [28507010](https://pubmed.ncbi.nlm.nih.gov/28507010/)]
15. Ishida R, Miller S. Localization vs. Internationalization. W3C. URL: <https://www.w3.org/International/questions/qa-i18n.en> [accessed 2022-05-25]
16. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. *J Diabetes Sci Technol* 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: [10.1177/1932296817713506](https://doi.org/10.1177/1932296817713506)] [Medline: [28560898](https://pubmed.ncbi.nlm.nih.gov/28560898/)]
17. Williams JP, Schroeder D. Popular glucose tracking apps and use of mHealth by Latinos with diabetes: review. *JMIR Mhealth Uhealth* 2015 Aug 25;3(3):e84 [FREE Full text] [doi: [10.2196/mhealth.3986](https://doi.org/10.2196/mhealth.3986)] [Medline: [26307533](https://pubmed.ncbi.nlm.nih.gov/26307533/)]
18. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. *Diabetes Technol Ther* 2012 Jun;14(6):492-496. [doi: [10.1089/dia.2011.0252](https://doi.org/10.1089/dia.2011.0252)] [Medline: [22524591](https://pubmed.ncbi.nlm.nih.gov/22524591/)]
19. Rodriguez JA, Singh K. The Spanish availability and readability of diabetes apps. *J Diabetes Sci Technol* 2018 May;12(3):719-724 [FREE Full text] [doi: [10.1177/1932296817749610](https://doi.org/10.1177/1932296817749610)] [Medline: [29291639](https://pubmed.ncbi.nlm.nih.gov/29291639/)]

20. Miller WR, Rollnick S. Motivational Interviewing, Third Edition: Helping People Change. New York, NY: Guilford Press; 2012.
21. McDaniel CC, Kavookjian J, Whitley HP. Telehealth delivery of motivational interviewing for diabetes management: A systematic review of randomized controlled trials. *Patient Educ Couns* 2022 Apr;105(4):805-820. [doi: [10.1016/j.pec.2021.07.036](https://doi.org/10.1016/j.pec.2021.07.036)] [Medline: [34366228](https://pubmed.ncbi.nlm.nih.gov/34366228/)]
22. Zimmermann G, Venkatesan A, Rawlings K, Scahill MD. Improved glycemic control with a digital health intervention in adults with type 2 diabetes: retrospective study. *JMIR Diabetes* 2021 Jun 02;6(2):e28033 [FREE Full text] [doi: [10.2196/28033](https://doi.org/10.2196/28033)] [Medline: [34075880](https://pubmed.ncbi.nlm.nih.gov/34075880/)]
23. Hood KK, Hilliard M, Piatt G, Ievers-Landis CE. Effective strategies for encouraging behavior change in people with diabetes. *Diabetes Manag (Lond)* 2015;5(6):499-510 [FREE Full text] [Medline: [30100925](https://pubmed.ncbi.nlm.nih.gov/30100925/)]
24. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 1998 Sep 12;352(9131):837-853. [doi: [10.1016/s0140-6736\(98\)07019-6](https://doi.org/10.1016/s0140-6736(98)07019-6)] [Medline: [9742976](https://pubmed.ncbi.nlm.nih.gov/9742976/)]
25. Turchioe MR, Heitkemper EM, Lor M, Burgermaster M, Mamykina L. Designing for engagement with self-monitoring: A user-centered approach with low-income, Latino adults with Type 2 Diabetes. *Int J Med Inform* 2019 Oct;130:103941 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.08.001](https://doi.org/10.1016/j.ijmedinf.2019.08.001)] [Medline: [31437618](https://pubmed.ncbi.nlm.nih.gov/31437618/)]
26. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TExT-MED): a randomized controlled trial. *Ann Emerg Med* 2014 Jun;63(6):745-54.e6. [doi: [10.1016/j.annemergmed.2013.10.012](https://doi.org/10.1016/j.annemergmed.2013.10.012)] [Medline: [24225332](https://pubmed.ncbi.nlm.nih.gov/24225332/)]
27. Lakerveld J, Palmeira AL, van Duinkerken E, Whitelock V, Peyrot M, Nouwen A. Motivation: key to a healthy lifestyle in people with diabetes? Current and emerging knowledge and applications. *Diabet Med* 2020 Mar 20;37(3):464-472. [doi: [10.1111/dme.14228](https://doi.org/10.1111/dme.14228)] [Medline: [31916283](https://pubmed.ncbi.nlm.nih.gov/31916283/)]
28. Glazier R, Bajcar J, Kennie N, Willson K. A systematic review of interventions to improve diabetes care in socially disadvantaged populations. *Diabetes Care* 2006 Jul;29(7):1675-1688. [doi: [10.2337/dc05-1942](https://doi.org/10.2337/dc05-1942)] [Medline: [16801602](https://pubmed.ncbi.nlm.nih.gov/16801602/)]
29. Oomen JS, Owen LJ, Suggs LS. Culture counts: why current treatment models fail Hispanic women with type 2 diabetes. *Diabetes Educ* 1999 Sep 03;25(2):220-225. [doi: [10.1177/014572179902500207](https://doi.org/10.1177/014572179902500207)] [Medline: [10531847](https://pubmed.ncbi.nlm.nih.gov/10531847/)]
30. Weech-Maldonado R, Morales LS, Elliott M, Spritzer K, Marshall G, Hays RD. Race/ethnicity, language, and patients' assessments of care in Medicaid managed care. *Health Serv Res* 2003 Jun;38(3):789-808 [FREE Full text] [doi: [10.1111/1475-6773.00147](https://doi.org/10.1111/1475-6773.00147)] [Medline: [12822913](https://pubmed.ncbi.nlm.nih.gov/12822913/)]
31. Ragsdale C, Wright J, Shokar G, Salaiz R, Shokar NK. Hispanic patient perspectives of the physician's role in obesity management. *J Clin Med Res* 2017 Feb;9(2):170-175 [FREE Full text] [doi: [10.14740/jocmr2868w](https://doi.org/10.14740/jocmr2868w)] [Medline: [28090233](https://pubmed.ncbi.nlm.nih.gov/28090233/)]
32. Joachim-Célestin M, Gamba-Maldonado T, Dos Santos H, Montgomery SB. Delivering the same intervention to Hispanic/Latinos with pre-diabetes and diabetes. Early evidence of success in a longitudinal mixed method study. *Inquiry* 2021 Nov 26;58:469580211055595 [FREE Full text] [doi: [10.1177/00469580211055595](https://doi.org/10.1177/00469580211055595)] [Medline: [34825596](https://pubmed.ncbi.nlm.nih.gov/34825596/)]

Abbreviations

DSMES: diabetes self-management education and support

HbA_{1c}: hemoglobin A_{1c}

mHealth: mobile health

Edited by A Mavragani; submitted 13.06.22; peer-reviewed by A Reviewer, M Barboza, L Quintana; comments to author 06.07.22; revised version received 26.07.22; accepted 15.11.22; published 07.12.22.

Please cite as:

Edwards C, Orellana E, Rawlings K, Rodriguez-Pla M, Venkatesan A

Changes in Glycemic Control Following Use of a Spanish-Language, Culturally Adapted Diabetes Program: Retrospective Study

JMIR Form Res 2022;6(12):e40278

URL: <https://formative.jmir.org/2022/12/e40278>

doi:[10.2196/40278](https://doi.org/10.2196/40278)

PMID:[36476397](https://pubmed.ncbi.nlm.nih.gov/36476397/)

©Caitlyn Edwards, Elisa Orellana, Kelly Rawlings, Mirta Rodriguez-Pla, Aarathi Venkatesan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The

complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Predicting Overweight and Obesity Status Among Malaysian Working Adults With Machine Learning or Logistic Regression: Retrospective Comparison Study

Jyh Eiin Wong¹, PhD; Miwa Yamaguchi², PhD; Nobuo Nishi², MD, PhD; Michihiro Araki², PhD; Lei Hum Wee^{1,3}, PhD

¹Centre for Community Health Studies, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

²National Institute of Health and Nutrition, National Institutes of Biomedical Innovation, Health and Nutrition, Tokyo, Japan

³Faculty of Health and Medical Sciences, School of Medicine, Taylor's University, Selangor, Malaysia

Corresponding Author:

Jyh Eiin Wong, PhD

Centre for Community Health Studies

Faculty of Health Sciences

Universiti Kebangsaan Malaysia

Jalan Raja Muda Abdul Aziz

Kuala Lumpur, 50300

Malaysia

Phone: 60 39289 ext 7683

Email: wjeiin@ukm.edu.my

Abstract

Background: Overweight or obesity is a primary health concern that leads to a significant burden of noncommunicable disease and threatens national productivity and economic growth. Given the complexity of the etiology of overweight or obesity, machine learning (ML) algorithms offer a promising alternative approach in disentangling interdependent factors for predicting overweight or obesity status.

Objective: This study examined the performance of 3 ML algorithms in comparison with logistic regression (LR) to predict overweight or obesity status among working adults in Malaysia.

Methods: Using data from 16,860 participants (mean age 34.2, SD 9.0 years; n=6904, 41% male; n=7048, 41.8% with overweight or obesity) in the Malaysia's Healthiest Workplace by AIA Vitality 2019 survey, predictor variables, including sociodemographic characteristics, job characteristics, health and weight perceptions, and lifestyle-related factors, were modeled using the extreme gradient boosting (XGBoost), random forest (RF), and support vector machine (SVM) algorithms, as well as LR, to predict overweight or obesity status based on a BMI cutoff of 25 kg/m².

Results: The area under the receiver operating characteristic curve was 0.81 (95% CI 0.79-0.82), 0.80 (95% CI 0.79-0.81), 0.80 (95% CI 0.78-0.81), and 0.78 (95% CI 0.77-0.80) for the XGBoost, RF, SVM, and LR models, respectively. Weight satisfaction was the top predictor, and ethnicity, age, and gender were also consistent predictor variables of overweight or obesity status in all models.

Conclusions: Based on multi-domain online workplace survey data, this study produced predictive models that identified overweight or obesity status with moderate to high accuracy. The performance of both ML-based and logistic regression models were comparable when predicting obesity among working adults in Malaysia.

(*JMIR Form Res* 2022;6(12):e40404) doi:[10.2196/40404](https://doi.org/10.2196/40404)

KEYWORDS

overweight; obesity; prediction; machine learning; logistic regression; etiology; algorithms; Malaysia; adults; predictive models; accuracy; working adults; surveillance

Introduction

Overweight and obesity are global health issues that are increasingly recognized as major public health concerns in low- and middle-income countries. In Malaysia, 1 in 2 adults, particularly those of working age (ie, aged 30 to 65 years), is either overweight or obese [1]. This is concerning, as obesity prevalence is rising at a very high rate (3.3%) in this country [2]. The increase in overweight and obesity is related to increases in noncommunicable diseases, the mortality rate, and health care costs, as well as decreases in productivity and economic growth [2-5].

Obesity is a chronic, relapsing, multifactorial disease that is attributable to individual or biological, psychological, sociocultural, local, and global environmental factors [6-8]. As obesity is largely preventable, understanding the determinants of and risk factors for obesity is important for the development of population-based strategies to prevent obesity. Identifying individuals at high risk of obesity enables early intervention to modify obesity risk factors. Conventional statistical methods, such as generalized linear or regression models with a low number of predictor variables, have been successful in identifying obesity [9]. However, given the complexity of the etiology of obesity, regression modeling may not be adept at disentangling nonlinear and interdependent relationships among factors for obesity prediction.

Machine learning (ML) is an advanced data analytical method that uses fine-tuned algorithms to characterize and predict outcomes by learning from data without being explicitly programmed to do so. As health data become more available and accessible, ML techniques are increasingly used to perform such complex tasks in obesity research as classifying and predicting obesity at individual and group levels [10-12]. ML techniques have advantages over regression modeling, as they are data driven and do not necessitate a priori assumptions, such as normality, linearity, and multicollinearity. In addition, ML techniques are capable of handling high-dimensional and complex data sources beyond numeric sources, and therefore may be able to provide new insights into unexplored predictor variables [9,13]. Thus, ML techniques are likely to be more accurate than regression models in obesity prediction [14].

A wide range of ML-based algorithms incorporating various predictors and risk factors, training set sizes, and degrees of implementation have been used to predict adult obesity [11,14]. The reported accuracy of ML algorithms to predict adult obesity as a binary outcome ranges broadly, from 0.59 to 0.97 for overall accuracy [15-24] and 0.51 to 0.99 for the area under the curve (AUC) [15,19,20,23,24]. A review suggested that ML-based models predicted childhood and adolescent obesity much better than linear regression [13]. However, studies that have compared the performance of different ML algorithms with regression in adult obesity have reported mixed findings. Some evidence

suggests superior performance for ML models compared to regression models [19,21], while some suggests similar or inferior performance [15,17,18,23]. These inconsistencies may partly be due to data quality, variable selection, and the use of different approaches to model fitting, parameter tuning, and validation among studies.

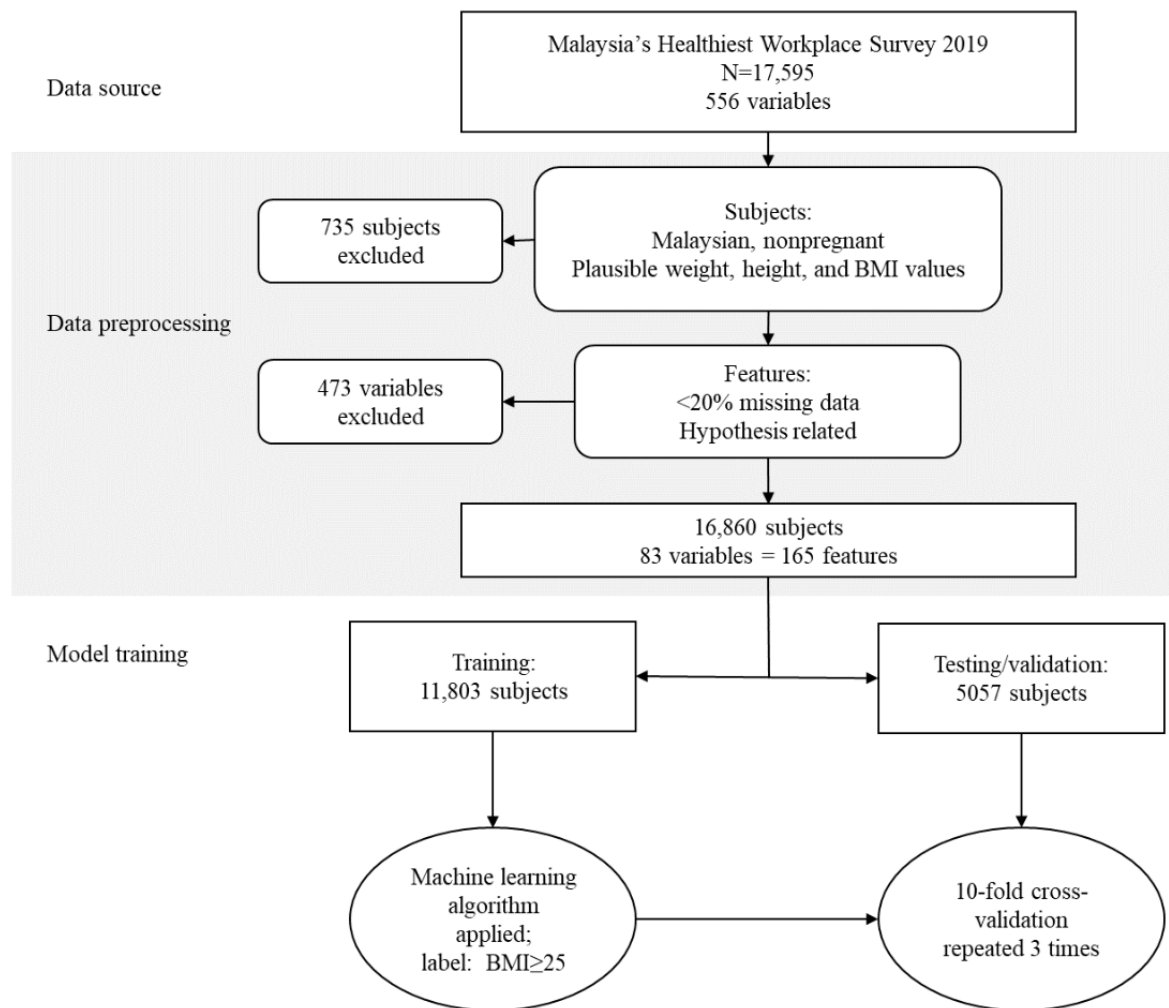
The Malaysia's Healthiest Workplace by AIA Vitality survey is a large, observational online survey of the health and well-being of Malaysian employees [25]. Since 2017 (with the exceptions of 2020 and 2021, because of the COVID-19 pandemic), this annual online workplace survey has collected comprehensive information on Malaysian employees' sociodemographic characteristics, physical and mental health, smoking and alcohol habits, physical activity, diet, musculoskeletal health, and work environment as a database to inform workplace interventions and improve productivity [25]. In this study, we propose an ML-based model to predict overweight and obesity status among employees in Malaysia based on multi-domain variables collected in this large survey. We evaluated the performance of 3 ML algorithms and compared them with logistic regression for the prediction of overweight and obesity status. We hypothesized that ML algorithms would outperform logistic regression models in predicting overweight and obesity status based on BMI.

Methods

Study Design and Data

This is a retrospective study of predictive model derivation using data from the Malaysia's Healthiest Workplace by AIA Vitality 2019 survey. This online survey, commissioned by AIA Malaysia and delivered in partnership with RAND Europe, was administered between May and August 2019. The survey, which has taken place annually in Malaysia from 2017 to 2019, aimed to determine workplace productivity and multi-domain factors that influence workplace productivity. Employees from small, medium, and large organizations were invited to answer a 40-minute employee survey questionnaire about their general health, lifestyle behaviors, mental health status, and work environment. The study rationale and methodology have been discussed in detail elsewhere [26-28].

The initial data set comprised data submitted by 17,595 participants from 230 companies. We initially included 16,931 participants resident in Malaysia for whom data were available for body weight and height. If they were women, participants were included if they were not pregnant. Participants with (1) body weight more than 200 kg, (2) height more than 200 cm, or (3) BMI values of more than 60 kg/m² or less than 14 kg/m² were deemed to have implausible values and were excluded from analysis. After excluding 71 participants who reported implausible weight, height, or BMI, the final data set included 16,860 of 16,931 participants (95.8%) (Figure 1).

Figure 1. Overview of data preprocessing, model development, and model evaluation.

Ethics Approval

The use of the data was approved by the Research Ethics Committee Universiti Kebangsaan Malaysia (JEP-2020-707). As the obtained pooled data were anonymized and deidentified, informed consent from the participants was not required. The study results were presented following the reporting guidelines and recommendations for ML [29,30].

Data Preprocessing

An overview of data preprocessing and model development is illustrated in Figure 1. Data preprocessing involved the selection of participants and variables (features) followed by mean substitution of missing data, one-hot encoding of categorical variables, and min-max scaling for data normalization.

Outcome Variable

The outcome of interest was overweight or obesity status, defined as a BMI of 25 kg/m² or more [31]. This was calculated by dividing the self-reported body weight (in kg) by the squared height (in m²). The cutoff of 25 was chosen as Southeast Asians are reported to have higher body fatness at a lower BMI than Europeans [32,33] and are therefore predisposed to elevated cardiovascular risk factors and other adverse effects of obesity

at lower BMI ranges (23 kg/m² to 25 kg/m²), as observed in local studies [34,35]. Further, a recent study suggested that a BMI of 24.8 kg/m² is an optimal BMI cutoff to define obesity among Malaysian adults based on percentage of body fat [36].

Predictor Variables

Initially, the data set consisted of 556 predictor variables. A total of 473 variables that contained redundant information or text information with more than 20% missing or nonapplicable data were removed from the data set. The reduced data set included 83 variables that were grouped into the following 4 main domains: sociodemographic characteristics, job characteristics, status perception, and lifestyle-related behaviors (the list of predictor variables is included in Multimedia Appendix 1).

Categorical variables (n=16) were one-hot encoded into binary variables. For instance, weight satisfaction was assessed by a categorical question that prompted participants to select 1 of 3 statements that best described how they felt about their current body weight. The participants indicated whether they (1) were happy with their weight, (2) were not happy with their weight but had no intention of losing or gaining weight, or (3) wanted to change their weight. This categorical variable was

subsequently encoded into 3 binary variables (ie, “weight_satisfaction_1,” “weight_satisfaction_2,” and “weight_satisfaction_3”). Finally, prediction models were trained and tested on the final 165 normalized variables. A total of 120 (73%) of these 165 predictor variables were binary (yes/no) variables.

Statistical Analysis Methods

Model Development

The R (version 3.6.1; R Software Foundation) package “caret” (version 6.0-90) was used for model training and validation [37]. Based on a random 70:30 split, a total of 11,803 participants, including 4934 (41.8%) with overweight or obesity, were used to train the model. The remaining 30% of the participants (5057/16,860) were used to predict the obesity outcome during model validation.

Three supervised, nonlinear ML classifiers were applied, namely extreme gradient boosting (XGBoost), random forest (RF), and a support vector machine (SVM). XGBoost is a tree-based ensemble algorithm that uses a boosting method to create multiple decision trees sequentially. The algorithm combines the predictions of weak decision trees to produce a more robust final model. Improvised on the gradient boosting framework, XGBoost is a popular learning algorithm due to its high predictive power and efficiency in handling continuous and categorical data using relatively low computational power [38]. RF is also an ensemble method but uses a bagging method to train multiple decision trees in parallel using random selection of predictors. The final model merges predictions from each decision tree to predict a class [39]. Finally, SVMs use a kernel-based algorithm to construct a decision boundary or hyperplane that best separates the data into 2 classes in n-dimensional space. SVMs use extreme cases, also known as support vectors, to create an optimal hyperplane that has the maximum margin between the vectors [40].

In this study, logistic regression (LR) was compared with the 3 ML models. Logistic regression is a part of the generalized linear model and is the conventional classifier for categorical outcome responses. The algorithm assumes a linear relationship between the predictor variables and the log odds (probability) of obesity as the outcome in this study. All predictor variables were included in the model, regardless of statistical significance, to maintain comparability across models. The goodness of fit of the logistic regression model was demonstrated by a McFadden R^2 value of 0.3452 and a Nagelkerke R^2 value of 0.3452. The probability produced by the logistic regression was subsequently assigned to a binary outcome (overweight/obese or not), based on the customary probability cutoff point of 0.5.

The details of the package, functions, and parameters used in this study are presented in [Multimedia Appendix 2](#). Using a grid search approach, the best combinations of parameters were employed for each algorithm. All models were tuned using 10-fold cross-validation repeated 3 times. Using the varImp function of the caret library, model-specific metrics were used to identify the best-performing predictors. To present the relative ranking of each predictor, the measures of importance for all models were scaled to have a maximum value of 100.

Model Evaluation

The final trained models were saved and restored for prediction using a separate test data set (n=5057) and for comparison with other models. Classification metrics were obtained from the confusion matrix (confusionMatrix) embedded in the caret package. A prediction of overweight or obesity status was considered a positive prediction. Performance was assessed by 4 main metrics (the first 3 metrics are limited in their discriminating power in selecting the best classifier [41], but they are the most common metrics used in the literature and are therefore presented for comparison with other studies): (1) accuracy, the proportion of correct predictions divided by the total number of instances evaluated; (2) sensitivity (also known as the true positive rate), the proportion of actual positives (ie, overweight or obese status) that were correctly predicted; (3) specificity (also known as the true negative rate), the proportion of actual negatives (ie, no overweight or obese status) that were correctly predicted; and (4) AUC, which represents a tradeoff between sensitivity and specificity and served as the main metric for model evaluation. AUC is extracted from the receiver operating characteristic (ROC) curve, which is the probability plot of the true positive rate (ie, sensitivity) against the false positive rate (ie, 1-specificity). An AUC above 0.5 indicates the model is better capable of distinguishing positives (ie, subjects with overweight or obesity) from negatives. In general, an AUC of 0.7 to <0.8 is considered acceptable, 0.8 to <0.9 excellent, and 0.9 or above outstanding predictive performance [42]. The ROCs and corresponding AUCs were computed and plotted with the pROC package.

The performance metrics of all predictive models are presented as point estimates with 95% CIs. For accuracy, sensitivity, and specificity, 95% CIs were calculated assuming a Gaussian distribution of the proportion. For AUCs, 95% CIs were derived through resampling with the bootstrap percentile method with 2000 repetitions. Model comparisons were made based on the 95% CIs of the 4 performance metrics.

Results

Study Characteristics

The analysis included 16,860 participants, of whom 41% (n=6904) were male and 41.8% (n=7048) had overweight or obese status. The male participants were significantly older, and the distributions for ethnicity, education level, marital status, occupation, individual monthly income, and obesity status were also significantly different by sex ($P<.001$ for all; [Multimedia Appendix 3](#)).

Model Comparisons

[Table 1](#) presents the predictive performance of the ML and logistic regression models. Among the 4 models, the RF and LR models had lower sensitivity but higher specificity. While XGBoost exhibited the best mean accuracy and AUC, overall accuracy was similar across all models based on the 95% CIs. The ROCs of the 4 models are illustrated in [Figure 2](#).

[Table 2](#) compares the performance of XGBoost and LR in predicting obesity by sex. For both algorithms, the models for female participants recorded higher specificity but lower

sensitivity than the models for male participants. Overall accuracy and AUC were similar across all 4 models, with the 2 algorithms showing no sex-specific differences in predictive performance.

The ranking of the most important predictors of the models is summarized in Figure 3. In order of importance, the top 4 predictor variables for the XGBoost ML model were weight satisfaction, ethnicity, age, and gender. For the LR model, the top predictor variables were weight satisfaction, physical health, age, and diet satisfaction.

Table 1. Performance of machine-learning algorithms and logistic regression in obesity prediction.

Metrics	Gradient boosting, mean (95% CI)	Random forest, mean (95% CI)	Support vector machine, mean (95% CI)	Logistic regression, mean (95% CI)
Accuracy ^a	0.73 (0.72-0.75)	0.73 (0.71-0.74)	0.72 (0.71-0.73)	0.71 (0.70-0.72)
Sensitivity ^a	0.67 (0.65-0.69)	0.60 (0.58-0.62)	0.65 (0.62-0.67)	0.56 (0.54-0.58)
Specificity ^a	0.78 (0.76-0.79)	0.82 (0.80-0.83)	0.77 (0.76-0.79)	0.82 (0.81-0.83)
Area under the curve ^b	0.81 (0.79-0.82)	0.80 (0.79-0.81)	0.80 (0.78-0.81)	0.78 (0.77-0.80)

^aIn these rows, 95% CIs were calculated assuming Gaussian distribution of the proportions.

^bIn this row, 95% CIs were derived through resampling with the bootstrap percentile method with 2000 repetitions.

Figure 2. Receiver operating characteristic curves with corresponding AUC values; AUC values for each model are also presented in Table 2. AUC: area under the curve.

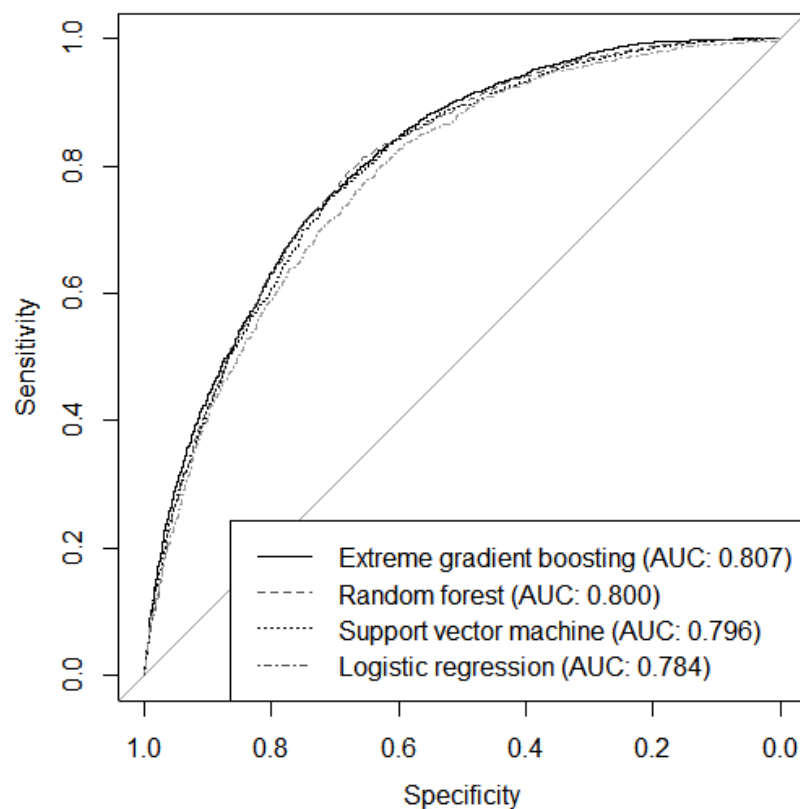


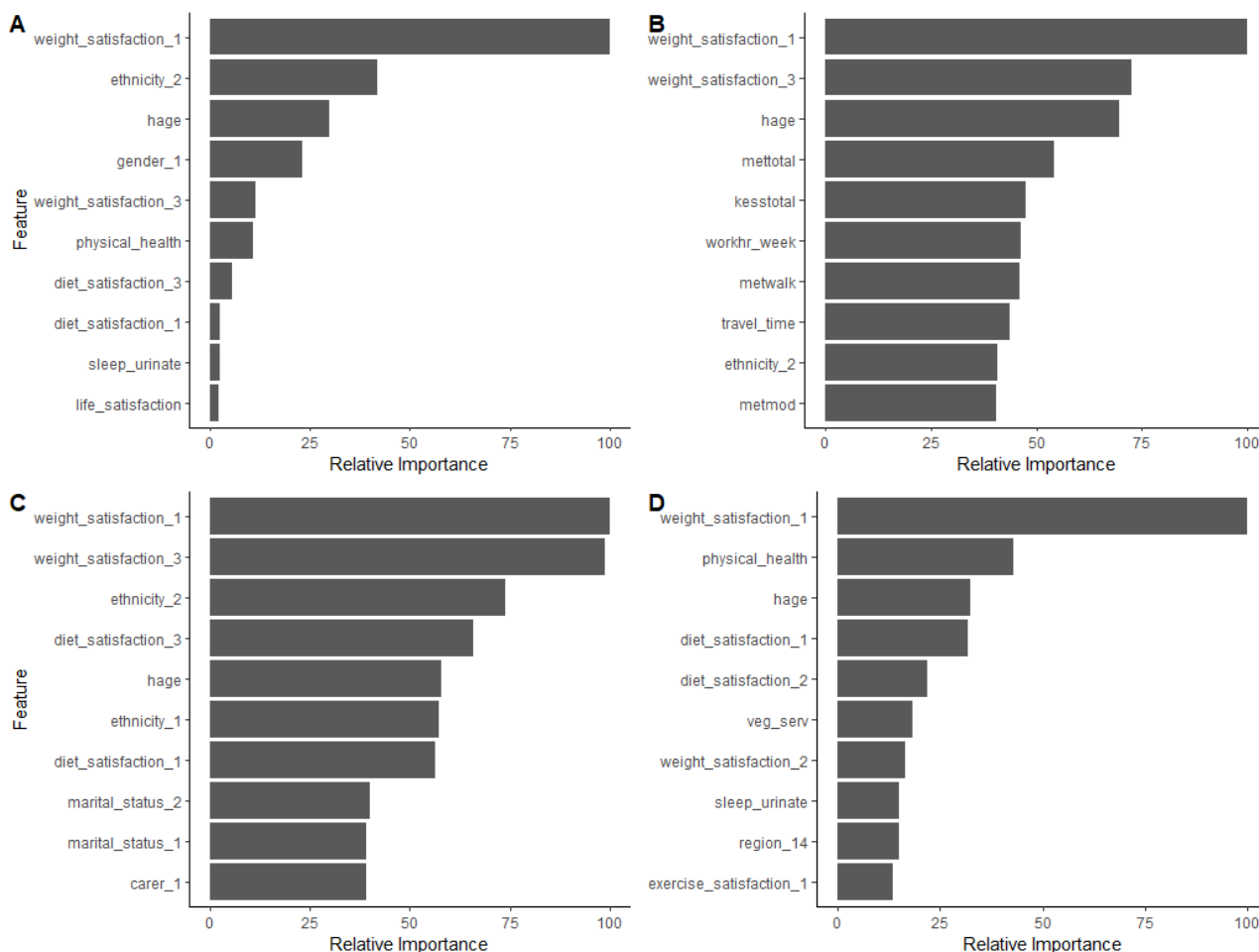
Table 2. Comparison of performance between machine learning and logistic regression in sex-specific obesity prediction.

Metrics	Gradient boosting, mean (95% CI)		Logistic regression, mean (95% CI)	
	Male participants	Female participants	Male participants	Female participants
Accuracy ^a	0.71 (0.69-0.73)	0.74 (0.72-0.75)	0.70 (0.68-0.72)	0.73 (0.71-0.74)
Sensitivity ^a	0.75 (0.73-0.78)	0.61 (0.58-0.63)	0.72 (0.69-0.75)	0.60 (0.57-0.63)
Specificity ^a	0.66 (0.63-0.69)	0.81 (0.80-0.83)	0.68 (0.65-0.71)	0.80 (0.78-0.81)
Area under the curve ^b	0.78 (0.76-0.80)	0.81 (0.79-0.82)	0.76 (0.74-0.78)	0.79 (0.77-0.80)

^aIn these rows, 95% CIs were calculated assuming Gaussian distribution of the proportions.

^bIn this row, 95% CIs were derived through resampling with the bootstrap percentile method with 2000 repetitions.

Figure 3. Variable importance plots of obesity predictors for extreme gradient boosting (A), random forest (B), support vector machine (C) and logistic regression (D) models. The top 10 predictors are shown for all models.



Discussion

Principal Results

This study applied various ML models and compared their performance to the performance of a conventional logistic regression model in predicting overweight or obesity status among working adults in Malaysia. Our results showed that ML and logistic regression had similarly acceptable or excellent predictive performance, as assessed by the metrics of accuracy (values ranged from 70% to 75%) and AUC (values ranged from 78% to 81%), for both the overall and sex-specific models.

Comparison With Prior Work

Our findings, based on data collected annually as part of a large-scale online survey of employees, compare favorably to those of a recent study by Thamrin et al [23] that also used a large Southeast Asian sample (N=618,898), in Indonesia. That study employed logistic regression, classification and regression trees, and a naive Bayes classifier for obesity prediction based on data for sociodemographic characteristics, diet, physical activity, lifestyle behaviors, and health status from the Indonesian Basic Health Research periodic survey. The study reported accuracy between 70.8% and 72.2% and an AUC between 0.75 and 0.80, which is comparable to the performance

of our models (mean accuracy 71%-73.3% and AUC 0.78-0.81). While there is no definite standard for acceptable accuracy, the models in our study recorded accuracy greater than 70% and AUC greater than 0.7, which is better than the accuracy and AUC of past models that used novel predictors, including genetics [20,24], detailed dietary intake [18,21], and objectively measured physical activity [15].

In this study, the overall performance of the ML models, namely XGBoost, RF, and SVM, was found to be similar to logistic regression, as indicated by the overlapping 95% CIs. This corroborates the findings of a systematic review of 71 studies, which concluded that ML did not offer greater performance benefits than logistic regression for clinical prediction models [43]. Specifically, for obesity prediction, Ferdowsy et al [17] employed 8 algorithms, in addition to logistic regression, in a data set that included 21 well-established risk factors for obesity, such as diet, physical activity, lifestyle behaviors, and disease history. Their study recorded the highest accuracy (97%) with the logistic regression model, which outperformed ML algorithms including k-nearest neighbor, RF, a multilayer perceptron, an SVM, a naive Bayes classifier, adaptive boosting, a decision tree, and a gradient boosting classifier for obesity prediction [17]. Kim et al [18] modeled the effects of 7 dietary factors on overweight or obesity status using data from the Korea National Health and Nutrition Examination Survey. That study showed that the predictive accuracy of logistic regression (0.62486) was higher than that of decision trees (0.54026) and similar to that of a deep neural network model of deep learning (0.62496). Taken together, comparative studies that deal with a small number of strong predictor variables [15,17,18,23] suggest that regression models are likely to perform better than, if not as well as, ML models in obesity prediction.

Another possible reason for this similar performance is that the observed relationships among the significant predictors of obesity in this sample may appear linear on the log-odd scale. Hence, logistic regression was not disadvantaged by assuming linearity in these predictors. In this study, we employed 3 nonlinear ML classifiers due to the fact that many variables, including intrapersonal and socioeconomic factors that affect body weight, such as age, sex, and gender, are nonlinear in nature [44]. However, it could be hypothesized that these nonlinear ML algorithms may have been less proficient at modeling the present data set because the data mostly consisted of binary variables (120/165, 73%).

It is important to acknowledge that different ML algorithms may fit and perform differently when used with different data sets. Guided by previous findings that obesity determinants are different for men and women [19,45], we developed separate, sex-stratified models for overweight or obesity status prediction. However, the predictive accuracy of the sex-specific models was similar to the overall or combined models. This suggests that separate prediction models for each sex are not warranted in this Malaysian adult population.

In terms of predictor variables, weight satisfaction appears to be a consistent, novel predictor in all predictive models, together with such well-established risk factors for obesity as ethnicity, age, and gender. Weight satisfaction is an attitudinal component

of body image, which reflects individuals' feelings and thoughts about their weight [46]. The variable "weight_satisfaction_1," which represents satisfaction or contentment with current body weight, appears to have had the most influential power in the trained model to predict overweight or obesity status (Figure 3).

This novel finding is consistent with previous studies showing that self-perception of body weight is an important determinant of weight management behaviors and lifestyle practices [47]. However, the relationships between weight satisfaction and weight-related behaviors are complex and multifaceted. Depending on sex, race, ethnicity, accuracy of weight perceptions, and psychological factors, weight satisfaction may promote positive diet and physical activity behaviors or lead to maladaptive or unhealthy weight-control or dieting behaviors [48-50]. As weight satisfaction and dissatisfaction appear to be mostly stable in adulthood [51,52], we posit that this subjective variable may be cognitively easier and more reliable to report than body weight and height among adults. This finding supports the usefulness of including weight satisfaction as a proxy for actual weight status in studies and e-surveys, where anthropometry measurements may not be available or feasible.

Strengths and Limitations

To the best of our knowledge, this is the first study to employ ML to predict overweight or obesity status in an adult working population in Malaysia. This study used rich data from a large annual survey that included a wide, multi-domain set of predictor variables in working adults with a broad range of ages (18 to 88 years) and occupations. Another strength of the study lay in its employment of advanced ML classifiers with careful cross-validation (to avoid model overfitting) and parameter optimization. The variable importance technique afforded novel insights into significant factors that are correlates of overweight or obesity status in a Malaysian working population.

This study was also limited in several ways. First, the study findings do not infer temporality or causality of the observed predictor-obesity relationships due to the use of a cross-sectional design. However, the findings suggest putative variables that could be explored using novel model interpretation techniques such as Shapley additive explanations [53] and could be considered for further testing in longitudinal or trial settings. Second, mislabeling of obesity was likely, due to the reliance on self-reported body weight and height to derive BMI as a surrogate measure of general obesity. Notably, the prevalence of individuals with overweight or obesity in this study (4934/11,803, 41.8%) was lower than the national prevalence of 50.1% [1]. Such errors, or noise, may have reduced the performance of the models. Therefore, the current findings represent conservative estimates of predictive accuracy. Finally, we acknowledge that the generalizability of our models is limited, as validation was based on testing data that came from the same sample. Validating the models with an external data set would more closely approximate the real performance of the prediction models. Future work is needed to confirm the external validity and reproducibility of the models in other data sets, such as the Malaysia's Healthiest Workplace surveys from 2018 or later.

Conclusions

Using a multi-domain set of predictors from a large online employee survey, we constructed models that were able to predict overweight or obesity status in a Malaysian working population with moderate to high accuracy. Weight satisfaction was the most prominent factor, followed by ethnicity, age, and gender, in differentiating individuals with overweight or obese status. Among the 3 ML models (XGBoost, RF, and SVM), XGBoost had the highest accuracy and AUC, but the overall performance of all ML-based models was similar to the logistic regression model for obesity prediction.

This study is complementary to and extends the growing literature showing that ML may be used to predict overweight or obesity status based on online survey data with reasonable accuracy. Besides unveiling distinctive factors that influence weight status in this Asian population, this work also produced potential models or algorithms that can be used to screen for overweight or obesity status in community settings, especially when body weight and height data are not available. A natural progression of this study would be to test the performance of the produced models in an external data set to establish the external validity of the findings.

Acknowledgments

The authors would like to thank AIA Malaysia for granting permission to use the data. The study data source was funded by AIA Malaysia (grants NN-2019-152 and NN-2021-004). Payment of the article processing fee for this publication was supported by the National Institutes of Biomedical Innovation, Health and Nutrition. JEW, MY, NN, and MA jointly conceived the study concept and design. LHW and MA contributed to data acquisition. JEW performed data processing, variable selection, and data analysis under the supervision of MA. JEW drafted the manuscript. All authors contributed to data and result interpretation. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Predictor variables used in the dataset (n=165).

[[DOCX File , 37 KB - formative_v6i12e40404_app1.docx](#)]

Multimedia Appendix 2

Description of software packages, methods and tuning parameters for model development.

[[DOCX File , 29 KB - formative_v6i12e40404_app2.docx](#)]

Multimedia Appendix 3

Study sample characteristics (n=16860).

[[DOCX File , 33 KB - formative_v6i12e40404_app3.docx](#)]

References

1. National Health and Morbidity Survey 2019: Non-communicable diseases, healthcare demand and healthy literacy. Volume I: NCDs – Non-Communicable Diseases: Risk Factors and other Health Problems. Ministry of Health Malaysia. 2020. URL: https://iku.moh.gov.my/images/IKU/Document/REPORT/NHMS2019/Report_NHMS2019-NCD_v2.pdf [accessed 2022-11-24]
2. Lobstein T, Brinsden H, Neveux M. World Obesity Atlas 2022. World Obesity Federation. 2022. URL: <https://www.worldobesity.org/resources/resource-library/world-obesity-atlas-2022> [accessed 2022-11-21]
3. Tremmel M, Gerdtham U, Nilsson PM, Saha S. Economic burden of obesity: a systematic literature review. *Int J Environ Res Public Health* 2017 Apr 19;14(4):435 [FREE Full text] [doi: [10.3390/ijerph14040435](https://doi.org/10.3390/ijerph14040435)] [Medline: [28422077](https://pubmed.ncbi.nlm.nih.gov/28422077/)]
4. Goettler A, Grosse A, Sonntag D. Productivity loss due to overweight and obesity: a systematic review of indirect costs. *BMJ Open* 2017 Oct 05;7(10):e014632 [FREE Full text] [doi: [10.1136/bmjopen-2016-014632](https://doi.org/10.1136/bmjopen-2016-014632)] [Medline: [28982806](https://pubmed.ncbi.nlm.nih.gov/28982806/)]
5. GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, et al. Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* 2017 Jul 06;377(1):13-27 [FREE Full text] [doi: [10.1056/NEJMoa1614362](https://doi.org/10.1056/NEJMoa1614362)] [Medline: [28604169](https://pubmed.ncbi.nlm.nih.gov/28604169/)]
6. Bray GA, Kim KK, Wilding JPH, World Obesity Federation. Obesity: a chronic relapsing progressive disease process. A position statement of the World Obesity Federation. *Obes Rev* 2017 Jul;18(7):715-723. [doi: [10.1111/obr.12551](https://doi.org/10.1111/obr.12551)] [Medline: [28489290](https://pubmed.ncbi.nlm.nih.gov/28489290/)]
7. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet* 2011 Aug 27;378(9793):804-814. [doi: [10.1016/S0140-6736\(11\)60813-1](https://doi.org/10.1016/S0140-6736(11)60813-1)] [Medline: [21872749](https://pubmed.ncbi.nlm.nih.gov/21872749/)]

8. Ford ND, Patel SA, Narayan KV. Obesity in low- and middle-income countries: burden, drivers, and emerging challenges. *Annu Rev Public Health* 2017 Mar 20;38:145-164 [[FREE Full text](#)] [doi: [10.1146/annurev-publhealth-031816-044604](https://doi.org/10.1146/annurev-publhealth-031816-044604)] [Medline: [28068485](#)]
9. DeGregory KW, Kuiper P, DeSilvio T, Pleuss JD, Miller R, Roginski JW, et al. A review of machine learning in obesity. *Obes Rev* 2018 May;19(5):668-685 [[FREE Full text](#)] [doi: [10.1111/obr.12667](https://doi.org/10.1111/obr.12667)] [Medline: [29426065](#)]
10. Birkin M, Wilkins E, Morris MA. Creating a long-term future for big data in obesity research. *Int J Obes (Lond)* 2019 Dec;43(12):2587-2592. [doi: [10.1038/s41366-019-0477-y](https://doi.org/10.1038/s41366-019-0477-y)] [Medline: [31641212](#)]
11. Chatterjee A, Gerdes MW, Martinez SG. Identification of risk factors associated with obesity and overweight-a machine learning overview. *Sensors (Basel)* 2020 May 11;20(9):2734 [[FREE Full text](#)] [doi: [10.3390/s20092734](https://doi.org/10.3390/s20092734)] [Medline: [32403349](#)]
12. Scheinker D, Valencia A, Rodriguez F. Identification of factors associated with variation in US county-level obesity prevalence rates using epidemiologic vs machine learning models. *JAMA Netw Open* 2019 Apr 05;2(4):e192884 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2019.2884](https://doi.org/10.1001/jamanetworkopen.2019.2884)] [Medline: [31026030](#)]
13. Colmenarejo G. Machine learning models to predict childhood and adolescent obesity: a review. *Nutrients* 2020 Aug 16;12(8):2466 [[FREE Full text](#)] [doi: [10.3390/nu12082466](https://doi.org/10.3390/nu12082466)] [Medline: [32824342](#)]
14. Safaei M, Sundararajan EA, Driss M, Boulila W, Shapi'i A. *Comput Biol Med* 2021 Sep;136:104754 [[FREE Full text](#)] [doi: [10.1016/j.compbiomed.2021.104754](https://doi.org/10.1016/j.compbiomed.2021.104754)] [Medline: [34426171](#)]
15. Cheng X, Lin S, Liu J, Liu S, Zhang J, Nie P, et al. Does physical activity predict obesity-a machine learning and statistical method-based analysis. *Int J Environ Res Public Health* 2021 Apr 09;18(8):3966 [[FREE Full text](#)] [doi: [10.3390/ijerph18083966](https://doi.org/10.3390/ijerph18083966)] [Medline: [33918760](#)]
16. Delnevo G, Mancini G, Rocchetti M, Salomoni P, Trombini E, Andrei F. The prediction of body mass index from negative affectivity through machine learning: a confirmatory study. *Sensors (Basel)* 2021 Mar 29;21(7):2361 [[FREE Full text](#)] [doi: [10.3390/s21072361](https://doi.org/10.3390/s21072361)] [Medline: [33805257](#)]
17. Ferdowsy F, Rahi KSA, Jabiullah MI, Habib MT. A machine learning approach for obesity risk prediction. *CRBS* 2021 Nov;2:100053. [doi: [10.1016/j.crbeha.2021.100053](https://doi.org/10.1016/j.crbeha.2021.100053)]
18. Kim H, Lim DH, Kim Y. Classification and prediction on the effects of nutritional intake on overweight/obesity, dyslipidemia, hypertension and type 2 diabetes mellitus using deep learning model: 4-7th Korea National Health and Nutrition Examination Survey. *Int J Environ Res Public Health* 2021 May 24;18(11):5597 [[FREE Full text](#)] [doi: [10.3390/ijerph18115597](https://doi.org/10.3390/ijerph18115597)] [Medline: [34073854](#)]
19. Lee BJ, Kim KH, Ku B, Jang J, Kim JY. Prediction of body mass index status from voice signals based on machine learning for automated medical applications. *Artif Intell Med* 2013 May;58(1):51-61. [doi: [10.1016/j.artmed.2013.02.001](https://doi.org/10.1016/j.artmed.2013.02.001)] [Medline: [23453267](#)]
20. Lee Y, Christensen JJ, Parnell LD, Smith CE, Shao J, McKeown NM, et al. Using machine learning to predict obesity based on genome-wide and epigenome-wide gene-gene and gene-diet interactions. *Front Genet* 2021;12:783845 [[FREE Full text](#)] [doi: [10.3389/fgene.2021.783845](https://doi.org/10.3389/fgene.2021.783845)] [Medline: [35047011](#)]
21. Selya AS, Anshutz D. Machine learning for the classification of obesity from dietary and physical activity patterns. In: Giabbanelli P, Mago V, Papageorgiou E, editors. *Advanced Data Analytics in Health. Smart Innovation, Systems and Technologies*, vol 93. Cham, Switzerland: Springer; 2018:77-97.
22. Taghiyev A, Altun A, Caglar S. A hybrid approach based on machine learning to identify the causes of obesity. *Journal of Control Engineering and Applied Informatics* 2020;22(2):56-66 [[FREE Full text](#)]
23. Thamrin SA, Arsyad DS, Kuswanto H, Lawi A, Nasir S. Predicting obesity in adults using machine learning techniques: an analysis of Indonesian basic health research 2018. *Front Nutr* 2021;8:669155 [[FREE Full text](#)] [doi: [10.3389/fnut.2021.669155](https://doi.org/10.3389/fnut.2021.669155)] [Medline: [34235168](#)]
24. Wang H, Chang S, Lin W, Chen C, Chiang S, Huang K, et al. Machine learning-based method for obesity risk evaluation using single-nucleotide polymorphisms derived from next-generation sequencing. *J Comput Biol* 2018 Dec;25(12):1347-1360. [doi: [10.1089/cmb.2018.0002](https://doi.org/10.1089/cmb.2018.0002)] [Medline: [30204480](#)]
25. AIA Vitality: The Healthiest Workplace. AIA Group. URL: <https://www.aia.com/language-masters/en/en/health-wellness/vitality/healthiest-workplace> [accessed 2022-11-24]
26. Toh B, Akmal A, Syed JS. Malaysia's healthiest workplace: AIA Vitality special report. *The EDGE Malaysia* 2017 Nov 20:1-16.
27. Malaysian Workforce: Sleepless and Overworked? AIA Group. URL: <https://www.aia.com.my/en/about-aia/media-centre/press-releases/2019/malaysian-workforce-sleepless-and-overworked.html> [accessed 2022-11-24]
28. Toh B, Akmal A, Syed JS. Malaysia's healthiest workplace AIA vitality special report: methodology and demographics. *The EDGE Malaysia* 2018 Dec 03:1-16.
29. Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res* 2016 Dec 16;18(12):e323 [[FREE Full text](#)] [doi: [10.2196/jmir.5870](https://doi.org/10.2196/jmir.5870)] [Medline: [27986644](#)]
30. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for reporting machine learning analyses in clinical research. *Circ Cardiovasc Qual Outcomes* 2020 Oct;13(10):e006556 [[FREE Full text](#)] [doi: [10.1161/CIRCOUTCOMES.120.006556](https://doi.org/10.1161/CIRCOUTCOMES.120.006556)] [Medline: [33079589](#)]

31. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. World Health Organization. 1995. URL: https://apps.who.int/iris/bitstream/handle/10665/37003/WHO_TRS_854.pdf [accessed 2022-11-21]
32. Chen KK, Wee S, Pang BWJ, Lau LK, Jabbar KA, Seah WT, et al. Relationship between BMI with percentage body fat and obesity in Singaporean adults - The Yishun Study. *BMC Public Health* 2021 Jun 01;21(1):1030 [FREE Full text] [doi: [10.1186/s12889-021-11070-7](https://doi.org/10.1186/s12889-021-11070-7)] [Medline: [34074272](https://pubmed.ncbi.nlm.nih.gov/34074272/)]
33. Deurenberg-Yap M, Schmidt G, van Staveren WA, Deurenberg P. The paradox of low body mass index and high body fat percentage among Chinese, Malays and Indians in Singapore. *Int J Obes Relat Metab Disord* 2000 Aug;24(8):1011-1017. [doi: [10.1038/sj.ijo.0801353](https://doi.org/10.1038/sj.ijo.0801353)] [Medline: [10951540](https://pubmed.ncbi.nlm.nih.gov/10951540/)]
34. Cheong KC, Yusoff AF, Ghazali SM, Lim KH, Selvarajah S, Haniff J, et al. Optimal BMI cut-off values for predicting diabetes, hypertension and hypercholesterolaemia in a multi-ethnic population. *Public Health Nutr* 2013 Mar;16(3):453-459. [doi: [10.1017/S1368980012002911](https://doi.org/10.1017/S1368980012002911)] [Medline: [22647482](https://pubmed.ncbi.nlm.nih.gov/22647482/)]
35. Zaher ZMM, Zambari R, Pheng CS, Muruga V, Ng B, Appannah G, et al. Optimal cut-off levels to define obesity: body mass index and waist circumference, and their relationship to cardiovascular disease, dyslipidaemia, hypertension and diabetes in Malaysia. *Asia Pac J Clin Nutr* 2009;18(2):209-216 [FREE Full text] [Medline: [19713180](https://pubmed.ncbi.nlm.nih.gov/19713180/)]
36. Aizuddin AN, Chan CM, Anwar AR, Ong YX, Chin K. Performance of body mass index in identifying obesity defined by body fat percentage and hypertension among Malaysian population: a retrospective study. *Int J Gen Med* 2021;14:3251-3257 [FREE Full text] [doi: [10.2147/IJGM.S316360](https://doi.org/10.2147/IJGM.S316360)] [Medline: [34267543](https://pubmed.ncbi.nlm.nih.gov/34267543/)]
37. Kuhn M. Building predictive models in R using the caret package. *J Stat Soft* 2008;28(5):1-26 [FREE Full text] [doi: [10.18637/jss.v028.i05](https://doi.org/10.18637/jss.v028.i05)]
38. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2016; San Francisco, CA p. 785-794. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
39. Breiman L. Random forests. *Mach Learn* 2001;45(1):5-32 [FREE Full text] [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
40. Vapnik VN. The Nature of Statistical Learning Theory. Berlin, Germany: Springer; 1995.
41. Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. *Int J Data Min Knowl Manag Process* 2015 Mar 31;5(2):01-11 [FREE Full text] [doi: [10.5121/ijdkp.2015.5201](https://doi.org/10.5121/ijdkp.2015.5201)]
42. Hosmer JD, Lemeshow S, Sturdivant R. Area under the receiver operating characteristic curve. In: Applied Logistic Regression, Third Ed. Hoboken, NJ: John Wiley & Sons; 2013:173-182.
43. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019 Jun;110:12-22. [doi: [10.1016/j.jclinepi.2019.02.004](https://doi.org/10.1016/j.jclinepi.2019.02.004)] [Medline: [30763612](https://pubmed.ncbi.nlm.nih.gov/30763612/)]
44. Dogbe W, Salazar-Ordóñez M, Gil JM. Disentangling the drivers of obesity: an analytical framework based on socioeconomic and intrapersonal factors. *Front Nutr* 2021 Mar 3;8:585318 [FREE Full text] [doi: [10.3389/fnut.2021.585318](https://doi.org/10.3389/fnut.2021.585318)] [Medline: [33791330](https://pubmed.ncbi.nlm.nih.gov/33791330/)]
45. Hammond R, Athanasiadou R, Curado S, Aphinyanaphongs Y, Abrams C, Messito MJ, et al. Predicting childhood obesity using electronic health records and publicly available data. *PLoS One* 2019;14(4):e0215571 [FREE Full text] [doi: [10.1371/journal.pone.0215571](https://doi.org/10.1371/journal.pone.0215571)] [Medline: [31009509](https://pubmed.ncbi.nlm.nih.gov/31009509/)]
46. Cash T, Smolak L. Body Image: A Handbook of Science, Practice, and Prevention. New York, NY: Guilford Press; 2012.
47. Haynes A, Kersbergen I, Sutin A, Daly M, Robinson E. A systematic review of the relationship between weight status perceptions and weight loss attempts, strategies, behaviours and outcomes. *Obes Rev* 2018 Mar;19(3):347-363 [FREE Full text] [doi: [10.1111/obr.12634](https://doi.org/10.1111/obr.12634)] [Medline: [29266851](https://pubmed.ncbi.nlm.nih.gov/29266851/)]
48. Blake CE, Hébert JR, Lee D, Adams SA, Steck SE, Sui X, et al. Adults with greater weight satisfaction report more positive health behaviors and have better health status regardless of BMI. *J Obes* 2013;2013:291371 [FREE Full text] [doi: [10.1155/2013/291371](https://doi.org/10.1155/2013/291371)] [Medline: [23862054](https://pubmed.ncbi.nlm.nih.gov/23862054/)]
49. Millstein RA, Carlson SA, Fulton JE, Galuska DA, Zhang J, Blanck HM, et al. Relationships between body size satisfaction and weight control practices among US adults. *Medscape J Med* 2008 May 19;10(5):119. [Medline: [18596944](https://pubmed.ncbi.nlm.nih.gov/18596944/)]
50. Kuk JL, Ardern CI, Church TS, Hebert JR, Sui X, Blair SN. Ideal weight and weight satisfaction: association with health practices. *Am J Epidemiol* 2009 Aug 15;170(4):456-463 [FREE Full text] [doi: [10.1093/aje/kwp135](https://doi.org/10.1093/aje/kwp135)] [Medline: [19546153](https://pubmed.ncbi.nlm.nih.gov/19546153/)]
51. Tiggemann M. Body image across the adult life span: stability and change. *Body Image* 2004 Jan;1(1):29-41. [doi: [10.1016/S1740-1445\(03\)00002-0](https://doi.org/10.1016/S1740-1445(03)00002-0)] [Medline: [18089139](https://pubmed.ncbi.nlm.nih.gov/18089139/)]
52. Quittkat HL, Hartmann AS, Düsing R, Buhlmann U, Vocks S. Body dissatisfaction, importance of appearance, and body appreciation in men and women over the lifespan. *Front Psychiatry* 2019;10:864 [FREE Full text] [doi: [10.3389/fpsy.2019.00864](https://doi.org/10.3389/fpsy.2019.00864)] [Medline: [31920737](https://pubmed.ncbi.nlm.nih.gov/31920737/)]
53. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Presented at: The 31st International Conference on Neural Information Processing Systems; Dec 4-9, 2017; Long Beach, CA.

Abbreviations

AUC: area under the curve
LR: logistic regression
ML: machine learning
RF: random forest
ROC: receiver operating characteristic curve
SVM: support vector machine
XGBoost: extreme gradient boosting

Edited by A Mavragani; submitted 20.06.22; peer-reviewed by V Kumar, M Rocchetti; comments to author 30.08.22; revised version received 09.10.22; accepted 11.10.22; published 07.12.22.

Please cite as:

Wong JE, Yamaguchi M, Nishi N, Araki M, Wee LH

Predicting Overweight and Obesity Status Among Malaysian Working Adults With Machine Learning or Logistic Regression: Retrospective Comparison Study

JMIR Form Res 2022;6(12):e40404

URL: <https://formative.jmir.org/2022/12/e40404>

doi: [10.2196/40404](https://doi.org/10.2196/40404)

PMID: [36476813](https://pubmed.ncbi.nlm.nih.gov/36476813/)

©Jyh Eiin Wong, Miwa Yamaguchi, Nobuo Nishi, Michihiro Araki, Lei Hum Wee. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

An Emotional Bias Modification for Children With Attention-Deficit/Hyperactivity Disorder: Co-design Study

Melvyn Zhang¹, MBBS; Vallabhajosyula Ranganath², PhD

¹Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore

²Anatomy, Office of Medical Education, Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore, Singapore

Corresponding Author:

Melvyn Zhang, MBBS

Family Medicine and Primary Care

Lee Kong Chian School of Medicine

Nanyang Technological University Singapore

11 Mandalay Road

Level 11, Family Medicine and Primary Care

Singapore, 308322

Singapore

Phone: 65 98556631

Email: melvynzhangweibin@gmail.com

Abstract

Background: Attention-deficit/hyperactivity disorder (ADHD) is one of the common neurodevelopment disorders. Children with ADHD typically have difficulties with emotional regulation. Previous studies have investigated the assessment for underlying emotional biases using the visual probe task. However, one of the significant limitations of the visual probe task is that it is demanding and repetitive over time. Previous studies have examined the use of gamification methods in addressing the limitations of the emotional bias visual probe task. There has also been increased recognition of the potential of participatory action research methods and how it could help to make the conceptualized interventions more relevant.

Objective: The primary aim of this study was to collate health care professionals' perspectives on the limitations of the existing visual probe task and to determine if gamification elements were viable to be incorporated into an emotional bias modification task.

Methods: A co-design workshop was conducted. Health care professionals from the Department of Development Psychiatry, Institute of Mental Health, Singapore, were invited to participate. Considering the COVID-19 pandemic and the restrictions, a web-based workshop was conducted. There were 3 main phases in the workshops. First, participants were asked to identify limitations and suggest potential methods to overcome some of the identified limitations. Second, participants were shown examples of existing gaming interventions in published literature and commercial stores. They were also asked to comment on the advantages and limitations of these interventions. Finally, participants were asked if gamification techniques would be appropriate.

Results: Overall, 4 health care professionals consented and participated. Several limitations were identified regarding the conventional emotional bias intervention. These included the nature of the task parameters, included stimulus set, and factors that could have an impact on the accuracy of responding to the task. After examining the existing ADHD games, participants raised concerns about the evidence base of some of the apps. They articulated that any developed ADHD game ought to identify the specific skill set that was targeted clearly. Regarding gamification strategies, participants preferred economic and performance-based gamification approaches.

Conclusions: This study has managed to elucidate health care professionals' perspectives toward refining a conventional emotional bias intervention for children with ADHD. In view of the repetitiveness of the conventional task, the suggested gamification techniques might help in influencing task adherence and reduce the attrition rates.

(*JMIR Form Res* 2022;6(12):e36390) doi:[10.2196/36390](https://doi.org/10.2196/36390)

KEYWORDS

emotional bias; cognitive biases; attention-deficit/hyperactivity disorder; ADHD; child psychiatry

Introduction

Background

Attention-deficit/hyperactivity disorder (ADHD) is one of the common neurodevelopment disorders. Children with ADHD typically display symptoms such as hyperactivity, impulsivity, and inattention [1]. These individuals have inherent underlying difficulties with emotional regulation [2]. The advances in experimental psychology have led to better understanding of emotional biases and methods by which these biases could be modified [3]. Previous studies have reported that emotional biases are most prevalent among those with the combined subtype of ADHD. Children with this subtype tend to have difficulties in comprehending the emotional states of others and recognizing facial emotions and emotional cues [2]. Previous studies have described how the visual probe or dot probe task is applied in the assessment and potentially in modifying these biases. In the visual probe task, individuals are presented with 2 stimuli on the screen simultaneously in the assessment phase. The presented stimuli are images showing different emotional cues, for example, an angry face or a neutral expression. The set of stimuli would disappear, and a probe would replace either stimulus. Individuals are required to indicate the position of the probe as rapidly as possible. For purposes of assessment, individuals may have positive emotional biases when they respond more readily to probes that replace the emotional stimulus instead of the neutral stimulus. For bias modification, the contingency of which stimulus is being replaced could be altered in a way that is similar to the application of the visual probe task for other psychiatric disorders [3]. Although there have been previous studies examining the effectiveness of cognitive bias modification for anxiety and depression among children and adolescents with other psychiatric disorders, for example, anxiety and depression [4,5], pioneering studies have been conducted in exploring emotional biases in the same population. In a recent study by Cremone et al [3], the authors reported the existence of emotional biases among children and adolescents with ADHD, and the amount of sleep affected the magnitude of the underlying emotional biases.

Although the visual probe task appears to be a relatively simple task to administer for the measurement of underlying emotional biases, and potentially for modifying biases, it is not without its inherent limitations. The numerous repeats that an individual needs to complete make such an intervention laborious. In recent years, serious games and gamification technologies have been considered for conventional psychological interventions [6,7]. It is believed that the use of these technologies would enhance engagement in the short and long term, promote self-empowerment, and improve existing skills [6]. As highlighted by Zhang et al [8], gamification, when applied to cognitive bias modification interventions, could help to reduce the repetitiveness of the game and increase motivation to train. As evident from previous study by Zhang et al [9], which involved conducting a series of participatory action research workshops involving health care professionals and patient service users, gamification could be used to address the limitations of the visual probe task and enhance the conventional task.

More recently, regarding ADHD, there have been further studies examining the evidence base of existing games for individuals with ADHD and the potential of gamification for serious games targeting individuals with ADHD. Jiang et al [10] conducted a scoping review of the currently available mobile games on several databases, and they reported that most of the existing games were focused on managing individuals with symptoms and less so on symptoms and diagnosis. Although the studies have shown an improvement in the performance of the children across the interventions, these studies were limited in several aspects, such as the inclusion of a small sample size and lack of a control group for comparison, and these factors influence the overall effectiveness of the study [10]. Nonetheless, the 19 studies highlighted the potential of considering a serious game approach when conceptualizing new interventions, and they also revealed that the most common techniques were that of having participants to respond to various cues, remembering details, or making association between different entities [10]. In another recent study by Sujar et al [11], the authors, having synthesized the evidence from literature, reported that future games that are designed for individuals with ADHD ought to consider the following aspects: the underlying mechanics of the game ought to be based on some form of cognitive exercise and therapeutic strategies that may be helpful include having levels of difficulties, a motivational element, time constraint for the task one has to undertake, and some form of reinforcement [11]. The findings reported by Sujar et al [11] are consistent with the aims and objectives of this study, which, as elaborated subsequently, aimed to seek the perspectives of health care professionals as an integral step in co-designing an intervention to modify emotional biases in children.

Objectives

There has also been increased recognition of the potential of participatory action research methods and how it could help to make the conceptualized interventions more relevant [12]. Therefore, the aim of this study was to use such a method in refining the conventional emotional bias visual probe task for children and adolescents. The primary aim of this study was to collate health care professionals' perspectives on the limitations of the existing visual probe task and to determine if gamification elements were viable to be incorporated into an emotional bias modification task. We sought to answer the following research questions: (1) What were health care professionals' perspectives on the conventional emotional bias task? (2) What were health care professionals' perspectives on existing gaming interventions for ADHD? (3) Would gamification be appropriate, and if appropriate, what strategies could be used?

Methods

Study Design

Principles of participatory action research were used for this study. A co-design workshop was conducted, and relevant key stakeholders (ie, health care professionals) were invited. Only health care professionals were invited, as they had knowledge of the psychiatric conditions and were best able to advise how the task could be modified, while adhering to the evidence base.

Study Setting

Health care professionals from the Department of Development Psychiatry, Institute of Mental Health, Singapore, were invited to participate. A diverse group of health care professionals was invited, and it included both psychiatrists and psychologists. We had originally planned to recruit occupational therapists also, but none of them expressed an interest in participating in the study.

Sample Size

On the basis of our previous protocol [13], we planned to recruit 8 participants. We managed to recruit 50% (4/8) of the projected participants for the focus group. However, there was a reduction in the number of participants recruited, mainly in part owing to the ongoing COVID-19 pandemic and the difficulties for individuals to commit to research studies as they have other clinical or administrative roles.

Details of the Co-design Workshops

Owing to the current COVID-19 pandemic and local governmental restrictions, we applied for the conduct of the workshop via the web. Therefore, all participants (4/4, 100%) who have had expressed interest were contacted separately to sign the informed consent form. The participants also had to complete a baseline demographic questionnaire individually. This questionnaire collated information regarding age, sex, and years of experience in treating children and adolescents with psychiatric disorders. Upon completion of the questionnaire, the principal investigators liaised with each of the participants separately to identify a common time for conducting the web-based workshop.

The workshop was subsequently conducted on May 19, 2021, via the web. Both principal investigators facilitated the workshop

and recorded the field notes. All participants (4/4, 100%) were informed of the study's rationale and the session's specific objectives. Participants were also informed that their responses and comments were confidential and that the session would be audio-recorded. All participants (4/4, 100%) were given participant numbers to identify themselves, to provide further anonymity. First, participants were shown an example of the visual probe task paradigm that has been traditionally used for cognitive bias modification. The example that participants viewed was based on the previous study by Zhang et al [9]. In that study, the specific nature of cognitive bias modification was that of attention bias modification, and it was applied to individuals with addictive disorders. Then, participants were shown how the traditional visual probe task paradigm is modified to become an emotional bias task paradigm (by modifying the visual cues presented). Then, they were asked to identify limitations and suggest potential methods to overcome some of the identified limitations regarding the emotional bias modification intervention. Subsequently, participants were shown examples of existing gaming interventions in published literature and commercial stores. Figure 1 provides an overview of some of the games that were shown to the participants.

They were asked to comment on the advantages and limitations of these interventions. Finally, participants were shown a list of gamification techniques, and each of the techniques was explained to them. The list of gamification techniques was based on the previous study by Hoffman et al [14]. The list of gamification techniques is described in Table 1, and this has also been published in the previous study by Zhang et al [9]. They were asked if the inclusion of such techniques would be appropriate.

Figure 1. Examples of games shown to participants.

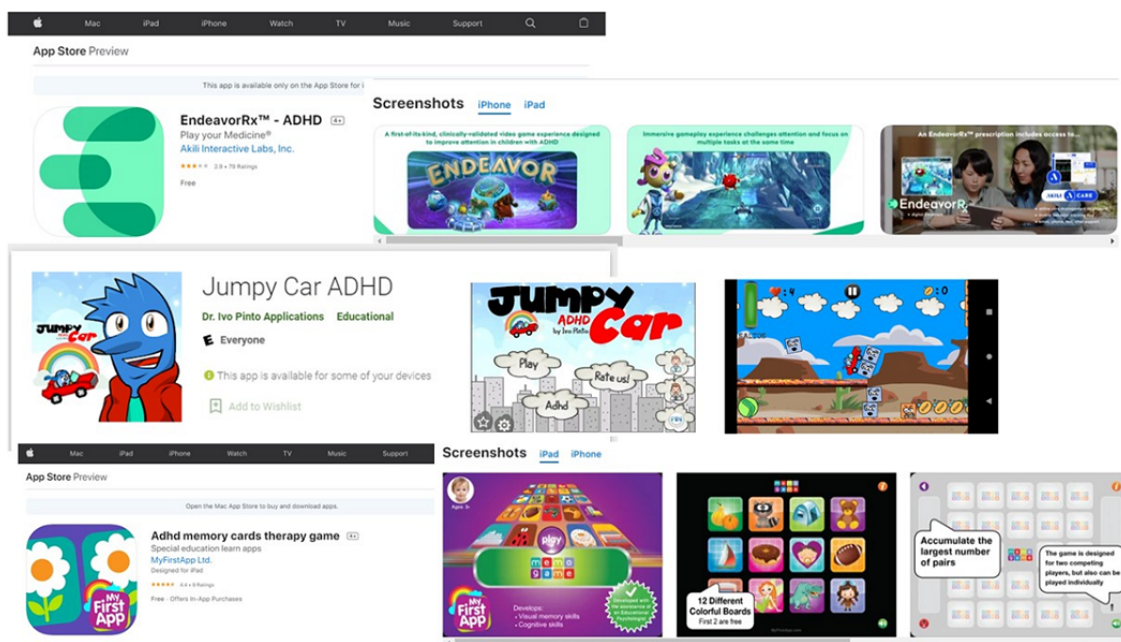


Table 1. Gamification techniques that were shown to participants.

Gaming approach	Description
Economic gamification techniques	
Marketplace and economies	Providing gamers with a web-based currency that allows them to deal in the game
Digital rewards	These include badges, game currency, game points, web-based goods, and powers or abilities
Real-world prizes	Provides gamers with options to exchange in-game credits for real-world prizes such as vouchers or other forms of goods and services
Social gamification techniques	
Avatar	Allows individuals to choose a web-based character to represent oneself
Agent	A web-based character that guides or provides instructions to the user
Competition	Allows individuals to compete with other players or with each other
Teams	Game that involves several individual players, allowing them to interact and form relationships
Parallel communication systems	Allows individuals to communicate with one another
Social pressure	Ability of game to pressurize individuals to perform in certain task, so that they will be invited to subsequent events
Performance-orientated techniques	
Feedback	Spoken, visual, or auditory feedback about user's performance
Levels	Information on the stage of a game one has attained
Secondary game objectives	Secondary goals that reward the player upon completion
Ranks of achievement	Measurement of character development
Leaderboards	Allows for comparisons with other players
Time pressure	Predetermined time limits for task completion
Embedding-focused techniques	
Narrative context	A storyboard or stories that guide the development of the character
3D environment	3D models of objects that parallel the real world

Data Analyses

The descriptive data (demographics) were summarized as means and SDs. The audio recording obtained from the workshop was transcribed verbatim. One of the principal investigators, MZ, performed the first transcription and developed the coding frame. To ensure the reliability of the coding frame that was adopted, two authors (MZ and RV) reviewed the transcripts and discussed the coding frame. This ensured that the process of intercoder consensus was adhered to. Codes that were identified were classified into categories and then reorganized into themes. The themes that were generated were subsequently reviewed and further refined. The underlying methodology used was in accordance with the previous recommendations by Braun and Clarke [15]. NVivo (version 12.0; QSR International) [16] was used for thematic analysis.

Ethics Approval

This study was approved by the ethical review board of the Nanyang Technological University Singapore (approval number IRB-2020-03-058). Informed consent was obtained from all the participants, and they were also informed that they could withdraw from the study at any time.

Data Management

No participant-related identifiers were captured on the hard-copy questionnaires. All the completed hard-copy forms and informed consent forms were stored in a secured facility. The audio recordings of the workshop were transferred to a local secured computer for storage, and the original recording was deleted from the recording device. The password of the local computer was changed frequently, and only the principal investigators were able to access the files. All the records and audio recordings will be maintained for a period of 6 years following the completion of the study. All participants were provided an inconvenience fee for their time and effort in participating in the study.

Results

Demographics

A total of 4 health care professionals consented and participated in the workshop, which was conducted on May 19, 2021. Of the 4 health care professionals, 1 (25%) was a child and adolescent psychiatrist and the remaining 3 (75%) were psychologists. The mean age was 43.5 (SD 5.74) years. Of the 4 participants, 1 (25%) was a man and the remaining 3 (75%) were women. The mean years of experience was 14.5 (SD 4.43)

years, ranging from 10 to 20 years. [Table 2](#) provides an overview of the baseline demographics of the participants.

Table 2. Overview of the baseline demographics of the participants (N=4).

Participant number	Age (years)	Nationality	Sex	Experience (years)	Race
1	52	Singaporean	Male	20	Chinese
2	42	Singaporean	Female	16	Chinese
3	40	Singaporean	Female	10	Chinese
4	40	Singaporean	Female	12	Indian

Phase 1

In phase 1, the participants identified several limitations regarding the nature of the conventional visual probe task used to assess emotional biases. These limitations were related to the nature of the task parameters, included stimulus set, and other

factors that could affect the accuracy of responding to the task. Participants also made recommendations regarding how the task could be improved. [Textbox 1](#) provides a summary of the verbatim comments of the participants for each of the identified themes.

Textbox 1. Themes related to the nature of the conventional visual probe task.

Theme 1

- Limitations—Issues related to task parameters
- Verbatim comments
 - “I suppose, theoretically, it sounds alright. I am a bit concerned about the speed at which the task goes.” [Participant 2]
 - “The design of the task, flashing multiple images in quick succession, can be distressing and could cause fatigue.” [Participant 1]
 - “I think the number of trials is ok. I think the speed I am not comfortable with.” [Participant 3]
 - “I am wondering about the practice trial. There are 8 trials. Do they go at the same speed as the actual 32? I don’t know whether existing emotional bias trials. When it is applied to children and adolescent. Could the speed be moderated? Or you could do half half. 8 which is lower speed to aid understanding the next 8 is at actual speed.” [Participant 2]
 - “I think just now my mind was a blank at the speed. Wow. What was there. Even I know what the instructions were and what I am supposed to do, I was not able to process when I saw the program.” [Participant 4]

Theme 2

- Limitations—Stimulus set
- Verbatim comments
 - “The other thing I wonder whether erm. In terms of the pictures, would it make a difference if you are used Asian versus Caucasian pictures. Would it make a difference?” [Participant 4]
 - “I am just wondering about the size of the photos. Some of the photos were quite big. Because the speed was quite fast, all I saw was a blur of colours. For this current picture, there was white space. For the previous example, all I saw was a blur of colours.” [Participant 4]
 - “I just want to echo the other respondent. If you could frame the picture like in the slide now, it homes the picture better. Rather than big white background. Black border, framing into a small visual field, might be more user-friendly.” [Participant 1]

Theme 3

- Limitation—Other confounders that may affect the accuracy of the task
- Verbatim comments
 - “Whether the speed would trigger impulsivity and so. Perhaps there is a confounder. Impulsivity or desire to put in an answer. Might confound with reaction time in measuring attentional biases. This is my gut feel. I do not know the research behind it.” [Participant 2]
 - “Children with ADHD do have comorbidities. Some of them have processing speed abnormalities. Could confound. Going back to your inclusion criteria you might want to think about.” [Participant 3]
 - “I am concerned about impulsivity. Because it is going so far, I suspect that it is still going to be quite novel for kids with ADHD. But novelty aside, I think if they don’t know what they are doing, they just guess. You probably get a big bias which is not what you are looking for.” [Participant 2]

Theme 4

- Recommendations pertaining to presentation of the stimulus set
- Verbatim comments
 - “For this current picture, there was white space. For the previous example, all I saw was a blur of colours.” [Participant 4]
 - “I just want to echo the other respondent. If you could frame the picture like in the slide now, it homes the picture better. Rather than big white background. Black border, framing into a small visual field, might be more user-friendly.” [Participant 1]

Phase 2

In this phase, participants were asked to provide their perspectives about the existing ADHD games. Several themes arose from the discussion. Participants highlighted that the games needed to be age-appropriate and clearly explain the specific skill sets that were targeted. Participants also raised

concerns regarding the scientific evidence base for some of the presented games. Participants also highlighted issues with existing games regarding their novelty and stated that the consideration of novelty was important when implementing games for individuals with ADHD.

Textbox 2 provides a summary of the verbatim comments of the participants for each of the identified themes.

Textbox 2. Verbatim comments of the participants for each of the identified themes.

Theme 1

- Being age-appropriate
- Verbatim comment
 - “Erm. I think these games appear to cater to different age groups. The animations. Some of them appear more sophisticated than other. For example, the first game appeared more sophisticated and complex. The third one seems to cater to younger population. So I think we need to be aware of how appealing it is to each age group of the child is in.” [Participant 1]

Theme 2

- Skill sets targeted
- Verbatim comments
 - “The second one, Jumpy car seems to be psychoeducational. Comprehensive package of game and psychoeducation. For psychoeducation, the second games seemed to be having more of that. For the other two, it seems to focus on skills and skills building. That is my first and most immediate observations of these games.” [Participant 1]
 - “The other thing I noted for the second one. Between the games they give you psycho education.” [Participant 4]
 - “I do quite like Jumpy car ADHD. I don’t have ADHD but I thought it is a good tool to engage them and helping them to understand what ADHD is. Of course, I have checked out the language and whether it is child friendly. Not sure about this memory card game. not sure what it was targeting about.” [Participant 2]
 - “This is number 3. I also like the jumpy car and I do have ADHD. I can’t download it. Just had a thought. Just now you mentioned about combining. Jumpy car just seemed to get through all the distractors and getting to the point destination. Wondering if you could combine that frustration with not being able to cross the hurdles. So, the distractions. When they cannot cross the distractions, they are somehow emotional. So, one way may be is to add on to that. Does that make sense? Instead of getting from A to B, that it. Just to get rid of the distractor or to recognise the distractions. Teach them to recognise their emotions. A cognitive component to it.” [Participant 3]
 - “Just wondering about the objective of doing emotional bias modification game for children with ADHD. So, what are we hoping the kids with kids with ADHD would become after the end of the game?” [Participant 2]
 - “Mine is along the same line. What is the game setup to do. Is it going to be specific for emotional biases, capturing emotional biases. So the game, how it is designed, needs to tackle that. It needs to be valid and also be specific. Intending to assess and moderate. Teach skills on how to minimise the emotional biases. Some of the games shared seemed to focus on different skill sets. Jumpy car was to ignore distractions. So there was actually no emotion. Not much emotional component to that. And then the third one. About ADHD memory game. More like on working memory. They may be. My sense is that they are targeting different deficiencies seen in ADHD. I think the game design would have to include situations in which they could trigger emotional reactions. Frustrations or disappointments. We do understand from past research. The brains of the child is different from adult. Different parts of the child brain is being activated when given a similar task, compared to adult brain. And there is also less. They also have more problem, for teenagers especially, in recognising neutral emotions, they tend to over interpret neutral faces as anger. If you want to have a game that addresses emotional biases, then we have to also take on this consideration as well. The game must be specific to you know. To target emotional situations for example.” [Participant 1]
 - “I think I may have misunderstood what you have asking. I think it is viable to convert to game. But like what participant 1 said, it has to be a situation that evokes the emotions. It is rather challenging now to kind of. Scope it in games, especially for our gen Z population right now. Just having situations taking example, game I, this is not going to really elicit that. Back to the previous point where you presented with ADHD games. Different ADHD games train different functions. Memory game would be working memory. Jumpy car would be impulsive control, or distraction. If you could add that emotional component to existing game. Would it be adding another layer? That is what I am trying to say? They run into situations and they do react. If there is a psycho edu component, what you can do about it and what you could do about it. Do I make sense now?” [Participant 3]
 - “I am also wondering how the kids can play this and how this can be translated to real life. Because a lot of times they play the game, the skills stay in the game. When I am out in the real world, everything in the game, the skills I learnt does not apply in the real world. Some way for the kids to know that the skills they applied can be applied to the real world. Maybe like a short film clip or someone demonstrating that skills can be used. Kids should know that this is something that can be used in the real world.” [Participant 4]

Theme 3

- Scientific evaluations
- Verbatim comment
 - “I suppose I am wondering that in terms of games that have been subjected to research evaluations, in terms of the follow-up period, whether the improvement have been sustained. I know that endeavour RX has some trials back. Data sustained attention. What sort of ratings are being used during the trial. We have a few of attention improvement games. We also try to include blinded and objective rating for example measuring brain waves to measure whether there are actual changes. I wonder about the follow-up period, sustainability in terms of improvement and whether there are any objective measures for these kinds of game. Ditto for those who does not have any research behind.” [Participant 2]

Theme 4

- Novelty and motivation to use
- Verbatim comments
 - “The other thing I noted for the second one. Between the games they give you psycho edu. If there is a skip function, I suspect that a lot of kids would skip and go to the game. They might not listen to it at all. If it is too repetitive and increasing in difficulty, kids would just give up especially when the novelty wears out. They try out the first few stages, but it is getting tougher, I just skip to something else.” [Participant 4]
 - “I am going to start that I don’t know enough about EB games. I wonder about. If we have a monkey going down the supermarket, ignoring negative emotional faces. number 1. I am thinking about novelty. After a while ADHD kids would lose interest. Number 2 is it going to have an impact on emotional dysregulation. Games are seemingly more friendly platform to engage kids with ADHD versus paper and pencil. There is definitive value in gamifying it. Effectiveness whether it does modify EB. How do we make it more interesting? Stage base. Can’t be monkey going down the aisle. They are going to get bored soon.” [Participant 3]

Phase 3

In this phase, the facilitator explained to the participants more about the common gamification strategies that have been used. Then, participants were asked to select the most appropriate gamification strategies that could be applied to the conventional

emotional bias modification task. Strategies such as economic and performance-based gamification were deemed to be more appropriate.

Textbox 3 provides a summary of the gamification techniques that the participants have selected and their justifications for their selection.

Textbox 3. Gamification techniques selected by the participants and their justifications for their selection.

Theme 1

- Economic approaches
- Verbatim comments
 - “Like in my work with children and parents with ADHD. Often talk about how children with ADHD are motivated by 3 different factors. First novelty, second interest and third competition. I am looking at the gaming approaches. Things like having economic gamification approach, having digital rewards or real-world rewards might be one.” [Participant 2]
 - “I think digital rewards will be helpful. A lot of kids have problems with delayed gratification. Some form of immediate reward would be helpful.” [Participant 1]
 - “I agree with what has been said as well. And also, I think the kids are motivated when they can buy some. They like to browse the store and see what they can buy.” [Participant 4]
- “I have been taking about novelty. Real world prizes might be more relevant. After a while, digital rewards might not be too attractive for them.” [Participant 2]

Theme 2

- Performance-based approaches
- Verbatim comments
 - “The other thing that I am looking at is performance orientated gaming approach. Where there is some sort of competition. Mindful that it is not too hard for the child. That the child is being ranked very far below. Competition, having that interest, economic gamification. Those two are ones to go with.” [Participant 2]
 - “A lot of local kids have a competitive edge as well. Having a game, allowing them to level up. A lot of reinforcement, feedback that reinforces the child behaviour would fare better.” [Participant 1]
 - “I agree with 1 and 2. The performance one, like what 2 said. Ranking low might result in them losing motivation.” [Participant 3]

Theme 3

- Inappropriate strategies
- Verbatim comment
 - “I have been taking about novelty. Real world prizes might be more relevant. After a while, digital rewards might not be too attractive for them. You asked about which techniques are not suitable. I do wonder about the team part. Where they play with other people or interact with other community. Some of the children I work with have bad experiences gaming in teams, where they have social conflicts. We do see some form of social conflicts in children with ADHD. That a question mark for me. There are ones who go onto form close knitted communities?” [Participant 2]

Theme 4

- Risk of gamification
- Verbatim comments
 - “I generally do not have any concerns about gamification as treatment modality. My concern is how accessible it is and its affordability as well. I think gamification has the benefit of allowing participant to play the game at his own time, under the supervision of caregiver or parent. Therapy could take place outside of the clinics. Probably more about accessibility and affordability. Probably more research studies to show that It works over the long term.” [Participant 1]
 - “I would always caveat it. It must be supervised and moderated by adults. ADHD population has problem moderating screen time.” [Participant 2]

Discussion

Principal Findings

This is the first study that has explored the use of participatory action research as a research design to refine the conventional emotional bias modification task. This study complements the ongoing research that has been highlighted in the *Introduction*

section, which demonstrated the presence of emotional biases among children and adolescents with ADHD. In our study involving health care professionals, several limitations were identified regarding the conventional emotional bias intervention. These included the nature of the task parameters, included stimulus set, and factors that could have an impact on the accuracy of responding to the task. After examining the existing ADHD games, participants raised concerns about the

evidence base of some of the apps. They articulated that any developed ADHD game ought to clearly identify the specific skill set that was targeted. When health care professionals were eventually asked to select gamification strategies that could improve the conventional intervention, they preferred economic and performance-based gamification approaches.

Some of the identified limitations of the existing conventional visual probe paradigm used for the assessment of emotional biases are congruent with the findings of previous studies. In our study, health care professionals identified limitations pertaining to the task parameters, particularly that which pertains to the speed of presentation of the stimulus set. This is congruent with the previous co-design workshop, by Zhang et al [9], involving health care professionals and patient participants, who similarly reported that the presentation of the stimulus was very rapid. In that study, the participants recommended a lengthy stimulus presentation time to allow them to process the set of stimulus images [9]. On the basis of the review by Zhang et al [9] about the paradigms for the visual probe task as applied for addictive disorders, it would be of important to have stimulus presented for both short and long stimulus intervals. This allows for assessing both attentional processes, namely, initial orientation and delayed disengagement. This should be considered when developing the task paradigm for assessing emotional biases among adolescents. Another limitation that was identified pertains to the nature of the images presented. Health care professionals recommended for standardization in the size of the images, whether images are presented with or without borders, and how much these presented images contrasted against the background. These identified limitations are crucial and need to be considered when designing the next iteration of apps that assess emotional biases.

One of the concerns raised by health care professionals about existing ADHD games was whether these games were based on validated frameworks and adhered to the evidence base. In our workshop, we presented participants with examples of ADHD games that are commercially available, some of which have been previously evaluated. In a recent review by Penuelas-Calvo et al [17], they examined video games for the assessment and intervention of individuals with ADHD. A total of 22 papers were identified, and they reported that the existing tools were influential in establishing whether individuals do have attentional issues as compared with the control. Unfortunately, the review was published after we conducted our focus group. Otherwise, the data obtained from the review would have helped to advance the discussion. Another concern highlighted during our focus group was the need for games to be specific in terms of identifying the skill sets they sought to target. In the examples we shared, it appeared that some of the commercially available apps were specific in terms of what they wanted to develop. Penuelas-Calvo et al [17], in their examination of the literature, reported that most of their identified video games have been developed based on previously validated tasks, for example, the performance task by Corner and the go-no-go task. Regarding video games that were interventional in nature, most of them focused on cognitive training, such as improving executive functioning, attention or working memory, reaction time, cognitive flexibility, or motor

ability. The responses obtained from our focus group and the findings by Penuelas-Calvo et al [17] further highlight the importance of computerized interventions on previously validated evidence-based tasks. It is also important to be mindful that when adapting the task to a computerized intervention or even a video game, the mechanism of the conventional task should not be altered.

Our participants also reported various gamification strategies that may benefit our emotional bias assessment task. Previous study has justified the importance of considering gamification strategies, given that they help with the improvement of engagement rates and reduction of dropouts [6]. The perceptions of our team of participants in the co-design workshop indicated that repetitive activities in the game with increased difficulty levels would potentially affect the performance, continuity, interest, and novelty of the game. Novelty, user-friendliness, and interactive game design concepts should be considered, so that the ADHD games can be made engaging effectively. One of the critical perspectives shared during the workshop was the need for clear instructions and directions in the gameplay. Otherwise, it is likely that children with ADHD may be frustrated with the game. Regarding the gamification strategies suggested, leader boards, digital rewards, and real-world prizes could engage the users and allow them to play the intervention multiple times. Our participants also advocated the consideration of economic approaches, given that they offer individuals a tangible, immediate reward. Despite these suggestions, we need to acknowledge that one of the major limitations of our study was that we have not considered the perspectives of children themselves and those of their caregivers. It is important to understand the perspectives of children, so that the intervention could be personalized to their needs. Although we agree with the previous recommendations by Penuelas-Calvo et al [17] that health care professionals should collaborate with computer engineers, we feel that apart from collaborating with an engineer to ensure a high-quality app that resembles the quality of commercial apps, it is far more critical to ensure that the app is personalized to the needs of the patients.

The main strength of this study was the use of co-design methods in the refinement of an existing evidence-based paradigm for the assessment of emotional biases. This helped to ensure that the eventual design is based on evidence but meets the potential needs of end users. This study also presented examples of existing ADHD apps to participants. In doing so, we were able to allow the participants to have a better in-depth understanding of existing apps and identify issues and limitations with existing apps. Despite these strengths, our study had several limitations. Although we initially planned to recruit a total of 8 participants, we managed to eventually recruit only 4 (50%) participants. The COVID-19 pandemic affected our ability to recruit participants, as they may have to attend to other clinical needs. In addition, although we had previously planned for a physical workshop, the COVID-19 pandemic prevented the execution of such a workshop owing to infection risks. Thus, we were limited to conducting a web-based workshop, which may have resulted in challenges among participants in responding. It would be ideal to also obtain insights from patient participants or their caregivers.

Conclusions

This study has managed to elucidate health care professionals' perspectives on the refinement of a conventional emotional bias intervention for individuals with ADHD. Taken together, gamification strategies could be applied to conventional emotional bias interventions. The findings from this study have

implications on the subsequent studies seeking the personalization and gamification of such apps. Although we have sampled the perspectives of health care providers, it remains necessary to discuss these perspectives with the intended sample group as the next step, to ensure that the future conceptualized app will be consistent with their needs.

Acknowledgments

MZ is supported by a grant under the Singapore Ministry of Health's National Medical Research Council (grant NMRC/Fellowship/0048/2017) for PhD training. The funding source was not involved in any part of this project. The project is funded by the Games for Health Innovation Centre (ALIVE) Serious Games Grant (SGG19/SN06), with the grant project titled *Gamified Emotional Bias Modification Intervention for Children with ADHD*. This study was made possible by a gift from the Estate of Irene Tan Liang Kheng.

Data Availability

All available data have been included in the manuscript.

Authors' Contributions

MZ and RV jointly conceptualized the study. MZ wrote the initial draft, which was revised by RV. All authors read and approved the manuscript before submission.

Conflicts of Interest

None declared.

References

1. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *Int J Epidemiol* 2014 Apr;43(2):434-442 [[FREE Full text](#)] [doi: [10.1093/ije/dyt261](https://doi.org/10.1093/ije/dyt261)] [Medline: [24464188](#)]
2. Shaw P, Stringaris A, Nigg J, Leibenluft E. Emotion dysregulation in attention deficit hyperactivity disorder. *Am J Psychiatry* 2014 Mar;171(3):276-293 [[FREE Full text](#)] [doi: [10.1176/appi.ajp.2013.13070966](https://doi.org/10.1176/appi.ajp.2013.13070966)] [Medline: [24480998](#)]
3. Cremona A, Lugo-Candelas CI, Harvey EA, McDermott JM, Spencer RM. Positive emotional attention bias in young children with symptoms of ADHD. *Child Neuropsychol* 2018 Nov 18;24(8):1137-1145 [[FREE Full text](#)] [doi: [10.1080/09297049.2018.1426743](https://doi.org/10.1080/09297049.2018.1426743)] [Medline: [29347861](#)]
4. Orchard F, Apetroaia A, Clarke K, Creswell C. Cognitive bias modification of interpretation in children with social anxiety disorder. *J Anxiety Disord* 2017 Jan;45:1-8 [[FREE Full text](#)] [doi: [10.1016/j.janxdis.2016.10.012](https://doi.org/10.1016/j.janxdis.2016.10.012)] [Medline: [27866085](#)]
5. de Voogd EL, de Hullu E, Burnett Heyes S, Blackwell SE, Wiers RW, Salemink E. Imagine the bright side of life: a randomized controlled trial of two types of interpretation bias modification procedure targeting adolescent anxiety and depression. *PLoS One* 2017;12(7):e0181147 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0181147](https://doi.org/10.1371/journal.pone.0181147)] [Medline: [28715495](#)]
6. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. *Front Psychiatry* 2017;7:215 [[FREE Full text](#)] [doi: [10.3389/fpsy.2016.00215](https://doi.org/10.3389/fpsy.2016.00215)] [Medline: [28119636](#)]
7. Strahler Rivero T, Herrera Nuñez LM, Uehara Pires E, Amodeo Bueno OF. ADHD rehabilitation through video gaming: a systematic review using PRISMA guidelines of the current findings and the associated risk of bias. *Front Psychiatry* 2015 Oct 22;6:151 [[FREE Full text](#)] [doi: [10.3389/fpsy.2015.00151](https://doi.org/10.3389/fpsy.2015.00151)] [Medline: [26557098](#)]
8. Zhang M, Ying J, Song G, Fung DS, Smith H. Gamified cognitive bias modification interventions for psychiatric disorders: review. *JMIR Ment Health* 2018 Oct 25;5(4):e11640 [[FREE Full text](#)] [doi: [10.2196/11640](https://doi.org/10.2196/11640)] [Medline: [30361194](#)]
9. Zhang M, Heng S, Song G, Fung DS, Smith HE. Co-designing a mobile gamified attention bias modification intervention for substance use disorders: participatory research study. *JMIR Mhealth Uhealth* 2019 Oct 03;7(10):e15871 [[FREE Full text](#)] [doi: [10.2196/15871](https://doi.org/10.2196/15871)] [Medline: [31584003](#)]
10. Jiang H, Natarajan R, Shuy YK, Rong L, Zhang MW, Vallabhajosyula R. The use of mobile games in the management of patients with attention deficit hyperactive disorder: a scoping review. *Front Psychiatry* 2022;13:792402 [[FREE Full text](#)] [doi: [10.3389/fpsy.2022.792402](https://doi.org/10.3389/fpsy.2022.792402)] [Medline: [35308884](#)]
11. Sújar A, Martín-Moratinos M, Rodrigo-Yanguas M, Bella-Fernández M, González-Tardón C, Delgado-Gómez D, et al. Developing serious video games to treat attention deficit hyperactivity disorder: tutorial guide. *JMIR Serious Games* 2022 Aug 01;10(3):e33884 [[FREE Full text](#)] [doi: [10.2196/33884](https://doi.org/10.2196/33884)] [Medline: [35916694](#)]

12. Zhang MW, Ying J. Incorporating participatory action research in attention bias modification interventions for addictive disorders: perspectives. *Int J Environ Res Public Health* 2019 Mar 06;16(5):822 [FREE Full text] [doi: [10.3390/ijerph16050822](https://doi.org/10.3390/ijerph16050822)] [Medline: [30845752](https://pubmed.ncbi.nlm.nih.gov/30845752/)]
13. Zhang M, Vallabhajosyula R, Fung D. Emotional bias modification for individuals with attention deficit hyperactivity disorder: protocol for a co-design study. *JMIR Res Protoc* 2020 Dec 23;9(12):e24078 [FREE Full text] [doi: [10.2196/24078](https://doi.org/10.2196/24078)] [Medline: [33355536](https://pubmed.ncbi.nlm.nih.gov/33355536/)]
14. Hoffmann A, Christmann CA, Bleser G. Gamification in stress management apps: a critical app review. *JMIR Serious Games* 2017 Jun 07;5(2):e13 [FREE Full text] [doi: [10.2196/games.7216](https://doi.org/10.2196/games.7216)] [Medline: [28592397](https://pubmed.ncbi.nlm.nih.gov/28592397/)]
15. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
16. NVivo (Version 12). QSR International Pty Ltd. 2018. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home?_ga=2.49973158.1287536700.1670422066-1667103760.1670422066
17. Peñuelas-Calvo I, Jiang-Lin LK, Girela-Serrano B, Delgado-Gomez D, Navarro-Jimenez R, Baca-Garcia E, et al. Video games for the assessment and treatment of attention-deficit/hyperactivity disorder: a systematic review. *Eur Child Adolesc Psychiatry* 2022 Jan;31(1):5-20 [FREE Full text] [doi: [10.1007/s00787-020-01557-w](https://doi.org/10.1007/s00787-020-01557-w)] [Medline: [32424511](https://pubmed.ncbi.nlm.nih.gov/32424511/)]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder

Edited by A Mavragani; submitted 12.01.22; peer-reviewed by J Zhu, Z Douba, C Shah; comments to author 23.08.22; revised version received 23.10.22; accepted 11.11.22; published 09.12.22.

Please cite as:

Zhang M, Ranganath V

An Emotional Bias Modification for Children With Attention-Deficit/Hyperactivity Disorder: Co-design Study

JMIR Form Res 2022;6(12):e36390

URL: <https://formative.jmir.org/2022/12/e36390>

doi: [10.2196/36390](https://doi.org/10.2196/36390)

PMID: [36485019](https://pubmed.ncbi.nlm.nih.gov/36485019/)

©Melvyn Zhang, Vallabhajosyula Ranganath. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 09.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Co-design and Development of EndoSMS, a Supportive Text Message Intervention for Individuals Living With Endometriosis: Mixed Methods Study

Kerry Anne Sherman¹, BSc (Hons1), MA, PhD; Melissa Jade Pehlivan¹, BPsych (Hons1), MRes; Anna Singleton², BSc (Hons1), MSc, PhD; Alexandra Hawkey³, BSc, MSc, PhD; Julie Redfern², BSc, BAppSc, PhD; Mike Armour^{3,4}, BHSc, BSc (Hons), PhD; Blake Dear⁵, BPsych (Hons), MPsychol (Clinical), PhD; Tanya Jane Duckworth⁶, BFA, BA Psych (Hons), MBMSc; Donna Ciccio⁷; Michael Cooper⁸, MBBS; Kelly Ann Parry⁹, BSc (Hons), PhD; Esther Gandhi⁵, BHuSBPsych (Hons); Sara A Imani⁵, BSc (Hons1)

¹Centre for Emotional Health, School of Psychological Sciences, Macquarie University, Sydney, Australia

²School of Health Sciences, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

³Translational Health Research Institute (THRI), Western Sydney University, Sydney, Australia

⁴NICM Health Research Institute, Western Sydney University, Sydney, Australia

⁵School of Psychological Sciences, Macquarie University, Sydney, Australia

⁶School of Psychology, Faculty of Medical and Health Sciences, University of Adelaide, Adelaide, Australia

⁷Endometriosis Australia, Sydney, Australia

⁸Royal Prince Alfred Hospital, Sydney, Australia

⁹Australian College of Physical Education, Sydney, Australia

Corresponding Author:

Kerry Anne Sherman, BSc (Hons1), MA, PhD

Centre for Emotional Health

School of Psychological Sciences

Macquarie University

Balaclava Rd

Sydney, 2109

Australia

Phone: 61 9850 6874

Email: kerry.sherman@mq.edu.au

Abstract

Background: Endometriosis, which affects 1 in 10 people assigned female at birth, is a chronic systemic inflammatory disease with a high symptom burden and adverse socioemotional impacts. There is a need for an accessible, cost-effective, and low-burden intervention to support individuals in managing their endometriosis condition.

Objective: This study aimed to co-design and evaluate the acceptability, readability, and quality of a bank of supportive SMS text messages (EndoSMS) for individuals with endometriosis.

Methods: In phase 1 of this mixed method design, 17 consumer representatives (individuals with endometriosis) participated across three 3-hour web-based (Zoom, Zoom Video Communications, Inc) focus groups. The transcripts were encoded and analyzed thematically. In phase 2, consumer representatives (n=14) and health care professionals (n=9) quantitatively rated the acceptability, readability, and appropriateness of the developed text messages in a web-based survey. All the participants initially completed a background survey assessing sociodemographic and medical factors.

Results: Consumer representatives demonstrated diverse sociodemographic characteristics (Mage=33.29), varying in location (metropolitan vs rural or regional), employment, and relationship and educational statuses. Participants reached a consensus regarding the delivery of 4 SMS text messages per week, delivered randomly throughout the week and in one direction (ie, no reply), with customization for the time of day and use of personal names. Seven main areas of unmet need for which participants required assistance were identified, which subsequently became the topic areas for the developed SMS text messages: emotional health, social support, looking after and caring for your body, patient empowerment, interpersonal issues, general endometriosis information, and physical health. Through a web-based survey, 371 co-designed SMS text messages were highly rated by consumers

and health care professionals as clear, useful, and appropriate for individuals with endometriosis. Readability indices (Flesch-Kincaid scale) indicated that the SMS text messages were accessible to individuals with a minimum of 7th grade high school education.

Conclusions: On the basis of the needs and preferences of a diverse consumer representative group, we co-designed EndoSMS, a supportive SMS text message program for individuals with endometriosis. The initial evaluation of the SMS text messages by consumer representatives and health professionals suggested the high acceptability and suitability of the developed SMS text messages. Future studies should further evaluate the acceptability and effectiveness of EndoSMS in a broader population of individuals with endometriosis.

(*JMIR Form Res* 2022;6(12):e40837) doi:[10.2196/40837](https://doi.org/10.2196/40837)

KEYWORDS

text message; intervention; co-design; development; endometriosis; SMS; mHealth; self-management; mobile phone

Introduction

Background

Globally, the inflammatory condition of endometriosis is estimated to affect 1 in 10 biological females [1]. Some countries report an even higher incidence, with estimates of prevalence in Australia being 1 in 9 over the reproductive lifetime [2]. Endometriosis is predominantly characterized by severe and chronic pelvic pain and painful periods (dysmenorrhea). Other common symptoms include painful sexual intercourse (dyspareunia), infertility, and abdominal bloating [3,4]. Migraine, fibromyalgia, irritable bowel syndrome, and chronic fatigue syndrome are frequently comorbid with endometriosis [5]. In the absence of a cure and with a chronic, high symptom burden, individuals living with endometriosis experience adverse effects on their socioemotional well-being and functioning [6-9] and poor overall well-being [10], characterized by diminished quality of life [11] and a high prevalence of psychological distress (eg, depression) [12,13]. Endometriosis-related impairments to physical and psychological functioning also impact the society through health care costs and lost productivity (approximately US \$20,898 per individual per year [14,15]). Given the high symptom burden and adverse socioemotional consequences of endometriosis, there is a need for supportive interventions to help address the psychosocial impacts of living with this condition [16].

As a chronic condition, endometriosis demands ongoing self-management of physical symptoms (eg, pain and sleep hygiene) [17], which represents a substantial psychological and emotional burden on the individual [18,19]. The need for self-management may also serve as a constant reminder of one's endometriosis diagnosis, exacerbating psychological distress [20]. From a psychological perspective, interventions designed to provide reassurance and support may assist in managing the self-regulatory aspects of living with endometriosis [21] and in diminishing psychological distress [22].

Previous research has highlighted that people with endometriosis have a tendency to be self-critical, particularly in relation to their body appearance and function [23], presenting a further barrier to maintaining good psychological health [24]. A growing body of literature indicates that self-compassion (the ability to view oneself in a kind, compassionate manner [25]) may act as a buffer in the face of adversity [26-29], reducing psychological distress and symptom burden among people with

endometriosis [30]. This finding is consistent with other domains of women's health (ie, breast cancer [31,32] and polycystic ovarian syndrome [33]). Perceived difficulties in forming and retaining social relationships [34-36] also contribute to high levels of psychological distress. Furthermore, endometriosis symptoms (eg, pelvic pain [3]) inhibit social functioning, particularly participation in social events [10,11,37].

Communicating with health professionals is another difficulty reported by many living with endometriosis [38,39]. This is in part because of a perceived normalization of endometriosis symptoms (eg, painful periods) by health professionals and associated feelings of menstrual stigma, which often leaves patients feeling unsupported in their endometriosis management [38,40,41]. For those actively seeking support, electronic, web-based resources are a preferred source of information [42]. However, despite there being websites from consumer-facing organizations [38] (eg, Endometriosis Australia) providing reliable information on a range of relevant topics [43], the feeling of being inadequately informed about endometriosis is commonly reported [41,42,44]. For many, not knowing where to find information or lacking the time to seek supportive information act as barriers to support [45]. Therefore, an intervention that helps direct those living with endometriosis to reliable web-based information on the nature and management of this condition is warranted.

In light of the many psychosocial challenges experienced by people with endometriosis, there have been worldwide calls for interventions to assist in managing difficulties with psychosocial functioning and to improve quality of life [16,46,47]. In particular, interventions providing tips and strategies to manage the psychological challenges of living with endometriosis that can be readily implemented by individuals living with endometriosis are highly valued. An approach ideally suited to achieve these aims are SMS text message interventions, entailing the delivery of brief SMS text messages through a mobile phone [48]. With the ubiquitous use of mobile phones worldwide [49], SMS text messages are a highly accessible, inexpensive, and convenient means of providing psychosocial interventions, as users can opt to receive messages at a preferred time [50]. SMS text messages have the added benefit of being "pandemic proof" in that they are not affected by government-imposed lockdowns or public health measures [22]. A key strength of SMS text messaging is its flexibility, allowing a range of content (within the 160-character limit), including psychoeducation, reminders (eg, to exercise or take medication), motivational messages,

self-care (eg, stress management tips), and informative links to health-related internet sites [51]. The simple language format used in texts also increases their accessibility, irrespective of the reading level [52-55]. The ability to send SMS text messages anonymously may help minimize stigma and barriers to accessing health care [56], which are particularly evident in geographically isolated underserved populations [57]. With the increasing global uptake of smartphones [49], particularly among younger adults [58], supportive interventions using SMS text messaging platforms are likely to be ideally suited for the endometriosis population (ie, typical mean age 30, SD 7.50 years [11]); however, to date, no known SMS text message intervention has been developed for this population.

Prior Work

Across a wide range of health contexts, SMS text message interventions have proven effective at facilitating health behavior change, such as smoking cessation [59-61], the adoption of exercise [61,62] and diet regimens [63], and in chronic disease prevention [61,63]. Critically, beyond a focus on health behavior change, SMS text message interventions are also being developed with the aim to provide psychological and emotional support [21,22,51,64]. For example, one psychologically supportive intervention designed to address the distress experienced by individuals living with the chronic condition diabetes uses SMS text messaging derived from positive psychology (eg, mindfulness and self-compassion) [65]. In this case, brief messages were designed to provide reassurance, encourage a positive outlook, and reflect on gratitude. Emerging evidence suggests that these psychologically focused supportive SMS text messaging interventions are very well tolerated and accepted [66], consistent with evidence from health behavior change-focused SMS text interventions [52,53].

Study Aims

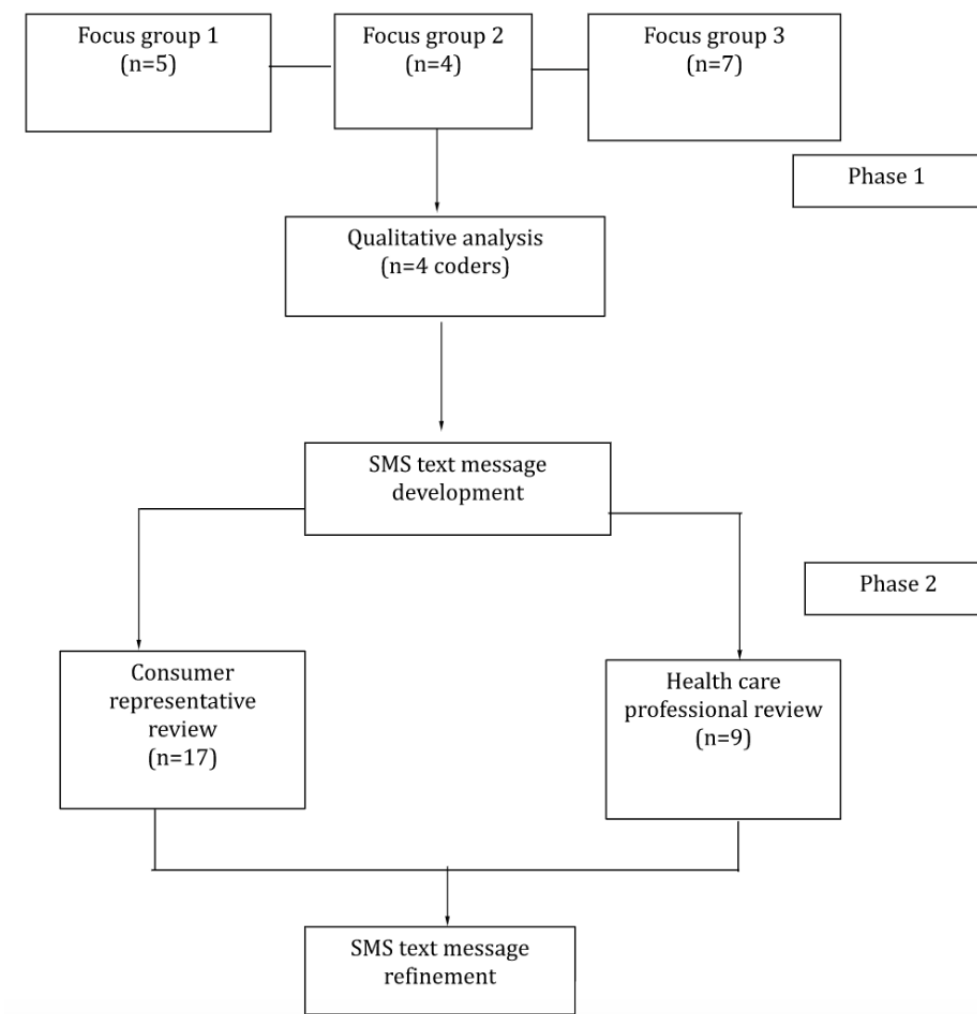
In sum, evidence across different health contexts suggests that SMS text messaging may be suitable to provide reassurance and support in managing the psychological and emotional self-regulatory aspects of living with endometriosis. Therefore,

the aim of this study was to identify the needs and preferences of consumers living with endometriosis to inform the development of a suitable supportive SMS text messaging intervention. The involvement of consumers in this manner, known as co-design, is regarded as best practice for the development of health-focused interventions, as it ensures the relevance of the intervention and its content to the target audience [67-69]. For more than a decade, it has been best practice for researchers to collaborate with consumers in research [70]. Co-designed SMS text messages have been found to improve health-promoting behaviors (eg, smoking cessation and dietary changes [63,71]) in individuals with chronic conditions. Additionally, this study aimed to evaluate the initial acceptability, readability, and quality of the co-designed SMS text messages, in preparation for further evaluation and use in future studies. To achieve these aims, this mixed methods study entailed two phases: (1) the identification of consumer SMS text messaging needs and preferences through focus groups and (2) the development and review of the SMS text message bank.

Methods

Study Design

A 2-phase mixed methods approach was used with consumer representatives and researcher-clinician collaborators (Figure 1). Phase 1 involved a series of web-based focus groups with consumer representatives to ascertain the needs and preferences for a supportive SMS text messaging intervention. The focus groups were also used to co-design the development of a bank of SMS text messages to ensure the suitability of the end product [67-69]. In phase 2, the developed SMS text message bank was evaluated quantitatively and qualitatively by consumer representatives and health care professionals. Reporting for the focus groups and intervention development was based on the Consolidated Criteria for Reporting Qualitative Research (Multimedia Appendix 1) [72] and Template for Intervention Description and Replication (Multimedia Appendix 2) [73], respectively.

Figure 1. Study design diagram.

Ethical Considerations

This study was approved by the Human Research Ethics Committee of Macquarie University (#52021963527729). All the participants provided web-based informed consent. The participants were informed that their confidentiality will be maintained, with only named investigators having access to their identifiable data. The participants were also informed that in the spirit of open science, their deidentified data would be made available in a public data store [74]. No incentives were provided for participation in the study.

Phase 1: Consumer Focus Groups

Using co-design principles, the focus groups aimed to determine the structure and content of the SMS text messages.

Participants

A convenience sample of 1331 individuals living with endometriosis who had participated in prior endometriosis-related research of the lead investigator (KAS) were invited via email to participate. The eligibility criteria were as follows: (1) the participants should be aged ≥ 18 years, (2) diagnosed with endometriosis (self-reported), and (3) proficient in English. A citizen collaborator (TD) of the research team was also invited to participate in one of the focus groups. Of the 1331 individuals invited, 89 (6.7%) potential participants

indicated their interest in the focus groups through return emails. Of these 89 individuals, 35 (39%) provided further details on their age and location (to facilitate the selection of a representative sample) and indicated their availability for the focus groups. Of the 35 participants who indicated availability for participation in the focus groups, 22 (63%) then completed a web-based informed consent form and a brief survey requesting further demographic (eg, age, education, employment, and location) and medical information (endometriosis treatments received and self-reported endometriosis severity, scored on a 5-point Likert scale from asymptomatic, 0, to severe, 4) via REDCap (Research Electronic Data Capture; Vanderbilt University). Consenting participants were then organized into 3 focus groups held over a week and spaced at least 1 day apart according to their preferred time and availability.

Facilitators

The focus groups were conducted by the lead investigator (KAS) and coauthor (AH), both of whom are experienced in qualitative research methods and research on women's health. KAS is a female professor of health psychology with a PhD in psychology. AH is a female research fellow in women's reproductive health with a PhD in critical public health. At the outset of the focus groups, KAS introduced herself as a professor with experience in women's health research and lived experience

with endometriosis, wanting to use her skills to give back to the endometriosis community. AH introduced herself as a researcher with interests in women's health conditions, particularly menstrual disorders. In all the focus groups, there was only 1 lead active facilitator, with the second facilitator (AH) joining one of the focus groups conducted by the first facilitator (KAS) for quality assurance reasons. Also present in the focus groups, although not active in discussions, were a female research assistant (MJP) who took field notes during discussions and a female student research intern (EH) who was available to provide technological support as needed.

Consumer Focus Groups

A series of 3 web-based (via Zoom, Zoom Video Communications, Inc) 3-hour workshop focus groups were held in mid-July 2021 to accommodate most participant availabilities. Each focus group commenced with a brief presentation (approximately 15 minutes) outlining the background and aim of the SMS text message program to be developed and the objectives of the session. During the workshop, the participants engaged in discussions focused on the preferred structure (eg, number of SMS text messages per week and text message delivery time of the day and days of the week) and content themes of the SMS text messaging program. Before these discussions, the participants were prompted regarding the SMS text message structure ("How many times would you like to receive a text?" "Which days of the week would you like to receive a text?" "Would you like to have a text that comes at a particular time of the day?" and "Would you like to be able to respond to the text messages?") and content themes. On the basis of prior endometriosis research [8,36,75-77] and SMS text messaging interventions developed by our research team [78,79], the participants were presented with five broad suggested topic areas along with example texts: (1) emotional health, (2) social support, (3) looking after and caring for your body, (4) general endometriosis information, and (5) physical health. The participants discussed the personal relevance of the proposed themes and were invited to suggest additional or alternate themes. The participants were also invited to draft their own example texts.

The 3 workshops were conducted in an iterative manner, building on the outcomes of the previous focus group sessions. At the conclusion of the third focus group, we reached data saturation, and, therefore, no more focus groups were needed. The focus groups were video recorded and transcribed using Zoom. The research intern (EG) reviewed the videos and corrected the transcripts for transcription errors. The participants were deidentified, were assigned an arbitrary number (eg, P1), and have been reported here accordingly to preserve their anonymity. Following the completion of the focus group transcription, four members of the research team (MJP, EG, SAI, and KAS) undertook a thematic analysis of these data [80] using a template approach [81] to identify the consumers' broad themes of preferred SMS text messaging. The analysis was guided by the six-stage [82] inductive method entailing: (1) familiarization with the whole data set, where all the coders read these data once and then reread them while actively searching for patterns and meaning through note-keeping; (2) separate generation of initial codes through meaningful

clustering of data; (3) comparison and refining of codes among coders to loosely identify potential themes—a process that was overseen by the senior member of the research team (KAS), ensuring consistency across codes and coders; (4) further discussion and refinement among the research team (MJP, EG, SAI, and KAS) for theme consolidation; (5) iterative naming of each theme, defining its singular focus, distinctness and extension of prior themes; and (6) integration of results through writing and selecting salient quotes.

Phase 2: Text Message Development, Review, and Refinement

Text Message Development

Following the identification of the preferred themes of SMS text message content, the research team (including those with lived experience—KAS and TD) drafted SMS text messages across all themes. The guiding principles in developing these SMS text messages were to keep the text within the 160-character limit, attain a readability grade of no more than 8 on the Flesh-Kincaid scale, and have an approximately even distribution of messages across all themes. General endometriosis information and other medical information were based on best practice guidelines, as summarized on the Endometriosis Australia and Jean Haile's Foundation websites, as well as clinical guidelines [83]. The SMS text messages were designed to provide support by drawing on the relevant psychological theory (Multimedia Appendix 3), such as the transactional model of stress [84] or self-compassion theory [25].

Participants

The research team, including clinicians (eg, gynecologists and psychologists) and academics, was invited to review the SMS text messages. All the consumer representatives and the citizen collaborator (TD) who participated in the focus groups, as well as the consumers who consented to the focus groups but were unable to attend them, were invited to review the SMS text messages via email.

Text Message Review

All the participants completed a web-based feedback survey via REDCap, with questions on participant demographics (eg, gender identity, profession, and years of experience) and approximately 40 draft SMS text messages (1-2 sentences each). Health care professionals reviewed the SMS text messages relevant to their area of expertise (eg, clinical psychologist researchers reviewed draft messages from the emotional health theme). Consumer representatives received a random selection of 30 draft SMS text messages, covering each of the different themes. For each draft text, the participants responded to 4 questions previously applied in SMS text messaging research [78], assessing the acceptability, readability, and appropriateness of the SMS text messages by rating them on a 5-point Likert scale (1=strongly disagree to 5=strongly agree; eg, "This message was easy to understand"). A free-text question for each item further gave the participants the opportunity to make suggestions for the improvement of the SMS text message. The scores on draft SMS text messages given by consumer representatives and health professionals were analyzed

separately, with descriptive statistics (ie, means and range) generated for each question. The SMS text messages that received a mean rating of ≤ 2 on any of the 3 questions were reviewed by the research team and either revised or removed. Free-text suggestions for improvement were collated for each SMS text message. Through this process of editing and deleting, the message content was refined.

Readability

All the refined SMS text messages were evaluated for readability using the validated Flesch-Kincaid grade level [85], indicating the required education level to read the material, ranging from grade 1 (easiest) to grade 12 (most difficult).

Results

Phase 1: Consumer Focus Groups

Descriptive Statistics

In total, 22 individuals living with endometriosis (including the citizen collaborator—TD) completed the background demographic and medical survey. Subsequently, 23% (5/22) of participants were unable to attend the 3 focus groups (approximately 6-8 participants in each session) owing to a lack of availability, work commitments, inclement weather, or illness. The participants were aged between 21 and 48 years (mean age 33.29, SD 9.11 years), reported an average diagnostic delay of 10.85 years (SD 8.03 years), and had diverse sociodemographic characteristics (Table 1).

Table 1. Characteristics of consumer representatives.

Characteristics	Focus group and reviewer participants (n=17), n (%)	Reviewer only participants (n=3), n (%)
Education		
High school education	4 (24)	1 (33)
Vocational or TAFE ^a	4 (24)	0 (0)
Undergraduate degree	5 (29)	1 (33)
Postgraduate degree	4 (24)	1 (33)
Employment		
Full time	11 (65)	3 (100)
Part time	2 (12)	0 (0)
Student	2 (12)	0 (0)
Unemployed	2 (12)	0 (0)
Location		
Metropolitan	11 (65)	2 (67)
Rural or regional	6 (35)	1 (33)
Relationship status		
Partnered	12 (71)	2 (67)
Not partnered	5 (29)	1 (33)
Treatment for endometriosis^b		
Surgery	11 (65)	2 (67)
Hormonal medication	8 (47)	3 (100)
Pain medication	13 (76)	3 (100)
Other	6 (35)	2 (67)
Endometriosis severity		
Asymptomatic	0 (0)	0 (0)
Mild	3 (18)	0 (0)
Moderate	5 (29)	1 (33)
Severe	9 (53)	2 (67)

^aTAFE: Technical and further education.

^bCategories not mutually exclusive.

Preferences for Content Themes

The 7 SMS text message themes and example quotes identified through template analysis are listed in [Textbox 1](#). All 5 of the proposed SMS text message themes (ie, general endometriosis information, physical health, emotional health, social support, and looking after and caring for your body) were endorsed by the focus group participants as being relevant and helpful. In

addition, two further SMS text message themes were identified: patient empowerment and interpersonal support. The consumers indicated that they require assistance to feel empowered in medical consultations and sought tips to help articulate their concerns so that health care professionals took them seriously. Tips and support regarding navigating intimate relationships, negotiating infertility and motherhood, and gaining support from loved ones and employers were also requested.

Textbox 1. Message content themes of focus groups with example quotes (focus groups consumers message content themes and quotes).

Emotional health

- “My mental health has been greatly affected...so even just having reminders that it’s okay to seek help, psychologically and it doesn’t mean that there isn’t a physical issue and it doesn’t take away from endo[metriosis], but it’s important to treat both aspects, because they do have a lot in common” (P14)
- “Anyone who lives with the chronic pain knows that we do have mental health issues, to some degree or another. So the idea of a text message taking the stigma out of reaching out for help would be a perfect hint to do that” (P13)

Social support

- “Tips on how to utilize your support network to the best of your ability could be good...it’s important to touch base when you’re feeling bad but even celebrating all your little wins with people too” (P8)
- “A reminder to jump on Endometriosis Australia or something and just read someone else’s story or a reminder just to check in” (P14)

Looking after and caring for your body

- “People do tend to push themselves and I think it’s very important to remind people that you know your limits. And you know what your body can and can’t handle so to be careful about pushing yourself and putting yourself through that” (P1)
- “Setting some time to properly relax and be more self aware, like body and mind consciousness, can really help your body” (P2)

Patient empowerment

- “How to advocate for yourself when you’re speaking to doctors because a lot of doctors will just say, ‘No, that’s not right’ and move onto what they think without really listening to what your experiences are with the particular treatments that they’re suggesting, and you end up down a rabbit hole of doing the same things over and over again without getting any anywhere” (P17)
- “One thing that I found really helpful was, I just take notes on my phone. And whenever I’m having symptoms, write those down because I have a bit of anxiety around going to the doctor. And so I kind of forget what you know my symptoms are or forget half of it” (P10)

Interpersonal issues

- “I definitely have unsupportive family members and it’s like, really hard to get them to, to talk to them about this [endometriosis]” (P2)
- “I think education or tips on how to manage and communicate with your employer about endo[metriosis]” (P16)
- “Painful sex...just kind of some tips around that, as well, because yeah it’s yeah, it’s a big thing and that can lead to like low self-esteem, confidence and things like that” (P2)
- “There’s not a lot of information about the impact of endometriosis on romantic relationships, and that it’s really important people are given information on how to navigate having endometriosis in a romantic relationship because it does have a strain particularly if it’s impacting sex” (P17)

General endometriosis information

- “I think it would be helpful just to kind of give some guidelines as to what’s actually true about the disease so that you know you’re kind of more equipped because when this all started, I was only 13” (P12)
- “I liked the idea of [general endometriosis information] being for beginners. But also I’ve had endometriosis for about 29 years and I’m still learning things about it” (P14)

Physical health

- “Definitely for me, I think, pain, like the physical side of it. Having tips on pain management. I think would be really useful” (P11)
- “Recipe tips because I know when I’m absolutely wrecked, the last thing I want to do is think about what to cook so if there’s a recipe suggestion that comes through...something that’s easy and quick because if you’re struggling, the last thing you want to do is cook for an hour” (P16)

Preferences for the Structure of the Text Message Program

The consumers could see the value in a SMS text messaging program that provides them with information on endometriosis and tips to manage their daily lives in a highly convenient and accessible way:

I've been searching for things to try to help me more in regard to learning more about endo[metriosis] and to actually have that information given to me, as opposed to me having to actually find it would be really good. [P5]

In particular, the consumers wanted to be linked up with web-based resources that they could access from their phone:

The whole point of this is that you are on your phone, so trying to give you links to other resources that are also on your phone could really help. [P16]

The consumers unanimously agreed that the most appropriate frequency would be 4 SMS text messages per week, covering a variety of topics:

I think the thing is if you are getting messages with like different topics or different things it's not necessarily going to feel like four a week. [P16]

A range of preferences regarding the timing of the SMS text message delivery was evident: randomly during the week or the weekend ("I'm happy for it [the day of the week] to be random. It doesn't really bother me." [P17]); a midweek text ("My first thought is that hump day Wednesday is a great day to get a pickup message from someone." [P13]); and a weekend text to assist with recovery after exerting themselves, for example, going out with friends and drinking ("I feel like I get more pain during the weekend sometimes depending what I do." [P2]).

Views on at what time of the day to receive texts were also mixed, with some consumers preferring to receive the texts at night ("Everyone's got stuff to do during the day. I think it's the nights that kind of get hard as well." [P11]) and others preferring a morning text to help prepare them for the day ("I'm definitely a night person myself, but I need a message in the morning because it's just hard to get out of bed when I feel rubbish." [P4]).

The consumers had differing views about whether they would prefer 1- or 2-way communication (ie, option to text back), with some indicating that 2-way communication would add an element of pressure:

I think I prefer not responding. Honestly, just so that I don't have that pressure. [P11]

Two-way communications from chatbots were also strongly disliked by the consumers:

I guess two way, but I want to know who is on the other side, if a robot is going to send a message back to me then I don't want to speak to a robot. [P2]

Following the discussions, all the focus group participants concluded that one-way communications may be the most practical. Regarding the formality of the texts, the consumers indicated that they would prefer causal greetings (eg, "hey") using their names with inclusive terminology (ie, avoid gender-specific pronouns).

A total of 376 messages were co-designed by the research team after the feedback from the focus group workshops, across seven topic areas, namely the topic areas initially suggested by the researchers—emotional health (n=52, 13.8%), looking after and caring for your body (n=52, 13.8%), social support (n=30, 8%), general endometriosis information (n=76, 20.2%), physical health (n=79, 21%)—and those suggested by the consumer representatives—patient empowerment (n=37, 9.8%) and interpersonal issues (n=50, 13.3%). On the basis of the feedback provided in the focus group workshops, the SMS text messages were designed to be delivered at a rate of 4 messages per week (on a mix of weekdays and weekends) and semicustomized to be delivered at their preferred time of the day (eg, morning, daytime, or nighttime), using their name.

Phase 2: Text Message Review and Refinement

A total of 14 participants from the consumer focus groups and an additional 3 consumer representatives completed the SMS text message review (Table 1). In addition, 9 health care professionals with an average of 10.11 (SD 6.86) years of experience completed the SMS text message review. Most of them identified as a female (6/9, 67%) and a health care researcher (6/9, 67%), with 33% (3/9) identifying as a researcher clinician.

The 376 co-designed messages were reviewed at least twice: once by a consumer representative and once by a health care professional. Overall, the consumers and health care professionals either agreed or strongly agreed that the SMS text messages were clear, useful, and appropriate for individuals with endometriosis (Table 2). In total, 7.7% (29/376) of messages received a score of 1 or 2 for at least one of the acceptability questions, warranting further review and revision by the research team.

For 41.8% (157/376) of the SMS text messages, the reviewers left free-text response comments suggesting improvements. Free-text feedback themes from health care professionals (120 in total) focused predominantly on website links and the grammar and clarity of the messages. Consumers' feedback (51 in total) focused on bolstering messages with resources (eg, links and tips) and ensuring that there was no insinuation of blame. On the basis of this feedback, of 376 messages, 6 (1.6%) messages were removed, 116 (30.9%) were revised to reflect reviewer feedback, and 1 (0.3%) new message was created, resulting in 371 co-designed SMS text messages, with an average Flesch-Kincaid readability score of 7 (providing high accessibility [86]). Table 3 displays the distribution of SMS text messages per theme, their Flesch-Kincaid grade, and example text messages for each.

Table 2. Consumer and health professional ratings of text messages.

Statements	Consumer representative ratings, mean (SD; range)	Health care professional ratings, mean (SD; range)
This message was easy to understand	4.40 (0.76; 2-5)	4.73 (0.46; 4-5)
The information provided in this message is useful	4.31 (0.82; 2-5)	4.57 (0.69; 1-5)
This message is appropriate for individuals with endometriosis	4.40 (0.78; 2-5)	4.63 (0.63; 1-5)

Table 3. Breakdown of readability and examples of SMS text messages for each theme.

Theme and example	Domain	Messages, n (%)	Flesch-Kincaid grade, mean (SD)
Emotional health		52 (14)	5.51 (2.58)
Hey <name>, it's okay to reach out and talk to someone about how you're feeling. Having endometriosis is tough and can impact every part of your life.	Social support enhancing		
Worried all the time? Why not try setting aside 10 minutes a day to write down whatever is troubling you, rather than letting it interrupt your day.	Coping strategy		
Social support		30 (8.1)	5.44 (2.77)
Hey <name>, remember to celebrate your daily wins with the people who support you!	Social support enhancing		
Don't feel like getting out and about lately? Why not try connecting with friends or family from the comfort of your own home. You could FaceTime, call, or chat online.	Coping strategy		
Looking after and caring for your body		49 (13.1)	6.31 (2.85)
Hey <name>, when was the last time you felt really relaxed? Try to do something you find relaxing today, whether it's reading a book or watching the sunset!	Coping strategy		
Hey <name>, everyone has their own limits. Try not to compare yourself with others. You are on your own journey!	Coping strategy		
Patient empowerment		37 (10)	6.55 (2.63)
Seeing your health care professional soon? It might help to keep a journal of your symptoms and bring this along with you to your appointment.	Coping strategy		
There's nothing wrong with getting a second or even third opinion on your condition. They can help you and your GP ^a better understand what's going on for you!	Education		
Interpersonal issues		51 (13.8)	7.06 (2.82)
Pain during sex can change throughout your cycle. Try and find a time that works best for you!	Education and coping strategy		
Hey <name>, having a bad flare up? Don't be afraid to ask your employer for some flexibility to work from home or some time off work.	Coping strategy		
General endometriosis information		75 (20.2)	7.97 (2.16)
Hi <name>, if you're experiencing heavy bleeding, you may also be feeling more tired than usual. Remember that it's okay to take time and rest.	Education		
Hi <name>, heard something about endometriosis that you aren't sure about? Double check the facts here <link to Endometriosis Australia fact buster information sheet>	Education		
Physical health		77 (20.8)	6.15 (1.95)
Hi <name>, if you're feeling okay today, why not try and get some exercise in? Just parking further away from the shops or work can get your steps up!	Symptom management		
Hi <name>, if you're in pain, you might find it hard to eat. Berries can be a great choice if you're wanting a snack, due to their anti-inflammatory effects.	Symptom management		

^aGP: general practitioner.

Discussion

Principal Findings

This study aimed to identify the needs and preferences of consumers living with endometriosis for an SMS text message–based supportive intervention. Consistent with prior research indicating a need for supportive interventions for endometriosis populations [16,20], our findings identified a strong desire for an SMS text message intervention that would provide both access to reliable information about endometriosis and its management and support. Consumers desired tips and strategies regarding physical and emotional health, social support and interpersonal issues, looking after and caring for their body, patient empowerment, and general endometriosis information. In particular, SMS text messaging was seen positively, with consumers expressing a desire to have supportive messages sent to their phones for easy access. It is likely that for individuals living with endometriosis who frequently experience fatigue and other high symptom burdens [3,5], having this supportive message readily available through the convenience of a mobile phone is particularly appealing. This points to the acceptability of mobile health interventions for individuals with endometriosis.

The need for psychological and emotional support was strongly evident in the focus groups, with 71% (5/7) of the themes reflecting this. Although various reviews have documented the psychological and socioemotional tolls associated with endometriosis [6,12,34,87,88], little research has been dedicated to developing supportive interventions that address these concerns [89,90]. To date, most of the psychological or supportive interventions developed for individuals with endometriosis are of low quality, lacking evidence-based protocols for replicability and being evaluated under nonrigorous conditions (eg, in the presence of confounding variables) [16,90]. In particular, our qualitative analysis revealed several distinct aspects of psychosocial functioning with which individuals with endometriosis required assistance. In line with prior research, individuals with endometriosis may have difficulty with self-compassion [23], and the consumers sought reminders to be kind to themselves, whether this be through physical acts (eg, self-care activities) or positive self-talk.

Similarly, consistent with extant literature [34–36], individuals with endometriosis expressed an interest in receiving support in maintaining social relationships through prompts to reach out or seek alternative methods of connection (eg, on the web) when facing debilitating symptoms. A need for support in managing interpersonal issues that went beyond ongoing relationship maintenance and general social support was also reported. In line with previous research [38,39], the participants reported a range of interpersonal conflicts as a result of endometriosis in their professional (eg, with employers and teachers) and more intimate relationships (eg, with family members and romantic partners), which they required assistance with. Furthermore, individuals with endometriosis expressed a desire to have more meaningful communications with health care professionals and sought support in advocating for their health care needs. This reflects prior research that documents

the difficulties individuals with endometriosis face in their relationships with health care professionals [38,39].

The consumers preferred that the SMS text messages be delivered in a variety of ways (eg, at different times during the day) to suit their diverse lives. This highlights a need for creative, adaptive, and tailorable intervention strategies. Another aspect that customization was required for was medication adherence. Some consumers expressed an interest in receiving SMS text messages designed to assist them in adhering to their medication for endometriosis. EndoSMS offers a novel solution, enabling tailoring to deliver messages at consumers' convenience with customization of the message content (eg, inclusion and exclusion of medication adherence texts).

Promisingly, the co-designed SMS text messages were rated highly overall in acceptability and suitability for an endometriosis population in the SMS text message review by both consumer representatives and health care professionals. Furthermore, the messages were considered easy to understand, with anyone with education until at least seventh grade able to read the messages. These preliminary results suggest that EndoSMS may be highly suitable for an endometriosis population; however, further evaluation in terms of the intervention's acceptability and effectiveness in a broader population of individuals with endometriosis is needed.

Strengths and Limitations

This mixed methods study co-designed and evaluated the acceptability of a supportive SMS text message intervention for individuals with endometriosis. To our knowledge, this is the first mobile health intervention to be designed for individuals with endometriosis. The strengths of this study include the involvement of consumer representatives in the co-design and evaluation of the intervention, ensuring tailoring to meet the needs of its end users in an appropriate and sensitive manner [67–69]. Furthermore, the inclusion of health care clinicians and researchers in the evaluation of the SMS text messages ensured the content validity and suitability for an endometriosis population. Consumer representatives demonstrated diverse sociodemographic characteristics, bolstering the generalizability of the EndoSMS program across different users. However, it should be noted that EndoSMS is designed for an Australian audience, and piloting of the program's feasibility is required to ensure its suitability for international use. In addition, although there was a high level of acceptability for the SMS text messages evident from the consumer representative ratings, it should be noted that many of these participants also informed the development of these messages, which perhaps explains the favorable ratings. Further research on the acceptability and helpfulness of the SMS text messages and EndoSMS program is needed to determine the suitability of the intervention for a broader endometriosis population. In particular, a pilot randomized controlled trial is needed to confirm that the SMS text messages are indeed supportive.

Conclusions

This study entailed the co-design of EndoSMS, a supportive SMS text message program for individuals with endometriosis, and investigation regarding its preliminary acceptability. By

ascertaining the needs and preferences of a diverse group of consumer representatives, EndoSMS was co-designed to better support individuals with endometriosis. Furthermore, the initial acceptability evaluation by consumers and health care

professionals was highly favorable. Further evaluation of EndoSMS is required to confirm its acceptability and effectiveness in a broader endometriosis population.

Acknowledgments

Funding for this project was granted by Endometriosis Australia.

Data Availability

The identified group-level data sets generated during and analyzed in this study are available from the Open Science Framework (<https://osf.io/m7bu8>).

Authors' Contributions

KAS and MJP wrote the first draft of this manuscript. KAS was the lead investigator on the Endometriosis Australia grant for this project. KAS and AH conducted the focus groups. MJP, EG, SAI, and KAS analyzed the transcripts from the focus groups. MJP, EG, SAI, KAS, and AS developed the bank of SMS text messages for review. All the authors edited the final manuscript.

Conflicts of Interest

Funding for this project was granted by Endometriosis Australia. JR was funded by the National Health and Medical Research Council Investigator Grant Leadership Level 2 (GNT2007946). AS received a PhD stipend provided by the University of Sydney, an Australian Government Research Training Program Scholarship, and a supplementary postgraduate research scholarship in breast cancer.

Multimedia Appendix 1

Consolidated Criteria for Reporting Qualitative Research checklist.

[[PDF File \(Adobe PDF File\), 482 KB - formative_v6i12e40837_app1.pdf](#)]

Multimedia Appendix 2

Template for Intervention Description and Replication checklist.

[[PDF File \(Adobe PDF File\), 207 KB - formative_v6i12e40837_app2.pdf](#)]

Multimedia Appendix 3

Example text messages across identified themes.

[[DOCX File , 19 KB - formative_v6i12e40837_app3.docx](#)]

References

1. Shafir A, Farland L, Shah D, Harris H, Kvaskoff M, Zondervan K, et al. Risk for and consequences of endometriosis: a critical epidemiologic review. *Best Pract Res Clin Obstet Gynaecol* 2018 Aug;51:1-15. [doi: [10.1016/j.bpobgyn.2018.06.001](https://doi.org/10.1016/j.bpobgyn.2018.06.001)] [Medline: [30017581](https://pubmed.ncbi.nlm.nih.gov/30017581/)]
2. Rowlands I, Abbott J, Montgomery G, Hockey R, Rogers P, Mishra G. Prevalence and incidence of endometriosis in Australian women: a data linkage cohort study. *BJOG* 2021 Mar;128(4):657-665. [doi: [10.1111/1471-0528.16447](https://doi.org/10.1111/1471-0528.16447)] [Medline: [32757329](https://pubmed.ncbi.nlm.nih.gov/32757329/)]
3. O'Hara R, Rowe H, Fisher J. Managing endometriosis: a cross-sectional survey of women in Australia. *J Psychosom Obstet Gynaecol* 2022 Sep;43(3):265-272. [doi: [10.1080/0167482X.2020.1825374](https://doi.org/10.1080/0167482X.2020.1825374)] [Medline: [33050751](https://pubmed.ncbi.nlm.nih.gov/33050751/)]
4. Armour M, Sinclair J, Ng CH, Hyman MS, Lawson K, Smith CA, et al. Endometriosis and chronic pelvic pain have similar impact on women, but time to diagnosis is decreasing: an Australian survey. *Sci Rep* 2020 Oct 01;10(1):16253 [[FREE Full text](#)] [doi: [10.1038/s41598-020-73389-2](https://doi.org/10.1038/s41598-020-73389-2)] [Medline: [33004965](https://pubmed.ncbi.nlm.nih.gov/33004965/)]
5. As-Sanie S, Black R, Giudice LC, Gray Valbrun T, Gupta J, Jones B, et al. Assessing research gaps and unmet needs in endometriosis. *Am J Obstet Gynecol* 2019 Aug;221(2):86-94. [doi: [10.1016/j.ajog.2019.02.033](https://doi.org/10.1016/j.ajog.2019.02.033)] [Medline: [30790565](https://pubmed.ncbi.nlm.nih.gov/30790565/)]
6. Marinho MC, Magalhaes TF, Fernandes LF, Augusto KL, Brilhante AV, Bezerra LR. Quality of life in women with endometriosis: an integrative review. *J Womens Health (Larchmt)* 2018 Mar;27(3):399-408. [doi: [10.1089/jwh.2017.6397](https://doi.org/10.1089/jwh.2017.6397)] [Medline: [29064316](https://pubmed.ncbi.nlm.nih.gov/29064316/)]
7. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, World Endometriosis Research Foundation Global Study of Women's Health consortium. Impact of endometriosis on quality of life and work

- productivity: a multicenter study across ten countries. *Fertil Steril* 2011 Aug;96(2):366-73.e8 [FREE Full text] [doi: [10.1016/j.fertnstert.2011.05.090](https://doi.org/10.1016/j.fertnstert.2011.05.090)] [Medline: [21718982](https://pubmed.ncbi.nlm.nih.gov/21718982/)]
8. Roomaney R, Kagee A. Salient aspects of quality of life among women diagnosed with endometriosis: a qualitative study. *J Health Psychol* 2018 Jun 17;23(7):905-916. [doi: [10.1177/1359105316643069](https://doi.org/10.1177/1359105316643069)] [Medline: [27091761](https://pubmed.ncbi.nlm.nih.gov/27091761/)]
 9. La Rosa VL, De Francis P, Barra F, Schiattarella A, Török P, Shah M, et al. Quality of life in women with endometriosis: a narrative overview. *Minerva Med* 2020 Feb;111(1):68-78. [doi: [10.23736/S0026-4806.19.06298-0](https://doi.org/10.23736/S0026-4806.19.06298-0)] [Medline: [31755667](https://pubmed.ncbi.nlm.nih.gov/31755667/)]
 10. Rush G, Misajon R. Examining subjective wellbeing and health-related quality of life in women with endometriosis. *Health Care Women Int* 2018 Mar;39(3):303-321. [doi: [10.1080/07399332.2017.1397671](https://doi.org/10.1080/07399332.2017.1397671)] [Medline: [29095116](https://pubmed.ncbi.nlm.nih.gov/29095116/)]
 11. Sullivan-Myers C, Sherman K, Beath A, Duckworth T, Cooper M. Delineating sociodemographic, medical and quality of life factors associated with psychological distress in individuals with endometriosis. *Hum Reprod* 2021 Jul 19;36(8):2170-2180. [doi: [10.1093/humrep/deab138](https://doi.org/10.1093/humrep/deab138)] [Medline: [34166496](https://pubmed.ncbi.nlm.nih.gov/34166496/)]
 12. van Barneveld E, Manders J, van Osch FH, van Poll M, Visser L, van Hanegem N, et al. Depression, anxiety, and correlating factors in endometriosis: a systematic review and meta-analysis. *J Womens Health (Larchmt)* 2022 Feb 01;31(2):219-230. [doi: [10.1089/jwh.2021.0021](https://doi.org/10.1089/jwh.2021.0021)] [Medline: [34077695](https://pubmed.ncbi.nlm.nih.gov/34077695/)]
 13. Estes S, Huisingh C, Chiuvè S, Petruski-Ivleva N, Missmer S. Depression, anxiety, and self-directed violence in women with endometriosis: a retrospective matched-cohort study. *Am J Epidemiol* 2021 May 04;190(5):843-852 [FREE Full text] [doi: [10.1093/aje/kwaa249](https://doi.org/10.1093/aje/kwaa249)] [Medline: [33184648](https://pubmed.ncbi.nlm.nih.gov/33184648/)]
 14. Armour M, Lawson K, Wood A, Smith CA, Abbott J. The cost of illness and economic burden of endometriosis and chronic pelvic pain in Australia: a national online survey. *PLoS One* 2019 Oct 10;14(10):e0223316 [FREE Full text] [doi: [10.1371/journal.pone.0223316](https://doi.org/10.1371/journal.pone.0223316)] [Medline: [31600241](https://pubmed.ncbi.nlm.nih.gov/31600241/)]
 15. Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum Reprod* 2012 May;27(5):1292-1299. [doi: [10.1093/humrep/des073](https://doi.org/10.1093/humrep/des073)] [Medline: [22422778](https://pubmed.ncbi.nlm.nih.gov/22422778/)]
 16. Van Niekerk L, Weaver-Pirie B, Matthewson M. Psychological interventions for endometriosis-related symptoms: a systematic review with narrative data synthesis. *Arch Womens Ment Health* 2019 Dec;22(6):723-735. [doi: [10.1007/s00737-019-00972-6](https://doi.org/10.1007/s00737-019-00972-6)] [Medline: [31081520](https://pubmed.ncbi.nlm.nih.gov/31081520/)]
 17. Leonardi M, Horne A, Vincent K, Sinclair J, Sherman K, Ciccia D, et al. Self-management strategies to consider to combat endometriosis symptoms during the COVID-19 pandemic. *Hum Reprod Open* 2020;2020(2):hoaa028 [FREE Full text] [doi: [10.1093/hropen/hoaa028](https://doi.org/10.1093/hropen/hoaa028)] [Medline: [32509977](https://pubmed.ncbi.nlm.nih.gov/32509977/)]
 18. Sherman KA, Koelmeyer L. Psychosocial predictors of adherence to lymphedema risk minimization guidelines among women with breast cancer. *Psychooncology* 2013 May 11;22(5):1120-1126. [doi: [10.1002/pon.3111](https://doi.org/10.1002/pon.3111)] [Medline: [22689156](https://pubmed.ncbi.nlm.nih.gov/22689156/)]
 19. Sherman KA, Koelmeyer L. The role of information sources and objective risk status on lymphedema risk-minimization behaviors in women recently diagnosed with breast cancer. *Oncology Nursing Forum* 2010 Dec 24;38(1):E27-E36. [doi: [10.1188/11.onf.e27-e36](https://doi.org/10.1188/11.onf.e27-e36)]
 20. O'Hara R, Rowe H, Fisher J. Self-management in condition-specific health: a systematic review of the evidence among women diagnosed with endometriosis. *BMC Womens Health* 2019 Jun 19;19(1):80 [FREE Full text] [doi: [10.1186/s12905-019-0774-6](https://doi.org/10.1186/s12905-019-0774-6)] [Medline: [31216998](https://pubmed.ncbi.nlm.nih.gov/31216998/)]
 21. Chu JT, Wadham A, Jiang Y, Whittaker R, Stasiak K, Shepherd M, et al. Effect of MyTeen SMS-based mobile intervention for parents of adolescents: a randomized clinical trial. *JAMA Netw Open* 2019 Sep 04;2(9):e1911120 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.11120](https://doi.org/10.1001/jamanetworkopen.2019.11120)] [Medline: [31509210](https://pubmed.ncbi.nlm.nih.gov/31509210/)]
 22. Agyapong VI, Hrabok M, Vuong W, Shalaby R, Noble JM, Gusnowski A, et al. Changes in stress, anxiety, and depression levels of subscribers to a daily supportive text message program (Text4Hope) during the COVID-19 pandemic: cross-sectional survey study. *JMIR Ment Health* 2020 Dec 18;7(12):e22423 [FREE Full text] [doi: [10.2196/22423](https://doi.org/10.2196/22423)] [Medline: [33296330](https://pubmed.ncbi.nlm.nih.gov/33296330/)]
 23. Sayer-Jones K, Sherman KA. Body image concerns in individuals diagnosed with benign gynaecological conditions: scoping review and meta-synthesis. *Health Psychol Behav Med* 2021 May 15;9(1):456-479 [FREE Full text] [doi: [10.1080/21642850.2021.1920949](https://doi.org/10.1080/21642850.2021.1920949)] [Medline: [34104570](https://pubmed.ncbi.nlm.nih.gov/34104570/)]
 24. Hughes M, Brown SL, Campbell S, Dandy S, Cherry MG. Self-Compassion and anxiety and depression in chronic physical illness populations: a systematic review. *Mindfulness* 2021 Mar 18;12(7):1597-1610. [doi: [10.1007/s12671-021-01602-y](https://doi.org/10.1007/s12671-021-01602-y)]
 25. Neff K. Self-compassion: an alternative conceptualization of a healthy attitude toward oneself. *Self Identity* 2003 Apr;2(2):85-101. [doi: [10.1080/15298860309032](https://doi.org/10.1080/15298860309032)]
 26. Diedrich A, Hofmann SG, Cuijpers P, Berking M. Self-compassion enhances the efficacy of explicit cognitive reappraisal as an emotion regulation strategy in individuals with major depressive disorder. *Behav Res Ther* 2016 Jul;82:1-10. [doi: [10.1016/j.brat.2016.04.003](https://doi.org/10.1016/j.brat.2016.04.003)] [Medline: [27152671](https://pubmed.ncbi.nlm.nih.gov/27152671/)]
 27. Ehret AM, Joormann J, Berking M. Self-compassion is more effective than acceptance and reappraisal in decreasing depressed mood in currently and formerly depressed individuals. *J Affect Disord* 2018 Jan 15;226:220-226. [doi: [10.1016/j.jad.2017.10.006](https://doi.org/10.1016/j.jad.2017.10.006)] [Medline: [28992586](https://pubmed.ncbi.nlm.nih.gov/28992586/)]
 28. Kirschner H, Kuyken W, Wright K, Roberts H, Brejcha C, Karl A. Soothing your heart and feeling connected: a new experimental paradigm to study the benefits of self-compassion. *Clin Psychol Sci* 2019 May 06;7(3):545-565 [FREE Full text] [doi: [10.1177/2167702618812438](https://doi.org/10.1177/2167702618812438)] [Medline: [32655984](https://pubmed.ncbi.nlm.nih.gov/32655984/)]

29. Leary MR, Tate EB, Adams CE, Allen AB, Hancock J. Self-compassion and reactions to unpleasant self-relevant events: the implications of treating oneself kindly. *J Pers Soc Psychol* 2007 May;92(5):887-904. [doi: [10.1037/0022-3514.92.5.887](https://doi.org/10.1037/0022-3514.92.5.887)] [Medline: [17484611](https://pubmed.ncbi.nlm.nih.gov/17484611/)]
30. Van Niekerk L, Johnstone L, Matthewson M. Predictors of self-compassion in endometriosis: the role of psychological health and endometriosis symptom burden. *Hum Reprod* 2022 Jan 28;37(2):264-273. [doi: [10.1093/humrep/deab257](https://doi.org/10.1093/humrep/deab257)] [Medline: [35102421](https://pubmed.ncbi.nlm.nih.gov/35102421/)]
31. Sherman KA, Przedziński A, Alcorso J, Kilby CJ, Elder E, Boyages J, et al. Reducing body image-related distress in women with breast cancer using a structured online writing exercise: results from the my changed body randomized controlled trial. *J Clin Oncol* 2018 Jul 01;36(19):1930-1940. [doi: [10.1200/jco.2017.76.3318](https://doi.org/10.1200/jco.2017.76.3318)]
32. Przedziński A, Sherman KA, Baillie A, Taylor A, Foley E, Stalgis-Bilinski K. My changed body: breast cancer, body image, distress and self-compassion. *Psychooncology* 2013 Aug;22(8):1872-1879. [doi: [10.1002/pon.3230](https://doi.org/10.1002/pon.3230)] [Medline: [23203842](https://pubmed.ncbi.nlm.nih.gov/23203842/)]
33. Van Niekerk LM, Bromfield H, Matthewson M. Physical and psychological correlates of self and body compassion in women with polycystic ovary syndrome. *J Health Psychol* 2022 Sep 06;27(11):2566-2580. [doi: [10.1177/13591053211059390](https://doi.org/10.1177/13591053211059390)] [Medline: [34865539](https://pubmed.ncbi.nlm.nih.gov/34865539/)]
34. Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, et al. The social and psychological impact of endometriosis on women's lives: a critical narrative review. *Hum Reprod Update* 2013;19(6):625-639. [doi: [10.1093/humupd/dmt027](https://doi.org/10.1093/humupd/dmt027)] [Medline: [23884896](https://pubmed.ncbi.nlm.nih.gov/23884896/)]
35. De Graaff AA, D'Hooghe T, Dunselman G, Dirksen C, Hummelshoj L, WERF EndoCost Consortium, et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum Reprod* 2013 Oct;28(10):2677-2685. [doi: [10.1093/humrep/det284](https://doi.org/10.1093/humrep/det284)] [Medline: [23847114](https://pubmed.ncbi.nlm.nih.gov/23847114/)]
36. Denny E. I never know from one day to another how I will feel: pain and uncertainty in women with endometriosis. *Qual Health Res* 2009 Jul 26;19(7):985-995. [doi: [10.1177/1049732309338725](https://doi.org/10.1177/1049732309338725)] [Medline: [19470614](https://pubmed.ncbi.nlm.nih.gov/19470614/)]
37. Yela DA, Quagliato ID, Benetti-Pinto CL. Quality of life in women with deep endometriosis: a cross-sectional study. *Rev Bras Ginecol Obstet* 2020 Feb 30;42(2):90-95 [FREE Full text] [doi: [10.1055/s-0040-1708091](https://doi.org/10.1055/s-0040-1708091)] [Medline: [32227324](https://pubmed.ncbi.nlm.nih.gov/32227324/)]
38. Young K, Fisher J, Kirkman M. Partners instead of patients: women negotiating power and knowledge within medical encounters for endometriosis. *Feminism Psychol* 2019 Jan 30;30(1):22-41. [doi: [10.1177/0959353519826170](https://doi.org/10.1177/0959353519826170)]
39. Grundström H, Alehagen S, Kjølhede P, Berterö C. The double-edged experience of healthcare encounters among women with endometriosis: a qualitative study. *J Clin Nurs* 2018 Jan 18;27(1-2):205-211. [doi: [10.1111/jocn.13872](https://doi.org/10.1111/jocn.13872)] [Medline: [28493635](https://pubmed.ncbi.nlm.nih.gov/28493635/)]
40. Pettersson A, Berterö CM. How women with endometriosis experience health care encounters. *Womens Health Rep (New Rochelle)* 2020 Nov 01;1(1):529-542 [FREE Full text] [doi: [10.1089/whr.2020.0099](https://doi.org/10.1089/whr.2020.0099)] [Medline: [33786519](https://pubmed.ncbi.nlm.nih.gov/33786519/)]
41. Armour M, Hawkey A, Micheal S, Diezel H, Chalmers K. 'A day to day struggle': a comparative qualitative study on experiences of women with endometriosis and chronic pelvic pain. *Fem Psychol* 2021;32(4):482-500 [FREE Full text] [doi: [10.21203/rs.3.rs-289745/v1](https://doi.org/10.21203/rs.3.rs-289745/v1)]
42. Roomaney R, Kagee A. Coping strategies employed by women with endometriosis in a public health-care setting. *J Health Psychol* 2016 Oct 10;21(10):2259-2268 [FREE Full text] [doi: [10.1177/1359105315573447](https://doi.org/10.1177/1359105315573447)] [Medline: [25769875](https://pubmed.ncbi.nlm.nih.gov/25769875/)]
43. Sbaffi L, King K. Living with endometriosis: the role of the internet in supporting the diagnosis and treatment process. *J Consum Health Internet* 2020 Dec 07;24(4):370-390. [doi: [10.1080/15398285.2020.1816803](https://doi.org/10.1080/15398285.2020.1816803)]
44. Moradi M, Parker M, Sneddon A, Lopez V, Ellwood D. Impact of endometriosis on women's lives: a qualitative study. *BMC Womens Health* 2014 Oct 04;14:123 [FREE Full text] [doi: [10.1186/1472-6874-14-123](https://doi.org/10.1186/1472-6874-14-123)] [Medline: [25280500](https://pubmed.ncbi.nlm.nih.gov/25280500/)]
45. Chu JT, Wang MP, Shen C, Viswanath K, Lam TH, Chan SS. How, when and why people seek health information online: qualitative study in Hong Kong. *Interact J Med Res* 2017 Dec 12;6(2):e24 [FREE Full text] [doi: [10.2196/ijmr.7000](https://doi.org/10.2196/ijmr.7000)] [Medline: [29233802](https://pubmed.ncbi.nlm.nih.gov/29233802/)]
46. Johnson NP, Hummelshoj L. Reply: consensus on current management of endometriosis. *Hum Reprod* 2013 Nov 05;28(11):3163-3164. [doi: [10.1093/humrep/det347](https://doi.org/10.1093/humrep/det347)] [Medline: [24014602](https://pubmed.ncbi.nlm.nih.gov/24014602/)]
47. Kuznetsov L, Dworzynski K, Davies M, Overton C, Guideline Committee. Diagnosis and management of endometriosis: summary of NICE guidance. *BMJ* 2017 Sep 06;358:j3935. [doi: [10.1136/bmj.j3935](https://doi.org/10.1136/bmj.j3935)] [Medline: [28877898](https://pubmed.ncbi.nlm.nih.gov/28877898/)]
48. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. *J Med Internet Res* 2013 Jun 14;15(6):e120 [FREE Full text] [doi: [10.2196/jmir.2600](https://doi.org/10.2196/jmir.2600)] [Medline: [23770578](https://pubmed.ncbi.nlm.nih.gov/23770578/)]
49. Wigginton C. Global mobile consumer trends: second edition. Deloitte. 2017. URL: <https://www2.deloitte.com/global/en/pages/technology-media-and-telecommunications/articles/gx-global-mobile-consumer-trends.html> [accessed 2022-05-11]
50. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. *Annu Rev Public Health* 2015 Mar 18;36:393-415 [FREE Full text] [doi: [10.1146/annurev-publhealth-031914-122855](https://doi.org/10.1146/annurev-publhealth-031914-122855)] [Medline: [25785892](https://pubmed.ncbi.nlm.nih.gov/25785892/)]
51. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. *J Med Internet Res* 2017 Aug 24;19(8):e295 [FREE Full text] [doi: [10.2196/jmir.7740](https://doi.org/10.2196/jmir.7740)] [Medline: [28838887](https://pubmed.ncbi.nlm.nih.gov/28838887/)]

52. Singleton AC, Raeside R, Partridge SR, Tat-Ko J, Che Mun Sum S, Hyun KK, et al. Supporting breast cancer survivors via text messages: reach, acceptability, and utility of EMPOWER-SMS. *J Cancer Surviv* 2022 Dec 09;16(6):1165-1175 [[FREE Full text](#)] [doi: [10.1007/s11764-021-01106-7](https://doi.org/10.1007/s11764-021-01106-7)] [Medline: [34505206](#)]
53. Redfern J, Santo K, Coorey G, Thakkar J, Hackett M, Thiagalingam A, et al. Factors influencing engagement, perceived usefulness and behavioral mechanisms associated with a text message support program. *PLoS One* 2016 Oct 14;11(10):e0163929 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0163929](https://doi.org/10.1371/journal.pone.0163929)] [Medline: [27741244](#)]
54. Waller K, Furber S, Bauman A, Allman-Farinelli M, van den Dolder P, Hayes A, et al. Effectiveness and acceptability of a text message intervention (DTEXT) on HbA1c and self-management for people with type 2 diabetes. A randomized controlled trial. *Patient Educ Couns* 2021 Jul;104(7):1736-1744. [doi: [10.1016/j.pec.2020.11.038](https://doi.org/10.1016/j.pec.2020.11.038)] [Medline: [33334634](#)]
55. Griffiths H. The acceptability and feasibility of using text messaging to support the delivery of physical health care in those suffering from a psychotic disorder: a review of the literature. *Psychiatr Q* 2020 Dec 24;91(4):1305-1316 [[FREE Full text](#)] [doi: [10.1007/s11126-020-09847-x](https://doi.org/10.1007/s11126-020-09847-x)] [Medline: [32970312](#)]
56. Naughton F, Jamison J, Sutton S. Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers. *Health Educ Res* 2013 Oct;28(5):911-922. [doi: [10.1093/her/cyt057](https://doi.org/10.1093/her/cyt057)] [Medline: [23640985](#)]
57. Dwyer A, de Almeida Neto A, Estival D, Li W, Lam-Cassettari C, Antoniou M. Suitability of text-based communications for the delivery of psychological therapeutic services to rural and remote communities: scoping review. *JMIR Ment Health* 2021 Feb 24;8(2):e19478 [[FREE Full text](#)] [doi: [10.2196/19478](https://doi.org/10.2196/19478)] [Medline: [33625373](#)]
58. Andone I, Błaszczewicz K, Eibes M, Trendafilov B, Montag C, Markowetz A. How age and gender affect smartphone usage. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016 Presented at: UbiComp '16: The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 12 - 16, 2016; Heidelberg Germany. [doi: [10.1145/2968219.2971451](https://doi.org/10.1145/2968219.2971451)]
59. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2019 Oct 22;10:CD006611 [[FREE Full text](#)] [doi: [10.1002/14651858.CD006611.pub5](https://doi.org/10.1002/14651858.CD006611.pub5)] [Medline: [31638271](#)]
60. Liao Y, Wu Q, Kelly BC, Zhang F, Tang Y, Wang Q, et al. Effectiveness of a text-messaging-based smoking cessation intervention ("Happy Quit") for smoking cessation in China: a randomized controlled trial. *PLoS Med* 2018 Dec 18;15(12):e1002713 [[FREE Full text](#)] [doi: [10.1371/journal.pmed.1002713](https://doi.org/10.1371/journal.pmed.1002713)] [Medline: [30562352](#)]
61. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. *JAMA* 2015;314(12):1255-1263. [doi: [10.1001/jama.2015.10945](https://doi.org/10.1001/jama.2015.10945)] [Medline: [26393848](#)]
62. Huffman JC, Harnedy LE, Massey CN, Carrillo A, Feig EH, Chung W, et al. A phone and text message intervention to improve physical activity in midlife: initial feasibility testing. *Health Psychol Behav Med* 2022 Mar 10;10(1):291-315 [[FREE Full text](#)] [doi: [10.1080/21642850.2022.2049796](https://doi.org/10.1080/21642850.2022.2049796)] [Medline: [35295924](#)]
63. Santo K, Hyun K, de Keizer L, Thiagalingam A, Hillis GS, Chalmers J, et al. The effects of a lifestyle-focused text-messaging intervention on adherence to dietary guideline recommendations in patients with coronary heart disease: an analysis of the TEXT ME study. *Int J Behav Nutr Phys Act* 2018 May 23;15(1):45 [[FREE Full text](#)] [doi: [10.1186/s12966-018-0677-1](https://doi.org/10.1186/s12966-018-0677-1)] [Medline: [29792202](#)]
64. Berrouguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [[FREE Full text](#)] [doi: [10.2196/jmir.5066](https://doi.org/10.2196/jmir.5066)] [Medline: [27287668](#)]
65. Jaser SS, Datye K, Morrow T, Sinisterra M, LeSturgeon L, Abadula F, et al. THRIVE! Positive psychology intervention to treat diabetes distress in teens with type 1 diabetes: rationale and trial design. *Contemp Clin Trials* 2020 Sep;96:106086 [[FREE Full text](#)] [doi: [10.1016/j.cct.2020.106086](https://doi.org/10.1016/j.cct.2020.106086)] [Medline: [32682996](#)]
66. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. *J Med Internet Res* 2017 Jul 21;19(8):e267 [[FREE Full text](#)] [doi: [10.2196/jmir.7023](https://doi.org/10.2196/jmir.7023)] [Medline: [28784594](#)]
67. Moser SC. Can science on transformation transform science? Lessons from co-design. *Curr Opin Environmental Sustainability* 2016 Jun;20:106-115. [doi: [10.1016/j.cosust.2016.10.007](https://doi.org/10.1016/j.cosust.2016.10.007)]
68. Brett J, Staniszevska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. *Health Expect* 2014 Oct;17(5):637-650 [[FREE Full text](#)] [doi: [10.1111/j.1369-7625.2012.00795.x](https://doi.org/10.1111/j.1369-7625.2012.00795.x)] [Medline: [22809132](#)]
69. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. *Health Res Policy Syst* 2020 Feb 11;18(1):17 [[FREE Full text](#)] [doi: [10.1186/s12961-020-0528-9](https://doi.org/10.1186/s12961-020-0528-9)] [Medline: [32046728](#)]
70. Archer F. Revision of the Joint NHMRC/AVCC Statement and Guidelines on Research Practice. *Australasian J Paramed* 2015 Jul 15;4(1). [doi: [10.33151/ajp.4.1.349](https://doi.org/10.33151/ajp.4.1.349)]
71. Klimis H, Marschner S, Von Huben A, Thiagalingam A, Chow CK. Predictors of smoking cessation in a lifestyle-focused text-message support programme delivered to people with coronary heart disease: an analysis from the tobacco exercise and diet messages (TEXTME) randomised clinical trial. *Tob Use Insights* 2020 Jan 28;13:1179173X2090148. [doi: [10.1177/1179173x20901486](https://doi.org/10.1177/1179173x20901486)]

72. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
73. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014 Mar 07;348(mar07 3):g1687. [doi: [10.1136/bmj.g1687](https://doi.org/10.1136/bmj.g1687)] [Medline: [24609605](https://pubmed.ncbi.nlm.nih.gov/24609605/)]
74. Co-design and developing EndoSMS: A supportive text message intervention for individuals with endometriosis. *OSF Registries*. URL: <https://osf.io/m7bu8> [accessed 2022-11-25]
75. Zarbo C, Brugnera A, Frigerio L, Malandrino C, Rabboni M, Bondi E, et al. Behavioral, cognitive, and emotional coping strategies of women with endometriosis: a critical narrative review. *Arch Womens Ment Health* 2018 Feb;21(1):1-13. [doi: [10.1007/s00737-017-0779-9](https://doi.org/10.1007/s00737-017-0779-9)] [Medline: [28932912](https://pubmed.ncbi.nlm.nih.gov/28932912/)]
76. Jones G, Jenkinson C, Kennedy S. The impact of endometriosis upon quality of life: a qualitative analysis. *J Psychosom Obstet Gynaecol* 2004 Jun 07;25(2):123-133. [doi: [10.1080/01674820400002279](https://doi.org/10.1080/01674820400002279)] [Medline: [15715035](https://pubmed.ncbi.nlm.nih.gov/15715035/)]
77. González-Echevarría AM, Rosario E, Acevedo S, Flores I. Impact of coping strategies on quality of life of adolescents and young women with endometriosis. *J Psychosom Obstet Gynaecol* 2019 Jun;40(2):138-145 [FREE Full text] [doi: [10.1080/0167482X.2018.1450384](https://doi.org/10.1080/0167482X.2018.1450384)] [Medline: [29648907](https://pubmed.ncbi.nlm.nih.gov/29648907/)]
78. Singleton A, Raeside R, Partridge SR, Hayes M, Maka K, Hyun KK, et al. Co-designing a lifestyle-focused text message intervention for women after breast cancer treatment: mixed methods study. *J Med Internet Res* 2021 Jun 14;23(6):e27076 [FREE Full text] [doi: [10.2196/27076](https://doi.org/10.2196/27076)] [Medline: [34125072](https://pubmed.ncbi.nlm.nih.gov/34125072/)]
79. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. *Eur J Prev Cardiol* 2014 Apr;21(4):492-499. [doi: [10.1177/2047487312449416](https://doi.org/10.1177/2047487312449416)] [Medline: [22605787](https://pubmed.ncbi.nlm.nih.gov/22605787/)]
80. Davidsen AS. Phenomenological approaches in psychology and health sciences. *Qual Res Psychol* 2013 Jul;10(3):318-339 [FREE Full text] [doi: [10.1080/14780887.2011.608466](https://doi.org/10.1080/14780887.2011.608466)] [Medline: [23606810](https://pubmed.ncbi.nlm.nih.gov/23606810/)]
81. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. *Qual Res Psychol* 2015 Apr 03;12(2):202-222 [FREE Full text] [doi: [10.1080/14780887.2014.955224](https://doi.org/10.1080/14780887.2014.955224)] [Medline: [27499705](https://pubmed.ncbi.nlm.nih.gov/27499705/)]
82. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher* 2020 May 01;42(8):846-854. [doi: [10.1080/0142159x.2020.1755030](https://doi.org/10.1080/0142159x.2020.1755030)]
83. Australian clinical practice guideline for the diagnosis and management of endometriosis. RANZCOG. URL: <https://ranzco.edu.au/> [accessed 2022-11-17]
84. Lazarus R, Folkman S. *Stress, Appraisal, and Coping*. Cham: Springer; 1984.
85. Flesch R. A new readability yardstick. *J Appl Psychol* 1948 Jun;32(3):221-233. [doi: [10.1037/h0057532](https://doi.org/10.1037/h0057532)] [Medline: [18867058](https://pubmed.ncbi.nlm.nih.gov/18867058/)]
86. Programme for the International Assessment of Adult Competencies, Australia. Australian Bureau of Statistics. 2013 Oct 9. URL: <https://www.abs.gov.au/statistics/people/education/programme-international-assessment-adult-competencies-australia/latest-release#cite-window1> [accessed 2022-11-17]
87. Gambadauro P, Carli V, Hadlaczy G. Depressive symptoms among women with endometriosis: a systematic review and meta-analysis. *Am J Obstet Gynecol* 2019 Mar;220(3):230-241. [doi: [10.1016/j.ajog.2018.11.123](https://doi.org/10.1016/j.ajog.2018.11.123)] [Medline: [30419199](https://pubmed.ncbi.nlm.nih.gov/30419199/)]
88. Pope CJ, Sharma V, Sharma S, Mazmanian D. A systematic review of the association between psychiatric disturbances and endometriosis. *J Obstet Gynaecol Can* 2015 Nov;37(11):1006-1015. [doi: [10.1016/s1701-2163\(16\)30050-0](https://doi.org/10.1016/s1701-2163(16)30050-0)] [Medline: [26629721](https://pubmed.ncbi.nlm.nih.gov/26629721/)]
89. Buggio L, Barbara G, Facchin F, Frattaruolo MP, Aimi G, Berlanda N. Self-management and psychological-sexological interventions in patients with endometriosis: strategies, outcomes, and integration into clinical care. *IJWH* 2017 May;Volume 9:281-293. [doi: [10.2147/ijwh.s119724](https://doi.org/10.2147/ijwh.s119724)]
90. Brooks T, Sharp R, Evans S, Baranoff J, Esterman A. Predictors of psychological outcomes and the effectiveness and experience of psychological interventions for adult women with chronic pelvic pain: a scoping review. *J Pain Res* 2020 May;Volume 13:1081-1102. [doi: [10.2147/jpr.s245723](https://doi.org/10.2147/jpr.s245723)]

Abbreviations

REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 13.07.22; peer-reviewed by HY Yoon, A Thiagalingam; comments to author 29.09.22; revised version received 17.10.22; accepted 01.11.22; published 09.12.22.

Please cite as:

Sherman KA, Pehlivan MJ, Singleton A, Hawkey A, Redfern J, Armour M, Dear B, Duckworth TJ, Ciccio D, Cooper M, Parry KA, Gandhi E, Imani SA

Co-design and Development of EndoSMS, a Supportive Text Message Intervention for Individuals Living With Endometriosis: Mixed Methods Study

JMIR Form Res 2022;6(12):e40837

URL: <https://formative.jmir.org/2022/12/e40837>

doi: [10.2196/40837](https://doi.org/10.2196/40837)

PMID: [36485029](https://pubmed.ncbi.nlm.nih.gov/36485029/)

©Kerry Anne Sherman, Melissa Jade Pehlivan, Anna Singleton, Alexandra Hawkey, Julie Redfern, Mike Armour, Blake Dear, Tanya Jane Duckworth, Donna Ciccio, Michael Cooper, Kelly Ann Parry, Esther Gandhi, Sara A Imani. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 09.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use of an Interactive Obesity Treatment Approach in Individuals With Severe Mental Illness: Feasibility, Acceptability, and Proposed Engagement Criteria

Ginger Nicol¹, MD; Madeline Jansen², MPH, MD; Rita Haddad³, MD; Amanda Ricchio³, BA; Michael D Yingling³, MS; Julia A Schweiger³, CCRC; Katie Keenoy⁴, MA; Bradley A Evanoff^{5*}, MPH, MD; John W Newcomer^{1,3*}, MD

¹Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States

²Division of Child and Adolescent Psychiatry, Department of Psychiatry, Los Angeles David Geffen School of Medicine, University of California, Los Angeles, CA, United States

³Department of Psychiatry, Washington University School of Medicine, St Louis, MO, United States

⁴Washington University School of Medicine, St. Louis, MO, United States

⁵Division of General Medical Sciences, Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO, United States

*these authors contributed equally

Corresponding Author:

Ginger Nicol, MD

Department of Psychiatry

Washington University School of Medicine

600 S. Taylor Ave.

Suite 121

St. Louis, MO, 63110

United States

Phone: 1 13143625939

Email: nicolg@wustl.edu

Abstract

Background: Digital and mobile health interventions are increasingly being used to support healthy lifestyle change, including in certain high-risk populations such as those with severe mental illnesses (SMIs). Life expectancy in this population lags 15 years behind counterparts in the general population, primarily due to obesity-related health conditions.

Objective: We tested the feasibility and usability of a 12-week interactive obesity treatment approach (iOTA) to adults with chronic SMIs (depression, bipolar disorder and schizophrenia spectrum disorder) receiving treatment in community settings. The iOTA incorporates short message service (SMS) text messages to supplement monthly in-person health coaching.

Methods: Factors hypothesized to be associated with weight change were illness severity and treatment engagement. Severe psychiatric symptoms were defined as baseline Clinical Global Impression severity score of >5. Criterion engagement was defined as a text messaging response rate >80% during the first 4 weeks of treatment. Disordered eating, assessed with the Loss of Control Over Eating Scores, was also evaluated. Participants provided qualitative data, further informing assessment of intervention feasibility, usability, and acceptability.

Results: A total of 26 participants were enrolled. The mean age was 48.5 (SD 15.67) years; 40% (10/26) were Black and 60% (15/26) female. Participants with lower symptom severity and adequate engagement demonstrated significantly decreased weight ($F_{1,16}=22.54, P<.001$). Conversely, high symptom severity and lower text message response rates were associated with trend-level increases in weight ($F_{1,7}=4.33, P=.08$). Loss-of-control eating was not observed to impact treatment outcome. Participants voiced preference for combination of live health coaching and text messaging, expressing desire for personalized message content.

Conclusions: These results demonstrate the feasibility of delivering an adapted iOTA to SMI patients receiving care in community settings and suggest testable criteria for defining sufficient treatment engagement and psychiatric symptom severity, two factors known to impact weight loss outcomes. These important findings suggest specific adaptations may be needed for optimal treatment outcomes in individuals with SMI.

(JMIR Form Res 2022;6(12):e38496) doi:[10.2196/38496](https://doi.org/10.2196/38496)

KEYWORDS

obesity; mentally ill people/persons; health services; mobile health

Introduction

The field of obesity medicine has embraced the concept that excess adiposity is a disease state with important neurobehavioral causes and consequences [1-3]. Thus, sustained weight management may not be achievable by altering energy balance so that the amount of energy consumed (eg, calories) is less than the amount of energy expended (eg, during exercise), as evidenced by emerging treatments that target mechanisms involving central reward and peripheral modulation of satiety neurocircuitry [4], insulin sensitivity [5,6], and inflammation [7]. Nonetheless, the primary tenets of successful weight loss remain deeply rooted in lifestyle change, with interventions including intensive coaching interactions, highly trained interventionists, and robust clinical infrastructure being most effective [8]. Behavioral phenotyping of individuals achieving long-term weight control suggests that the ability to sustain negative energy balance with a combination of dietary restriction and high levels of physical activity, along with frequent self-monitoring, are indicators of long-term success [9]. However, individuals who sustain clinically significant weight loss also have greater distress over body image and are more likely to engage in disordered eating behaviors [10]. Psychiatric comorbidity is common in treatment-seeking populations [11,12], with eating disorders being more prevalent in high-weight individuals than in those with normal or low weight [13]. These complicating factors contribute to delayed or suboptimal treatment engagement and response [14-16].

Patients with chronic severe mental illness (SMI) are at high risk for developing obesity and related adverse health conditions, attributable to treatment with obesogenic medications [17], sedentary lifestyle [18], and unbalanced dietary intake [19]. Individuals with SMI express a preference for mobile access to behavioral treatments [20,21] but report unique challenges to engagement in lifestyle changes that are specific to the type and severity of their behavioral health symptoms [22,23]. Behavioral interventions to reverse preexisting obesity in chronic SMI demonstrate modest effectiveness during active intervention [22,24,25], but longer term benefits are attenuated or reduced, in part due to limited engagement [16,26-28]. As in the general population, more frequent contacts and longer intervention periods are associated with better adherence and long-term effectiveness [29,30]. However, staffing and other cost requirements (eg, gym memberships) may impact large-scale implementation efforts [22,25,31-33]. Incorporation of mobile health (mHealth) components into lifestyle interventions specifically adapted for people with SMIs treated in community settings may improve scalability, engagement and long-term maintenance effects.

Interactive obesity treatment approaches (iOTAs) employing telephone or digital strategies have been used to engage lower-income and underrepresented communities in lifestyle changes that promote healthy body weight. For example, the Be Fit Be Well intervention, based on the well-known Dietary Approaches to Stop Hypertension program, produced weight

loss at 24 months by extending health coaching with automated telephone messaging and internet-based self-guided content to increase engagement [34]. Be Fit Be Well has become a platform for further iOTA adaptations in underrepresented groups [35-38], where access to internet and web-enabled devices may be limited. A further adaptation of Be Fit Be Well, the Working for You iOTA for lower-income hospital workers, amplified quarterly face-to-face health coach interactions with daily interactive, semiautomated short message service (SMS) text messaging [39-41]. The use of low-cost SMS technology makes the Working for You intervention an ideal iOTA for people living with SMIs, who may have limited access to smart or mobile devices or the internet [21,42,43].

In this study, we tested the usability of the Working for You iOTA in settings where patients with SMIs are most likely to engage in psychosocial rehabilitation and mental health treatment—outpatient community clinical settings [22,25]. We sought to establish criterion-level treatment engagement, defining characteristics that would allow for operationalizing engagement criteria in a scaled intervention and identifying participants less likely to engage, hypothesizing that illness severity and SMS response rate would be associated with weight change. The new operationalized inclusion criteria were then applied to evaluate whether symptom severity and early intervention engagement would impact treatment outcome. We also obtained acceptability and usability data to guide future treatment adaptation.

Methods

Participants

Individuals ages 16 to 75 years who were actively engaged in outpatient community mental health clinic or clubhouse programming in the St. Louis and surrounding urban, suburban, and rural areas [44-46] were eligible for participation. Clubhouses are community-based programs that provide structured daytime programming for psychosocial rehabilitation through supported educational, vocational, and social activities, referring to participants as members rather than clients or patients [47].

A priori study inclusion criteria were BMI >28 and diagnosis of a severe and persistent mental illness (recurrent major depressive disorder, schizophrenia spectrum disorder, or bipolar disorder) confirmed with medical record review. Known eating disorder diagnosis, active substance use disorder, and acute suicidality were exclusionary.

Ethics Approval

This study was approved by the Washington University in St. Louis institutional review board (protocol number 201706118). Capacity to provide informed consent was confirmed by assessing basic medical literacy using the Rapid Estimate of Adult Literacy in Medicine–Short Form (REALM-SF), a 7-item word recognition test, with a score of 5 or higher being consistent with an 8th grade reading level and minimal ability

to understand medical terminology [48] and the University of California, San Diego, Brief Assessment of Capacity to Consent (UBACC) [49], a 10-item scale that includes questions focusing on understanding of the information concerning a specific research protocol as an indicator of decisional capacity.

Description of the Parent Intervention

The parent iOTA was developed for low-wage hospital workers participating in the Working for You study [22]. Behavioral goals, scripts for the text messages, and counseling approach in the parent iOTA Working for You were developed based on previous iterations of the intervention [39,41] and previously reported effective weight-control interventions with low-income individuals and the challenges they face in terms of behavior change [50]. Specifically, low-income individuals struggle with access to care due to transportation issues reducing engagement.

In the Working for You study, participants met one-to-one with a health coach on a quarterly basis. At the first study visit, the health coach met with participants to obtain written informed consent, review the individual's health risk assessment, and choose up to 3 behavior change goals related to principles of

energy balance. Goals were set based on behaviors identified as those (1) in highest need of change, (2) for which the participant has high self-efficacy and readiness for change, (3) for which the participant identifies few change barriers, and (4) that fulfill the intended impact on energy balance. Goals domains are based on simple, concrete behavior changes known to be effective based on empirical evidence of link to energy balance/weight and relevance to low-income populations [39,41]. Subsequent in-person health coaching visits occurred on a quarterly basis and were designed to review goal progress, problem-solve barriers to behavior change using motivational interviewing principles, and revise goal selections as needed.

Table 1 presents the options for goals from which participants could choose. Participants received text messages 5 days per week that are directly linked to goals selected in coaching sessions. Automated weekly SMS check-ins prompted a weekly reply with weight and progress toward selected goals. Participants were offered an automatic opportunity to increase the level of their selected goals if they were successful at meeting their dietary or physical activity goal target for 2 weeks in a row.

Table 1. Goal categories and associated health tip text messages.

Goal category	Representative text
Activity goals	
Steps	It's easy to add steps to your day! Take the stairs instead of the elevator; get off the bus a stop early.
Brisk activity	Is your brisk activity brisk enough? A good test: You should still be able to talk easily but singing puts you out of breath.
Dietary goals	
Sugar-sweetened beverages	Sugary drinks include juice, regular soda, sports/energy drinks, & sugar-sweetened teas/coffees. Cutting back is good. Zero is best.
Healthy breakfast	A protein bar and piece of fruit can be an easy on-the-go breakfast. Choose bars with at least 10 grams of protein and 200 or fewer calories.
Purchased meals	If you're craving fast food, choose a lower-calorie option, like a small burger, or small fries, or small bean burrito.
Purchased snacks	Save some money and calories by packing healthy snacks in your lunch instead of buying them at vending machines and gift shops.
Free food	Sometimes we eat free food just because it's free. Slow down and decide if the calories are really worth it.
Eat meals at home	Don't like to cook? Start simple. Try sandwiches, whole-grain cereal, canned low-sodium beans, or quick-cook brown rice. Whatever works.
Low-fat dairy	2% milk is a good choice over whole milk. Even better is 1% or skim milk. Work your way down. You can do this!
Fruits & Vegetables	Eating a lot of fruits & veggies can help keep hunger away and your weight in check. Work up to 5 or more servings a day.
Vegetables	Stock up on frozen vegetables, so you can just grab them out of the freezer when you need them. Plus, they taste great and are just as healthy as fresh.
Whole grains	Boost your whole grains with popcorn. Buy kernels and pop them in a brown paper bag in the microwave. Use a dash of powder seasoning for taste.
High-fat meats	Beans & lentils are a great substitution for meat. You can make a lot of different dishes with them, and they're cheap & filling.
High-calorie snacks	Choose healthy snacks under 200 calories like 20 to 25 nuts, a banana & 1 low-fat string cheese, or baby carrots & 2 tablespoons of hummus.
Screen time snacks	Instead of automatically grabbing a snack while watching TV, have some unsweetened tea, a diet soda, or zero-calorie fizzy water instead.
Added calories	Tacos, burritos & nachos can be a minefield of added calories. Replace high-calorie toppings with tomatoes, salsa, lettuce, onions, and jalapenos.
Total calories	Think of calories like money in a checking account. You can spend them any way you want, as long as they balance out at your calorie goal over time.
Portion control	Eating slowly is a great way to feel full with smaller portions. It gives your stomach time to tell your brain when it's had enough.
Dietary self-monitoring	Be honest with yourself when writing down what you're eating & drinking. Food logs help you the most when they're as accurate as possible.

Modifications to the Parent Intervention

Participants in this study underwent 12 weeks of treatment consisting of either monthly one-on-one in-person visits for participants seen in the outpatient clinic setting or monthly group sessions to deliver educational content, directly followed by brief one-to-one goal-setting for participants seen in the clubhouse setting. To more closely parallel the chronic care model in which most SMI patients receive care, participants in this study met monthly with health coaches. Additional treatment adaptations included adding weekly phone check-ins as needed, elements of problem-solving therapy (coaching, behavioral modeling and shaping, rehearsing and providing feedback on new behaviors, and positive reinforcement of desired health

behaviors) [51] and motivational interviewing (exploring ambivalence, assessing confidence in ability to change, and shoring up self-efficacy for health behaviors) [52] in monthly in-person meetings to address barriers to behavior change. No modifications were made to the program health goals or related SMS-delivered health tips.

As in the parent intervention, participants were prompted weekly to respond via SMS text with their weight and progress toward health goals. During monthly in-person coaching meetings, the health coach reviewed progress toward goals, problem solved barriers to goal achievement, and assisted participants in selecting new goals, if necessary, at each monthly meeting based on mastery, preference, and energy balance priority. Session 1

consisted of a weigh-in and discussion of the relationship between health behaviors, health goals, and body weight (10 minutes), followed by discussion of energy balance and self-monitoring (30 minutes). The session concluded with a review of health goals, revising and scaling difficulty as appropriate (10 minutes). Subsequent visits followed the same format, with the 30-minute discussion including session-specific content (session 2: meal planning and nutrition, session 3: physical activity, and session 4: skills for long-term health behavior change). The order of the content delivered in the sessions was determined based on prior behavioral weight loss treatment development experience in youth and adults with SMI treated in community settings, where qualitative efforts identified the greatest health behavior knowledge and skill deficits in this population were understanding nutrition and healthy meal planning [46,53].

Defining Characteristics Associated With Intervention Engagement

The primary objective of this study was to evaluate factors associated with engagement, a primary indicator of successful weight loss in behavioral lifestyle interventions [8,54]. We evaluated illness severity, treatment engagement (measured by weekly text messaging response rate), and loss-of-control eating as potential exclusion criteria. Severe psychiatric symptoms were defined as baseline clinician-administered Clinical Global Impression (CGI) [55] severity score of >5 (1=not at all ill, 2=borderline mentally ill, 3=mildly ill, 4=moderately ill, 5=markedly ill, 6=severely ill, 7=among the most extremely ill patients). Engagement was defined as text messaging response rate of >80% over the first 4 weeks of treatment.

The presence of loss-of-control eating was assessed with the 7-item Loss of Control Over Eating Score (LOCES) [56], which has demonstrated good test-retest reliability at 4 weeks (Cohen $d=0.82$, $P<.001$), strong content validity and internal consistency, factor structure, and convergent and discriminant validity [57]. The assessment was administered by a study clinician at the initial study visit. Participants were asked to rate how often they experience behavioral (“I found myself eating despite negative consequences”) and dissociative (“my eating felt like a ball rolling down a hill that just kept going and going”) symptoms associated with loss-of-control eating behavior in the prior 4 weeks (1=never, 2=rarely/once weekly, 3=sometimes/2-3 times per week, 4=often/4-5 times per week, 5=always/daily). Questions were asked in a conversational way, and feelings of shame or guardedness in responding were normalized and validated to minimize bias in reporting symptoms. Care setting (clubhouse versus community mental health clinic) was also evaluated.

Usability, Acceptability, and Feasibility

We evaluated participant experiences with the in-person and text-messaging aspects of the intervention, specifically assessing satisfaction with treatment content, visit and text messaging

frequency, and usability of the text messaging portion of the intervention. At the end of the 12-week study, participants completed a 5-question treatment satisfaction survey. Questions were based on the Contextual Technology Adaptation Process (CTAP) developed by Lyon and colleagues [58].

The CTAP model is based on user-centered design and implementation science principles and incorporates mixed quantitative and qualitative assessments considering aspects of the technology under study (eg, complexity, intended frequency of use), contexts in which the technology will be used (eg, user types and experiences, organizational setting, culture, and policies), and resources available for adaptation efforts (eg, time, money to devote to programming). CTAP involves 5 phases, including initial assessment in relevant contexts, testing of unadapted technology, adaptation, retesting, and sustained iterative assessment and adaptation processes.

This study used early-phase CTAP approaches to assess the acceptability of the unadapted technology in a new user population (adults with SMI) and treatment settings (clubhouse or community mental health clinic). Our CTAP questionnaire consisted of 5 questions focused on satisfaction and acceptability of the (1) overall program, (2) in-person health coaching visits, (3) goal options, (4) text message responses, and (5) health tip text message content. Response options were 1=very unsatisfied or unhelpful to 5=very satisfied or helpful. Each domain included 3 corresponding open-ended questions: What did you like? What did you dislike? What would you change?

Analytic Approach

Repeated measures analysis of covariance was used to test for the effect of time on weight change, including 2-level factors for membership in included/excluded participant group (Clinical Global Impression-measured illness severity at baseline, text message response rate <80% in the first 4 weeks), and treatment setting, as well as an exploratory covariate representing score on the LOCES. Finally, we evaluated whether treatment setting for monthly in-person (one-to-one versus group) sessions influenced weight change over time. Significance was set at $P<.05$ using a 2-tailed test. CTAP responses were tabulated and presented as frequency and percentage followed by a representative quotation from corresponding open-ended questions.

Results

Participant Characteristics

A total of 26 participants were recruited for the study (6/26, 24% schizophrenia; 17/26, 68% mood disorder). The mean age of the overall population was 48.5 (SD 15.67) years; 60% (15/26) were white and 62% (16/26) female (Table 2). One participant was excluded per protocol for an alcohol use disorder relapse during study participation.

Table 2. Participant demographics and characteristics at enrollment.

Characteristic	Total (n=25)	Clubhouse (group), (n=12)	Outpatient clinic (individual), (n=13)
Demographics			
Age (years), median (IQR)	50.0 (39.0-59.5)	55.0 (46.5-63.8)	48.0 (24.5-53.0)
Male, n (%)	10.0 (40.0)	5 (41.7)	5 (38.5)
White, n (%)	15 (60.0)	6 (50.0)	9 (69.2)
Hispanic, n (%)	2 (8.0)	0 (0.0)	2 (15.4)
Clinical assessments			
Weight (lbs), mean (SD)	224.0 (195.0-291.0)	227.0 (208.0-287.8)	220.0 (181.0-300.0)
CGI-S ^a , median (IQR)	4.0 (3.0-4.0)	4.0 (3.3-4.0)	4.0 (3.0-4.5)
LOCES ^b , median (IQR)	18.0 (12.0-20.0)	19.5 (18.0-28.5)	16.0 (10.5-18.5)
Primary psychiatric diagnosis, n (%)			
Schizophrenia	6 (24.0)	3 (25.0)	3 (23.1)
Bipolar disorder	12 (48.0)	8 (66.7)	4 (30.8)
MDD ^c	5 (20.0)	1 (8.3)	4 (30.8)
ADHD ^d	1 (4.0)	0 (0.0)	1 (7.7)
ASD ^e	1 (4.0)	0 (0.0)	1 (7.7)

^aCGI-S: Clinical Global Impression–Severity.

^bLOCES: Loss of Control Over Eating Scale.

^cMDD: Major depressive disorder.

^dADHD: Attention deficit-hyperactivity disorder.

^eASD: Autism spectrum disorder.

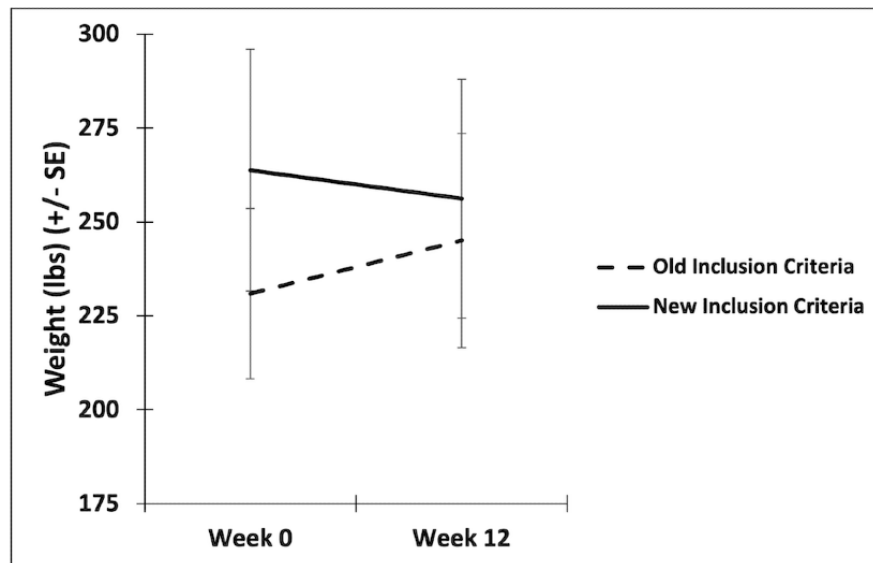
Engagement Characteristics

Of the 26 participants who enrolled, 13 received in-person health coaching in a group-based clubhouse setting (mean -5.4 , SD 6.9 lbs) and 12 received one-to-one health coaching with a study interventionist (mean 3.9 , SD 20 lbs). One participant lost eligibility due to relapse in substance use. In the pooled treatment group, mean weight change was -0.7 (SD 15.4) lbs (min -20.5 lbs, max 58 lbs). Eight participants met the exclusion criteria under evaluation; 2 participants had a Clinical Global Impression score >5 ; mean weight change during treatment in this group was 32.9 (SD 25.1) lbs. Six participants had a text message response rate $<80\%$; mean weight change in this group was 8.1 (SD 3.8) lbs.

Using repeated measures analysis of covariance (Figure 1), a significant interaction was observed between group (eg, prior to and following application of revised inclusion criteria) and time ($F_{1,23}=17.98$, $P<.001$), explained by lower symptom severity and $>80\%$ text-messaging response rate in the first 4 weeks of participation ($n=18$) exhibiting a significant decrease in weight ($F_{1,16}=22.54$, $P<.001$). Participants with high symptom severity and low treatment engagement ($n=8$) had a trend-level increase in weight ($F_{1,7}=4.33$, $P=.08$).

We also tested the interactive effect of time and treatment setting, and LOCES score on change in weight over 12 weeks of iOTA. No significant interactions were observed with treatment setting ($F_{1,23}=2.22$, $P=.15$) or LOCES score ($F_{1,22}=0.02$, $P=.90$).

Figure 1. Intervention-related change in weight before and after adding engagement and illness severity inclusion criteria with standard error (SE) bars.



Treatment Acceptability and Satisfaction

CTAP questions regarding acceptability and satisfaction with the intervention (Table 3) and optional, open-ended questions regarding likes, dislikes, and aspects of the intervention participants would want to change (Table 4) were administered at the final study visit using a paper version of the questionnaire (Multimedia Appendix 1). Quotes were transcribed from the handwritten responses. The majority (23/26, 92%) of participants reported general satisfaction with the treatment program. Those who reported a mild level of dissatisfaction (2/26, 8%) reported that they experienced the program positively but wanted it to be longer. Participants in general (23/26, 92%) reported a moderate to high level of satisfaction with the in-person health coaching visits, with those reporting neutral or mild dissatisfaction attributing this to a desire for more phone or text

message access to the health coach. In terms of in-person treatment content, most (24, 96%) participants felt the health goals offered in the unadapted program were helpful but expressed a preference for the option to personalize health goals.

The weekly SMS prompts to track weight and goal progress were viewed positively (23/26, 92%) in that they provided an easy option for accountability and self-monitoring. However, to be associated with the selected health goals, responses included prescribed tags consisting of 3-5 letters in all caps (eg, participants were reminded to type in LBS before their weekly weight) were difficult to remember; omitting the response tag or responding with an incorrect tag generated error messages that may have resulted in abandonment of the check-in process. Finally, participants in general felt the text message content was helpful but expressed a desire for customizable text messages relevant to personalized goals.

Table 3. Contextual Technology Adaptation Process questions.

Question	Very unsatisfied/unhelpful	Somewhat unsatisfied/unhelpful	Neutral	Somewhat satisfied/helpful	Very satisfied/helpful
How satisfied were you with the overall program?	— ^a	2 (7.7)	1 (3.8)	4 (15.4)	19 (73.1)
How satisfied were you with the amount of contact with the health coach?	—	1 (3.8)	1 (3.8)	5 (19.2)	18 (69.2)
How helpful were the health goal options offered in the program?	—	—	1 (3.8)	7 (26.9)	17 (65.4)
How helpful was the weekly weight check-in via text message?	1 (3.8)	1 (3.8)	1 (3.8)	5 (19.2)	18 (69.2)
How helpful were the daily health tip text messages?	1 (3.8)	2 (7.7)	2 (7.7)	6 (23.1)	15 (57.7)

^aNot applicable.

Table 4. Representative quotes from open-ended Contextual Technology Adaptation Process questions.

Representative verbatim quotes in response to: What did you like? What did you dislike? What would you change?	Acceptability score associated with representative quote
"I'm very thankful to have been in this program. It improved my mental and physical state."	Somewhat satisfied
"The program is too short and needs to be longer. One year would be good."	Very satisfied
"I liked the encouragement and interaction [with] my health coach, and the helpful information I received"	Very satisfied
"Would like more calls or texts with the health coach."	Very satisfied
"I liked the options of health goals that were offered & meeting in person—I felt motivated to make changes."	Very satisfied
"More customizable goals."	Very satisfied
"The accountability reminder to weigh and stay on track every week was helpful."	Very helpful
"Responding was confusing. If I used the wrong tag to answer, I got an error message."	Somewhat helpful
"The texts are kinda like having a little voice in the back of my head reminding me of stuff!"	Very helpful
"...More inspirational, motivational, educational texts that are longer and more frequent."	Neutral

Discussion

Principal Findings

In this study, we explored the feasibility of delivering an interactive text messaging weight loss intervention to adults with SMIs treated in community clinical settings. Our initial hypothesis, driven by results from previous studies of weight loss interventions in SMI populations [16,59], was that illness severity would play an important role in engagement. Secondly, we hypothesized that an mHealth intervention would be well accepted in this population, potentially increasing engagement and improving outcomes. We also aimed to establish minimal engagement criteria for future study by characterizing participants most and least likely to respond to this iOTA based on severity of illness [15], early engagement [60,61], and loss-of-control eating [62,63], all items that have been identified as barriers to obesity treatment success in people with SMIs. Finally, we aimed to collect usability data to identify areas for additional treatment adaptation specific to people with SMIs.

In this usability test, engagement, measured by percentage response to SMS messages within the first month of participation, and lower severity of illness at baseline predicted change in weight at 12 weeks. Qualitative data on user experience and user satisfaction indicated that participants felt positively toward text messaging. Participants still felt that in-person visits were important but would also be acceptable if done remotely via telehealth or text. Like prior studies testing usability of health behavior change apps [64], participants with SMIs wanted more personalized goals and texts and simplified ways to respond to accountability check-ins. Longer term engagement was also preferred by many and is consistent with recommended length of treatment [31,65].

Comparison With Prior Work

It is well known that obesity and mental illness are highly comorbid, with most patients seeking obesity treatment having two or more treatable diagnoses [11,12]. Further, individuals who receive appropriate treatment to stabilize psychiatric symptoms experience benefits of obesity treatments and have

similar outcomes to those without psychiatric illnesses [66]. However, clinical populations with mental health conditions have much higher rates of obesity [67,68] and increased risk for cardiometabolic conditions like diabetes and hypertension than the general population [69-71]. They also exhibit attenuated response to unadapted obesity treatments, which is thought to be related to numerous hampering factors that decrease engagement, including low motivation and cognitive difficulties [22,53], socioeconomic disadvantages such as limited transportation [72], and psychosocial disadvantages preventing engagement in clinical services that could be reduced with mHealth approaches [73]. People with SMIs are amenable to mHealth interventions for weight management [74] and already have access to and familiarity with low-tech methods like SMS [21]. Combined interventions that include in-person visits with mHealth extenders for coaching might be reasonably implemented in settings where monthly or more frequent in-person visits are part of the existing care structure.

Limitations

This study is subject to important limitations; namely, the sample size is small and relatively heterogeneous in terms of diagnosis, care setting, and baseline weight and age. Thus, adequate power to detect between-group differences in subgroup analyses is limited. It should also be noted that the inclusion of patients engaged in services only limits the generalizability of the results; people with SMI are often disengaged from clinical care for a variety of reasons having to do with disparity and disadvantage [75]. As a result, individuals who are not engaged in care may have needs and preferences that could be uniquely addressed by a digital health intervention [76] but who are not reached by recruiting only from populations encountered in clinical care settings. Additionally, while treatment content and procedures were consistent regardless of treatment setting, feasibility of implementation in each setting was not formally assessed. Further study is needed to evaluate additional adaptation needs specific to treatment setting, including qualitative assessment of implementation barriers and facilitators from perspectives of decision-makers, clinicians, and patients.

Conclusions

These data are relevant to future study design considerations and support further testing of specific exclusion criteria for defining treatment engagement and psychiatric symptom control. Future studies applying these selection criteria may be important in evaluating the effect of this iOTA in SMI but will need to consider methods for reaching individuals who may benefit most from mHealth interventions—those who are disengaged from care, from underrepresented or disadvantaged populations,

or who live in rural areas. Finally, more comprehensive symptom assessments may be needed to understand the effect of eating disorder symptoms on weight change outcomes in this populations. Nonetheless, these results demonstrate the feasibility of delivering an adapted iOTA intervention to SMI patients receiving care in clubhouse and community mental health clinic settings and suggest testable criteria for defining sufficient treatment engagement and psychiatric symptom severity.

Acknowledgments

The authors would like to thank the clinical leadership, clinicians, staff, and consumers at BJC Behavioral Health, Independence Center and Places for People in St. Louis Missouri, and Fellowship House and Thriving Mind South Florida in Miami, FL. This study was funded by the Washington University School of Medicine (WUSM) Center for Diabetes Translational Research; grant P30DK092950 from the National Institute of Diabetes and Digestive and Kidney Diseases, the WUSM Center for Brain Research in Mood Disorders, and grant UL1 TR000448 from the WUSM Dissemination & Implementation Research Core and the WUSM Institute of Clinical and Translational Sciences. The WUSM Mobile Health Research Core and the Dissemination Implementation Research Core, both part of the WUSM Institute of Clinical and Translational Sciences, the Psychiatry Resident Research Education Program (R25 MH112473), and the Washington University Applied Health Behavior Research Program also contributed to this work.

Authors' Contributions

As principal investigator, GN was responsible for conception and design of the study and involved in data acquisition and analysis and interpretation of results. MJ participated in the interpretation of results and was involved in drafting and revising the manuscript. RH participated in the acquisition and analysis of data as a health coach in the study, assisted in interpretation of results, and contributed to drafting and revising the manuscript. AR helped to design the study, functioned as the primary health coach interventionist in the study, and was involved in the acquisition of data and interpretation of results and critical review and revision of the manuscript. MDY was the primary data analyst for the study and was involved in the interpretation of results and drafting and revising the manuscript. JAS supervised day-to-day study conduct, participating in aspects of study design, data acquisition, and management; interpretation of results; and critically reviewing and revising the manuscript. KK participated in the acquisition and analysis of data as a health coach, assisted in interpretation of results, and assisted in critical review and revision of the manuscript. BAE created the parent interactive obesity treatment approach as the principal investigator of the Working for You study and participated extensively in early modifications of the intervention and interpretation of results and writing and revising the manuscript. JWN played a key role in helping to design the study and make early modifications to the iOTA intervention and was extensively involved in data acquisition, analysis, and interpretation of results and drafting and critically reviewing and revising the manuscript.

Conflicts of Interest

GN has received grant support from the National Institutes of Health (NIH), Health Resources and Services Administration, Barnes Jewish Hospital Foundation, McDonnell Center for Systems Neuroscience, and Usona Institute (drug only) and has served as a consultant for Alkermes Inc, Otsuka, and Sunovion. BAE has received funding from the NIH and the Centers for Disease Control and Prevention. JWN has received grant support from the NIH and the Substance Abuse and Mental Health Services Administration; served as a consultant for Alkermes Inc, Intra-cellular Therapies Inc, Sunovion, and Merck; and served on a Data Safety Monitoring Board for Amgen. MJ, AR, JAS, MDY, RH, and KK have no disclosures.

Multimedia Appendix 1

Behavioral assessment questionnaire.

[[PDF File \(Adobe PDF File\), 156 KB - formative_v6i12e38496_app1.pdf](#)]

References

1. Vainik U, Baker TE, Dadar M, Zeighami Y, Michaud A, Zhang Y, et al. Neurobehavioral correlates of obesity are largely heritable. *Proc Natl Acad Sci U S A* 2018 Sep 11;115(37):9312-9317 [[FREE Full text](#)] [doi: [10.1073/pnas.1718206115](https://doi.org/10.1073/pnas.1718206115)] [Medline: [30154161](https://pubmed.ncbi.nlm.nih.gov/30154161/)]
2. Wyckoff EP, Evans BC, Manasse SM, Butryn ML, Forman EM. Executive functioning and dietary intake: neurocognitive correlates of fruit, vegetable, and saturated fat intake in adults with obesity. *Appetite* 2017 Apr 01;111:79-85 [[FREE Full text](#)] [doi: [10.1016/j.appet.2016.12.039](https://doi.org/10.1016/j.appet.2016.12.039)] [Medline: [28042040](https://pubmed.ncbi.nlm.nih.gov/28042040/)]

3. Vainik U, Dagher A, Dubé L, Fellows LK. Neurobehavioural correlates of body mass index and eating behaviours in adults: a systematic review. *Neurosci Biobehav Rev* 2013 Mar;37(3):279-299 [FREE Full text] [doi: [10.1016/j.neubiorev.2012.11.008](https://doi.org/10.1016/j.neubiorev.2012.11.008)] [Medline: [23261403](https://pubmed.ncbi.nlm.nih.gov/23261403/)]
4. Hollmann M, Hellrung L, Pleger B, Schlögl H, Kabisch S, Stumvoll M, et al. Neural correlates of the volitional regulation of the desire for food. *Int J Obes (Lond)* 2012 May;36(5):648-655. [doi: [10.1038/ijo.2011.125](https://doi.org/10.1038/ijo.2011.125)] [Medline: [21712804](https://pubmed.ncbi.nlm.nih.gov/21712804/)]
5. Nicol GE, Yingling MD, Flavin KS, Schweiger JA, Patterson BW, Schechtman KB, et al. Metabolic effects of antipsychotics on adiposity and insulin sensitivity in youths: a randomized clinical trial. *JAMA Psychiatry* 2018 Aug 01;75(8):788-796 [FREE Full text] [doi: [10.1001/jamapsychiatry.2018.1088](https://doi.org/10.1001/jamapsychiatry.2018.1088)] [Medline: [29898210](https://pubmed.ncbi.nlm.nih.gov/29898210/)]
6. Correll CU, Manu P, Olshanskiy V, Napolitano B, Kane JM, Malhotra AK. Cardiometabolic risk of second-generation antipsychotic medications during first-time use in children and adolescents. *JAMA* 2009 Oct 28;302(16):1765-1773 [FREE Full text] [doi: [10.1001/jama.2009.1549](https://doi.org/10.1001/jama.2009.1549)] [Medline: [19861668](https://pubmed.ncbi.nlm.nih.gov/19861668/)]
7. Kowalchuk C, Kanagasundaram P, McIntyre WB, Belsham DD, Hahn MK. Direct effects of antipsychotic drugs on insulin, energy sensing and inflammatory pathways in hypothalamic mouse neurons. *Psychoneuroendocrinology* 2019 Nov;109:104400. [doi: [10.1016/j.psyneuen.2019.104400](https://doi.org/10.1016/j.psyneuen.2019.104400)] [Medline: [31404896](https://pubmed.ncbi.nlm.nih.gov/31404896/)]
8. LeBlanc E, Patnode C, Webber E, Redmond N, Rushkin M, O'Connor EA. Behavioral and pharmacotherapy weight loss interventions to prevent obesity-related morbidity and mortality in adults: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 2018 Sep 18;320(11):1172-1191. [doi: [10.1001/jama.2018.7777](https://doi.org/10.1001/jama.2018.7777)] [Medline: [30326501](https://pubmed.ncbi.nlm.nih.gov/30326501/)]
9. Wing RR, Phelan S. Long-term weight loss maintenance. *Am J Clin Nutr* 2005 Jul;82(1 Suppl):222S-225S. [doi: [10.1093/ajcn/82.1.222S](https://doi.org/10.1093/ajcn/82.1.222S)] [Medline: [16002825](https://pubmed.ncbi.nlm.nih.gov/16002825/)]
10. Feller S, Müller A, Mayr A, Engeli S, Hilbert A, de Zwaan M. What distinguishes weight loss maintainers of the German Weight Control Registry from the general population? *Obesity (Silver Spring)* 2015 May;23(5):1112-1118 [FREE Full text] [doi: [10.1002/oby.21054](https://doi.org/10.1002/oby.21054)] [Medline: [25864871](https://pubmed.ncbi.nlm.nih.gov/25864871/)]
11. Dreber H, Reynisdottir S, Angelin B, Hemmingsson E. Who is the treatment-seeking young adult with severe obesity: a comprehensive characterization with emphasis on mental health. *PLoS One* 2015 Dec 22;10(12):e0145273 [FREE Full text] [doi: [10.1371/journal.pone.0145273](https://doi.org/10.1371/journal.pone.0145273)] [Medline: [26694031](https://pubmed.ncbi.nlm.nih.gov/26694031/)]
12. Dawes AJ, Maggard-Gibbons M, Maher AR, Booth MJ, Mlake-Lye I, Beroes JM, et al. Mental health conditions among patients seeking and undergoing bariatric surgery: a meta-analysis. *JAMA* 2016 Jan 12;315(2):150-163. [doi: [10.1001/jama.2015.18118](https://doi.org/10.1001/jama.2015.18118)] [Medline: [26757464](https://pubmed.ncbi.nlm.nih.gov/26757464/)]
13. Duncan AE, Ziobrowski HN, Nicol G. The prevalence of past 12-month and lifetime DSM-IV eating disorders by BMI category in US men and women. *Eur Eat Disord Rev* 2017 May 27;25(3):165-171. [doi: [10.1002/erv.2503](https://doi.org/10.1002/erv.2503)] [Medline: [28127825](https://pubmed.ncbi.nlm.nih.gov/28127825/)]
14. Hayes JF, Hoggatt KJ, Breland JY. Mental health diagnostic patterns in behavioral weight loss program attendees. *Mil Med* 2020 Aug 14;185(7-8):e1263-e1270 [FREE Full text] [doi: [10.1093/milmed/usaa046](https://doi.org/10.1093/milmed/usaa046)] [Medline: [32328624](https://pubmed.ncbi.nlm.nih.gov/32328624/)]
15. Janney CA, Greenberg JM, Moin T, Kim HM, Holleman RG, Hughes M, et al. Does mental health influence weight loss in adults with prediabetes? Findings from the VA Diabetes Prevention Program. *Gen Hosp Psychiatry* 2018;53:32-37. [doi: [10.1016/j.genhosppsy.2018.04.001](https://doi.org/10.1016/j.genhosppsy.2018.04.001)] [Medline: [29729586](https://pubmed.ncbi.nlm.nih.gov/29729586/)]
16. Nicol GE, Kolko R, Lenze EJ, Yingling MD, Miller JP, Ricchio AR, et al. Adiposity, hepatic triglyceride, and carotid intima media thickness during behavioral weight loss treatment in antipsychotic-treated youth: a randomized pilot study. *J Child Adolesc Psychopharmacol* 2019 Aug;29(6):439-447 [FREE Full text] [doi: [10.1089/cap.2018.0120](https://doi.org/10.1089/cap.2018.0120)] [Medline: [30994376](https://pubmed.ncbi.nlm.nih.gov/30994376/)]
17. Newcomer JW. Second-generation (atypical) antipsychotics and metabolic effects: a comprehensive literature review. *CNS Drugs* 2005;19 Suppl 1:1-93. [doi: [10.2165/00023210-200519001-00001](https://doi.org/10.2165/00023210-200519001-00001)] [Medline: [15998156](https://pubmed.ncbi.nlm.nih.gov/15998156/)]
18. Daumit GL, Clark JM, Steinwachs DM, Graham CM, Lehman A, Ford DE. Prevalence and correlates of obesity in a community sample of individuals with severe and persistent mental illness. *J Nerv Ment Dis* 2003 Dec;191(12):799-805. [doi: [10.1097/01.nmd.0000100923.20188.2d](https://doi.org/10.1097/01.nmd.0000100923.20188.2d)] [Medline: [14671456](https://pubmed.ncbi.nlm.nih.gov/14671456/)]
19. McCreddie RG, Scottish Schizophrenia Lifestyle Group. Diet, smoking and cardiovascular risk in people with schizophrenia: descriptive study. *Br J Psychiatry* 2003 Dec;183:534-539. [doi: [10.1192/bjp.183.6.534](https://doi.org/10.1192/bjp.183.6.534)] [Medline: [14645025](https://pubmed.ncbi.nlm.nih.gov/14645025/)]
20. Seko Y, Kidd S, Wiljer D, McKenzie K. Youth mental health interventions via mobile phones: a scoping review. *Cyberpsychol Behav Soc Netw* 2014 Sep;17(9):591-602. [doi: [10.1089/cyber.2014.0078](https://doi.org/10.1089/cyber.2014.0078)] [Medline: [25007383](https://pubmed.ncbi.nlm.nih.gov/25007383/)]
21. Young AS, Cohen AN, Niv N, Nowlin-Finch N, Oberman RS, Olmos-Ochoa TT, et al. Mobile phone and smartphone use by people with serious mental illness. *Psychiatr Serv* 2020 Mar 01;71(3):280-283. [doi: [10.1176/appi.ps.201900203](https://doi.org/10.1176/appi.ps.201900203)] [Medline: [31744429](https://pubmed.ncbi.nlm.nih.gov/31744429/)]
22. Daumit GL, Dickerson FB, Wang N, Dalcin A, Jerome GJ, Anderson CAM, et al. A behavioral weight-loss intervention in persons with serious mental illness. *N Engl J Med* 2013 Apr 25;368(17):1594-1602 [FREE Full text] [doi: [10.1056/NEJMoa1214530](https://doi.org/10.1056/NEJMoa1214530)] [Medline: [23517118](https://pubmed.ncbi.nlm.nih.gov/23517118/)]
23. Casagrande SS, Jerome GJ, Dalcin AT, Dickerson FB, Anderson CA, Appel LJ, et al. Randomized trial of achieving healthy lifestyles in psychiatric rehabilitation: the ACHIEVE trial. *BMC Psychiatry* 2010 Dec 13;10:108 [FREE Full text] [doi: [10.1186/1471-244X-10-108](https://doi.org/10.1186/1471-244X-10-108)] [Medline: [21144025](https://pubmed.ncbi.nlm.nih.gov/21144025/)]

24. Green CA, Yarborough BJH, Leo MC, Yarborough MT, Stumbo SP, Janoff SL, et al. The STRIDE weight loss and lifestyle intervention for individuals taking antipsychotic medications: a randomized trial. *Am J Psychiatry* 2015 Jan;172(1):71-81 [FREE Full text] [doi: [10.1176/appi.ajp.2014.14020173](https://doi.org/10.1176/appi.ajp.2014.14020173)] [Medline: [25219423](https://pubmed.ncbi.nlm.nih.gov/25219423/)]
25. Robinson DG, Schooler NR, Correll CU, John M, Kurian BT, Marcy P, et al. Psychopharmacological treatment in the RAISE-ETP study: outcomes of a manual and computer decision support system based intervention. *Am J Psychiatry* 2018 Feb 01;175(2):169-179 [FREE Full text] [doi: [10.1176/appi.ajp.2017.16080919](https://doi.org/10.1176/appi.ajp.2017.16080919)] [Medline: [28945118](https://pubmed.ncbi.nlm.nih.gov/28945118/)]
26. Brightman L, Dunne A, Huang HC. Mental illness as a predictor of patient attendance and anthropometric changes: observations from an Australian publicly funded obesity management service. *Australas Psychiatry* 2021 Jun;29(3):266-271. [doi: [10.1177/1039856220953713](https://doi.org/10.1177/1039856220953713)] [Medline: [32910691](https://pubmed.ncbi.nlm.nih.gov/32910691/)]
27. Müller M, Nett PC, Borbély YM, Buri C, Stirnimann G, Laederach K, et al. Mental illness has a negative impact on weight loss in bariatric patients: a 4-year follow-up. *J Gastrointest Surg* 2019 Feb;23(2):232-238. [doi: [10.1007/s11605-018-3903-x](https://doi.org/10.1007/s11605-018-3903-x)] [Medline: [30091038](https://pubmed.ncbi.nlm.nih.gov/30091038/)]
28. Berkowitz RI, Marcus MD, Anderson BJ, Delahanty L, Grover N, Kriska A, TODAY Study Group. Adherence to a lifestyle program for youth with type 2 diabetes and its association with treatment outcome in the TODAY clinical trial. *Pediatr Diabetes* 2018 Mar;19(2):191-198 [FREE Full text] [doi: [10.1111/pedi.12555](https://doi.org/10.1111/pedi.12555)] [Medline: [28664624](https://pubmed.ncbi.nlm.nih.gov/28664624/)]
29. Naslund JA, Whiteman KL, McHugo GJ, Aschbrenner KA, Marsch LA, Bartels SJ. Lifestyle interventions for weight loss among overweight and obese adults with serious mental illness: a systematic review and meta-analysis. *Gen Hosp Psychiatry* 2017 Jul;47:83-102 [FREE Full text] [doi: [10.1016/j.genhosppsych.2017.04.003](https://doi.org/10.1016/j.genhosppsych.2017.04.003)] [Medline: [28807143](https://pubmed.ncbi.nlm.nih.gov/28807143/)]
30. McGinty EE, Gudzone KA, Dalcin A, Jerome GJ, Dickerson F, Gennusa J, et al. Bringing an effective behavioral weight loss intervention for people with serious mental illness to scale. *Front Psychiatry* 2018;9:604 [FREE Full text] [doi: [10.3389/fpsy.2018.00604](https://doi.org/10.3389/fpsy.2018.00604)] [Medline: [30515109](https://pubmed.ncbi.nlm.nih.gov/30515109/)]
31. Green CA, Yarborough BJH, Leo MC, Stumbo SP, Perrin NA, Nichols GA, et al. Weight maintenance following the STRIDE lifestyle intervention for individuals taking antipsychotic medications. *Obesity (Silver Spring)* 2015 Oct;23(10):1995-2001 [FREE Full text] [doi: [10.1002/oby.21205](https://doi.org/10.1002/oby.21205)] [Medline: [26334929](https://pubmed.ncbi.nlm.nih.gov/26334929/)]
32. Alvarez-Jiménez M, Martínez-García O, Pérez-Iglesias R, Ramírez ML, Vázquez-Barquero JL, Crespo-Facorro B. Prevention of antipsychotic-induced weight gain with early behavioural intervention in first-episode psychosis: 2-year results of a randomized controlled trial. *Schizophr Res* 2010 Jan;116(1):16-19. [doi: [10.1016/j.schres.2009.10.012](https://doi.org/10.1016/j.schres.2009.10.012)] [Medline: [19896336](https://pubmed.ncbi.nlm.nih.gov/19896336/)]
33. Bartels SJ, Pratt SI, Aschbrenner KA, Barre LK, Naslund JA, Wolfe R, et al. Pragmatic replication trial of health promotion coaching for obesity in serious mental illness and maintenance of outcomes. *Am J Psychiatry* 2015 Apr;172(4):344-352 [FREE Full text] [doi: [10.1176/appi.ajp.2014.14030357](https://doi.org/10.1176/appi.ajp.2014.14030357)] [Medline: [25827032](https://pubmed.ncbi.nlm.nih.gov/25827032/)]
34. Bennett GG, Warner ET, Glasgow RE, Askew S, Goldman J, Ritzwoller DP, et al. Obesity treatment for socioeconomically disadvantaged patients in primary care practice. *Arch Intern Med* 2012 Apr 9;172(7):565-574 [FREE Full text] [doi: [10.1001/archinternmed.2012.1](https://doi.org/10.1001/archinternmed.2012.1)] [Medline: [22412073](https://pubmed.ncbi.nlm.nih.gov/22412073/)]
35. National Institutes of Health. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. *Obes Res* 1998 Sep;6 Suppl 2:51S-209S [FREE Full text] [Medline: [9813653](https://pubmed.ncbi.nlm.nih.gov/9813653/)]
36. Anderson LM, Quinn TA, Glanz K, Ramirez G, Kahwati LC, Johnson DB, Task Force on Community Preventive Services. The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: a systematic review. *Am J Prev Med* 2009 Oct;37(4):340-357. [doi: [10.1016/j.amepre.2009.07.003](https://doi.org/10.1016/j.amepre.2009.07.003)] [Medline: [19765507](https://pubmed.ncbi.nlm.nih.gov/19765507/)]
37. Special Committee on Health, Productivity, and Disability Management, American College of Occupational and Environmental Medicine. Healthy workforce/healthy economy: the role of health, productivity, and disability management in addressing the nation's health care crisis: why an emphasis on the health of the workforce is vital to the health of the economy. *J Occup Environ Med* 2009 Jan;51(1):114-119. [doi: [10.1097/JOM.0b013e318195dad2](https://doi.org/10.1097/JOM.0b013e318195dad2)] [Medline: [19136880](https://pubmed.ncbi.nlm.nih.gov/19136880/)]
38. Dodson E, Lovegreen S, Elliott M, Haire-Joshu D, Brownson R. Worksite policies and environments supporting physical activity in midwestern communities. *Am J Health Promot* 2008 Oct;23(1):51-55. [doi: [10.4278/ajhp.07031626](https://doi.org/10.4278/ajhp.07031626)] [Medline: [18785375](https://pubmed.ncbi.nlm.nih.gov/18785375/)]
39. Tabak RG, Strickland JR, Stein RI, Dart H, Colditz GA, Kirk B, et al. Development of a scalable weight loss intervention for low-income workers through adaptation of interactive obesity treatment approach (iOTA). *BMC Public Health* 2018 Nov 16;18(1):1265 [FREE Full text] [doi: [10.1186/s12889-018-6176-0](https://doi.org/10.1186/s12889-018-6176-0)] [Medline: [30445939](https://pubmed.ncbi.nlm.nih.gov/30445939/)]
40. Stein RI, Strickland JR, Tabak RG, Dale AM, Colditz GA, Evanoff BA. Design of a randomized trial testing a multi-level weight-control intervention to reduce obesity and related health conditions in low-income workers. *Contemp Clin Trials* 2019 Apr;79:89-97 [FREE Full text] [doi: [10.1016/j.cct.2019.01.011](https://doi.org/10.1016/j.cct.2019.01.011)] [Medline: [30664943](https://pubmed.ncbi.nlm.nih.gov/30664943/)]
41. Tabak RG, Strickland JR, Kirk B, Colvin R, Stein RI, Dart H, et al. Pilot test of an interactive obesity treatment approach among employed adults in a university medical billing office. *Pilot Feasibility Stud* 2020;6:57 [FREE Full text] [doi: [10.1186/s40814-020-00599-w](https://doi.org/10.1186/s40814-020-00599-w)] [Medline: [32355567](https://pubmed.ncbi.nlm.nih.gov/32355567/)]
42. Newport F. The New Era of Communication Among Americans. 2014 Nov 10. URL: <http://news.gallup.com/poll/179288/new-era-communication-americans.aspx> [accessed 2022-10-24]
43. Aschbrenner KA, Naslund JA, Shevenell M, Kinney E, Bartels SJ. A pilot study of a peer-group lifestyle intervention enhanced with mhealth technology and social media for adults with serious mental illness. *J Nerv Ment Dis* 2016 Dec;204(6):483-486 [FREE Full text] [doi: [10.1097/NMD.0000000000000530](https://doi.org/10.1097/NMD.0000000000000530)] [Medline: [27233056](https://pubmed.ncbi.nlm.nih.gov/27233056/)]

44. Nicol GE, Morrato EH, Johnson MC, Campagna E, Yingling MD, Pham V, et al. Best practices: implementation of a glucose screening program based on diffusion of innovation theory methods. *Psychiatr Serv* 2011 Jan;62(1):12-14. [doi: [10.1176/ps.62.1.pss6201_0012](https://doi.org/10.1176/ps.62.1.pss6201_0012)] [Medline: [21209293](https://pubmed.ncbi.nlm.nih.gov/21209293/)]
45. Nicol GE, Campagna EJ, Garfield LD, Newcomer JW, Parks JJ, Morrato EH. The role of clinical setting and management approach in metabolic testing among youths and adults treated with antipsychotics. *Psychiatr Serv* 2016 Jan;67(1):128-132 [FREE Full text] [doi: [10.1176/appi.ps.201400428](https://doi.org/10.1176/appi.ps.201400428)] [Medline: [26325456](https://pubmed.ncbi.nlm.nih.gov/26325456/)]
46. Haddad R, Badke D'Andrea C, Ricchio A, Evanoff B, Morrato EH, Parks J, et al. Using Innovation-Corps (I-Corps™) methods to adapt a mobile health (mhealth) obesity treatment for community mental health settings. *Front Digit Health* 2022;4:835002 [FREE Full text] [doi: [10.3389/fdgh.2022.835002](https://doi.org/10.3389/fdgh.2022.835002)] [Medline: [35721796](https://pubmed.ncbi.nlm.nih.gov/35721796/)]
47. McKay C, Nugent KL, Johnsen M, Eaton WW, Lidz CW. A systematic review of evidence for the clubhouse model of psychosocial rehabilitation. *Adm Policy Ment Health* 2018 Jan;45(1):28-47 [FREE Full text] [doi: [10.1007/s10488-016-0760-3](https://doi.org/10.1007/s10488-016-0760-3)] [Medline: [27580614](https://pubmed.ncbi.nlm.nih.gov/27580614/)]
48. Arozullah AM, Yarnold PR, Bennett CL, Soltysik RC, Wolf MS, Ferreira RM, et al. Development and validation of a short-form, rapid estimate of adult literacy in medicine. *Med Care* 2007 Nov;45(11):1026-1033. [doi: [10.1097/MLR.0b013e3180616c1b](https://doi.org/10.1097/MLR.0b013e3180616c1b)] [Medline: [18049342](https://pubmed.ncbi.nlm.nih.gov/18049342/)]
49. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. *Arch Gen Psychiatry* 2007 Aug;64(8):966-974. [doi: [10.1001/archpsyc.64.8.966](https://doi.org/10.1001/archpsyc.64.8.966)] [Medline: [17679641](https://pubmed.ncbi.nlm.nih.gov/17679641/)]
50. Glasgow RE, Askew S, Purcell P, Levine E, Warner ET, Stange KC, et al. Use of RE-AIM to address health inequities: application in a low-income community health center based weight loss and hypertension self-management program. *Transl Behav Med* 2013 Jun 01;3(2):200-210 [FREE Full text] [doi: [10.1007/s13142-013-0201-8](https://doi.org/10.1007/s13142-013-0201-8)] [Medline: [23750180](https://pubmed.ncbi.nlm.nih.gov/23750180/)]
51. Nezu A, D'Zurilla T. *Problem-Solving Therapy: A Positive Approach To Clinical Intervention*. Berlin: Springer Publishing Company; 2006.
52. Barrett S, Begg S, O'Halloran P, Kingsley M. Integrated motivational interviewing and cognitive behaviour therapy for lifestyle mediators of overweight and obesity in community-dwelling adults: a systematic review and meta-analysis. *BMC Public Health* 2018 Oct 05;18(1):1160 [FREE Full text] [doi: [10.1186/s12889-018-6062-9](https://doi.org/10.1186/s12889-018-6062-9)] [Medline: [30290793](https://pubmed.ncbi.nlm.nih.gov/30290793/)]
53. Nicol G, Worsham E, Haire-Joshu D, Duncan A, Schweiger J, Yingling M, et al. Getting to more effective weight management in antipsychotic-treated youth: a survey of barriers and preferences. *Child Obes* 2016 Feb;12(1):70-76 [FREE Full text] [doi: [10.1089/chi.2015.0076](https://doi.org/10.1089/chi.2015.0076)] [Medline: [26788619](https://pubmed.ncbi.nlm.nih.gov/26788619/)]
54. Morrato EH, Lennox LA, Sendro ER, Schuster AL, Pincus HA, Humensky J, et al. Scale-up of the Accrual to Clinical Trials (ACT) network across the Clinical and Translational Science Award Consortium: a mixed-methods evaluation of the first 18 months. *J Clin Transl Sci* 2020 Jun 30;4(6):515-528 [FREE Full text] [doi: [10.1017/cts.2020.505](https://doi.org/10.1017/cts.2020.505)] [Medline: [33948228](https://pubmed.ncbi.nlm.nih.gov/33948228/)]
55. Guy W. *NCDEU Assessment Manual for Psychopharmacology*. Rockville: Alcohol, Drug Abuse and Mental Health Administration; 1976:217-222.
56. Stefano EC, Wagner AF, Mond JM, Cicero DC, Latner JD. Loss of Control Over Eating Scale (LOCES): validation in undergraduate men and women with and without eating disorder symptoms. *Eat Behav* 2016 Dec;23:137-140. [doi: [10.1016/j.eatbeh.2016.09.005](https://doi.org/10.1016/j.eatbeh.2016.09.005)] [Medline: [27679970](https://pubmed.ncbi.nlm.nih.gov/27679970/)]
57. Latner JD, Mond JM, Kelly MC, Haynes SN, Hay PJ. The Loss of Control Over Eating Scale: development and psychometric evaluation. *Int J Eat Disord* 2014 Sep;47(6):647-659. [doi: [10.1002/eat.22296](https://doi.org/10.1002/eat.22296)] [Medline: [24862351](https://pubmed.ncbi.nlm.nih.gov/24862351/)]
58. Lyon AR, Wasse JK, Ludwig K, Zachry M, Bruns EJ, Untzter J, et al. The contextualized technology adaptation process (CTAP): optimizing health information technology to improve mental health systems. *Adm Policy Ment Health* 2016 May;43(3):394-409. [doi: [10.1007/s10488-015-0637-x](https://doi.org/10.1007/s10488-015-0637-x)] [Medline: [25677251](https://pubmed.ncbi.nlm.nih.gov/25677251/)]
59. Nicol G, Worsham E, Haire-Joshu D, Duncan A, Schweiger J, Yingling M, et al. Getting to more effective weight management in antipsychotic-treated youth: a survey of barriers and preferences. *Child Obes* 2016 Feb;12(1):70-76 [FREE Full text] [doi: [10.1089/chi.2015.0076](https://doi.org/10.1089/chi.2015.0076)] [Medline: [26788619](https://pubmed.ncbi.nlm.nih.gov/26788619/)]
60. Aschbrenner KA, Naslund JA, Shevenell M, Mueser KT, Bartels SJ. Feasibility of behavioral weight loss treatment enhanced with peer support and mobile health technology for individuals with serious mental illness. *Psychiatr Q* 2016 Sep;87(3):401-415 [FREE Full text] [doi: [10.1007/s11126-015-9395-x](https://doi.org/10.1007/s11126-015-9395-x)] [Medline: [26462674](https://pubmed.ncbi.nlm.nih.gov/26462674/)]
61. Olmos-Ochoa TT, Niv N, Hellemann G, Cohen AN, Oberman R, Goldberg R, et al. Barriers to participation in web-based and in-person weight management interventions for serious mental illness. *Psychiatr Rehabil J* 2019 Sep;42(3):220-228. [doi: [10.1037/prj0000363](https://doi.org/10.1037/prj0000363)] [Medline: [31081651](https://pubmed.ncbi.nlm.nih.gov/31081651/)]
62. Bartholdy S, Dalton B, O'Daly OG, Campbell IC, Schmidt U. A systematic review of the relationship between eating, weight and inhibitory control using the stop signal task. *Neurosci Biobehav Rev* 2016 May;64:35-62. [doi: [10.1016/j.neubiorev.2016.02.010](https://doi.org/10.1016/j.neubiorev.2016.02.010)] [Medline: [26900651](https://pubmed.ncbi.nlm.nih.gov/26900651/)]
63. Stunkard A, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res* 1985 Jan;29(1):71-83 [FREE Full text] [doi: [10.1016/0022-3999\(85\)90010-8](https://doi.org/10.1016/0022-3999(85)90010-8)]

64. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophr Bull* 2014 Nov;40(6):1244-1253. [doi: [10.1093/schbul/sbu033](https://doi.org/10.1093/schbul/sbu033)] [Medline: [24609454](https://pubmed.ncbi.nlm.nih.gov/24609454/)]
65. McGinty EE, Sridhara S. Potentially preventable medical hospitalizations among Maryland residents with mental illness, 2005-2010. *Psychiatr Serv* 2014 Jul;65(7):951-953. [doi: [10.1176/appi.ps.201300323](https://doi.org/10.1176/appi.ps.201300323)] [Medline: [26037005](https://pubmed.ncbi.nlm.nih.gov/26037005/)]
66. Kinzl JF, Schrattecker M, Traweger C, Mattesich M, Fiala M, Biebl W. Psychosocial predictors of weight loss after bariatric surgery. *Obes Surg* 2006 Dec;16(12):1609-1614. [doi: [10.1381/096089206779319301](https://doi.org/10.1381/096089206779319301)] [Medline: [17217637](https://pubmed.ncbi.nlm.nih.gov/17217637/)]
67. Allison DB, Loebel AD, Lombardo I, Romano SJ, Siu CO. Understanding the relationship between baseline BMI and subsequent weight change in antipsychotic trials: effect modification or regression to the mean? *Psychiatry Res* 2009 Dec 30;170(2-3):172-176. [doi: [10.1016/j.psychres.2008.10.007](https://doi.org/10.1016/j.psychres.2008.10.007)] [Medline: [19897253](https://pubmed.ncbi.nlm.nih.gov/19897253/)]
68. Dean AG, Arner TG, Sunki GG, Friedman R, Lantinga M, Sangam S, et al. Epi Info™, a database and statistics program for public health professionals. Atlanta, GA, USA: Centers for Disease Control and Prevention; 2011. URL: <http://www.cdc.gov/epiinfo/> [accessed 2022-07-07]
69. Colton CW, Manderscheid RW. Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. *Prev Chronic Dis* 2006 Apr;3(2):A42 [FREE Full text] [Medline: [16539783](https://pubmed.ncbi.nlm.nih.gov/16539783/)]
70. Newcomer JW, Hennekens CH. Severe mental illness and risk of cardiovascular disease. *JAMA* 2007 Oct 17;298(15):1794-1796. [doi: [10.1001/jama.298.15.1794](https://doi.org/10.1001/jama.298.15.1794)] [Medline: [17940236](https://pubmed.ncbi.nlm.nih.gov/17940236/)]
71. Mitchell AJ, Vancampfort D, Sweers K, van Winkel R, Yu W, De Hert M. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders: a systematic review and meta-analysis. *Schizophr Bull* 2013 Mar;39(2):306-318 [FREE Full text] [doi: [10.1093/schbul/sbr148](https://doi.org/10.1093/schbul/sbr148)] [Medline: [22207632](https://pubmed.ncbi.nlm.nih.gov/22207632/)]
72. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol* 2014 Mar 25;14(1):42 [FREE Full text] [doi: [10.1186/1471-2288-14-42](https://doi.org/10.1186/1471-2288-14-42)] [Medline: [24669751](https://pubmed.ncbi.nlm.nih.gov/24669751/)]
73. Polillo A, Voineskos AN, Foussias G, Kidd SA, Sav A, Hawley S, et al. Using digital tools to engage patients with psychosis and their families in research: survey recruitment and completion in an early psychosis intervention program. *JMIR Ment Health* 2021 May 31;8(5):e24567 [FREE Full text] [doi: [10.2196/24567](https://doi.org/10.2196/24567)] [Medline: [34057421](https://pubmed.ncbi.nlm.nih.gov/34057421/)]
74. Naslund JA, Aschbrenner KA, Barre LK, Bartels SJ. Feasibility of popular m-health technologies for activity tracking among individuals with serious mental illness. *Telemed J E Health* 2015 Mar;21(3):213-216 [FREE Full text] [doi: [10.1089/tmj.2014.0105](https://doi.org/10.1089/tmj.2014.0105)] [Medline: [25536190](https://pubmed.ncbi.nlm.nih.gov/25536190/)]
75. Pedersen SL, Lindstrom R, Powe PM, Louie K, Escobar-Viera C. Lack of representation in psychiatric research: a data-driven example from scientific articles published in 2019 and 2020 in the *American Journal of Psychiatry*. *Am J Psychiatry* 2022 May;179(5):388-392. [doi: [10.1176/appi.ajp.21070758](https://doi.org/10.1176/appi.ajp.21070758)] [Medline: [35491569](https://pubmed.ncbi.nlm.nih.gov/35491569/)]
76. Buck B, Chander A, Ben-Zeev D. Clinical and demographic predictors of engagement in mobile health vs. clinic-based interventions for serious mental illness. *J Behav Cogn Ther* 2020 Apr;30(1):3-11 [FREE Full text] [doi: [10.1016/j.jbct.2020.03.004](https://doi.org/10.1016/j.jbct.2020.03.004)]

Abbreviations

CTAP: Contextual Technology Adaptation Process
iOTA: interactive obesity treatment approach
LOCES: Loss of Control Over Eating Scale
mHealth: mobile health
NIH: National Institutes of Health
SMI: severe mental illness
SMS: short message service
WUSM: Washington University School of Medicine

Edited by A Mavragani; submitted 04.04.22; peer-reviewed by R Benson, I Wilson; comments to author 01.08.22; revised version received 16.08.22; accepted 02.09.22; published 13.12.22.

Please cite as:

Nicol G, Jansen M, Haddad R, Ricchio A, Yingling MD, Schweiger JA, Keenoy K, Evanoff BA, Newcomer JW

Use of an Interactive Obesity Treatment Approach in Individuals With Severe Mental Illness: Feasibility, Acceptability, and Proposed Engagement Criteria

JMIR Form Res 2022;6(12):e38496

URL: <https://formative.jmir.org/2022/12/e38496>

doi: [10.2196/38496](https://doi.org/10.2196/38496)

PMID: [36512399](https://pubmed.ncbi.nlm.nih.gov/36512399/)

©Ginger Nicol, Madeline Jansen, Rita Haddad, Amanda Ricchio, Michael D Yingling, Julia A Schweiger, Katie Keenoy, Bradley A Evanoff, John W Newcomer. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Detailed Versus Simplified Dietary Self-monitoring in a Digital Weight Loss Intervention Among Racial and Ethnic Minority Adults: Fully Remote, Randomized Pilot Study

Michele L Patel¹, PhD; Angel E Cleare²; Carly M Smith²; Lisa Goldman Rosas³, MPH, PhD; Abby C King^{1,3}, PhD

¹Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Palo Alto, CA, United States

²Stanford University, Stanford, CA, United States

³Department of Epidemiology & Population Health, Stanford University School of Medicine, Stanford, CA, United States

Corresponding Author:

Michele L Patel, PhD

Stanford Prevention Research Center

Department of Medicine

Stanford University School of Medicine

3180 Porter Drive

Palo Alto, CA, 94304

United States

Phone: 1 650 549 7047

Email: michele.patel@stanford.edu

Abstract

Background: Detailed self-monitoring (or *tracking*) of dietary intake is a popular and effective weight loss approach that can be delivered via digital tools, although engagement declines over time. Simplifying the experience of self-monitoring diet may counteract this decline in engagement. Testing these strategies among racial and ethnic minority groups is important as these groups are often disproportionately affected by obesity yet underrepresented in behavioral obesity treatment.

Objective: In this 2-arm pilot study, we aimed to evaluate the feasibility and acceptability of a digital weight loss intervention with either detailed or simplified dietary self-monitoring.

Methods: We recruited racial and ethnic minority adults aged ≥ 21 years with a BMI of 25 kg/m² to 45 kg/m² and living in the United States. The Pacific time zone was selected for a fully remote study. Participants received a 3-month stand-alone digital weight loss intervention and were randomized 1:1 to either the *detailed* arm that was instructed to self-monitor all foods and drinks consumed each day using the Fitbit mobile app or to the *simplified* arm that was instructed to self-monitor only red zone foods (foods that are highly caloric and of limited nutritional value) each day via a web-based checklist. All participants were instructed to self-monitor both steps and body weight daily. Each week, participants were emailed behavioral lessons, action plans, and personalized feedback. In total, 12 a priori benchmarks were set to establish feasibility, including outcomes related to reach, retention, and self-monitoring engagement (assessed objectively via digital tools). Acceptability was assessed using a questionnaire. Weight change was assessed using scales shipped to the participants' homes and reported descriptively.

Results: The eligibility screen was completed by 248 individuals, of whom 38 (15.3%) were randomized, 18 to *detailed* and 20 to *simplified*. At baseline, participants had a mean age of 47.4 (SD 14.0) years and BMI of 31.2 (SD 4.8) kg/m². More than half (22/38, 58%) were identified as Hispanic of any race. The study retention rate was 92% (35/38) at 3 months. The *detailed* arm met 9 of 12 feasibility benchmarks, while the *simplified* arm met all 12. Self-monitoring engagement was moderate to high (self-monitoring diet: median of 49% of days for *detailed*, 97% for *simplified*; self-monitoring steps: 99% for *detailed*, 100% for *simplified*; self-monitoring weight: 67% for *detailed*, 80% for *simplified*). Participants in both arms reported high satisfaction, with 89% indicating that they would recommend the intervention. Weight change was -3.4 (95% CI -4.6 to -2.2) kg for *detailed* and -3.3 (95% CI -4.4 to -2.2) kg for *simplified*.

Conclusions: A digital weight loss intervention that incorporated either detailed or simplified dietary self-monitoring was feasible, with high retention and engagement, and acceptable to racial and ethnic minority adults.

Trial Registration: ASPREDICTED #66674; <https://aspredicted.org/ka478.pdf>

KEYWORDS

weight loss; obesity; behavioral intervention; self-monitoring; race; ethnicity; digital health; diet tracking; engagement; randomized controlled trial; RCT; mobile phone

Introduction

Background

Obesity is a pervasive health concern in the United States that disproportionately affects many racial and ethnic minority groups [1]. For example, Hispanic adults have an obesity prevalence of 45%, whereas non-Hispanic Black adults have a prevalence of 50% [1]. Excess weight increases the risk of chronic diseases [2] and adverse events from COVID-19 [3]; racial and ethnic minority groups often face a higher risk for these diseases than non-Hispanic White adults or experience greater risk at lower BMI levels [4,5].

Behavioral weight loss interventions are the gold standard for treating obesity [6], with weight loss outcomes of up to 8% by 6 months [7]. These interventions involve diet and exercise goals, frequent counseling, and behavioral strategies. However, few studies have prioritized enrolling racial and ethnic minority groups or creating tailored interventions for these groups [8,9]. The lack of adequate representation of these groups limits generalizability, making it unclear whether the interventions are indeed suitable for populations who might benefit the most from weight loss. Digital health interventions that deliver treatment remotely have the potential to reach broad populations who may otherwise not be able to access behavioral obesity treatment and can minimize barriers to enrollment in these studies. Stand-alone interventions (ie, those without human counseling), in particular, offer greater scalability and can facilitate modest weight loss [10].

A core component of behavioral weight loss interventions is self-monitoring dietary intake [11,12]. The most popular approach, according to recent systematic reviews [13,14], involves detailed self-monitoring of all foods and drinks consumed every day, along with the portion sizes and caloric value of those items. The act of self-monitoring allows individuals to pay attention to their behavior and gain feedback on how specific actions impact their weight. Self-regulation theories, including Social Cognitive Theory, posit that behavior change occurs through the comparison of one's behavior to one's goals or past performance [15,16]. Indeed, self-monitoring dietary intake is among the strongest predictors of behavior change [17] and weight loss [18]. Moreover, using digital tools for self-monitoring promotes engagement, compared with paper-based self-monitoring [13]. The advantages of using digital platforms for self-monitoring include immediate personalized feedback, reduced tracking time with automatic calculations of total caloric intake and nutrients, and high portability of mobile health tools, which increases the likelihood of use while reducing retrospective reporting errors.

However, studies that examined the trajectory of dietary self-monitoring engagement have consistently observed prominent declines over the course of an intervention [19-22],

which suggests that this detailed dietary self-monitoring approach is often perceived as burdensome. The burden of dietary self-monitoring has been defined by Turner-McGrievy et al [23] as time-intensive and active efforts to initiate, use, and record eating events. As such, simplified or abbreviated approaches to self-monitoring dietary intake are needed to counteract this decline in engagement. These simplified approaches could lower burden by making it easier or quicker to self-monitor one's diet, which has the potential to sustain engagement levels for longer periods. They could also increase the proportion of people who feel confident self-monitoring by requiring fewer health literacy and numeracy skills. Examples of specific types of simplified strategies include self-monitoring only less routine, less nutritious, or highly caloric foods; these approaches were recommended by experts in behavioral weight management via a Delphi study [24].

Several weight loss studies have tested simplified dietary self-monitoring approaches [23,25-42]. Five of these studies were randomized controlled trials (RCTs) that empirically compared a simplified to detailed self-monitoring approach [25-29]. In all 5 studies, the detailed arms involved daily self-monitoring of all foods consumed along with their caloric intake, whereas the simplified arms varied in their approach, and involved daily self-monitoring of only highly caloric foods via an investigator-designed app [25], photos of foods consumed via the MealLogger commercial app [26], bites of food consumed via a wearable Bite Counter device [27], dietary lapses via an investigator-designed app [28], or 8 weeks of detailed self-monitoring via a paper diary then transitioning to self-monitoring with checklists of estimates of portion size and fat content rather than tracking calories [29].

No studies have tested a simplified versus detailed diet self-monitoring approach among adults from US racial and ethnic minority groups (either 100% of the sample or analyzing outcomes by race or ethnicity). Of the 5 aforementioned RCTs, the proportion of participants from a racial or ethnic minority group ranged from 19% to 21% or was not reported. This is an important gap to address, given the need to develop effective obesity treatments that can be disseminated to racially and ethnically diverse populations. One challenge to detailed dietary self-monitoring is that nutrition databases used in mobile platforms are sometimes missing food items, including foods specific to a geographic region [43] or different types of ethnic foods [44]. As a result, when users have to create many new food entries, self-monitoring becomes more effortful, which could lead to steeper declines in engagement. Understanding both engagement in and the acceptability of detailed and simplified dietary self-monitoring approaches among racial and ethnic minority groups is therefore vital for creating an intervention that meets the needs of its users.

Objectives

To address this gap, the Spark Pilot Study aimed to evaluate the feasibility and acceptability of a stand-alone, digital behavioral weight loss intervention with either *detailed* or *simplified* dietary self-monitoring in a racially and ethnically diverse sample of US adults. In particular, we examined the reach to our target population, feasibility of our study procedures (eg, study retention) and interventions (eg, engagement in self-monitoring), acceptability of the interventions, and descriptive accounts of our exploratory outcome variables at baseline, 1 month, and 3 months. We conducted a fully remotely delivered trial with all interventions, assessments, and study procedures conducted remotely. This approach allows for greater reach and recruitment speed and may be more acceptable to racial and ethnic minority groups than in-person methods [8,45]. If feasibility and acceptability are established, the next step would be to conduct a fully powered trial to compare the efficacy of a detailed versus simplified dietary self-monitoring strategy among our target population.

Methods

Study Design Overview

The Spark Pilot Study was a 2-arm, parallel group, randomized pilot study of a 3-month digital weight loss intervention. Participants were randomized to either the detailed dietary self-monitoring arm (*detailed*) or the simplified dietary self-monitoring arm (*simplified*). We followed the Consolidated Standards of Reporting Trials (CONSORT) guidelines for pilot studies [46]. Given the feasibility and acceptability aims, it was not considered by the funding agency to be a clinical trial; thus, we did not preregister on ClinicalTrials.gov, and instead, we preregistered on AsPredicted, a platform for preregistering research studies (#66674) [47].

Ethics Approval

All study procedures and human subject research ethics were approved by the Stanford University School of Medicine Institutional Review Board (protocol #59400; approval date: January 27, 2021).

Participant Compensation

The participants provided written informed consent before enrollment in the study. They were compensated a maximum of US \$50 (via e-gift cards) for their completion of assessments, as follows: US \$20 at 1 month, US \$20 at 3 months, and an additional US \$10 for completion of all 4 dietary recalls. The study data are deidentified.

Participants

Inclusion criteria were adults aged ≥ 21 years who self-identified as a member of at least one US racial or ethnic minority group; who had a BMI of 25.0 kg/m² to 45.0 kg/m², which corresponds to having overweight or obesity [2]; who owned a smartphone and had access to a personal email account; who were willing to install the Fitbit mobile app on their phone; who were proficient in the English language; who were living in the United States in the Pacific time zone; and who were interested in losing weight through behavioral strategies. The exclusion criteria

were concurrent enrollment in another weight management intervention; loss of ≥ 4.5 kg (ie, 10 lbs) in the past 6 months; current use of a weight loss medication; prior or planned bariatric surgery; current, recent, or planned pregnancy during the study period; currently breastfeeding; living with someone else participating in the study; inability to engage in moderate forms of physical activity akin to brisk walking (assessed via the Physical Activity Readiness Questionnaire for Everyone [PAR-Q+]; [48]); potential contraindications to losing weight due to a serious medical condition (eg, cancer or dementia) or medication; a history of an eating disorder or cardiovascular event or uncontrolled diabetes mellitus that could predispose an individual to be better suited for a more intensive or different type of intervention; and investigator discretion for safety reasons.

Recruitment

We conducted the study using a fully remote trial design. Enrollment occurred on a rolling basis until our target sample size was met. Participants were recruited from May to June 2021 via free, remote strategies, including ResearchMatch (a web-based US national registry of volunteers interested in participating in research), an institute-specific diabetes registry, and Nextdoor (Nextdoor Holdings, Inc, a web-based neighborhood networking service). Recruitment materials included a brief description of the study and eligibility criteria, along with a link to a web-based screening questionnaire on REDCap (Research Electronic Data Capture; Vanderbilt University), a secure, web-based software platform [49]. Individuals who completed the web-based screen and were eligible and interested were prompted to provide contact information. An automated email was then sent to these individuals to prompt them to sign up for a remote baseline visit via a scheduling website, Calendly, and watch an orientation video. The orientation video was 11 minutes in length and provided an overview of the study objectives, the purpose of randomization, expectations for study participation, and a study timeline. The orientation video was intended to heighten research literacy and trust and increase knowledge of the time commitment and activities involved with study participation to help individuals decide whether the study is a good fit for them and ultimately increase study retention [50,51].

Study Procedures

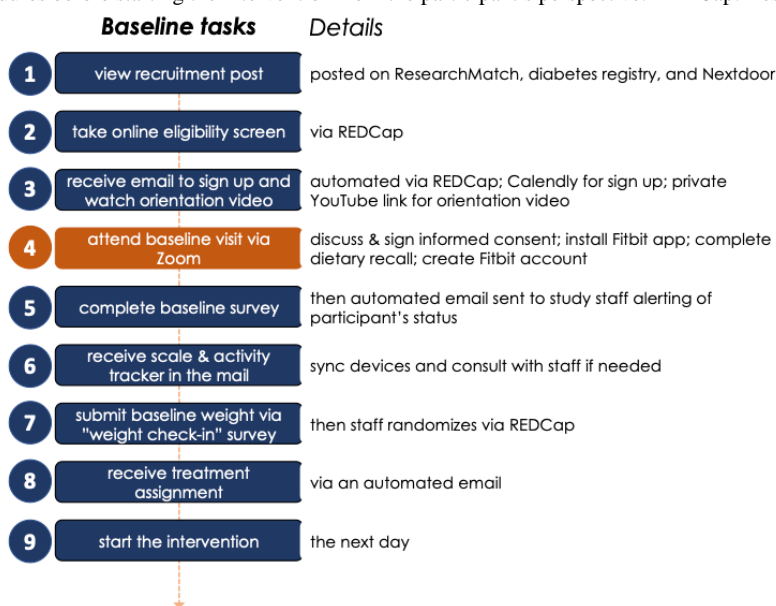
Baseline Tasks

An overview of remote tasks completed before randomization is shown in Figure 1. Briefly, during the 1- to 1.5-hour remote baseline visit held via Zoom videoconference, trained study personnel reviewed the purpose and procedures of the study with participants, assessed the capacity to give consent, and obtained informed consent using REDCap's electronic signature feature. Participants also completed a dietary recall assessment. The study personnel created a unique Fitbit account for each participant, and participants installed the Fitbit app and set it up with the sections, called *tiles*, they were going to use (eg, keep the *steps* tile and hide the *track water* tile). At the end of the baseline visit, the participants received a link to the baseline survey. Once completed, study personnel shipped to the participants' homes a Fitbit activity tracker (Inspire 2) and a

Withings e-scale (Withing Body) from their respective consumer websites. Once both these devices were received, an email was sent to participants that provided information on syncing their devices and prompted them to weigh themselves the following day using a standardized protocol (see the *Data Collection* section). The principal investigator (MLP) then randomized each participant to a treatment arm using REDCap's randomization module with stratification for gender (men, women, and other gender). The allocation sequence was

generated using Microsoft Excel's random number generator and stored in REDCap. To note, during the period between enrollment (the baseline visit) and randomization, participants were asked to maintain their current health behaviors. Once randomized, participants received an automated email with their treatment assignment, intervention details, and goal sheet and were instructed to start the intervention the next day. Study personnel were available for troubleshooting via phone or email to assist in syncing devices.

Figure 1. Remote study procedures before starting the intervention from the participant's perspective. REDCap: Research Electronic Data Capture.



Data Collection

Study assessments were conducted remotely at baseline, 1 month, and 3 months. Survey questions were administered in English via REDCap. Data collection ended on October 8, 2021. At each assessment time point, participants received an automated email instructing them to complete a weight check-in survey and a general survey, both administered via REDCap.

Weight was obtained using a commercially available electronic scale (Withings Body), and participants were instructed to follow standardized procedures [52] at the 3 assessments. Specifically, participants were instructed to weigh themselves on the scale in the morning before eating or drinking and after emptying their bladder; place the scale on a hard surface; remove all articles of clothing and accessories; step on the scale and record the value; and repeat it 2 more times for a total of 3 weight measurements. Collecting weight via commercial scales has demonstrated high concordance with weights measured in a clinic setting [53]. Initially, we intended for participants to have their weight automatically synchronize from the scale to the Fitbit app via Wi-Fi or Bluetooth, but the scale we originally intended to use was removed from the commercial store shortly before the study started, and the institutional review board did not approve of the use of the wireless features of the replacement e-scale by the time recruitment ended. Therefore, in accordance with our protocol, we asked participants to manually input the weight value into the weight check-in survey. Reminders to complete these tasks were sent via SMS text message or email.

No blinding occurred for treatment assignment, assessment, or analysis, owing to resource constraints. Participants were informed before enrollment about the 2 types of treatments to which they may be assigned and were, by design, not blinded to treatment. The study data were collected and managed using REDCap.

Intervention

Overview

All participants received a 12-week behavioral weight loss intervention. This intervention was a stand-alone treatment, meaning that no human counseling was provided. It included the following empirically supported components: a weight loss goal, 3 domains to self-monitor daily (steps [54,55], body weight [56], and diet [14]) via digital tools [13], weekly tailored feedback on self-monitoring behaviors [57], and weekly behavioral lessons and action plans [17,58,59]. These components are supported by Social Cognitive Theory [16] and are intended to promote behavior change through increased self-regulation, self-efficacy, and outcome expectations.

Core Components

Fitbit Mobile App

We used a commercially available app from Fitbit that is free on iPhone and Android platforms. The app was set up with the help of study personnel to reflect the weight loss goals and only the self-monitoring components (tracking steps and weight and tracking diet if in the *detailed* arm) to which each participant

was assigned (in the form of app tiles). In-app graphical feedback allowed participants to view their self-monitoring progress in real time.

Weight Loss Goal

All participants received a goal of 5% weight loss by 3 months. This goal is consistent with obesity treatment guidelines [6] and equates to a weekly weight loss goal of 0.45 kg to 0.91 kg (ie, 1 to 2 lbs) per week, depending on the initial weight.

Self-monitoring Steps and Adaptive Step Goal

The participants were instructed to self-monitor their step count daily using a wrist-worn Fitbit activity tracker (Inspire 2). In conjunction with this self-monitoring goal, a daily step goal was given that adapted weekly based on progress. The initial week's step goal was based on the participant's baseline scores on the Godin Leisure-Time Exercise Questionnaire (GLTEQ) leisure score index [60,61], with scores ranging from 0 to 13 (interpreted as *insufficiently active*) assigned to a goal of 5000 steps per day (n=10), scores of 14 to 23 (*moderately active*) assigned to 7000 steps per day (n=7), and scores ≥ 24 (*active*) assigned to 10,000 steps per day (n=21).

Starting in the second week of the intervention, an adaptive step goal was given based on an algorithm adapted from previous studies [62,63], whereby the 60th percentile of the past week's daily step counts, rounded up to the nearest multiple of 25, was assigned as the subsequent week's daily step goal. For example, a week with daily steps of 5000, 5100, 6000, 6500, 7000, 8200, and 8500 would result in a step goal of 6800 per day in the subsequent week. The Fitbit activity tracker synced with the Fitbit app to allow participants to view their progress toward the step goal.

Self-monitoring Weight

The participants were instructed to self-monitor their body weight daily using a commercially available scale (Withings Body). In a change to our protocol (see *Study Procedures*), we asked participants to manually input the weight value into the Fitbit app each day. In the app, participants could view their progress toward their weight goal.

Tailored Feedback

Each week, participants received an email with tailored, automated feedback pertaining to their progress on their 3 assigned self-monitoring goals (eg, the number of days tracked steps last week), their dietary and step behaviors (eg, "took an average of 9,525 steps per day" last week), and weight change in the past week along with overall weight change since the start of the study. The study team generated feedback by first retrieving data weekly from Fitabase (Small Steps Lab, LLC), a software data management platform. Then, they inputted these data into an Excel spreadsheet that contained all active participants and their goals and semiautomatically generated the progress report using Microsoft Word's Mail Merge feature. The feasibility of this approach was demonstrated in a prior trial by the investigative team [19].

Reminders of Goals

Each week, participants were reminded of their 6 goals (ie, 3 self-monitoring goals along with their personalized weight loss goal, step goal, and either calorie goal or red zone food goal, depending on their treatment assignment). These reminders were attached to the end of each week's progress report, sent via email.

Behavioral Skills Training

Each week, participants received a separate email with theory-informed skills training materials that include structured behavioral lessons on nutrition and physical activity, as well as corresponding action plans. These materials were adapted from gold standard weight loss curricula [64,65]. The lessons included topics such as reading nutrition labels and promoting physical activity (see [Textbox 1](#) for details). Embedded in this email was a link to a brief action plan survey (set up on Qualtrics) that incorporated motivational interviewing [66,67] and problem-solving strategies [68]. Specifically, participants were prompted each week to reflect on their current behaviors and areas for change, generate actionable steps to change related to the week's lesson, identify confidence in doing so, brainstorm potential barriers, and support people. See [Multimedia Appendix 1](#) for a screen recording of an action plan. Reminders were sent via email to participants who had not yet completed that week's action plan 4 days after the initial email.

Textbox 1. Topics of weekly lessons.

- Week 1: Overview of the Spark weight loss program (why self-monitor; losing 5% weight)
- Week 2: Green zone foods
- Week 3: Importance of physical activity
- Week 4: Reading food labels
- Week 5: Reducing sugar
- Week 6: Portion control
- Week 7: Eating out
- Week 8: Preparing meals at home; emotional eating
- Week 9: Social support; environmental cues
- Week 10: Overcoming barriers to physical activity
- Week 11: Weight loss maintenance; slippery slope; relapse prevention

Experimental Component

Participants were randomized to either the *detailed* arm or *simplified* arm, which varied only in their type of dietary self-monitoring.

Detailed Self-monitoring of Diet

Participants randomized to the *detailed* arm were instructed to self-monitor their dietary intake daily via the Fitbit mobile app. This app allows users to track their calories of all foods and beverages consumed using a built-in nutritional database, barcode scanner, or manual entry of individual recipes and to graphically view their change in caloric intake. The participants received a daily calorie goal, with a minimum of 1200 calories (kcal) per day for women and 1500 kcal per day for men, based on national guidelines [2].

Simplified Self-monitoring of Diet

Participants randomized to the *simplified* arm were instructed to self-monitor daily only highly caloric foods consumed that have limited nutritional value, referred to as *red zone foods*. This approach is derived from the Traffic Light Diet [69,70], which posits that limiting this category of foods will reduce intake and variety of highly caloric foods, thereby reducing the total caloric intake. Each morning, the participants in this arm were sent an automated email with a link to a brief web-based checklist (on REDCap). This checklist comprised a list of common red zone foods (ie, foods containing many calories and few nutrients). Participants were asked to select the red zone foods eaten the previous day from the list. Foods were grouped by category (eg, beverages, meat or poultry, and desserts or sweets) for easy navigation, and other red zone foods could be entered if they did not fit any item on the checklist. A goal of eating no more than 3 to 5 red zone foods per day was provided. No calorie goal was provided. Completion of this task was expected to take 2 to 5 minutes per day, compared with an estimated 34 minutes per day for tracking the calories of all foods eaten [71].

The purpose of using a simplified dietary self-monitoring strategy is to reduce the burden of self-monitoring one's dietary intake and reduce health literacy and numeracy barriers, with the hope that this would extend engagement over longer periods while still remaining potent or impactful enough to promote healthy dietary change and subsequent weight loss.

Measures

Baseline Characteristics

In the baseline survey, we collected data on participant sociodemographics, clinical characteristics (eg, smoking status, prediabetes, type 2 diabetes, or hypertension), and type of smartphone. The frequency of self-monitoring in the month before study enrollment was assessed separately for weight, steps, and diet (ie, tracking calories) using a 7-point scale ranging from *several times per day* to *never* [72].

We assessed several psychosocial variables only at the baseline time point; these variables could be examined as potential moderators of intervention effects in a fully powered efficacy trial. In particular, health literacy was assessed via the 6-item Newest Vital Sign (NVS [73]), which was administered orally

by study personnel over Zoom. Scores could range from 0 to 6, with scores of 0 to 3 indicate limited health literacy and scores of 4 to 6 indicate adequate health literacy [74]. All other psychosocial measures were administered through a survey. The occurrence of 16 negative life events (eg, death of a close relative or loss of job) over the past 12 months was assessed, and a composite score was created with a possible range of 0 to 16, with higher scores indicating a greater number of negative life events experienced. Weight bias internalization, which occurs when applying negative stereotypes about one's weight to oneself resulting in self-critical thoughts, was assessed using the modified 10-item Weight Bias Internalization Scale (WBIS [75,76]); possible scores ranged from 0 to 7, with higher scores indicating greater weight bias internalization. Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI [77]); a composite score was created with a possible range from 0 to 21, with higher scores reflecting worse sleep quality. A categorical measure was also created that classified scores >5 as *poor sleep* and those ≤5 as *good sleep*. Social support for eating habits was assessed by asking 10 questions each about support from family and friends [78]; for each, two 5-item subscales related to encouragement and discouragement were created, each with a possible range from 5 to 25, with higher scores indicating more encouragement or discouragement. Similarly, social support for exercise was assessed using 10 questions each about participation from family and friends [78]; for each, a composite score was created with possible scores ranging from 10 to 50, with higher scores indicating greater support.

Feasibility Outcomes

Overview

Our primary outcomes pertained to assessing the feasibility of the study procedures and engagement with the intervention. Before the start of the study, we set benchmarks that, if met, would indicate successful feasibility; these are preregistered on AsPredicted [47]. If benchmarks were not met, we brainstormed the reasons why and possible solutions for modifying the study procedures or intervention components.

Reach

Reach was assessed by examining the enrollment rate (ie, the proportion of participants who were enrolled and randomized out of the individuals who were eligible after completing the web-based screen) as well as by evaluating the number of participants recruited from each recruitment strategy (eg, Nextdoor, ResearchMatch, and the recruitment speed, ie, the number of participants enrolled per week). We also assessed the reasons for ineligibility among those who took the web-based screen.

Retention

The retention rate was operationalized as the percentage of participants who submitted a weight entry via the weight check-in survey at 1 month and 3 months and was compared with our a priori benchmark of 80% retention. We also assessed whether completers differed from noncompleters in terms of baseline sociodemographic, clinical, or psychosocial measures.

Survey and Dietary Recall Completion

The rate of completion of the web-based survey, as well as of the dietary recalls, was assessed and compared with our a priori benchmark of 80% completion at each time point.

Self-monitoring Engagement

We used Fitabase to retrieve real-time, objective, self-monitoring data on steps, body weight, and calories. We examined the median and IQR of the percentage of days of self-monitoring weight, steps, and diet over the course of the 84-day intervention. For the *simplified* arm, the self-monitoring diet was operationalized as submitting the red zone foods checklist on a given day (via a REDCap survey), whereas for the *detailed* arm, it was operationalized as a self-monitoring diet via the Fitbit app, counting only days with ≥ 800 calories recorded, a threshold used in previous app-based studies [19,20,25]. Each self-monitoring engagement metric was compared with our a priori benchmark of self-monitoring 75% of intervention days. Engagement in self-monitoring steps was operationalized as wearing the Fitbit activity tracker and recording at least 1 step on a given day.

In post hoc analyses, we assessed whether any participant self-monitored 100% of the days or 0% of the days. We assessed contamination by examining the number of participants who self-monitored something they were not instructed to self-monitor, which was operationalized as the number of *simplified* arm participants who self-monitored their diet via the Fitbit app (even though they were supposed to self-monitor their diet only via the red zone foods checklist).

Other Intervention Engagement: Action Plans

We objectively assessed action plan completion via Qualtrics and examined the median (IQR) percentage of action plans completed, with 100% indicating completion of all 11 weekly action plans. We also assessed the percentage of participants who received a reminder to complete the action plan each week and the overall reminder success rate (completion rate among those who received the reminder).

Other Intervention Engagement: Lessons

Using self-report in the 3-month survey, we asked participants to indicate which lessons they read and to indicate up to 3 lessons that they found most helpful. We examined the median (IQR) of lessons read.

Other Intervention Engagement: Feedback Emails

Using self-report in the 3-month survey, participants reported the frequency with which they read their progress reports, with 4 response options, including *weekly, less than 1 time per week, less than 1 time per month, and never*.

Other Feasibility Metrics: Timing of Baseline Procedures

To characterize the length of the recruitment and baseline period, we assessed the mean number of days elapsed between the web-based eligibility screen, remote baseline visit, and randomization.

Other Feasibility Metrics: Survey Time and Modality

To provide an estimate of participant burden, we assessed the median number of minutes it took to complete the 3 surveys

and dietary recalls. At each time point, we asked participants to report the type of device (eg, laptop, desktop, tablet, or mobile phone) on which they completed the survey.

Acceptability Outcomes

We assessed the acceptability of the intervention through a series of questions in the 3-month survey, including “Would you recommend the Spark weight loss program to a friend who is trying to lose weight?” with *yes* and *no* response options. We assessed satisfaction with the Fitbit app for self-monitoring foods (shown to the *detailed* arm) or with the web-based checklist for self-monitoring red zone foods (shown to the *simplified* arm), with 6 responses ranging from *extremely dissatisfied* to *extremely satisfied* and an additional response option for *I never tracked my foods*. To understand the comprehensiveness of the Fitbit nutrition database and the list of red zone foods, we asked how likely the Fitbit app (for the *detailed* arm) or the red zone foods checklist (for the *simplified* arm) was to have the foods they typically eat, with 6 response options ranging from *extremely unlikely* to *extremely likely*. We also assessed the helpfulness of the 10 intervention components (listed in the *Results* section), with 5 response options ranging from *not at all helpful* to *extremely helpful*.

Exploratory Outcomes

Weight Change

Weight change from baseline to 1 and 3 months was assessed separately for each treatment arm. We also examined the proportion of participants who achieved clinically significant weight loss of $\geq 3\%$ or $\geq 5\%$ from baseline by 3 months [2,79], as well as the proportion of participants who achieved $\geq 2\%$ weight loss by 1 month, which has been considered an indicator of early success in past research [80-82].

Caloric Intake

We examined changes in caloric intake using the Automated Self-Administered 24-hour (ASA24) Dietary Assessment Tool (version 2020), which is a free web-based tool developed by the National Cancer Institute [83]. We asked participants to complete a total of four 24-hour dietary recalls, including 2 at baseline and 2 at 3 months (for each time point, 1 on a weekday and 1 on a weekend day). Dietary recalls were available in English or Spanish (1 participant completed them in Spanish). We sent up to 4 reminders via email or SMS text message per time point to request the completion of these recalls. We excluded recalls with outliers of daily caloric intake reported as < 600 kcal or > 4400 kcal for women and < 650 kcal or > 5700 kcal for men [83]. To calculate caloric intake at each time point, we calculated the mean of the weekday recall and the weekend-day recall; if only one recall was available at a given time point, we used that value.

Physical Activity

We collected both a self-report and an objective measure of physical activity.

The GLTEQ self-report measure [60,61] assesses the frequency of different types of exercise (strenuous, moderate, and mild or light) for more than 15 minutes during one’s free time during the past week. Strenuous activities were described as those

where one's "heart beats rapidly" (eg, strenuous: running, jogging, or swimming), moderate activities such as "not exhausting" (eg, fast walking or tennis), and mild activities such as "minimal effort" (eg, yoga or easy walking). A leisure score index was then created using the following formula: $(\text{strenuous} \times 9) + (\text{moderate} \times 5) + (\text{light} \times 3)$, with higher scores indicating more frequent exercise. To assess weekly moderate to vigorous physical activity (MVPA), a composite score was created using the same procedures but excluding the light activities; from this MVPA score index, scores of ≥ 24 units were interpreted as *active* and scores < 24 were considered *insufficiently active* [84].

We assessed step count objectively using the Fitbit Inspire 2 activity tracker. We operationalized the baseline step count as the average of the first 7 days of the intervention (week 1) and 3-month step count as the average of the last 7 days (week 12). Fitbit activity tracker devices have shown acceptable to excellent validity for step measurements [85].

Psychosocial Factors

Seven psychosocial measures were assessed at multiple time points via a survey.

Self-efficacy for dietary change was assessed using the 8-item Weight Efficacy Lifestyle Questionnaire Short-Form (WEL-SF [86]); possible scores ranged from 0 to 80, with higher scores indicating greater self-efficacy for making changes to one's eating behavior in a variety of contexts, such as when tired or when in a social situation. Self-efficacy for exercise was assessed using the 12-item Self Efficacy and Exercise Habits Survey [87], possible scores ranged from 1 to 5, with higher scores reflecting greater self-efficacy for exercising in situations such as "after a long, tiring day at work."

Three facets of motivation were assessed via the 15-item Treatment Self-Regulation Questionnaire (TSRQ) [88] with the prompt "The reason I want to achieve a healthier weight is..." The Amotivation subscale assesses the degree of lack of motivation, the Controlled Motivation subscale assesses the extent of feeling guilty or feeling external pressure to achieve this goal, and the Autonomous Motivation subscale assesses internal motivation for doing so. All subscales ranged from 1 to 7, with higher scores indicating greater levels of that construct.

Self-regulation for eating was assessed using the 18-item Three-Factor Eating Questionnaire-R18 (TFEQ-R18) [89]; 3 subscales assessed cognitive restraint (consciously restricting food intake), uncontrolled eating (feeling out of control when eating), and emotional eating (eating in response to negative emotions), with composite scores ranging from 0% to 100%; higher scores were reflective of greater levels of that construct.

Perceived stress was assessed using the 10-item Perceived Stress Scale (PSS-10) [90]; possible scores ranged from 0 to 40, with higher scores indicating greater perceived stress over the last month.

Outcome expectations were assessed at baseline and 1 month via a 17-item measure [91] of the extent to which individuals expected 16 various factors (eg, energy and eating habits) to change over the next 3 months due to participation in the Spark

weight loss program. Participants also selected the 1 benefit that is most important to them. Outcome realizations were assessed at 3 months via a 17-item measure [91] that assessed the degree to which individuals felt that these 16 factors had changed owing to participation in the weight loss program. Possible scores of the composite ranged from 0 to 160, with higher scores indicating higher expectations about (or higher success of) the weight loss program.

Finally, self-efficacy for self-monitoring dietary intake was assessed at only 1 month since it could serve as a mediator of intervention effects. We administered a 14-item measure with the prompt: "a number of situations are described below that can make it hard to [track your food] (*shown to Detailed arm*) OR [track your Red Zone Foods] (*shown to Simplified arm*) on a daily basis" in contexts such as "during weekends" and "during vacations" [92]; possible scores ranged from 0% to 100%, with higher scores indicative of greater self-efficacy in self-monitoring one's dietary intake.

Statistical Analysis

To align with recommendations for conducting pilot studies [93-95], the purpose of our Spark Pilot Study was to assess the feasibility and acceptability of interventions and study procedures. As a result, no power analysis was conducted. The sample size of 38 participants was selected to obtain data on recruitment, randomization, and retention procedures while meeting the budget and timeline constraints. Given the higher costs than those anticipated of a study device and software, we reduced our sample size from 40 to 38 participants.

We used descriptive statistics to assess baseline characteristics, reach, feasibility outcomes, and acceptability outcomes, stratified by treatment arm, and compared them to our a priori feasibility benchmarks, when applicable. We assessed patterns of self-monitoring engagement over time by treatment arm; we report engagement data using median and IQR given their nonnormal distribution. In additional exploratory analyses, we used Spearman rank correlation coefficients to examine the relationship between self-monitoring engagement and weight change over 3 months. To examine whether any baseline variables differed by retention status (completers vs noncompleters), we used Pearson chi-square tests for categorical variables, ANOVA for continuous variables, and Fisher exact tests for small cell counts.

To capture within-arm weight change over time, we used intent-to-treat, linear mixed models via SAS PROC MIXED (SAS Institute) with an unstructured covariance matrix and restricted maximum likelihood estimates while assuming missing data at random; this approach is intended to be used in a future efficacy trial where weight change is the primary outcome. We used chi-square tests to assess the proportion of participants achieving clinically significant weight loss ($\geq 3\%$ and $\geq 5\%$ at 3 months; $\geq 2\%$ at 1 month); we used an intent-to-treat approach that included all participants. Noncompleters (ie, those who were missing a weight value at a relevant assessment time point) were assumed to have not met the clinical threshold. All other secondary outcomes were reported descriptively via change scores over time among completers. As this study was not designed to assess efficacy,

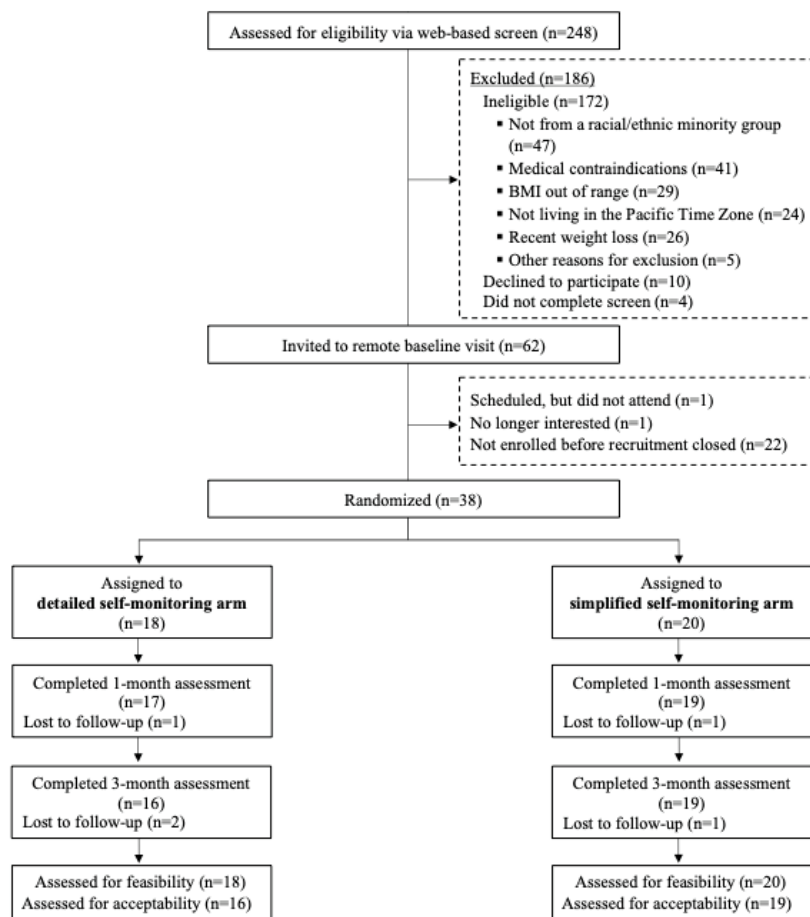
we did not report between-group differences in any outcome, as recommended [93,94]. The analyses were conducted using SAS Studio (SAS Institute). Given the study's small sample size and time constraints, we updated our preregistered protocol and no longer examined potential baseline moderators of intervention effects or conducted qualitative interviews with a subset of participants. The data analysis was conducted in 2022.

Results

Overview

Figure 2, the CONSORT diagram [46], illustrates the flow of participants through the Spark Pilot Study. The web-based eligibility screen was taken by 248 individuals, of whom 25% (n=62) were eligible and invited to the remote baseline session. We enrolled 38 participants, 18 of whom were randomized to the *detailed* arm and 20 randomized to the *simplified* arm. Randomization for the study began on May 31, 2021, and data collection ended on October 8, 2021.

Figure 2. Consolidated Standards of Reporting Trials (CONSORT) flow diagram.



Baseline Characteristics

By design, all the participants self-identified as members of US racial or ethnic minority groups. Specifically, 58% (22/38) of our sample was Hispanic of any race, 32% (12/38) were non-Hispanic Asian or Native Hawaiian or Pacific Islander, 8% (3/38) were non-Hispanic Black, and 2% (1/38) were non-Hispanic American Indian or Alaska Native (see Table 1 for all baseline characteristics).

Participants lived in 4 US states in the western region of the country and were predominantly women (32/38, 84%) with an age of 47.4 (SD 14.0; range 23-78) years and mean BMI of 31.2

(SD 4.8) kg/m². In the month before the start of the study, 68% (26/38) of participants were self-monitoring steps to some degree and 87% (33/38) were self-monitoring weight, whereas only 34% (13/38) were self-monitoring dietary intake. Approximately one-quarter (11/38, 29%) of participants reported having prediabetes or type 2 diabetes and a similar (9/38, 24%) percentage had hypertension. The most important benefit participants marked as hoping to achieve from the intervention was *physical shape and appearance* (17/38, 45%), *confidence and well-being* (6/38, 16%), and *weight* (5/38, 13%). Most (31/38, 82%) participants had adequate health literacy, and almost half (18/38, 47%) reported poor sleep.

Table 1. Baseline demographic, clinical, and psychosocial characteristics of Spark Pilot Study participants.

	Total (N=38)	Detailed self-monitoring (n=18)	Simplified self-monitoring (n=20)
Demographic and clinical variables			
Age (years), mean (SD)	47.4 (14.0)	44.6 (12.5)	49.9 (15.2)
Weight (kg), mean (SD)	81.2 (14.9)	83.6 (16.8)	79.0 (13.1)
BMI (kg/m ²), mean (SD)	31.2 (4.8)	32.1 (4.8)	30.5 (4.8)
BMI category, n (%)			
Overweight (25-29.9 kg/m ²)	18 (47)	8 (44)	10 (50)
Obesity (30-45.0 kg/m ²)	20 (53)	10 (56)	10 (50)
Gender, n (%)			
Man	6 (16)	2 (11)	4 (20)
Woman	32 (84)	16 (89)	16 (80)
Marital status, n (%)			
Married or living with partner	26 (68)	9 (50)	17 (85)
Not married or living with partner	12 (32)	9 (50)	3 (15)
Race and ethnicity, n (%)			
Hispanic (any race)	22 (58)	11 (61)	11 (55)
Non-Hispanic Asian or Native Hawaiian or Pacific Islander	12 (32)	7 (39)	5 (25)
Non-Hispanic Black	3 (8)	0 (0)	3 (15)
Non-Hispanic American Indian or Alaska Native	1 (3)	0 (0)	1 (5)
Education, n (%)			
Less than college graduate	16 (42)	6 (33)	10 (50)
College graduate (4 years) or more	22 (58)	12 (67)	10 (50)
Employment status, n (%)			
Employed	28 (74)	15 (83)	13 (65)
Not employed	10 (26)	3 (17)	7 (35)
Annual household income (US \$), n (%)			
0-49,999	5 (13)	3 (17)	2 (10)
50,000-99,999	13 (34)	6 (33)	7 (35)
100,000 or greater	18 (47)	7 (39)	11 (55)
Unknown	2 (5)	2 (11)	0 (0)
Self-monitoring of diet frequency before enrollment, n (%)			
Daily	0 (0)	0 (0)	0 (0)
1 to 6 times per week	4 (11)	2 (11)	2 (10)
<1 time per week	9 (24)	5 (28)	4 (20)
Never	25 (64)	11 (61)	14 (70)
Self-monitoring of weight frequency before enrollment, n (%)			
Daily	5 (13)	1 (6)	4 (20)
1 to 6 times per week	14 (37)	5 (28)	9 (45)
<1 time per week	14 (37)	8 (44)	6 (30)
Never	5 (13)	4 (22)	1 (5)
Self-monitoring of steps frequency before enrollment, n (%)			

	Total (N=38)	Detailed self-monitoring (n=18)	Simplified self-monitoring (n=20)
Daily	13 (34)	2 (11)	11 (55)
1 to 6 times per week	7 (18)	6 (33)	1 (5)
<1 time per week	6 (16)	2 (11)	4 (20)
Never	12 (32)	8 (44)	4 (20)
Smoking status, n (%)			
Never smoker	33 (87)	15 (83)	18 (90)
Former smoker	3 (8)	1 (6)	2 (10)
Current smoker	2 (5)	2 (11)	0 (0)
Prediabetes or type 2 diabetes, n (%)	11 (29)	7 (39)	4 (20)
Hypertension, n (%)	9 (24)	5 (28)	4 (13)
Type of smartphone, n (%)			
iPhone	25 (66)	13 (72.2)	12 (60)
Android	13 (34)	5 (27.8)	8 (40)
Psychosocial variables			
Limited health literacy (NVS ^a), n (%)	7 (18)	2 (11)	5 (25)
Negative life events in past 4 months, mean (SD)	2.5 (2.6)	2.9 (3.1)	2.1 (2.1)
Weight bias internalization (WBIS ^b), mean (SD)	3.5 (1.6)	4.3 (1.4)	2.7 (1.3)
Sleep quality (PSQI^c), mean (SD)	6.4 (3.7)	6.9 (3.8)	6.1 (3.7)
Poor sleep, n (%)	18 (47)	11 (61)	7 (35)
Social support for eating habits (SSEH^d), mean (SD)			
Encouragement—family	11.2 (4.9)	10.9 (4.5)	11.5 (5.3)
Encouragement—friends	8.3 (4.6)	8.5 (4.8)	8.1 (4.5)
Discouragement—family	11.9 (3.8)	12.4 (4.2)	11.5 (3.5)
Discouragement—friends	8.3 (3.4)	8.2 (3.8)	8.5 (3.2)
Social support for exercise (SSES^e), mean (SD)			
Family participation	25.9 (12.3)	21.5 (10.7)	29.8 (12.6)
Friend participation	17.8 (8.4)	17.4 (7.7)	18.2 (9.2)

^aNVS: Newest Vital Sign.

^bWBIS: Weight Bias Internalization Scale.

^cPSQI: Pittsburgh Sleep Quality Index.

^dSSEH: Social Support for Eating Habits.

^eSSES: Social Support for Exercise Survey.

Feasibility Outcomes

Overview

Table 2 provides an overview of the feasibility findings compared with our a priori benchmarks of success. Benchmarks

were met for all 12 metrics in the *simplified* arm versus 9 of the 12 metrics in the *detailed* arm.

Table 2. Feasibility outcomes compared with a priori benchmarks.

Feasibility metric	Detailed self-monitoring (n=18)	Simplified self-monitoring (n=20)	Benchmark (%)
Intervention Engagement over 3 months, median (IQR)			
Days self-monitoring weight (%)	67 (2-82)	80 (38-100)	75
Days self-monitoring steps (%)	99 (58-100)	100 (99-100)	75
Days self-monitoring diet via red zone foods survey (%)	N/A ^a	97 (86-100)	75
Days self-monitoring diet via Fitbit app (%)	49 (5-67)	N/A	75
Action plans completed (%)	95 (55-100)	100 (91-100)	80
Lessons reviewed ^b (%)	100 (45-100)	100 (95-100)	80
Retention, n (%)			
1 month	17 (94)	19 (95)	80
3 months	16 (89)	19 (95)	80
Survey completion, n (%)			
Baseline	18 (100)	20 (100)	80
1 month	17 (94)	19 (95)	80
3 months	16 (89)	19 (95)	80
Dietary recall completion, n (%)^c			
Baseline	17 (94)	19 (95)	80
3 months	12 (67)	17 (85)	80
Weight change from baseline (kg), mean (SD; 95% CI)			
1 month	-1.85 (1.96; -2.79 to -0.91)	-1.59 (1.94; -2.48 to -0.71)	— ^d
3 months	-3.41 (2.52; -4.62 to -2.20)	-3.29 (2.44; -4.41 to -2.18)	—
Weight change from baseline (%), mean (SD; 95% CI)			
1 month	-2.13 (2.75; -3.45 to -0.81)	-2.07 (2.71; -3.30 to -0.83)	—
3 months	-4.12 (2.80; -5.47 to -2.78)	-4.20 (2.71; -5.43 to -2.96)	—
Clinically significant weight loss from baseline, n (%)^e			
≥2% weight loss at 1 month	9 (50)	9 (45)	—
≥3% weight loss at 3 months	9 (50)	14 (70)	—
≥5% weight loss at 3 months	8 (44)	7 (35)	—

^aN/A: not applicable. N/A because the specific arm was not assigned to self-monitor this item.

^bAnalyses assume that individuals who did not complete the 3-month survey did not review any progress reports or lessons.

^cParticipants were asked to complete 2 ASA24 dietary recalls (1 week day and 1 weekend day) at each time point. If 0 or 1 completed, it is marked as not completed for that time point.

^dIndicates no a priori benchmark was set.

^eAnalyses assumed that individuals who did not submit a weight entry at a given follow-up time point did not achieve clinically meaningful weight loss.

Reach

We achieved an enrollment rate of 61% (38 randomized out of 62 eligible patients after screening). The main sources of recruitment were ResearchMatch (20/38, 53%), an institute-specific diabetes registry (14/38, 37%), and Nextdoor (4/38, 11%).

We achieved an average recruitment speed of 10 participants per week, with all participants being enrolled within 1 month.

The most common reasons for ineligibility were not identifying as a racial or ethnic minority group member (47/172, 27.3%), having a medical contraindication (41/172, 23.8%), having a BMI <25 kg/m² or >45 kg/m² (29/172, 17%), experiencing recent weight loss of 5% (26/172, 15.1%), and not living in the US Pacific time zone (24/172, 13.9%).

Retention

Retention rates surpassed our a priori benchmark of 80% at both 1 month (95% retention) and 3 months (92% retention).

Completers (n=35) at 3 months differed from noncompleters (n=3) on several baseline factors, including reporting fewer negative life events (P=.03) and lower perceived stress (P=.01), amotivation (P=.02), uncontrolled eating (P=.003), and discouragement (and encouragement) from friends in making dietary change (P=.004 and P=.03), as well as their history of self-monitoring dietary intake (P=.02), with 71% (25/35) of completers never tracking diet in the past month compared with 0% of noncompleters (data not shown).

Survey and Dietary Recall Completion

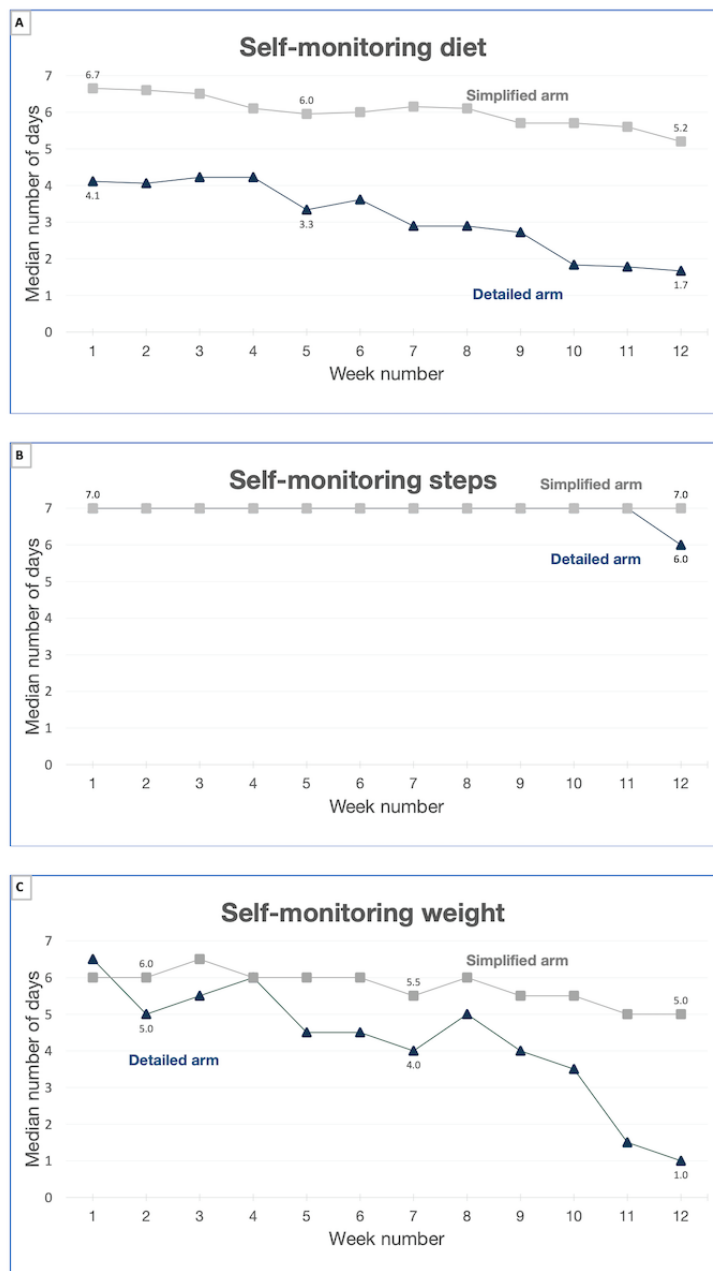
The survey completion rates met our a priori benchmark of 80% at all 3 time points for both the *detailed* and *simplified* arms. Dietary recall (ie, ASA24) completion rates met the a priori

benchmark of 80% at the baseline assessment for both *detailed* and *simplified* arms, but at 3 months, only the *simplified* arm met this benchmark (85% completion vs 67% for *detailed*).

Self-monitoring Engagement

The *detailed* arm did not meet benchmarks for 2 of 3 self-monitoring metrics, with a median percent day of self-monitoring of 49% for diet (ie, all foods eaten via the Fitbit app), 67% for weight, and 99% for steps. The *simplified* arm met benchmarks for all 3 self-monitoring metrics, with a median percent day of self-monitoring of 97% for diet (ie, red zone foods via checklist), 80% for weight, and 100% for steps (see [Figure 3](#) for engagement over time).

Figure 3. Self-monitoring engagement over the 12-week intervention by arm.



High and Low Engagement

Among all 38 participants, several self-monitored every day in the intervention, including 8 (21%) participants who did so for diet tracking, 23 (61%) for step tracking, and 7 (18%) for weight tracking. Of the 8 participants who self-monitored their diet for all days, 7 (87%) were in the *simplified* arm. In addition, some participants did not self-monitor. Specifically, in the *detailed* arm, of the 18 participants, 4 (22%) participants never tracked diet, while 3 (17%) never tracked steps, and 4 (22%) never tracked weight. Three participants did not track any domains. In the *simplified* arm, all (n=20) participants tracked diet and steps to some degree, although 1 (5%) never tracked weight.

Contamination

We found that 2 *simplified* arm participants self-monitored their diet via the Fitbit app (when they were instructed to track only red zone foods), one of whom did so for 9 days and then stopped, and another who did so for 99% of days. Both participants self-monitored the red zone foods for 100% of days.

Other Intervention Engagement

Action Plans

Action plan completion rates met our a priori benchmark of 80%, with a median of 95% completed in *detailed* and 100% completed in *simplified*. Among all participants, 58% (22/38) completed all 11 action plans, while only 1 participant completed none. Reminders were sent each week to participants who had not yet completed their action plans, with 9% (3/34) of participants receiving reminders for action plan 1 to 45% (17/38) receiving reminders for action plan 10. Reminders appeared to be helpful in prompting the completion of action plans, with an overall reminder success rate of 39%.

Lessons

The rates of lessons read met our a priori benchmark of 80%, with a median of 100% in both arms. Many participants (26/38, 68%) reported reading all 11 lessons, while no participants reported reading none (though 3 participants did not answer this question because they did not complete the 3-month survey). The lessons with the most votes for *most helpful* were on the topics of *weight loss maintenance* (34% of participants selected this in their top 3), *reading food labels* (26%), and *portion control* (26%).

Feedback Emails

We did not gather objective data on which feedback emails were read, but we did assess the self-reported frequency of doing so

over the 3-month intervention. Many participants (14/18, 78% of the *detailed* arm; 18/20, 90% of the *simplified* arm) reported reading their progress reports weekly.

Other Feasibility Metrics

Timing of Baseline Procedures

The mean number of days elapsed between the eligibility screen and the remote baseline visit was 7 (SD 5) days. A mean of 16 (SD 4) days elapsed from the baseline visit to randomization (as participants waited for their devices to arrive in the mail).

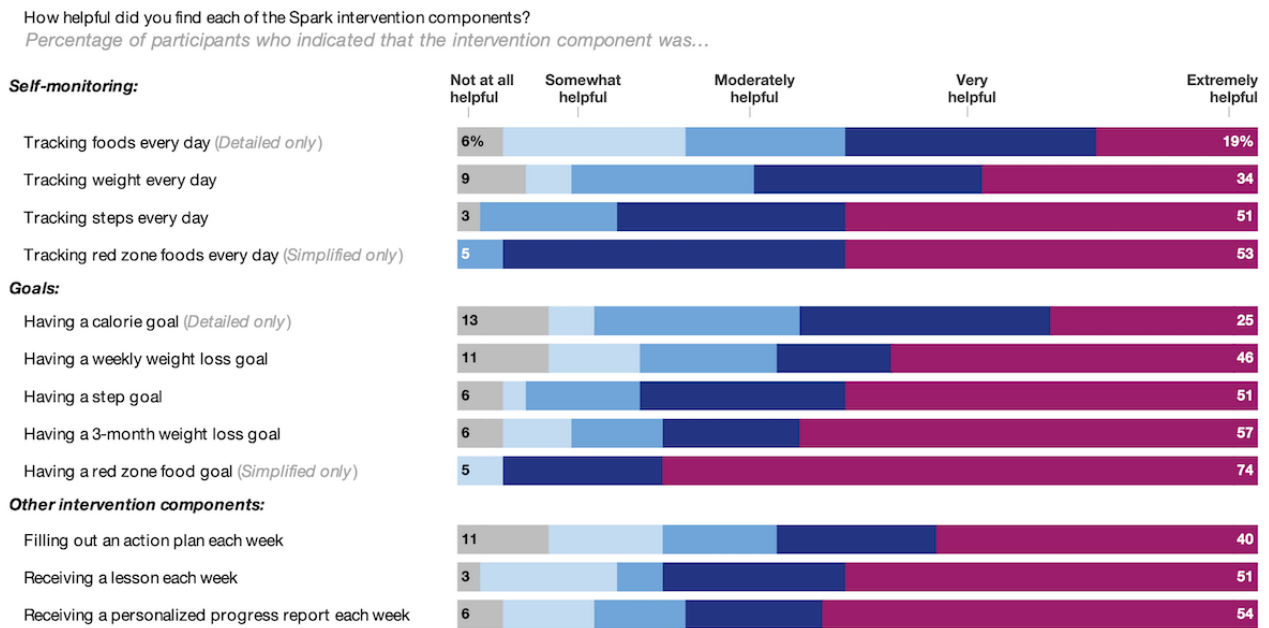
Survey Time and Modality

The median number of minutes to complete the ASA24 dietary recall was 20 minutes: 29 minutes for the baseline survey, 16 minutes for the 1-month survey, and 32 minutes for the 3-month survey. Across all time points, the majority (77/109, 70.6%) of participants reported taking the survey on a laptop or desktop computer, while 24.7% (27/109) reported using a smartphone and 4.5% (5/109) reported using a tablet device. Many (21/36, 58%) participants completed the surveys on different modalities at different time points.

Acceptability Outcomes

Most participants (34/38, 89% overall; 16/18, 88% in the *detailed* arm and 18/20, 90% in the *simplified* arm) indicated that they would recommend the Spark Pilot weight loss program to a friend who is trying to lose weight. Among survey completers (n=35), roughly half (9/16, 56%) of the participants in the *detailed* arm indicated being *somewhat* to *extremely* satisfied with the Fitbit app for tracking foods, compared with almost all (18/19, 95%) participants in the *simplified* arm who used the web-based checklist for tracking red zone foods. When asked about how likely the assigned self-monitoring platform was to have the foods participants typically eat, 38% (6/16) of participants in the *detailed* arm indicated *moderately* or *somewhat* unlikely when reflecting on the Fitbit app, compared with 11% (1/19) of participants in the *simplified* arm who reflected on the red zone foods checklist. Each of the intervention components was rated by most participants as *very* or *extremely* helpful, with “tracking red zone foods every day” rated as such by 95% of the participants in the *simplified* arm, whereas “tracking foods every day” was rated as such by 56% of the participants in the *detailed* arm (see [Figure 4](#) for all ratings).

Figure 4. Helpfulness ratings for each intervention component. This chart depicts the participants’ satisfaction with the 10 intervention components. The proportion of participants who selected “not at all helpful” or “extremely helpful” is shown.



Exploratory Outcomes

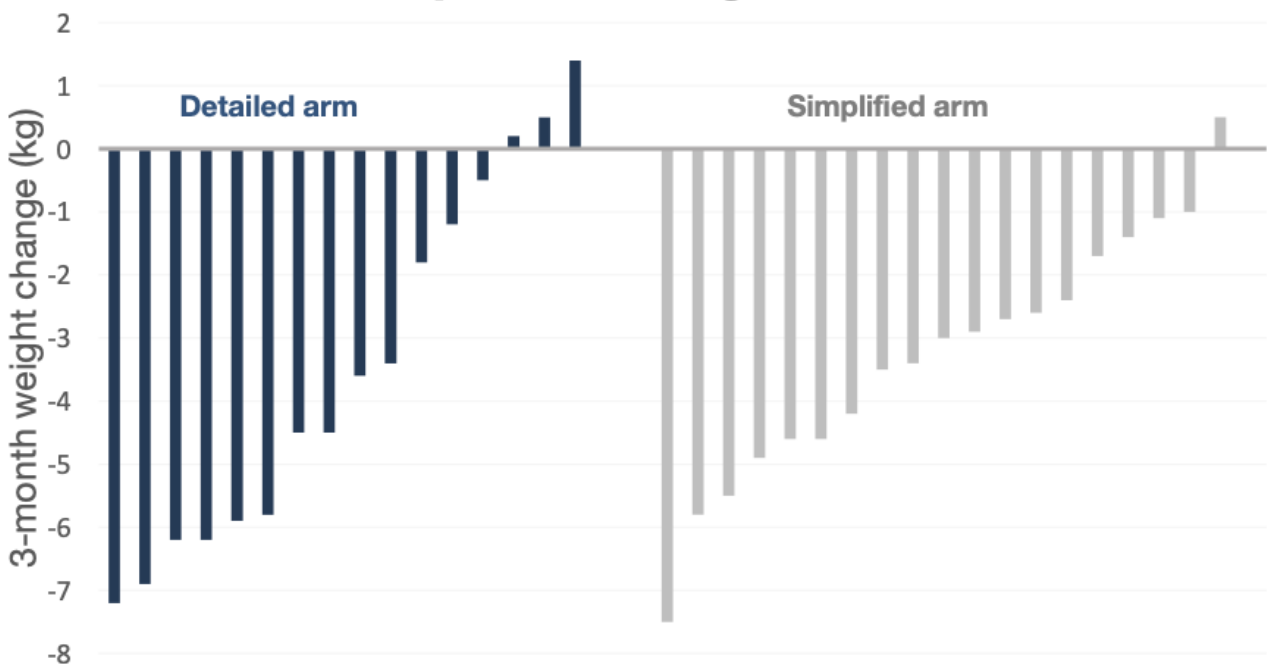
Weight Change

At 1 month, mean weight change from baseline was -1.85 (95% CI -2.79 to -0.91) kg in the *detailed* arm and -1.59 (95% CI -2.48 to -0.71) kg in the *simplified* arm. At 3 months, weight change from baseline was -3.41 (95% CI -4.62 to -2.20) kg

in the *detailed* arm and -3.29 (95% CI -4.41 to -2.18) kg in the *simplified* arm. Over one-third of the participants achieved $\geq 5\%$ weight loss at 3 months (8/18, 44% in the *detailed* arm; 7/20, 35% in the *simplified* arm). The percent weight change outcomes are presented in Table 2. The individual participant weight change is presented via waterfall plots in Figure 5. See Multimedia Appendix 2 for the rest of the exploratory outcomes.

Figure 5. Waterfall plots showing individual participant weight change in kilograms from baseline to 3 months in the detailed arm (left) and in the simplified arm (right). Each bar represents an individual participant. Three participants did not complete the 3-month assessment and are not represented in the figure.

Waterfall plots of weight measurement



Relation Between Self-monitoring and Weight Change

In the *detailed* arm, there were no significant associations between 3-month percent weight change and engagement in any of the self-monitoring metrics (self-monitoring steps: $r_s=0.08$, $P=.78$; self-monitoring weight: $r_s=-0.11$, $P=.69$; self-monitoring diet: $r_s=0.18$, $P=.52$). In the *simplified* arm, the 3-month percent weight change was associated with self-monitoring weight ($r_s=-0.46$, $P=.048$), but not with the other self-monitoring metrics (self-monitoring steps: $r_s=-0.42$, $P=.07$; self-monitoring diet: $r_s=0.06$, $P=.80$).

Caloric Intake

Descriptively, both the *detailed* and *simplified* arms saw reductions in caloric intake over time (-403 kcal/day and -364 kcal/day, respectively).

Physical Activity

In week 1, the mean step count was high (8257 steps), with a range of 2651 to 15,264 steps. Descriptively, physical activity increased in the *simplified* arm over 3 months, with a mean 15.3 (SD 12.7) unit increase in the GLTEQ leisure score index, a 24% increase in the number of participants achieving the *active* category of MVPA, and a mean increase of 250 (SD 3162) steps. These improvements were not observed in the *detailed* arm (D GLTEQ leisure score index: mean 0.9, SD 22.9; D in percentage of *active* participants: -29.5% ; D steps: mean -1306 , SD 2444 steps).

In a post hoc exploratory analysis of completers, those assigned to the initial 5000 step goal ($n=10$) had a mean of 6316 steps per day by the end of the intervention, compared with 9384 steps per day among those with the initial 7000 step goal ($n=7$) and 9203 steps per day among those with the initial 10,000 step goal ($n=21$).

Psychosocial Factors

Descriptively, improvements from baseline to 1 month and 3 months were found for both arms in self-efficacy for dietary change, self-efficacy for exercise, and all self-regulation subscales (Multimedia Appendix 2). Motivation scores on all 3 subscales showed little change over time. Perceived stress levels remained high throughout the intervention, whereas outcome realization scores at 3 months were lower than outcome expectations assessed at baseline and 1 month.

Discussion

Principal Findings and Lessons Learned

In this pilot study, we established the feasibility and acceptability of a 3-month stand-alone digital weight loss intervention in a racially or ethnically diverse sample of adults that involved either detailed or simplified dietary self-monitoring. We observed high retention, moderate to high engagement in self-monitoring, high satisfaction with the weight

loss program, and clinically meaningful short-term weight loss. Importantly, we were able to reach our target population of racial and ethnic minority adults, who are underrepresented in clinical trials of behavioral weight management interventions despite having disproportionate obesity rates. We also found that using a fully remote trial design in which recruitment, intervention, and assessment procedures were all performed remotely enabled greater geographic diversity and more rapid recruitment than is typically possible in research studies involving in-person procedures. To further achieve these recruitment and retention targets, we offered an array of times, including evenings and weekend days, to meet with participants via Zoom for the baseline visit, which was the only face-to-face interaction between the study personnel and participants. All other assessments and intervention procedures were performed via email or text message communication.

The feasibility was assessed based on whether a priori benchmarks were met. Specifically, the *simplified* arm successfully met all 12 a priori feasibility benchmarks, whereas the *detailed* arm met 9 of 12 benchmarks. Domains that were not met for the latter arm include engagement rates for self-monitoring weight and diet (67% and 49% of days, respectively, vs the 75% of days benchmark) and the completion rate of the 3-month dietary recall (67% vs 80% benchmark). To guide further modifications of our intervention and provide suggestions for other studies, we have compiled challenges and proposed solutions in Table 3.

Although we observed higher engagement in dietary self-monitoring in the *simplified* arm than in the *detailed* arm (97% vs 49% of days, respectively), as well as higher satisfaction and helpfulness ratings, the magnitudes of weight loss were similar in both arms (3.3 kg in the *simplified* arm and 3.4 kg in the *detailed* arm). A larger, longer-term intervention study would provide further information on whether such differences would at some point reflect differential weight loss between the 2 arms. When designing digital strategies to promote engagement, both the perceived ease of use and usefulness of such approaches are needed, according to the Technology Acceptance Model [96] and supported by qualitative feedback [97]. It is likely that the simplified dietary self-monitoring approach resulted in high engagement because of meeting both of these needs. Concurrently, the detailed approach may have been harder to use and more time-intensive (eg, 6/16, 38% of participants in the *detailed* arm reported that the app was unlikely to have foods they typically eat, which would have required the manual creation of new foods), which likely resulted in lower engagement. However, whether such distinctions between the 2 approaches result in measurable effects across longer weight loss intervention periods remains unclear. As noted earlier, these findings are meant to be hypothesis generating, as we were not able to detect significant weight loss effects between the arms.

Table 3. Challenges encountered in the Spark Pilot Study and potential solutions.

Challenge	Purpose of addressing it	Potential solutions
Did not meet a priori benchmark		
Detailed arm: engagement in self-monitoring dietary intake via Fitbit app	Suboptimal weight loss can occur with declining engagement rates over time	Use an app with a built-in reminder to track diet or set up automated reminders via a software program
Detailed arm: engagement in self-monitoring weight via e-scale possibly due to switching from manual entry vs automated collection of weight values	Suboptimal weight loss can occur with declining engagement rates over time	Use the smart scale features of a scale to enhance ease of use and avoid manual entry of weight data into an app; request institutional review board approval for multiple scales before the start of recruitment if purchasing devices via a consumer website
Detailed arm: completion of dietary recalls	To gain a more accurate understanding of dietary change	Set clear expectations of time and purpose of the recall; provide greater incentives for completion; drop or switch to a simpler recall
Other issues identified		
Treatment contamination: 2 simplified arm participants self-monitored all foods eaten via Fitbit app (while also self-monitoring red zone foods)	Minimizes ability to detect differences in feasibility, acceptability, and (eventually) efficacy between conditions when participants engage in intervention content intended for a different condition	Clarify upfront participants' understanding of what to do and what not to do as part of their intervention. State the importance of adhering to the intervention content of one's condition and not engaging in activities intended for other conditions. In addition, conduct interviews with participants to understand reasons for engaging in these activities.
Three participants never self-monitored any domain	It is helpful to uncover the reason for no engagement (eg, owing to technical issues, unclear expectations set, life circumstances, or other)	Provide additional support from study personnel for setting up devices within 1 to 2 days of randomization if not yet synchronizing; set up prompts through a variety of channels to check in on whether any technical issues have been encountered; set clear expectations before enrollment of time commitment
Some confusion about when to start the weight loss program	To avoid dissatisfaction and dropout	During the baseline visit, provide an estimate of the average time before starting the weight loss program to receive the devices (in our study, a mean of 16 days elapsed owing to time for personnel to order devices and delivery time); explain that devices might arrive on different days, and they should wait until receiving an email to begin the weight loss program

Comparison With Prior Work

In total, 4 of the 5 prior RCTs comparing detailed versus simplified self-monitoring approaches [25,26,28,29] found similar weight loss between arms at the end of the intervention; the other trial observed greater weight loss in the detailed arm [27]. In all 5 trials, engagement rates in dietary self-monitoring were similar between arms, which is contrary to our finding that engagement rates appeared meaningfully higher in the *simplified* arm than in the *detailed* arm. These differences may be explained by the specific type of simplified dietary self-monitoring approach that we tested (ie, a web-based checklist of red zone foods). It is likely that participants found this strategy particularly easy to use, which is supported by our finding of high ratings of satisfaction and helpfulness. To aid in ease of use, the *simplified* arm had a daily reminder to track their red zone foods by way of receiving an automated email from REDCap each morning to complete their checklist. The use of reminders has been noted as a helpful strategy in promoting self-monitoring engagement [98]. Among prior weight loss studies, the only other interventions with high engagement (>80%) in simplified dietary self-monitoring strategies, to our knowledge, included those that assessed rates on a weekly rather than daily time frame [33,36], although many studies did not report engagement rates. Among all of these

trials, the results should be interpreted with caution because of their small sample sizes or unreported engagement rates.

Engagement in dietary self-monitoring in the *detailed* arm (49% of days) was comparable with engagement rates in other weight loss trials; a 2021 systematic review found that 58% (11/19) of interventions with dietary self-monitoring achieved average engagement rates of $\geq 50\%$ of days [13]. It is possible that incorporating reminders to track diet at the end of the day if not yet done could enhance engagement, as suggested by our previous trial that found a median of 77% (IQR 27%-96%) days of self-monitoring in a 3-month intervention that used a commercial app (MyFitnessPal) that had built-in reminders to track diet [19]. The addition of these reminders should be tested empirically.

Compared with other stand-alone digital weight loss interventions, the magnitude of weight loss found in our study was comparable with many studies [10,25,27,29,36,99], higher than some [26,41,42,100], and lower than a recent fully automated, large-scale pragmatic trial comprising mostly non-Hispanic White participants [101]. This is an important area of research, as many barriers exist to attending in-person treatment, such as living far from a medical clinic or treatment program [102], having childcare or caregiver responsibilities, or having fluctuating working schedules. In addition, not all

individuals who want to lose weight are ready or able to commit to treatment during weekly counseling sessions. Thus, stand-alone interventions have the potential to meet the weight loss needs of many individuals by producing modest weight loss [10], while simultaneously lowering the intensity of participation and offering greater scalability and reach than interventions with counselor support, given widespread accessibility and lower personnel costs. Digital interventions have also produced reductions in attrition in clinical trials [103], likely because they afford participants greater flexibility in engaging in treatment and impose fewer transportation and time constraints. With the widespread uptake of smartphones among US adults, including among racial and ethnic minorities (eg, 79% among Hispanic or Latinx adults and 80% among Black adults [104]), stand-alone digital health interventions have the potential to reach broad populations who may otherwise not be able to access behavioral obesity treatment.

To our knowledge, few past studies on behavioral weight management interventions have been conducted entirely remotely [28,105,106]. Two recent studies switched from an in-person to a fully remote format due to the COVID-19 pandemic and reported how outcomes differed by format: the weight loss maintenance study by Leahey et al [45] observed higher treatment attendance when done remotely versus in-person and excellent retention at study assessments for up to 18 months in both formats. They found that Hispanic participants, in particular, preferred remote sessions. In a 16-week weight loss intervention conducted by Ross et al [107], similar high levels of weight loss and self-monitoring engagement were achieved between a completely remote cohort and a cohort that started in person and transitioned to a remote setting at 11 weeks.

Strengths and Limitations

A strength of the Spark Pilot Study is the focus on racial and ethnic minority adults who are underrepresented in clinical trials of behavioral weight management interventions despite having disproportionate obesity rates. Other strengths include high retention, collection of objective self-monitoring data via digital tools, and use of an RCT design to assess the acceptability and feasibility of the intervention and of study procedures, as well as the successful implementation of a fully remote trial format (spanning from recruitment, onboarding during the baseline visit, mailing devices to participants' homes, randomization, assessments, and intervention delivery) that will be replicated in a future efficacy trial. As the COVID-19 pandemic has prompted other researchers to swiftly transition their in-person research studies to remote platforms, we hope that our study will serve as one example for how to leverage remote technologies and project management tools.

This study has several limitations. First, we were unable to use and test the smart scale features of the Withings scale we provided to participants because of a global chip shortage that made the commercial scale unavailable, followed by institutional review board delays in approving a different commercial scale. Second, given the limited number of personnel, the principal investigator and study staff were not blinded to treatment allocation. However, because the intervention was stand-alone,

the study team did not interact face-to-face with the participants beyond the baseline visit, and automated surveys and templated email messages were used to collect weight and other outcomes. Third, we did not capture metrics for ease of use or perceived burden or time intensity of the 2 dietary self-monitoring approaches, although questions on satisfaction and helpfulness, as well as the objective engagement rates help inform the acceptability of these approaches. Fourth, we recruited only 3 non-Hispanic Black participants, all of whom were randomized to the simplified arm. Further efforts to reach this demographic, who face the highest obesity prevalence, should be considered. Relatedly, only 16% of the participants were men, with twice as many randomized to the simplified arm than the detailed arm. Recruiting a larger sample size and stratifying by gender should help to minimize the potential for gender to be a confounder in a future efficacy trial. Fifth, because of the pilot study's focus on feasibility and acceptability and small sample size, we did not assess outcomes by race and ethnicity status, which would be an important step in a larger efficacy trial to determine if results vary by race or ethnicity. Finally, contamination occurred among 2 of the simplified arm participants in that they also self-monitored caloric intake when instructed not to do so; minimizing contamination will be particularly important in a future efficacy trial because it could impact the magnitude of weight change and lead to an inaccurate interpretation of findings.

Future Research

With the feasibility and efficacy established and minimal refinements needed, a fully powered efficacy trial can now be conducted to test between-group differences in weight loss between the *detailed* and *simplified* self-monitoring arms. This future trial would replicate most of the study procedures used in this pilot study; the main changes would include extending the intervention length, recruiting nationwide in the United States, expanding the sample size to have sufficient power to examine efficacy outcomes, and investigating moderators of treatment response, such as whether certain racial and ethnic groups respond better to a simplified versus detailed self-monitoring approach. None of the prior 5 RCTs that compared detailed to simplified self-monitoring enrolled more than 100 participants (mean 60), and none extended beyond 6 months; thus, trials with larger sample sizes and longer trial durations are needed to more clearly assess efficacy, moderators of treatment response, and maintenance of weight loss. One trial, AGILE, is currently being conducted among 608 young adults testing a detailed versus simplified dietary self-monitoring approach as part of a factorial design in a mobile health intervention (ClinicalTrials.gov: NCT04922216). The findings from this study will advance our understanding of which approach is more suitable for a young adult population.

It is also important to replicate this research question in the context of a more intensive behavioral weight loss treatment that involves frequent counseling [2]. Although self-monitoring engagement rates were found in a systematic review to be similar in stand-alone interventions versus those with counseling [13], it is possible that empirically testing this question in an intensive treatment context would result in a different conclusion. It would also be worthwhile to continue to develop

and test new simplified dietary self-monitoring approaches, such as using artificial intelligence to estimate the caloric intake of foods [108], which may increase accuracy while decreasing burden; comparing different simplified approaches to one another in a clinical trial (eg, self-monitor red zone foods vs dietary lapses vs photos of food) is also needed to determine the strategy with the greatest engagement, highest satisfaction, and largest magnitude of weight loss. Furthermore, using a simplified dietary self-monitoring approach for the duration of an intervention versus using it only if low engagement in detailed dietary self-monitoring occurs is another area to be empirically tested [23,24]. Finally, whether both passive and active forms of self-monitoring impact weight loss similarly should be investigated, given that passive forms of monitoring (eg, wearing an activity tracker) may decrease burden by leveraging automatic data collection from a device, whereas active forms of self-monitoring require the user to volitionally record information somewhere.

Taken together, continuing to enhance the potency of digital weight loss interventions is advantageous to address the obesity epidemic on a large scale. Evaluating different intensities, frequencies, and formats of dietary self-monitoring will help refine these weight loss programs and ultimately maximize clinically meaningful weight loss.

Conclusions

Given that the positive impacts of dietary self-monitoring on weight loss have been repeatedly demonstrated [13,14,109,110], what needs to follow is how to prolong self-monitoring engagement to enhance and sustain weight loss. Simplified dietary self-monitoring strategies are explicitly designed to be easier to use than detailed approaches, but a question remains as to whether they can be potent enough to produce clinically meaningful weight loss. We found that a simplified self-monitoring strategy consisting of tracking only foods that are high in calories and low in nutritional value via a daily web-based checklist resulted in high engagement, high ratings of acceptability, and clinically meaningful weight loss—indicators that both perceived ease of use and usefulness were achieved. These findings were obtained in the context of a fully remote intervention among racial and ethnic minority adults. This pilot study laid the foundation for conducting a long-term, fully powered trial to compare the efficacy of a simplified versus detailed dietary self-monitoring approach in this context. If deemed effective in a subsequent efficacy trial, this lower-intensity stand-alone intervention has the potential to serve as a first-line treatment strategy for this population. In addition, fully remote study procedures could serve as a model for researchers seeking to broaden their reach and access to similar behavioral interventions.

Acknowledgments

This work was supported by the Stanford Clinical and Translational Science Award to Spectrum (UL1TR003142), awarded to MLP and ACK. The Clinical and Translational Science Award program is led by the National Center for Advancing Translational Sciences at the National Institutes of Health (NIH). MLP is currently supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the NIH under award number K23DK129805. ACK was supported in part by a Stanford Catalyst Program grant award. The content is the sole responsibility of the authors and does not necessarily represent the official views of the NIH.

Data Availability

The data set used in this study is available from the corresponding author (MLP) upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screen recording of an example action plan.

[[MOV File , 24754 KB - formative_v6i12e42191_app1.mov](#)]

Multimedia Appendix 2

Exploratory outcomes over time, by treatment arm.

[[PDF File \(Adobe PDF File\), 161 KB - formative_v6i12e42191_app2.pdf](#)]

Multimedia Appendix 3

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1163 KB - formative_v6i12e42191_app3.pdf](#)]

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb(360):1-8 [[FREE Full text](#)] [Medline: [32487284](#)]

2. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation* 2014 Jun 24;129(25 Suppl 2):S102-S138 [FREE Full text] [doi: [10.1161/01.cir.0000437739.71477.ee](https://doi.org/10.1161/01.cir.0000437739.71477.ee)] [Medline: [24222017](https://pubmed.ncbi.nlm.nih.gov/24222017/)]
3. Townsend MJ, Kyle TK, Stanford FC. Outcomes of COVID-19: disparities in obesity and by ethnicity/race. *Int J Obes (Lond)* 2020 Sep;44(9):1807-1809 [FREE Full text] [doi: [10.1038/s41366-020-0635-2](https://doi.org/10.1038/s41366-020-0635-2)] [Medline: [32647359](https://pubmed.ncbi.nlm.nih.gov/32647359/)]
4. Zhu Y, Sidell MA, Arterburn D, Daley MF, Desai J, Fitzpatrick SL, et al. Racial/ethnic disparities in the prevalence of diabetes and prediabetes by BMI: patient outcomes research to advance learning (PORTAL) multisite cohort of adults in the U.S. *Diabetes Care* 2019 Dec;42(12):2211-2219 [FREE Full text] [doi: [10.2337/dc19-0532](https://doi.org/10.2337/dc19-0532)] [Medline: [31537541](https://pubmed.ncbi.nlm.nih.gov/31537541/)]
5. Fei K, Rodriguez-Lopez JS, Ramos M, Islam N, Trinh-Shevrin C, Yi SS, et al. Racial and ethnic subgroup disparities in hypertension prevalence, New York city health and nutrition examination survey, 2013-2014. *Prev Chronic Dis* 2017 Apr 20;14:E33 [FREE Full text] [doi: [10.5888/pcd14.160478](https://doi.org/10.5888/pcd14.160478)] [Medline: [28427484](https://pubmed.ncbi.nlm.nih.gov/28427484/)]
6. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Behavioral weight loss interventions to prevent obesity-related morbidity and mortality in adults: US preventive services task force recommendation statement. *JAMA* 2018 Sep 18;320(11):1163-1171. [doi: [10.1001/jama.2018.13022](https://doi.org/10.1001/jama.2018.13022)] [Medline: [30326502](https://pubmed.ncbi.nlm.nih.gov/30326502/)]
7. Wadden TA, Tronieri JS, Butryn ML. Lifestyle modification approaches for the treatment of obesity in adults. *Am Psychol* 2020;75(2):235-251 [FREE Full text] [doi: [10.1037/amp0000517](https://doi.org/10.1037/amp0000517)] [Medline: [32052997](https://pubmed.ncbi.nlm.nih.gov/32052997/)]
8. Rosenbaum DL, Piers AD, Schumacher LM, Kase CA, Butryn ML. Racial and ethnic minority enrollment in randomized clinical trials of behavioural weight loss utilizing technology: a systematic review. *Obes Rev* 2017 Jul;18(7):808-817. [doi: [10.1111/obr.12545](https://doi.org/10.1111/obr.12545)] [Medline: [28524643](https://pubmed.ncbi.nlm.nih.gov/28524643/)]
9. Haughton CF, Silfee VJ, Wang ML, Lopez-Cepero AC, Estabrook DP, Frisard C, et al. Racial/ethnic representation in lifestyle weight loss intervention studies in the United States: a systematic review. *Prev Med Rep* 2018 Mar;9:131-137 [FREE Full text] [doi: [10.1016/j.pmedr.2018.01.012](https://doi.org/10.1016/j.pmedr.2018.01.012)] [Medline: [29616185](https://pubmed.ncbi.nlm.nih.gov/29616185/)]
10. Tang JC, Abraham C, Greaves CJ, Nikolaou V. Self-directed interventions to promote weight loss: a systematic review and meta-analysis. *Health Psychol Rev* 2016 Sep;10(3):358-372. [doi: [10.1080/17437199.2016.1172979](https://doi.org/10.1080/17437199.2016.1172979)] [Medline: [27091296](https://pubmed.ncbi.nlm.nih.gov/27091296/)]
11. Baker RC, Kirschenbaum DS. Self-monitoring may be necessary for successful weight control. *Behav Ther* 1993 Feb;24(3):377-394. [doi: [10.1016/s0005-7894\(05\)80212-6](https://doi.org/10.1016/s0005-7894(05)80212-6)]
12. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. *J Am Diet Assoc* 2011 Jan;111(1):92-102 [FREE Full text] [doi: [10.1016/j.jada.2010.10.008](https://doi.org/10.1016/j.jada.2010.10.008)] [Medline: [21185970](https://pubmed.ncbi.nlm.nih.gov/21185970/)]
13. Patel ML, Wakayama LN, Bennett GG. Self-monitoring via digital health in weight loss interventions: a systematic review among adults with overweight or obesity. *Obesity (Silver Spring)* 2021 Mar;29(3):478-499. [doi: [10.1002/oby.23088](https://doi.org/10.1002/oby.23088)] [Medline: [33624440](https://pubmed.ncbi.nlm.nih.gov/33624440/)]
14. Raber M, Liao Y, Rara A, Schembre SM, Krause KJ, Strong L, et al. A systematic review of the use of dietary self-monitoring in behavioural weight loss interventions: delivery, intensity and effectiveness. *Public Health Nutr* 2021 Dec;24(17):5885-5913 [FREE Full text] [doi: [10.1017/S136898002100358X](https://doi.org/10.1017/S136898002100358X)] [Medline: [34412727](https://pubmed.ncbi.nlm.nih.gov/34412727/)]
15. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychol Bull* 1982 Jul;92(1):111-135. [Medline: [7134324](https://pubmed.ncbi.nlm.nih.gov/7134324/)]
16. Bandura A. Social cognitive theory of self-regulation. *Organizational Behav Human Decision Process* 1991 Dec;50(2):248-287. [doi: [10.1016/0749-5978\(91\)90022-1](https://doi.org/10.1016/0749-5978(91)90022-1)]
17. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701 [FREE Full text] [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
18. Hartmann-Boyce J, Johns DJ, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. *Obes Rev* 2014 Jul;15(7):598-609 [FREE Full text] [doi: [10.1111/obr.12165](https://doi.org/10.1111/obr.12165)] [Medline: [24636238](https://pubmed.ncbi.nlm.nih.gov/24636238/)]
19. Patel ML, Hopkins CM, Brooks TL, Bennett GG. Comparing self-monitoring strategies for weight loss in a smartphone app: randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Feb 28;7(2):e12209 [FREE Full text] [doi: [10.2196/12209](https://doi.org/10.2196/12209)] [Medline: [30816851](https://pubmed.ncbi.nlm.nih.gov/30816851/)]
20. Turner-McGrievy GM, Dunn CG, Wilcox S, Boutté AK, Hutto B, Hoover A, et al. Defining adherence to mobile dietary self-monitoring and assessing tracking over time: tracking at least two eating occasions per day is best marker of adherence within two different mobile health randomized weight loss interventions. *J Acad Nutr Diet* 2019 Sep;119(9):1516-1524 [FREE Full text] [doi: [10.1016/j.jand.2019.03.012](https://doi.org/10.1016/j.jand.2019.03.012)] [Medline: [31155473](https://pubmed.ncbi.nlm.nih.gov/31155473/)]
21. Thomas JG, Bond DS, Raynor HA, Papandonatos GD, Wing RR. Comparison of smartphone-based behavioral obesity treatment with gold standard group treatment and control: a randomized trial. *Obesity (Silver Spring)* 2019 Apr;27(4):572-580 [FREE Full text] [doi: [10.1002/oby.22410](https://doi.org/10.1002/oby.22410)] [Medline: [30779333](https://pubmed.ncbi.nlm.nih.gov/30779333/)]

22. Butryn ML, Martinelli MK, Crane NT, Godfrey K, Roberts SR, Zhang F, et al. Counselor surveillance of digital self-monitoring data: a pilot randomized controlled trial. *Obesity (Silver Spring)* 2020 Dec;28(12):2339-2346 [FREE Full text] [doi: [10.1002/oby.23015](https://doi.org/10.1002/oby.23015)] [Medline: [33098278](https://pubmed.ncbi.nlm.nih.gov/33098278/)]
23. Turner-McGrievy G, Yang C, Monroe C, Pellegrini C, West D. Is burden always bad? Emerging low-burden approaches to mobile dietary self-monitoring and the role burden plays with engagement. *J Technol Behav Sci* 2021 Apr 09;6(3):447-455 [FREE Full text] [doi: [10.1007/s41347-021-00203-9](https://doi.org/10.1007/s41347-021-00203-9)]
24. Krukowski RA, Harvey J, Borden J, Stansbury ML, West DS. Expert opinions on reducing dietary self-monitoring burden and maintaining efficacy in weight loss programs: a Delphi study. *Obes Sci Pract* 2022 Aug;8(4):401-410 [FREE Full text] [doi: [10.1002/osp4.586](https://doi.org/10.1002/osp4.586)] [Medline: [35949285](https://pubmed.ncbi.nlm.nih.gov/35949285/)]
25. Nezami BT, Hurley L, Power J, Valle CG, Tate DF. A pilot randomized trial of simplified versus standard calorie dietary self-monitoring in a mobile weight loss intervention. *Obesity (Silver Spring)* 2022 Mar;30(3):628-638 [FREE Full text] [doi: [10.1002/oby.23377](https://doi.org/10.1002/oby.23377)] [Medline: [35146942](https://pubmed.ncbi.nlm.nih.gov/35146942/)]
26. Dunn CG, Turner-McGrievy GM, Wilcox S, Hutto B. Dietary self-monitoring through calorie tracking but not through a digital photography app is associated with significant weight loss: the 2smart pilot study—a 6-month randomized trial. *J Acad Nutr Diet* 2019 Sep;19(9):1525-1532. [doi: [10.1016/j.jand.2019.03.013](https://doi.org/10.1016/j.jand.2019.03.013)] [Medline: [31155474](https://pubmed.ncbi.nlm.nih.gov/31155474/)]
27. Turner-McGrievy GM, Wilcox S, Boutté A, Hutto BE, Singletary C, Muth ER, et al. The Dietary Intervention to Enhance Tracking with Mobile devices (DIET Mobile) study: a 6-month randomized weight loss trial. *Obesity (Silver Spring)* 2017 Aug;25(8):1336-1342 [FREE Full text] [doi: [10.1002/oby.21889](https://doi.org/10.1002/oby.21889)] [Medline: [28600833](https://pubmed.ncbi.nlm.nih.gov/28600833/)]
28. Pagoto S, Tulu B, Waring ME, Goetz J, Bibeau J, Divito J, et al. Slip buddy app for weight management: randomized feasibility trial of a dietary lapse tracking app. *JMIR Mhealth Uhealth* 2021 Apr 01;9(4):e24249 [FREE Full text] [doi: [10.2196/24249](https://doi.org/10.2196/24249)] [Medline: [33792547](https://pubmed.ncbi.nlm.nih.gov/33792547/)]
29. Helsel DL, Jakicic JM, Otto AD. Comparison of techniques for self-monitoring eating and exercise behaviors on weight loss in a correspondence-based intervention. *J Am Diet Assoc* 2007 Oct;107(10):1807-1810. [doi: [10.1016/j.jada.2007.07.014](https://doi.org/10.1016/j.jada.2007.07.014)] [Medline: [17904942](https://pubmed.ncbi.nlm.nih.gov/17904942/)]
30. Tate DF, Lutes LD, Bryant M, Truesdale KP, Hatley KE, Griffiths Z, et al. Efficacy of a commercial weight management program compared with a do-it-yourself approach: a randomized clinical trial. *JAMA Netw Open* 2022 Aug 01;5(8):e2226561 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.26561](https://doi.org/10.1001/jamanetworkopen.2022.26561)] [Medline: [35972742](https://pubmed.ncbi.nlm.nih.gov/35972742/)]
31. Tate DF, Quesnel DA, Lutes L, Hatley KE, Nezami BT, Wojtanowski AC, et al. Examination of a partial dietary self-monitoring approach for behavioral weight management. *Obes Sci Pract* 2020 Aug;6(4):353-364 [FREE Full text] [doi: [10.1002/osp4.416](https://doi.org/10.1002/osp4.416)] [Medline: [32874670](https://pubmed.ncbi.nlm.nih.gov/32874670/)]
32. Hales SB, Schulte EM, Turner TF, Malcolm R, Wojtanowski AC, Rethorst C, et al. Pilot evaluation of a personalized commercial program on weight loss, health outcomes, and quality of life. *Transl Behav Med* 2021 Dec 14;11(12):2091-2098. [doi: [10.1093/tbm/ibab110](https://doi.org/10.1093/tbm/ibab110)] [Medline: [34479369](https://pubmed.ncbi.nlm.nih.gov/34479369/)]
33. Bennett GG, Steinberg D, Askew S, Levine E, Foley P, Batch BC, et al. Effectiveness of an app and provider counseling for obesity treatment in primary care. *Am J Prev Med* 2018 Dec;55(6):777-786 [FREE Full text] [doi: [10.1016/j.amepre.2018.07.005](https://doi.org/10.1016/j.amepre.2018.07.005)] [Medline: [30361140](https://pubmed.ncbi.nlm.nih.gov/30361140/)]
34. Damschroder LJ, Lutes LD, Kirsh S, Kim HM, Gillon L, Holleman RG, et al. Small-changes obesity treatment among veterans: 12-month outcomes. *Am J Prev Med* 2014 Nov;47(5):541-553. [doi: [10.1016/j.amepre.2014.06.016](https://doi.org/10.1016/j.amepre.2014.06.016)] [Medline: [25217098](https://pubmed.ncbi.nlm.nih.gov/25217098/)]
35. Ben Neriah D, Geliebter A. Weight loss following use of a smartphone food photo feature: retrospective cohort study. *JMIR Mhealth Uhealth* 2019 May 29;7(6):e11917 [FREE Full text] [doi: [10.2196/11917](https://doi.org/10.2196/11917)] [Medline: [31199300](https://pubmed.ncbi.nlm.nih.gov/31199300/)]
36. Crane MM, Lutes LD, Ward DS, Bowling JM, Tate DF. A randomized trial testing the efficacy of a novel approach to weight loss among men with overweight and obesity. *Obesity (Silver Spring)* 2015 Dec;23(12):2398-2405 [FREE Full text] [doi: [10.1002/oby.21265](https://doi.org/10.1002/oby.21265)] [Medline: [26727117](https://pubmed.ncbi.nlm.nih.gov/26727117/)]
37. Lin P, Wang Y, Levine E, Askew S, Lin S, Chang C, et al. A text messaging-assisted randomized lifestyle weight loss clinical trial among overweight adults in Beijing. *Obesity (Silver Spring)* 2014 May;22(5):E29-E37 [FREE Full text] [doi: [10.1002/oby.20686](https://doi.org/10.1002/oby.20686)] [Medline: [24375969](https://pubmed.ncbi.nlm.nih.gov/24375969/)]
38. Olson R, Wipfli B, Thompson SV, Elliot DL, Anger WK, Bodner T, et al. Weight control intervention for truck drivers: the SHIFT randomized controlled trial, United States. *Am J Public Health* 2016 Sep;106(9):1698-1706. [doi: [10.2105/AJPH.2016.303262](https://doi.org/10.2105/AJPH.2016.303262)] [Medline: [27463067](https://pubmed.ncbi.nlm.nih.gov/27463067/)]
39. Luley C, Blaik A, Götz A, Kicherer F, Kropf S, Isermann B, et al. Weight loss by telemonitoring of nutrition and physical activity in patients with metabolic syndrome for 1 year. *J Am Coll Nutr* 2014;33(5):363-374. [doi: [10.1080/07315724.2013.875437](https://doi.org/10.1080/07315724.2013.875437)] [Medline: [25105874](https://pubmed.ncbi.nlm.nih.gov/25105874/)]
40. Tanaka K, Sasai H, Wakaba K, Murakami S, Ueda M, Yamagata F, et al. Professional dietary coaching within a group chat using a smartphone application for weight loss: a randomized controlled trial. *J Multidiscip Healthc* 2018;11:339-347 [FREE Full text] [doi: [10.2147/JMDH.S165422](https://doi.org/10.2147/JMDH.S165422)] [Medline: [30038502](https://pubmed.ncbi.nlm.nih.gov/30038502/)]
41. Whitelock V, Kersbergen I, Higgs S, Aveyard P, Halford JC, Robinson E. A smartphone based attentive eating intervention for energy intake and weight loss: results from a randomised controlled trial. *BMC Public Health* 2019 May 21;19(1):611 [FREE Full text] [doi: [10.1186/s12889-019-6923-x](https://doi.org/10.1186/s12889-019-6923-x)] [Medline: [31113400](https://pubmed.ncbi.nlm.nih.gov/31113400/)]

42. Beeken RJ, Leurent B, Vickerstaff V, Wilson R, Croker H, Morris S, et al. A brief intervention for weight control based on habit-formation theory delivered through primary care: results from a randomised controlled trial. *Int J Obes (Lond)* 2017 Feb;41(2):246-254 [FREE Full text] [doi: [10.1038/ijo.2016.206](https://doi.org/10.1038/ijo.2016.206)] [Medline: [27867204](https://pubmed.ncbi.nlm.nih.gov/27867204/)]
43. Chen J, Berkman W, Bardouh M, Ng CY, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. *Nutrition* 2019 Jan;57:208-216. [doi: [10.1016/j.nut.2018.05.003](https://doi.org/10.1016/j.nut.2018.05.003)] [Medline: [30184514](https://pubmed.ncbi.nlm.nih.gov/30184514/)]
44. Liefers JRL, Arocha JF, Grindrod K, Hanning RM. Experiences and perceptions of adults accessing publicly available nutrition behavior-change mobile apps for weight management. *J Acad Nutr Diet* 2018 Feb;118(2):229-39.e3. [doi: [10.1016/j.jand.2017.04.015](https://doi.org/10.1016/j.jand.2017.04.015)] [Medline: [28625662](https://pubmed.ncbi.nlm.nih.gov/28625662/)]
45. Leahey T, Pham J, Denmat Z, Jenkins K, Harris - Starling C, Gilder C, et al. Feasibility of online behavioral clinical trials: the future of weight management randomized clinical trials? *Obesity Sci Pract* 2022 May 18 [FREE Full text] [doi: [10.1002/osp4.615](https://doi.org/10.1002/osp4.615)]
46. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016 Oct 24;355:i5239 [FREE Full text] [doi: [10.1136/bmj.i5239](https://doi.org/10.1136/bmj.i5239)] [Medline: [27777223](https://pubmed.ncbi.nlm.nih.gov/27777223/)]
47. Patel M, Rosas LG, King A. Spark pilot study: weight loss intervention for racial/ethnic minorities (#66674). Wharton Credibility Lab. 2022 Jun 23. URL: <https://aspredicted.org/ka478.pdf> [accessed 2022-11-30]
48. Warburton D, Jamnik V, Bredin S, Gledhill N. The physical activity readiness questionnaire for everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X+). *Health Fitness J Canada* 2011;4(2):3-17 [FREE Full text] [doi: [10.14288/hfjc.v4i2.103](https://doi.org/10.14288/hfjc.v4i2.103)]
49. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
50. Jake-Schoffman DE, Brown SD, Baiocchi M, Bibeau JL, Daubenmier J, Ferrara A, et al. Methods-motivational interviewing approach for enhanced retention and attendance. *Am J Prev Med* 2021 Oct;61(4):606-617 [FREE Full text] [doi: [10.1016/j.amepre.2021.04.005](https://doi.org/10.1016/j.amepre.2021.04.005)] [Medline: [34544560](https://pubmed.ncbi.nlm.nih.gov/34544560/)]
51. Goldberg JH, Kiernan M. Innovative techniques to address retention in a behavioral weight-loss trial. *Health Educ Res* 2005 Aug;20(4):439-447. [doi: [10.1093/her/cyg139](https://doi.org/10.1093/her/cyg139)] [Medline: [15598664](https://pubmed.ncbi.nlm.nih.gov/15598664/)]
52. Krukowski R, Ross K. Measuring weight with electronic scales in clinical and research settings during the coronavirus disease 2019 pandemic. *Obesity (Silver Spring)* 2020 Jul;28(7):1182-1183 [FREE Full text] [doi: [10.1002/oby.22851](https://doi.org/10.1002/oby.22851)] [Medline: [32339394](https://pubmed.ncbi.nlm.nih.gov/32339394/)]
53. Ross KM, Wing RR. Concordance of in-home "Smart" scale measurement with body weight measured in-person. *Obes Sci Pract* 2016 Jun;2(2):224-248 [FREE Full text] [doi: [10.1002/osp4.41](https://doi.org/10.1002/osp4.41)] [Medline: [27499884](https://pubmed.ncbi.nlm.nih.gov/27499884/)]
54. Yen H, Chiu H. The effectiveness of wearable technologies as physical activity interventions in weight control: a systematic review and meta-analysis of randomized controlled trials. *Obes Rev* 2019 Oct;20(10):1485-1493. [doi: [10.1111/obr.12909](https://doi.org/10.1111/obr.12909)] [Medline: [31342646](https://pubmed.ncbi.nlm.nih.gov/31342646/)]
55. Larsen R, Wagner V, Korffitsen C, Keller C, Juhl CB, Langberg H, et al. Effectiveness of physical activity monitors in adults: systematic review and meta-analysis. *BMJ* 2022 Jan 26;376:e068047 [FREE Full text] [doi: [10.1136/bmj-2021-068047](https://doi.org/10.1136/bmj-2021-068047)] [Medline: [35082116](https://pubmed.ncbi.nlm.nih.gov/35082116/)]
56. Zheng Y, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: a systematic literature review. *Obesity (Silver Spring)* 2015 Feb;23(2):256-265 [FREE Full text] [doi: [10.1002/oby.20946](https://doi.org/10.1002/oby.20946)] [Medline: [25521523](https://pubmed.ncbi.nlm.nih.gov/25521523/)]
57. Sherrington A, Newham JJ, Bell R, Adamson A, McColl E, Araujo-Soares V. Systematic review and meta-analysis of internet-delivered interventions providing personalized feedback for weight loss in overweight and obese adults. *Obes Rev* 2016 Jun;17(6):541-551 [FREE Full text] [doi: [10.1111/obr.12396](https://doi.org/10.1111/obr.12396)] [Medline: [26948257](https://pubmed.ncbi.nlm.nih.gov/26948257/)]
58. Harkin B, Webb TL, Chang BP, Prestwich A, Conner M, Kellar I, et al. Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychol Bull* 2016 Feb;142(2):198-229. [doi: [10.1037/bul0000025](https://doi.org/10.1037/bul0000025)] [Medline: [26479070](https://pubmed.ncbi.nlm.nih.gov/26479070/)]
59. Frie K, Hartmann-Boyce J, Jebb SA, Aveyard P. Effectiveness of a self-regulation intervention for weight loss: a randomized controlled trial. *Br J Health Psychol* 2020 Sep;25(3):652-676. [doi: [10.1111/bjhp.12436](https://doi.org/10.1111/bjhp.12436)] [Medline: [32489005](https://pubmed.ncbi.nlm.nih.gov/32489005/)]
60. -. Godin leisure-time exercise questionnaire. *Med Sci Sports Exerc* 1997;29(Supplement):36-38. [doi: [10.1097/00005768-199706001-00009](https://doi.org/10.1097/00005768-199706001-00009)]
61. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Can J Appl Sport Sci* 1985 Sep;10(3):141-146. [Medline: [4053261](https://pubmed.ncbi.nlm.nih.gov/4053261/)]
62. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. *PLoS One* 2013;8(12):e82901 [FREE Full text] [doi: [10.1371/journal.pone.0082901](https://doi.org/10.1371/journal.pone.0082901)] [Medline: [24349392](https://pubmed.ncbi.nlm.nih.gov/24349392/)]

63. Adams MA, Todd M, Angadi SS, Hurley JC, Stecher C, Berardi V, et al. Adaptive goals and reinforcement timing to increase physical activity in adults: a factorial randomized trial. *Am J Prev Med* 2022 Feb;62(2):e57-e68 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2021.09.014](https://doi.org/10.1016/j.amepre.2021.09.014)] [Medline: [35000693](#)]
64. Curricula and handouts. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/diabetes/prevention/lifestyle-program/curriculum.html> [accessed 2022-11-11]
65. Look AHEAD Research Group, Pi-Sunyer X, Blackburn G, Brancati FL, Bray GA, Bright R, et al. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. *Diabetes Care* 2007 Jun;30(6):1374-1383 [[FREE Full text](#)] [doi: [10.2337/dc07-0048](https://doi.org/10.2337/dc07-0048)] [Medline: [17363746](#)]
66. Miller WR, Rollnick S. *Motivational Interviewing Preparing People to Change Addictive Behavior*. New York, United States: Guilford press; 1991.
67. Patel ML, Wakayama LN, Bass MB, Breland JY. Motivational interviewing in eHealth and telehealth interventions for weight loss: a systematic review. *Prev Med* 2019 Sep;126:105738. [doi: [10.1016/j.ypmed.2019.05.026](https://doi.org/10.1016/j.ypmed.2019.05.026)] [Medline: [31153917](#)]
68. Perri MG, Nezu AM, McKelvey WF, Shermer RL, Renjilian DA, Viegner BJ. Relapse prevention training and problem-solving therapy in the long-term management of obesity. *J Consult Clin Psychol* 2001 Aug;69(4):722-726. [Medline: [11550740](#)]
69. Epstein LH, Paluch RA, Beecher MD, Roemmich JN. Increasing healthy eating vs. reducing high energy-dense foods to treat pediatric obesity. *Obesity (Silver Spring)* 2008 Feb;16(2):318-326 [[FREE Full text](#)] [doi: [10.1038/oby.2007.61](https://doi.org/10.1038/oby.2007.61)] [Medline: [18239639](#)]
70. Epstein L. A brief history and future of the traffic light diet. *Curr Dev Nutr* 2022 Sep;6(9):nzac120 [[FREE Full text](#)] [doi: [10.1093/cdn/nzac120](https://doi.org/10.1093/cdn/nzac120)] [Medline: [36157846](#)]
71. Harvey J, Krukowski R, Priest J, West D. Log often, lose more: electronic dietary self-monitoring for weight loss. *Obesity (Silver Spring)* 2019 Mar;27(3):380-384 [[FREE Full text](#)] [doi: [10.1002/oby.22382](https://doi.org/10.1002/oby.22382)] [Medline: [30801989](#)]
72. Linde JA, Rothman AJ, Baldwin AS, Jeffery RW. The impact of self-efficacy on behavior change and weight change among overweight participants in a weight loss trial. *Health Psychol* 2006 May;25(3):282-291. [doi: [10.1037/0278-6133.25.3.282](https://doi.org/10.1037/0278-6133.25.3.282)] [Medline: [16719599](#)]
73. The Newest Vital Sign: a health literacy assessment tool. Pfizer. URL: <https://www.pfizer.com/products/medicine-safety/health-literacy/nvs-toolkit> [accessed 2022-11-30]
74. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. *Ann Fam Med* 2005;3(6):514-522 [[FREE Full text](#)] [doi: [10.1370/afm.405](https://doi.org/10.1370/afm.405)] [Medline: [16338915](#)]
75. Durso LE, Latner JD. Understanding self-directed stigma: development of the weight bias internalization scale. *Obesity (Silver Spring)* 2008 Nov;16 Suppl 2:S80-S86 [[FREE Full text](#)] [doi: [10.1038/oby.2008.448](https://doi.org/10.1038/oby.2008.448)] [Medline: [18978768](#)]
76. Lee MS, Dedrick RF. Weight bias internalization scale: psychometric properties using alternative weight status classification approaches. *Body Image* 2016 Jun;17:25-29. [doi: [10.1016/j.bodyim.2016.01.008](https://doi.org/10.1016/j.bodyim.2016.01.008)] [Medline: [26907913](#)]
77. 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. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](#)]
78. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. *Prev Med* 1987 Nov;16(6):825-836. [doi: [10.1016/0091-7435\(87\)90022-3](https://doi.org/10.1016/0091-7435(87)90022-3)] [Medline: [3432232](#)]
79. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, Look AHEAD Research Group. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. *Diabetes Care* 2011 Jul;34(7):1481-1486 [[FREE Full text](#)] [doi: [10.2337/dc10-2415](https://doi.org/10.2337/dc10-2415)] [Medline: [21593294](#)]
80. Patel ML, Hopkins CM, Bennett GG. Early weight loss in a standalone mHealth intervention predicting treatment success. *Obes Sci Pract* 2019 Jun;5(3):231-237 [[FREE Full text](#)] [doi: [10.1002/osp4.329](https://doi.org/10.1002/osp4.329)] [Medline: [31275596](#)]
81. Unick JL, Leahey T, Kent K, Wing RR. Examination of whether early weight loss predicts 1-year weight loss among those enrolled in an internet-based weight loss program. *Int J Obes (Lond)* 2015 Oct;39(10):1558-1560 [[FREE Full text](#)] [doi: [10.1038/ijo.2015.89](https://doi.org/10.1038/ijo.2015.89)] [Medline: [25982792](#)]
82. Unick JL, Ross KM, Wing RR. Factors associated with early non-response within an internet-based behavioural weight loss program. *Obes Sci Pract* 2019 Aug;5(4):324-332 [[FREE Full text](#)] [doi: [10.1002/osp4.341](https://doi.org/10.1002/osp4.341)] [Medline: [31452917](#)]
83. Automated self-administered 24-hour (ASA24®) dietary assessment tool. NIH National Cancer Institute. URL: <https://epi.grants.cancer.gov/asa24/> [accessed 2022-11-30]
84. Amireault S, Godin G. The Godin-Shepherd leisure-time physical activity questionnaire: validity evidence supporting its use for classifying healthy adults into active and insufficiently active categories. *Percept Mot Skills* 2015 Apr;120(2):604-622. [doi: [10.2466/03.27.PMS.120v19x7](https://doi.org/10.2466/03.27.PMS.120v19x7)] [Medline: [25799030](#)]
85. Germini F, Noronha N, Borg Debono V, Abraham Philip B, Pete D, Navarro T, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: systematic review of the literature. *J Med Internet Res* 2022 Jan 21;24(1):e30791 [[FREE Full text](#)] [doi: [10.2196/30791](https://doi.org/10.2196/30791)] [Medline: [35060915](#)]
86. Clark MM, Abrams DB, Niaura RS, Eaton CA, Rossi JS. Self-efficacy in weight management. *J Consult Clin Psychol* 1991 Oct;59(5):739-744. [doi: [10.1037//0022-006x.59.5.739](https://doi.org/10.1037//0022-006x.59.5.739)] [Medline: [1955608](#)]

87. Sallis J, Pinski R, Grossman R, Patterson T, Nader P. The development of self-efficacy scales for healthrelated diet and exercise behaviors. *Health Educ Res* 1988;3(3):283-292 [FREE Full text] [doi: [10.1093/her/3.3.283](https://doi.org/10.1093/her/3.3.283)]
88. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. *Health Educ Res* 2007 Oct;22(5):691-702. [doi: [10.1093/her/cyl148](https://doi.org/10.1093/her/cyl148)] [Medline: [17138613](https://pubmed.ncbi.nlm.nih.gov/17138613/)]
89. Karlsson J, Persson LO, Sjöström L, Sullivan M. Psychometric properties and factor structure of the Three-Factor Eating Questionnaire (TFEQ) in obese men and women. Results from the Swedish Obese Subjects (SOS) study. *Int J Obes Relat Metab Disord* 2000 Dec;24(12):1715-1725. [doi: [10.1038/sj.ijo.0801442](https://doi.org/10.1038/sj.ijo.0801442)] [Medline: [11126230](https://pubmed.ncbi.nlm.nih.gov/11126230/)]
90. Cohen S, Kamarck T, Mermelstein R. *Measuring Stress A Guide for Health and Social Scientists*. Oxfordshire, England, UK: Oxford University Press; 1997.
91. Neff K, King A. Exercise program adherence in older adults: the importance of achieving one's expected benefits. *Med Exerc Nutr Health* 1995;4:355-362.
92. Bandura A. *Guide for constructing self-efficacy scales*. In: *Self-Efficacy Beliefs of Adolescent*. Charlotte, North Carolina, United States: Information Age Publishing; 2005.
93. Freedland KE. Pilot trials in health-related behavioral intervention research: problems, solutions, and recommendations. *Health Psychol* 2020 Oct;39(10):851-862. [doi: [10.1037/hea0000946](https://doi.org/10.1037/hea0000946)] [Medline: [32614197](https://pubmed.ncbi.nlm.nih.gov/32614197/)]
94. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res* 2011 May;45(5):626-629 [FREE Full text] [doi: [10.1016/j.jpsychires.2010.10.008](https://doi.org/10.1016/j.jpsychires.2010.10.008)] [Medline: [21035130](https://pubmed.ncbi.nlm.nih.gov/21035130/)]
95. Pilot studies: common uses and misuses. National Center for Complementary and Integrative Health. URL: <https://www.nccih.nih.gov/grants/pilot-studies-common-uses-and-misuses> [accessed 2022-11-30]
96. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989 Sep;13(3):319-340 [FREE Full text] [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
97. Sharpe EE, Karasouli E, Meyer C. Examining factors of engagement with digital interventions for weight management: rapid review. *JMIR Res Protoc* 2017 Oct 23;6(10):e205 [FREE Full text] [doi: [10.2196/resprot.6059](https://doi.org/10.2196/resprot.6059)] [Medline: [29061557](https://pubmed.ncbi.nlm.nih.gov/29061557/)]
98. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuisen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. *J Med Internet Res* 2017 Aug 01;19(8):e277 [FREE Full text] [doi: [10.2196/jmir.7288](https://doi.org/10.2196/jmir.7288)] [Medline: [28765103](https://pubmed.ncbi.nlm.nih.gov/28765103/)]
99. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. *Obes Rev* 2017 Apr;18(4):450-459. [doi: [10.1111/obr.12492](https://doi.org/10.1111/obr.12492)] [Medline: [28187246](https://pubmed.ncbi.nlm.nih.gov/28187246/)]
100. 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 [FREE Full text] [doi: [10.7326/M13-3005](https://doi.org/10.7326/M13-3005)] [Medline: [25402403](https://pubmed.ncbi.nlm.nih.gov/25402403/)]
101. Thomas JG, Panza E, Espel-Huynh HM, Goldstein CM, O'Leary K, Benedict N, et al. Pragmatic implementation of a fully automated online obesity treatment in primary care. *Obesity (Silver Spring)* 2022 Aug;30(8):1621-1628. [doi: [10.1002/oby.23502](https://doi.org/10.1002/oby.23502)] [Medline: [35894075](https://pubmed.ncbi.nlm.nih.gov/35894075/)]
102. Ariel-Donges AH, Gordon EL, Dixon BN, Eastman AJ, Bauman V, Ross KM, et al. Rural/urban disparities in access to the National Diabetes Prevention Program. *Transl Behav Med* 2020 Dec 31;10(6):1554-1558 [FREE Full text] [doi: [10.1093/tbm/ibz098](https://doi.org/10.1093/tbm/ibz098)] [Medline: [31228199](https://pubmed.ncbi.nlm.nih.gov/31228199/)]
103. Pirotta S, Joham A, Hochberg L, Moran L, Lim S, Hindle A, et al. Strategies to reduce attrition in weight loss interventions: a systematic review and meta-analysis. *Obes Rev* 2019 Oct;20(10):1400-1412. [doi: [10.1111/obr.12914](https://doi.org/10.1111/obr.12914)] [Medline: [31347759](https://pubmed.ncbi.nlm.nih.gov/31347759/)]
104. Mobile Fact Sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2022-11-30]
105. Pagoto SL, Schroeder MW, Xu R, Waring ME, Groshon L, Goetz JM, et al. A Facebook-delivered weight loss intervention using open enrollment: randomized pilot feasibility trial. *JMIR Form Res* 2022 May 06;6(5):e33663 [FREE Full text] [doi: [10.2196/33663](https://doi.org/10.2196/33663)] [Medline: [35522466](https://pubmed.ncbi.nlm.nih.gov/35522466/)]
106. West DS, Krukowski RA, Stansbury ML, Ogden D, Borden J, Harvey JR. Examining weekly facilitated group sessions and counselor-crafted self-monitoring feedback on treatment outcome in digital weight control: a pilot factorial study. *Obes Sci Pract* 2022 Aug;8(4):433-441 [FREE Full text] [doi: [10.1002/osp4.585](https://doi.org/10.1002/osp4.585)] [Medline: [35949286](https://pubmed.ncbi.nlm.nih.gov/35949286/)]
107. Ross KM, Carpenter CA, Arroyo KM, Shankar MN, Yi F, Qiu P, et al. Impact of transition from face-to-face to telehealth on behavioral obesity treatment during the COVID-19 pandemic. *Obesity (Silver Spring)* 2022 Apr;30(4):858-863. [doi: [10.1002/oby.23383](https://doi.org/10.1002/oby.23383)] [Medline: [35037410](https://pubmed.ncbi.nlm.nih.gov/35037410/)]
108. Mortazavi BJ, Gutierrez-Osuna R. A review of digital innovations for diet monitoring and precision nutrition. *J Diabetes Sci Technol* 2021 Sep 01:19322968211041356. [doi: [10.1177/19322968211041356](https://doi.org/10.1177/19322968211041356)] [Medline: [34467803](https://pubmed.ncbi.nlm.nih.gov/34467803/)]
109. Berry R, Kassavou A, Sutton S. Does self-monitoring diet and physical activity behaviors using digital technology support adults with obesity or overweight to lose weight? A systematic literature review with meta-analysis. *Obes Rev* 2021 Oct;22(10):e13306. [doi: [10.1111/obr.13306](https://doi.org/10.1111/obr.13306)] [Medline: [34192411](https://pubmed.ncbi.nlm.nih.gov/34192411/)]
110. Cavero-Redondo I, Martinez-Vizcaino V, Fernandez-Rodriguez R, Saz-Lara A, Pascual-Morena C, Álvarez-Bueno C. Effect of behavioral weight management interventions using lifestyle mHealth self-monitoring on weight loss: a systematic

review and meta-analysis. *Nutrients* 2020 Jul 03;12(7):1977 [FREE Full text] [doi: [10.3390/nu12071977](https://doi.org/10.3390/nu12071977)] [Medline: [32635174](https://pubmed.ncbi.nlm.nih.gov/32635174/)]

Abbreviations

ASA24: Automated Self-Administered 24-Hour
CONSORT: Consolidated Standards of Reporting Trials
GLTEQ: Godin Leisure-Time Exercise Questionnaire
MVPA: moderate to vigorous physical activity
NVS: Newest Vital Sign
PAR-Q+: Physical Activity Readiness Questionnaire for Everyone
PSQI: Pittsburgh Sleep Quality Index
PSS-10: 10-item Perceived Stress Scale
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
TFEQ-R18: 18-item Three-Factor Eating Questionnaire-R18
TSRQ: Treatment Self-Regulation Questionnaire
WBIS: Weight Bias Internalization Scale
WEL-SF: Weight Efficacy Lifestyle Questionnaire Short-Form

Edited by A Mavragani; submitted 25.08.22; peer-reviewed by G Aljuraiban, R Krukowski, R Bannor; comments to author 25.10.22; revised version received 29.10.22; accepted 04.11.22; published 13.12.22.

Please cite as:

Patel ML, Cleare AE, Smith CM, Rosas LG, King AC

Detailed Versus Simplified Dietary Self-monitoring in a Digital Weight Loss Intervention Among Racial and Ethnic Minority Adults: Fully Remote, Randomized Pilot Study

JMIR Form Res 2022;6(12):e42191

URL: <https://formative.jmir.org/2022/12/e42191>

doi: [10.2196/42191](https://doi.org/10.2196/42191)

PMID: [36512404](https://pubmed.ncbi.nlm.nih.gov/36512404/)

©Michele L Patel, Angel E Cleare, Carly M Smith, Lisa Goldman Rosas, Abby C King. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Understanding the Role of Support in Digital Mental Health Programs With Older Adults: Users' Perspective and Mixed Methods Study

Judith Borghouts¹, PhD; Elizabeth V Eikey^{2,3}, PhD; Cinthia De Leon¹, MPH; Stephen M Schueller^{4,5}, PhD; Margaret Schneider⁶, PhD; Nicole A Stadnick^{7,8,9}, MPH, PhD; Kai Zheng⁵, PhD; Lorraine Wilson¹⁰, MSW; Damaris Caro¹⁰, MA; Dana B Mukamel¹, PhD; Dara H Sorkin¹, PhD

¹Department of Medicine, University of California, Irvine, Irvine, CA, United States

²Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, United States

³The Design Lab, University of California, San Diego, La Jolla, CA, United States

⁴Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

⁵Department of Informatics, University of California, Irvine, Irvine, CA, United States

⁶Department of Public Health, University of California, Irvine, Irvine, CA, United States

⁷Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States

⁸Dissemination and Implementation Science Center, Altman Clinical and Translational Research Institute, University of California, San Diego, La Jolla, CA, United States

⁹Child and Adolescent Services Research Center, San Diego, CA, United States

¹⁰Department of Health and Human Services, County of Marin, San Rafael, CA, United States

Corresponding Author:

Elizabeth V Eikey, PhD

Herbert Wertheim School of Public Health and Human Longevity Science

University of California, San Diego

9500 Gilman Dr

La Jolla, CA, 92093

United States

Phone: 1 949 438 1337

Email: eeikey@health.ucsd.edu

Abstract

Background: Digital mental health interventions have the potential to increase mental health support among isolated older adults. However, the older adult population can experience several barriers to accessing and using digital health resources and may need extra support to experience its benefits.

Objective: This paper aimed to understand what older adults experience as an important aspect of support during engagement in a digital mental health program. The program entailed 3 months of staff support to participate in digital literacy training and engage with the digital mental health platform myStrength, which offers support for a range of mental health challenges, including depression and anxiety.

Methods: A total of 30 older adults participated in surveys and interviews to assess their experience of participating in a digital mental health program provided by county mental health services. As part of the program, participants attended 4 classes of digital literacy training, had access to the digital mental health platform myStrength for 2 months with staff support (and 10 months after the program without support), and received support from program staff during the entire 3-month program. Survey data were analyzed using descriptive statistics, and interview data were analyzed using thematic analysis.

Results: A thematic analysis of the interview data revealed that participants valued ongoing support in 3 main areas: technical support to assist them in using technology, guided support to remind them to use myStrength and practice skills they had learned, and social support to enable them to connect with others through the program. Furthermore, participants reported that social connections was the most important aspect of the program and that they were mainly motivated to participate in the program because it was recommended to them by trusted others such as a community partner or because they believed it could potentially help others.

Conclusions: Our findings can be used to inform the design of future digital mental health programs for older adults who may have unique support needs in terms of dedicated technical support and ongoing guided support to use technology and social support to increase social connectedness.

(*JMIR Form Res* 2022;6(12):e43192) doi:[10.2196/43192](https://doi.org/10.2196/43192)

KEYWORDS

older adults; mental health; digital mental health intervention; human support

Introduction

Background

Mental health is an increasing concern among the older adult population. Even before the global COVID-19 pandemic, social isolation among older adults (ie, those older than 60 years) was a significant public health issue [1], and depression was estimated to be a more common health condition among older adults than dementia [2]. In addition to making it more difficult to access in-person care, the context of COVID-19 has raised concerns such as anxiety and loneliness [3]. According to a 2020 survey conducted by the American Association of Retired Persons Foundation and the United Health Foundation, the pandemic has caused an epidemic of loneliness and social isolation among the older adult population, with two-thirds of respondents saying that they experienced social isolation and 66% saying that their anxiety increased during the pandemic [4].

Potential of Digital Mental Health Interventions

Digital mental health interventions (DMHIs) have the potential to remotely increase access to mental health support. Mental health support refers to any type of support aimed at protecting or promoting psychosocial well-being [5], such as therapy, education, and health promotion programs. Recent reviews have demonstrated the effectiveness of digital health interventions in reducing social isolation and depression in older adults [6-9]. However, older adults may experience several barriers to accessing and using digital health resources, such as lack of knowledge and awareness of digital health resources, little exposure to and experience with technology in general, and lack of confidence in using a computer [10]. Rather than increasing access to care, these barriers to digital health access may instead result in larger health care inequalities [11-13].

To better understand older adults' perspectives toward the concept of mental health technologies, the study by Andrews et al [14] held a discussion group with older adults and enabled them to interact with various mental health apps and websites during a session to increase awareness. The authors found that there were overall positive attitudes toward mental health technologies. However, although study participants were aware of mental health websites, awareness of these websites did not motivate their actual use [14].

Implementation of DMHIs for the Older Adult Population

Exposure and awareness alone appear to be insufficient for older adults to engage with DMHIs. A recent report by the Substance Abuse Mental Health Services Administration highlighted the

need for support and digital training when offering DMHIs to older adults [15]. It is well known that guided DMHIs (eg, in the form of an internet-based or remote therapist or coach) generally have higher engagement than self-guided interventions [16], and this may be especially true for the older adult population. For example, Egede [17] found that web-based cognitive behavioral therapy was effective in improving depression and anxiety symptoms among older adults, but this was applicable only when a therapist was present to provide guidance. If an individual coach is not available, peer support can also be an effective and cost-effective type of support for web-based cognitive behavioral therapy interventions for older adults [18].

Research Question

Prior recommendations to provide support have been mainly evaluated from an operational (eg, staffing considerations) perspective, with the aim of facilitating successful implementation of an intervention and achieving desired outcomes, rather than from a user's perspective. For example, several recent studies have focused on the effectiveness of DMHIs with older adult participants and have provided support to older adults in the form of therapists, health professionals, research assistants, and coaches [3,17,19]. Support improves the effectiveness of DMHIs, and positive and concrete support messages have been associated with better clinical mental health outcomes [20]. However, few studies have assessed how and whether different types of support affect older adults' user experiences. There are different types of possible support [21,22] that may be more or less suitable depending on the type of barriers they are supposed to address. For example, although text-based reminders may be suitable for people who tend to forget to use an intervention [22], they may not be sufficient to address a lack of motivation, social isolation, or low digital literacy (ie, a person's ability to find, evaluate, and communicate information through digital platforms).

It is important to consider not only what types of support older adults need but also what they value, as participants' experiences affect real-world use and can be crucial in whether they would continue to engage with a digital intervention beyond a research setting.

This paper reports on a study situated within a digital mental health program with older adults. The aim of this study was to evaluate older adults' experiences with the program and specify what they experienced as important aspects of support while they were taking part in the program. The research question we aimed to address was as follows: What do older adults experience as important factors of support during engagement in a digital mental health program?

Methods

Overview

Thirty older adults participated in a digital mental health program for 3 months. The aim of the program was to engage English-speaking and Spanish-speaking isolated older adults with technology and to improve their well-being and sense of social connectedness. Participants experienced several types of isolation. First, everyone experienced some degree of social isolation as the program took place during the global COVID-19 pandemic. Second, older adults were selected based on geographic isolation, and finally, Spanish-speaking participants were also selected based on perceived isolation by program staff due to language, culture, and generational barriers.

Participants received digital literacy training as part of the program. Participants were given free access to the digital mental health platform myStrength for 1 year, and those who did not have the required resources to use myStrength were provided with tablets and an internet connection.

myStrength Description

Participants were given 1-year access to myStrength, a digital and mobile-based mental health platform [23]. myStrength was chosen as the technology based on feedback from an exploratory study of multiple mental health apps with older adults in Marin County. The platform offers evidence-based resources to provide support for a range of mental health challenges, including depression, anxiety, stress, chronic pain, drug or alcohol recovery, and insomnia. It offers tools such as dialectical and cognitive behavioral therapy courses, acceptance and commitment therapy materials, mood and sleep tracking, short videos teaching mindfulness, community forums, motivational quotes, spiritual resources, activities ranging from 2 to 10 minutes, and a library of more than 1600 mental health resources. Participants can select user interests when signing up and receive personalized recommendations aligned with these interests. The participants could access any of the resources as many times as they wanted.

During the program, participants received weekly phone calls from the program staff to check in and to encourage them to use the platform. In addition, participants received at least one in-person visit from nurses to help them access and use technology. myStrength is offered in both English and Spanish. The Spanish version offered some content such as videos in Spanish, but the platform was not fully translated into Spanish (eg, the menu to navigate to content remaining in English).

Study Design and Procedure

To address our research question, we conducted a mixed method study, in which participants took part in 3 surveys and an interview during the program, and 1 follow-up survey 8 months after the program (the survey instruments and interview protocol are included in [Multimedia Appendix 1](#)). The purpose of the surveys was to assess participants' satisfaction with the digital

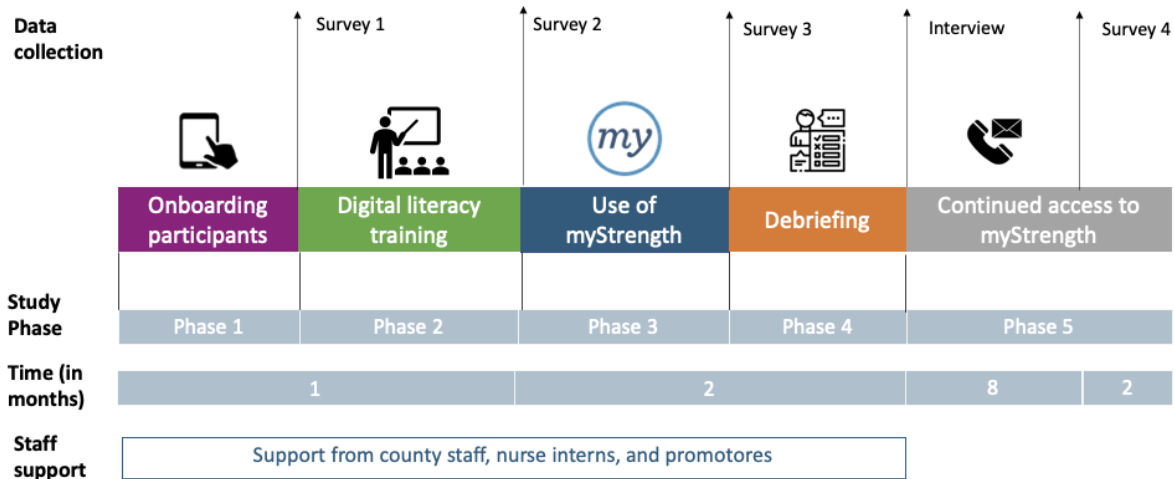
literacy training, myStrength, and the program overall and to measure any improvement in outcomes throughout the program. The purpose of the interviews was to gain a more in-depth understanding of participants' experiences with the program and the support provided, as well as any barriers they experienced in participating.

Older adults were invited to participate in a digital mental health program initiated by Marin County Behavioral Health and Recovery Services. Marin County is located in Northern California; its largest ethnic groups are White (non-Hispanic) and Hispanic, with a population of approximately 259,000 people, of which approximately 22% are aged ≥ 65 years [24]. The program was part of the Help@Hand project, a state-wide evaluation project in California that aims to understand how digital therapeutics fit within the public mental health care system. The specific aim of Marin County's program was to engage isolated older adults (both English- and Spanish-speaking older adults) with technology and to enhance their well-being and sense of social connectedness. Program participants were offered a series of 4 digital literacy classes to enhance their technology skills and were then given free access to the digital mental health platform, myStrength. The program took place from March to June 2021 and consisted of five phases:

1. Phase 1: Onboarding participants by giving them access to required resources, such as a device and an internet connection, to be able to participate in the program (4 weeks)
2. Phase 2: Providing digital literacy training consisting of four 2-hour web-based classes to establish or improve participants' digital literacy skills (4 weeks)
3. Phase 3: Providing participants free access to the myStrength platform and nurse promotores support (8 weeks)
4. Phase 4: Debriefing participants at the end of the program and conducting surveys and interviews (8 weeks)
5. Phase 5: Providing continued free access to myStrength after completion of the program (10 months)

When referring to the program, we refer to phases 1 to 4 as the 3-month period during which staff support was provided to the participants.

There were 5 key data collection points ([Figure 1](#)): 1 survey at the beginning of the program (phase 1), 1 survey after the digital literacy training and before myStrength was made available (between phases 2 and 3), 1 survey and 1 semistructured interview at the end of older adults' participation in the program (phase 4), and 1 follow-up survey 8 months after the end of the program (phase 5). The first 3 surveys were completed on the web or over the phone, and interviews and follow-up surveys were conducted over the phone. For each data collection point, the research staff contacted participants via phone to remind them to complete the survey or interview for a maximum of five attempted phone calls. Participants could complete the surveys and interviews in either English or Spanish.

Figure 1. Study timeline, showing the different study phases and data collection points.

The data collection instruments were developed in English. The first 3 surveys and the interview guide were translated into Spanish by an external translation vendor. These translations were reviewed for accuracy and readability by 3 Spanish-speaking members of the research team. For example, vendor translation sometimes used formal words that are not typically used in everyday language, or the intent of a question was lost. These translations were updated to better encapsulate the original intent of the questions. The fourth follow-up survey was translated directly by the Spanish-speaking members of the research team.

Digital literacy training is offered in both English and Spanish. The training consisted of four classes on the following topics: (1) computer basics, (2) internet basics, (3) email basics, and (4) myStrength. All digital literacy classes were voluntary for both participants and program staff.

Because of the study taking place during the COVID-19 pandemic, digital literacy classes had to take place remotely over Zoom videoconferencing. However, because the participant population had little prior technology experience, Marin County decided to provide tablets and internet connections for those participants who needed it, and each participant received at least one in-person visit by the program staff to ensure that all participants were able to access the web-based training and engage in the program. To support participants throughout the program, Marin County partnered with volunteer nurse interns from 2 local universities and *promotores* (volunteers from the North Marin Community Service's Promotores Program that work closely with the Spanish-speaking community in Marin County). Marin County chose to partner with nurse interns and promotores because they were trained and experienced community partners; furthermore, promotores were trusted community leaders in the Spanish-speaking community. The promotores network works to bridge and connect the community with health services and community engagement (for more information about the program, see the North Marin website [25]). The support provided by the interns and promoters included the following:

- Technical support, such as helping participants get connected to Wi-Fi,

- support during digital literacy training, such as teaching participants how to use their devices and Zoom app to participate in digital literacy classes,
- checking in with participants about their participation in the program, and
- connecting participants to resources and mental health services.

Participants were also offered the option of having a staff member visit them in person or provide remote support. Support was ongoing throughout the program, with weekly check-in calls from staff, nurse interns and promotores, and in-person visits as needed for participants who experienced more challenges with technology. The number of visits ranged from 1 (all participants had at least one in-person visit) to 10 visits.

Recruitment

To help recruit older adult participants for the program, the county approached existing community-based organizations and agencies working with isolated adults, such as West Marin Senior Services, Meals on Wheels, and organizations that work closely with the Spanish-speaking older adult community. They also approached divisions in the county that already organized health programs for older adults, as they had the infrastructure and relationships in place to reach the older adult population.

Recruitment was primarily done by sharing information about the program in partnership with West Marin Senior Services and the Telehealth Equity Project in the Marin County Division of Aging. In addition, information was shared by the promotores within their network to recruit Spanish-speaking older adults. One of the county staff members was a *promotora* who had full access to the network. These partners also helped identify isolated older adults.

The program was specifically intended for isolated English- and Spanish-speaking older adults, as they were identified by county staff as underserved populations in the county. Participants expressed interest in the program by completing a screener form to assess their eligibility ([Multimedia Appendix 2](#)). This form can be completed on the web or over the phone in English or Spanish. Participants had to be 60 years or older (the threshold age for older adults as defined by the United

Nations [26]) and be able to speak, read, and write in either English or Spanish. If participants had any mental or physical challenges that could limit their participation in the program, such as hearing or vision loss, they were contacted by a staff member to determine their eligibility. Program participants were expected to participate in surveys and interviews as part of the program. Marin County aimed to recruit 15 English- and 15 Spanish-speaking isolated older adults for the program, motivated by perceived capacity and the number of people the county staff were able to handle overall and recruit.

Participants

A total of 30 participants were enrolled in the study. The first survey (phase 1) was completed by 28 participants, the second survey (between phases 2 and 3) was completed by 25 participants, the third survey (phase 4) was completed by 23 participants, the fourth survey (after phase 4) was completed by 18 participants, and 24 participants participated in the interview.

One participant did not complete the demographic survey. Among the 29 participants who provided demographic data,

their ages ranged from 60 to 89 years, with a mean age of 72 (SD 7.8) years, and 27 (93%) participants identified as female. Furthermore, 28% (8/29) identified as non-Hispanic White, 24% (7/29) as Mexican or Mexican American, 17% (5/29) as Central American, 10% (3/29) as South American, 10% (3/29) as other White, 3% (1/29) as other Hispanic or Latinx, 3% (1/29) as Black or African American, and 3% (1/29) as other or African. In addition, 48% (14/29) of participants indicated that English was their preferred language, and 45% (13/29) of participants reported Spanish as their preferred language. One participant indicated a preference for Vietnamese, and 1 participant did not indicate a preference. One Spanish-speaking participant did not have reading proficiency and completed the survey verbally with staff assistance.

Measures

Overview

Table 1 presents an overview of the study measures and the points at which the measures were assessed. Next, we describe these measures in detail.

Table 1. The measures of the study^a.

Study measures	Survey 1 (phase 1)	Survey 2 (between phase 2 and 3)	Survey 3 (phase 4)	Interview (phase 4)	Survey 4 (phase 5)
Important aspects about technology		✓		✓	
Activities people would like to do using (mental health) technology		✓		✓	
Experience with myStrength			✓	✓	✓
Use of myStrength			✓	✓	✓
Other resources and strategies used to support mental health and well-being				✓	
Overall experience with or impact of the program			✓	✓	✓
Experience with program staff				✓	
(If applicable) Reasons for dropping out of the program early				✓	
(If applicable) Reasons for nonuse of myStrength			✓	✓	✓
Digital literacy, mental health concerns	✓				
Loneliness and social isolation	✓	✓	✓		
Other health outcomes (eg, mental health improvements, and stigma reduction)				✓	

^aThe columns indicate at which data collection points the measures were assessed.

Digital Literacy

At the start of the program (phase 1), participants were asked to rate 2 statements related to digital literacy (eg, "I am confident using technology to look up information") taken from the Mental Health Literacy Scale [27]. They were asked to rate these statements on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Participants were also asked about their technology use on the screener form. The purpose of these questions was to assess the extent to which support was needed for the onboarding participants.

Loneliness

Loneliness was measured at 3 time points: at the beginning of the program (phase 1), after digital literacy training (phase 2), and at the end of the program (phase 4). Loneliness was measured using the Three-Item Loneliness Scale [28], a shortened version of the University of California, Los Angeles Loneliness Scale [29]. Participants were asked to rate 3 statements related to loneliness on a 3-point Likert scale ranging from hardly ever (1) to often (3; eg, "How often do you feel left out?") with a total added score ranging from 3 to 9. People with a score of 6 or higher were grouped as lonely [30].

Social Isolation

Social isolation was measured at three time points: at the beginning of the program (phase 1), after digital literacy training

(phase 2), and at the end of the program (phase 4). The social interaction subscale of the Modified Duke Social Support Index [31,32] was used to measure social isolation. The 6-item scale asked participants about the frequency of social interactions in the past week (eg, 'How many times in the past week did you talk with friends or relatives on the telephone?') and how many people in their local area they could depend on. The total score ranges from 0 to 17. Individuals were considered socially isolated if they scored 6 points or less [30,33].

Mental Health

At the beginning of the program (phase 1), participants were asked a multiple-choice question on whether they had ever experienced or were currently experiencing any mental health challenges. Participants indicated that they had not experienced any challenges, had been diagnosed with a mental health condition, had not been diagnosed but were experiencing challenges, or could self-describe their experience. The survey explained that many different terms could be used to refer to these experiences, including mental health challenges, emotional distress, and psychological disorders.

Important Aspects of Mental Health Apps and Activities People Would Like to Do Using (Mental Health) Technology

Before engaging with myStrength (before phase 3), participants were asked what they considered important aspects of mental health technologies and what they would like to do using technology. For each of these questions, they were instructed to "select all that apply" from a list of 9 and 8 options, respectively, or give an answer in their own words. These lists were based on common needs from the literature regarding mental health technology [16], as well as common answers given by community members in prior studies of the Help@Hand project. Participants were also asked to elaborate on the important aspects and any barriers to use during the interviews.

User Experience of myStrength

After engaging with myStrength (after phase 3), participants were asked about their experience using myStrength (eg, their satisfaction using myStrength, its perceived usefulness, ease of use, and cultural competency). Participants were asked to rate 15 statements (eg, "I think myStrength is easy to use") on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5). The survey items were based on the Unified Theory of Acceptance and Use of Technology questionnaire [34], which is used to evaluate technology acceptance and adoption.

Continued Use and Experience of myStrength

Participants were contacted for a follow-up survey 8 months after the program was concluded (phase 5). The aim of the final survey was to learn more about participants' continued use and experience of myStrength after the end of the program and staff support and asked questions similar to survey 3 (phase 4) applied to the period after the staff support had ended. Survey questions assessed people's self-reported use of, and intention to use, myStrength, reasons for use and nonuse, challenges using

myStrength, people's perception of myStrength, facilitating conditions that would make it easier to use myStrength, and the overall impact of the program.

Interviews

Participants were invited for an interview at the end of their participation in the program (phase 4). The aim of the interviews was to gain a more in-depth understanding of participants' experiences with the program. The research team developed an interview guide with discussions and sample questions. The topics covered in the interview guide included the following:

1. Overall experience with the program
2. Health outcomes (eg, mental health improvements and stigma reduction)
3. Experience with program staff
4. Other resources and strategies used to support mental health and well-being
5. If applicable, reasons for dropping out of the program early

Ethics Approval

The study was approved by the University of California, Irvine, Institutional Review Board (review number 20195406).

Participant Consent and Compensation

Before each survey and interview, the participants were emailed a study information sheet that was reviewed and approved by the Institutional Review Board. For interviews and surveys over the phone, the sheet was reviewed at the start of the phone call with an opportunity for participants to ask questions. Participants were asked for their permission to audio-record the interviews at the start of the phone call; participants were still able to participate if they declined to be recorded. In addition, participants were asked to sign a participation agreement and a device user agreement by Marin County at the start of the program. The participation and device user agreements were reviewed and approved by the county's legal department. All collected research data were stored securely and confidentially on a password-protected secure server. Participants received a US \$10 gift card for each survey that they completed and a US \$20 gift card for completing the interview. The survey took about 20 minutes to complete, and the interview took approximately 40 minutes.

Data Analysis

We analyzed the survey data using descriptive statistics, such as frequency counts. The analyses were performed using SAS statistical software (version 9.4; SAS Institute Inc) [35]. Changes in loneliness and social isolation before and after the program were assessed using paired 2-tailed *t* tests.

Audio recordings of the interviews were transcribed verbatim in the language they were conducted in. The transcripts were analyzed by 2 members of the research team (one of whom was fluent in both English and Spanish) using thematic analysis. The themes were derived as follows: the interview protocol was developed to ascertain the overall experience with the program based on some collectively agreed upon outcomes among the research team, the county, and the broader collaboration that supported the Help@Hand Project; thus, these constructs (eg, myStrength user experience and mental health improvements)

were used as a starting point for the thematic analysis, and interview questions were grouped based on construct (eg, all questions related to barriers to use of myStrength were grouped together). However, the interview also allowed for emergent ideas either not mapped directly to the outcomes of the project or high enough that important subthemes could emerge (eg, in the case of myStrength use). Transcripts were first independently reviewed by each research member. The 2 members then met to discuss patterns and notes and iteratively developed a codebook that fit within the topic categories and captured the patterns identified in both English and Spanish transcripts. Spanish transcripts were then translated into English by an external translation vendor, and transcripts were coded by 1 researcher and discussed with the other research team member. The desktop version of the qualitative analysis platform Dedoose [36] was used to code the transcripts.

Results

Overview

The following section presents an overview of the survey and interview results. The results are organized by theme rather than by method (ie, survey vs interview). Themes were based on interview themes (eg, all findings related to myStrength user experience were grouped together). Throughout this section, we report the number of people who provided a particular survey answer and the total number of people who responded to the survey question.

Motivation to Participate in the Program

Participants reported that they became aware of the program primarily through word of mouth and recommendations from others. The reasons for deciding to participate were as follows: (1) participants trusted the program based on recommendations from others who had participated in other programs offered by Marin County, (2) they wanted to contribute to a program that may help other people, and (3) they were interested in learning something new.

Before engaging with myStrength, participants, overall, reported positive expectations about myStrength on the survey: they believed that it could be useful in their everyday life (18/26, 69%) and improve their mental health (15/16, 94%), and around half (14/26, 54%) of them believed that it might meet their mental health needs. However, during the interviews at the end of the program, it was revealed that not everyone knew beforehand what the app did or had a clear image of what “digital mental health” was: “I had no idea what it was about, so I had no expectations. I didn’t know about it. She told me about it, and so it’s hard to have expectations on something you don’t know about.”

Support During the Program

Technical Support

Although participants were excited about participating in the program, they had varying degrees of readiness to begin the program. For example, before starting the program, 48% (11/23) of participants reported that they did not have the required devices to use myStrength and participate in the digital literacy

classes; these participants were given a tablet by Marin County to participate in the program. In addition, 21% (6/28) of participants needed support to access Wi-Fi, and 21% (6/28) of participants said they had never accessed the internet before. The technical support available through the program was experienced by participants as very useful, as expressed during the interviews: “The one-on-one technical [support]...[opened the door to new technology].”

Support to Improve Digital Literacy

Participants indicated different levels of confidence in using technology in the surveys, and some not only needed support to access technology but also needed support to learn how to use technology. Almost half (12/26, 46%) of the participants reported that they were not confident about using technology at the start of the program. Furthermore, 76% (19/25) of participants did not know how to download a mobile app on their phone, and 62% (16/26) of them reported they were not comfortable using technology to support their well-being: “Todo era nuevo para mí. No sabía manejar estas cosas. Simplemente, el celular... (Everything was new to me. I didn’t know how to handle these things. Simply, the cell phone...)”

However, participants were interested in learning new technology skills and improving their digital literacy. Specifically, participants were keen to learn how to use smartphones (24/26, 92%), the internet (24/26, 92%), computers (22/26, 85%), and how to use technology to support their mental health (22/26, 85%).

Guided Support

Overall, 80% (24/30) of participants participated in at least one of 4 digital literacy classes. Even though participants had positive experiences with digital literacy training and gained digital skills, not everyone felt confident enough after the training to use web-based tools on their own for mental health. Participants reported feeling overwhelmed by the digital literacy training at times and the variety of options they were exposed to and noted the challenges of trying to teach and accommodate different needs in a group setting:

Well, for a long time, I found it quite confusing. And somehow I guess I was feeling negative at first and thinking, Oh, dear...It’s sort of a new speak, you know? [laughs]...But later, I began to see some of the value in it [digital literacy training] and realized it wasn’t really the way I was feeling it would be...Well, it’s certainly helped to be able to be in touch with family. And so, that was lovely.

To apply the skills they learned, participants not only needed human support at the beginning of the program but also continued support during the program. This ongoing support was experienced as helpful throughout the program:

Oh, I thought it was great. The nurse—Yes, she was wonderful. She just went overboard to do the training as best she could...Oh, yes, they were very willing to do whatever I needed help with...She was continually there. And if I needed her to help me with something, she went out of her way to come over and help me at

that particular time...[helped] in person...Yes, exactly. If I needed help, I need someone right here.

A total of 5 participants reported that they did not have the support of family and friends to remind them or help out if they had any technical difficulties, and staff, therefore, checked in with participants one on one throughout the program. This personal support was experienced as very helpful: “The one-on-one contact was very good, excellent.”

Social Component

Using Technology to Connect to Others

Social connection was an important aspect of the program for participants: 75% (21/28) of participants reported that they wanted to learn how to use technology to connect with friends and family. This was echoed during the interviews as follows:

I think it's WhatsApp or something like that?...Yeah, they, my friends have been telling me, so that way you don't have to pay long distance phone calls...So that's what I remember. So I think, you know, it's good to have that.

Impact of Program on Loneliness and Social Connectedness

A total of 38% (11/29) of participants reported experiencing mental health challenges at the beginning of the program. During the interviews, participants described how the global coronavirus pandemic had affected them, including increased isolation, lack of social connection and activities (including evolving practices, tensions in personal safety and risk vs social connection and mental health), safety measures, and even some having experienced the virus. In addition, many participants reported feeling isolated or lonely. For some, health conditions impacted feelings of isolation:

Sí, ahorita con la pandemia, sí me siento sola. Y como que sí me da miedo quedarme sola porque, primero, hay tantas cosas que me han pasado... (Yes, right now with the pandemic, I do feel alone. And yes I'm afraid of being left alone because, firstly, there are so many things that have happened to me...)

Overall, 78% (18/23) of participants indicated on the survey at the end of the program that they felt more connected to others because of the program. To compare any changes in loneliness and social isolation during the program, we considered 22 participants who answered the survey questions on loneliness and social isolation at baseline and at the end of the program. There was a statistically significant decrease in loneliness from the beginning of the program (mean 6.6, SD 1.7) to the end (mean 5.5, SD 2.1; $t_{21}=3.04$; $P=.006$). A total of 17 participants scored high on loneliness at the beginning of the program; at the end of the program, only 9 participants scored high on loneliness.

There was also a statistically significant decrease in social isolation from the beginning of the program (mean 7.4, SD 4.0) to the end (mean 9.2, SD 5.0; $t_{21}=-3.11$; $P=.005$). A total of 7 participants were considered socially isolated at the beginning

of the program and 4 were considered socially isolated at the end of the program.

Participants expanded on feelings of connection during the interviews. For example, interacting with other participants during digital literacy classes helped with loneliness, and it improved feelings of connectedness and purpose:

Yes, it [program] did help [impact feelings of connectedness]...Well, it just—I took some classes on Zoom, and going through the program, myStrength, yeah, it broadened my atmosphere a little bit, a lot I can say...

Y esto me ayuda porque ya me distrae, es algo como que puedo hacer. Porque no tengo ganas de nada, ni de limpiar la casa ni de...Ahorita no tenía ganas pero de nada. Y con esto sí me ha ayudado mucho. (And this helps me because it distracts me, it's like something I can do. Because I have no desire for anything, not to clean the house or...Right now I did not feel like doing anything. And with this it has helped me a lot.)

Participants also felt connected because they learned how to use technology to reach out to family and friends during the digital literacy classes: “I was pretty much at home most of the time and alone, so that was nice to be able to get into the technology, and reach out to more people.” Participants were satisfied with the program and hoped that more programs like this would be offered in the future.

Digital Literacy Training Experience

Most (19/24, 79%) participants were satisfied with the digital literacy training, and participants reported various benefits from participating in the training; they (13/17, 76%) reported that they were more likely to use technology and felt more socially connected (12/19, 63%), and the percentage of participants who felt confident using technology increased from 31% (8/26) before training to 73% (19/26) after training.

myStrength Use and Experience

Overall, participants had a positive experience with myStrength: they recommended myStrength (18/23, 78%), found it useful (17/23, 74%), and found it easy to use (15/23, 65%). Benefits of myStrength were that it changed how they thought about mental health and helped them recognize symptoms:

Maybe it made me more aware so that when I got in this bad mood a couple weeks ago, that...I didn't know. [chuckles] One of the sections of myStrength was—I think it was Depression...And I've never had depression. So, I don't know what it's like. But I think it was last week, all of a sudden for about two days, I think I felt what might be depression...And so, I mean to go back and check that out.

In addition, 70% (16/23) of participants indicated that they had used myStrength for 2 months at the end of the program, and 52% (12/23) had said they used it several times a week or daily. Furthermore, 57% (13/23) participants self-reported that they had stopped using myStrength at some point during the study. The main reasons participants mentioned for not using

myStrength were that they had no time (4/13, 31%), health reasons such as chronic fatigue (4/13, 31%), technical issues (3/13, 23%), and they did not think they needed it (2/13, 15%): “No funciona mucho, falla el Internet. Por eso a veces no lo ocupo. (It doesn’t work much, the Internet fails. That’s why sometimes I don’t use it.)”

Although (18/23, 78%) participants agreed that myStrength valued cultural differences, several Spanish-speaking participants mentioned that many aspects of the Spanish version were available only in English:

Pero cuando yo entré a myStrength, todo es en inglés...porque, como te digo, si sería en español, entonces iría más rápido al video...Ahorita estoy como poquito a poquito...en la tablet, son en español las instrucciones. Pero cuando voy a myStrength, todo sale en inglés...lo que yo digo de ellos es que si estuviera en español el programa, la mayoría de los programas, o sea, en inglés y en español, sería maravilloso. (But when I entered myStrength, everything is in English...because, as I told you, if it would be in Spanish, then I would go faster to the video...Right now I’m like little by little...on the tablet, the instructions are in Spanish. But when I go to myStrength, everything comes out in English...what I say about them is that if the program were in Spanish, most of the programs, that is, in English and Spanish, it would be wonderful.)

When thinking about mental health technology beyond myStrength, important aspects for participants were that the data should be kept private (20/28, 71%), the app would be free (20/28, 71%), and the app would not have a negative effect on their device, such as draining their battery (19/28, 68%).

Continued Use of myStrength: 8-Month Follow-up

Eight months after the program and staff support ended, 16 (53%) of the 30 participants completed the follow-up survey. During the program, participants overall had a positive experience with myStrength, reported high use, and reported that they were more likely to use technology to support their well-being; only 1 participant reported that they were still using myStrength at the time of the follow-up survey, even though participants still had free access to the platform. Furthermore, 75% (12/16) of the participants reported that they experienced technical issues that prevented them from continuing to use myStrength. Technical issues included forgetting how to use a tablet, smartphone, or computer and not knowing how to log back into the platform after being logged out. In addition, 19% (3/16) of the participants also reported significant life events, such as illness or injury, contributing to nonuse. Although 50% (8/16) participants reported that they intended to use myStrength again in the future, 44% (7/16) of participants also indicated that ongoing support would make it much easier to use myStrength.

Discussion

Principal Findings

This paper sought to share what older adults consider important factors of support when engaging in a digital mental health program. We found that participants valued different types of support to stay engaged in the program: technical support, guided support to be reminded to engage with the program, and social support. Social connections were the most important aspect of the program for participants. In addition, even though participants were excited to participate in the program, they did not have a clear image of digital mental health and were mostly motivated to participate in the program because it was recommended by someone they trusted, such as a community partner with whom they were familiar or peers who had participated in similar programs offered by the county. Next, we discuss these findings in more detail and the implications they may have for the future design of digital mental health programs for older adults.

Technical Support and Guidance

Although the program showed high engagement and an improvement in loneliness and social connectedness, it is important to understand these results in the context of considerable hands-on staff support, rather than solely attributing these improvements to mental health technology. When contacted 8 months after the program—and thus the support—had ended, most respondents had stopped using myStrength.

Nonuse of a digital health app does not always have to be negative [37]; for example, users’ needs may have changed, or they may have received sufficient benefits from the app and no longer need it. Therefore, it is important to look beyond user engagement and understand the reasons behind the potential nonuse of a DMHI. In our study, participants indicated that they intended to use myStrength beyond the program but experienced several barriers in doing so. For example, in the follow-up survey, respondents noted that they encountered technical issues and were unable to easily overcome them without any technical assistance. This is consistent with previous work that found technical issues to be a common barrier to engagement with DMHIs [16], and technical assistance has been shown to be an important aspect in facilitating engagement [38,39].

Providing continued support from a trained person was shown to be important in reminding older adults to use the technology and troubleshoot in case of technical issues during the program. Furthermore, it was important to provide ongoing personal support. In a separate publication [40], we discuss the role of support in the program from a staff perspective, who confirmed that, in particular, participants with lower digital literacy needed considerable support and reminders in between classes to practice what they had learned during the digital literacy training and to use the technology.

These findings indicate that when integrating DMHIs into existing programs, typical measures of success, such as efficacy or changes in mental health outcomes, may not be sufficient to achieve a long-term beneficial impact. The real-world success

of DMHIs in older adults may depend on the presence of intensive and ongoing support. In addition, this support may have benefits for both sustaining engagement and alleviating social isolation.

To ensure that this resource-rich support is sustainable, it is important to assess support needs and build wrap-around services around DMHIs for older adult populations. For example, there may be an opportunity to build and maintain close partnerships with local agencies that already work closely with older adult communities, such as home nursing services.

The facilitating factor of personal guidance is widely known in the digital mental health literature, and guided DMHIs typically have higher engagement than unguided DMHIs. For example, 2 studies that evaluated a self-guided DMHI found that participants regularly forgot to use the intervention [41,42]. In contrast, studies that compared versions of an intervention with and without guidance (in the form of emails, phone calls, or text reminders by a human coach) found that guided use increased user engagement [43-45]. In addition, prior studies found that participants preferred whether an intervention was integrated into therapy or an existing program rather than being offered as a standalone tool [43,46,47].

However, human support is resource intensive and may not always be feasible. Therefore, different types of support may be used to address the different types of challenges people experience. For example, for older adults with higher digital literacy, email and text reminders could remind people of the technology at regular intervals; automated reminders have been found to facilitate engagement with the general population and have been experienced positively [16]. In addition, a larger digital literacy initiative may be needed to improve skills in overcoming technical issues, such as formal training classes or more informal walk-in sessions to troubleshoot common technical issues. Digital literacy training in this study increased participants' confidence in using technology and the likelihood of using technology and may further increase their ability to engage in digital programs. In addition, these classes can provide an opportunity to learn how to use technology beyond a mental health tool, such as communication technologies, to connect with friends and family, which can impact social connectedness. Individual human support can be dedicated to those needing more personalized support or motivating encouragement.

Social Connectedness Beyond DMHIs

Both feelings of loneliness and feelings of being socially isolated significantly decreased from before to after the program. This decrease may indicate that participation in the program reduced social isolation in some participants. Participants indicated that they liked being able to connect with people such as other participants, digital training teachers, county staff, nurses, and promotores during the program, and some even took part in the program because they wanted to help others in improving the program for future deployment among older adults.

This is in line with prior research highlighting that offering social opportunities is an important facilitator of DMHI use among older adults [48]. For example, a study evaluating a physical activity program with older adults offered participants

the option to complete activities in a group setting with other participants [48]. Those who chose to complete activities as a group reported greater social connectedness and improvements in their mental health, compared with those who chose to do the activities on their own.

Given the great value placed on a social component, previous work has recommended that DMHIs should incorporate a feature or ability to connect with others [49]. However, although our study participants did consider a social component to be important when considering mental health technologies, these interactions may not necessarily have to be through the DMHI: participants in our study also valued being able to connect with others through the program and the digital literacy classes by connecting with staff members and loved ones by learning communication technologies such as WhatsApp and Zoom. This indicates that different types of social connections may be important: connections with other community members who may or may not have shared mental health experiences, connections with family and friends, and connections with supporting staff.

Future digital mental health programs could host sessions where participants may be able to meet others taking part in the program and connect with others and may focus their digital literacy initiatives on teaching skills on how to connect and communicate with loved ones.

Endorsement and Outreach by Community Partners

Our study also revealed the importance of collaborating with community partners who have a close relationship with the community to get a digital mental health program that starts with older adults. Participants primarily indicated that they became aware through word of mouth and decided to participate because they trusted the person who recommended it to them. Efforts to recruit and engage participants without the assistance of partners were largely ineffective. For instance, promotores helped recruit almost all Spanish-speaking older adults, whereas other partners were instrumental in recruiting English-speaking older adults. This finding is consistent with prior work that highlighted that community-based partnerships can be crucial for populations who may face barriers to accessing care [50]. For example, a previous case study with older adults indicated that social workers could be suitable community partners because of their strong commitment to improve services for older adults, direct practical experience with barriers to care, and their skills for building collaborative relationships [51].

Promotores may have also played an important role in engaging Spanish-speaking participants throughout the program, as they have the cultural sensitivity and experience of the community. There is a need for mental health programs to be culturally appropriate [52], which ideally includes cultural vetting of recruitment, professional staff, and program materials. The fact that not all parts of myStrength were available in Spanish may have formed a barrier for Spanish-speaking participants in our study to fully engage with the platform. Nevertheless, an exploratory study in Marin County found that older adults still preferred myStrength over other mental health products, notably given the lack of suitable alternatives. This highlights the need to develop mental health programs that address the diverse

cultural and linguistic backgrounds of the underserved communities.

Technology adoption is often influenced by the beliefs and recommendations of people close to the user, such as family, friends, or health providers [53]. This may be especially true for an older adult population who may have had little exposure to technology and may lack a mental model of what digital mental health entails. Digital mental health is a novel concept: it may be difficult to explain a digital mental health program to people who have little prior experience with technology and for them to envision how a mental health app may work. Therefore, participants may have been less likely to participate if it was advertised as a digital mental health program without endorsement from the people they trusted. Partnering with local organizations that have already established trust in communities can help reach specific populations.

Community partners can further play an important role in sustained engagement, as indicated by the finding that a large number of the participants discontinued myStrength after the program and support had ended. There may be opportunities in the future to make this type of support and community engagement an essential part of future staff job roles. In addition, prior work with a vulnerable youth population recommended integrating mental health services into services that youth are already receiving, such as schools [50]. There may be an opportunity to integrate aspects of a digital mental health program into existing home and community services that older adults are already receiving.

Limitations

One limitation of our study pertains to missing data as not all participants completed each survey or survey question. Engaging older adults in research activities is a known challenge [54], and participants in our study at times could be overwhelmed by various program activities. Future programs should take this into consideration and keep survey instruments as brief as possible and offer different methods to complete surveys and interviews, for example, on the web, over the phone, or when possible, in person.

In addition, most of the participants were female. Men are often underrepresented in older adult health research programs, and

future efforts may need to tailor recruitment strategies to increase male participation, for example, by featuring men in promotional material [55].

The digital mental health program studied in this paper was particularly focused on isolated (eg, geographically, culturally, and socially) older adults in specific communities with little access to mental health resources; therefore, caution must be exercised when generalizing the results. That said, we expect our findings to generalize to isolated older adults with low digital literacy.

The program offered staff support and myStrength to all participants and there was no comparison group. We did not know if the participants would have experienced similar benefits in the absence of this intervention.

Finally, this study was conducted during the COVID-19 pandemic, which may have exacerbated the level of isolation. It also added the challenge of providing support and digital literacy training remotely when participants were already struggling to get on the web in the first place. Future programs may offer in-person classes and provide more in-person support.

Conclusions

DMHIs have the potential to improve mental health among older adults; however, factors such as low digital literacy and lack of support can form a barrier to starting and sustaining engagement. We found that ongoing support (eg, technical support, guided support, and social support) not only may be important from an implementation perspective to retain people in a digital health intervention but also may be more crucial in improving people's experience and supporting their mental health than the technology itself. Having a social component and being able to participate in a program with other people may address loneliness and motivate people to enroll in a program that can help them stay connected with family and friends and improve their technical skills.

Future research should keep this in mind and provide a social support component, whenever possible. For example, web-based Zoom sessions can be held to connect with others using the platform, or a DMHI can provide the option to provide individual support and guidance from staff or coaches.

Acknowledgments

This work was funded by the Help@Hand Project (agreement 417-ITS-UCI-2019), a project overseen by the California Mental Health Service Authority (CalMHSA). CalMHSA reviewed the manuscript to ensure its confidentiality. The information or content and conclusions presented here are those of the authors and should not be construed as the official position or policy of nor should any endorsements be inferred by the participating Help@Hand Counties or CalMHSA.

Data Availability

The survey instruments and interview protocol are included as supplementary materials in [Multimedia Appendix 1](#). The data collected and analyzed during the study are not publicly available as they were not funding requirements, and the study participants did not consent for their data to be shared publicly.

Conflicts of Interest

SMS has received consulting payments from Otsuka Pharmaceuticals and Trusst (K Health) and is a member of the Headspace Scientific Advisory Board, for which he receives compensation. The authors have no further interests to declare.

Multimedia Appendix 1

The survey instruments and interview protocol.

[[DOCX File, 135 KB - formative_v6i12e43192_app1.docx](#)]

Multimedia Appendix 2

A screener form was used to assess participants' eligibility to participate in the program.

[[DOCX File, 41 KB - formative_v6i12e43192_app2.docx](#)]

References

1. Smith ML, Steinman LE, Casey EA. Combatting social isolation among older adults in a time of physical distancing: the COVID-19 social connectivity paradox. *Front Public Health* 2020 Jul 21;8(403):1-9 [FREE Full text] [doi: [10.3389/fpubh.2020.00403](https://doi.org/10.3389/fpubh.2020.00403)] [Medline: [32850605](https://pubmed.ncbi.nlm.nih.gov/32850605/)]
2. Allan CE, Valkanova V, Ebmeier KP. Depression in older people is underdiagnosed. *Practitioner* 2014 May;258(1771):19-13. [Medline: [25065018](https://pubmed.ncbi.nlm.nih.gov/25065018/)]
3. Cosco TD, Fortuna K, Wister A, Riadi I, Wagner K, Sixsmith A. COVID-19, social isolation, and mental health among older adults: a digital catch-22. *J Med Internet Res* 2021 May 06;23(5):e21864 [FREE Full text] [doi: [10.2196/21864](https://doi.org/10.2196/21864)] [Medline: [33891557](https://pubmed.ncbi.nlm.nih.gov/33891557/)]
4. The Pandemic Effect: A Social Isolation Report. American Association of Retired Persons Foundation. 2020 Oct 6. URL: <https://connect2affect.org/wp-content/uploads/2020/10/The-Pandemic-Effect-A-Social-Isolation-Report-AARP-Foundation.pdf> [accessed 2022-10-04]
5. Mental health and psychosocial support. United Nations High Commissioner for Refugees. URL: <https://emergency.unhcr.org/entry/49304/mental-health-and-psychosocial-support> [accessed 2022-11-08]
6. Ibarra F, Baez M, Cernuzzi L, Casati F. A systematic review on technology-supported interventions to improve old-age social wellbeing: loneliness, social isolation, and connectedness. *J Healthc Eng* 2020 Jul 12;2020:2036842 [FREE Full text] [doi: [10.1155/2020/2036842](https://doi.org/10.1155/2020/2036842)] [Medline: [32765823](https://pubmed.ncbi.nlm.nih.gov/32765823/)]
7. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. *SSM Popul Health* 2022 Mar;17:101020 [FREE Full text] [doi: [10.1016/j.ssmph.2021.101020](https://doi.org/10.1016/j.ssmph.2021.101020)] [Medline: [35024424](https://pubmed.ncbi.nlm.nih.gov/35024424/)]
8. Xiang X, Sun Y, Smith S, Lai PH, Himle J. Internet-based cognitive-behavioral therapy for depression: a feasibility study for homecare older adults. *Innov Aging* 2020;4(Suppl 1):371. [doi: [10.1093/geroni/igaa057.1194](https://doi.org/10.1093/geroni/igaa057.1194)]
9. Xiang X, Wu S, Zuverink A, Tomasino KN, An R, Himle JA. Internet-delivered cognitive behavioral therapies for late-life depressive symptoms: a systematic review and meta-analysis. *Aging Ment Health* 2020 Aug;24(8):1196-1206 [FREE Full text] [doi: [10.1080/13607863.2019.1590309](https://doi.org/10.1080/13607863.2019.1590309)] [Medline: [30913898](https://pubmed.ncbi.nlm.nih.gov/30913898/)]
10. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. *Digit Health* 2020 Feb 11;6:2055207620905422 [FREE Full text] [doi: [10.1177/2055207620905422](https://doi.org/10.1177/2055207620905422)] [Medline: [32110429](https://pubmed.ncbi.nlm.nih.gov/32110429/)]
11. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. *Health Educ Behav* 2015 Apr;42(2):202-209 [FREE Full text] [doi: [10.1177/1090198114547815](https://doi.org/10.1177/1090198114547815)] [Medline: [25156311](https://pubmed.ncbi.nlm.nih.gov/25156311/)]
12. Ibrahim H, Liu X, Zariffa N, Morris AD, Denniston AK. Health data poverty: an assailable barrier to equitable digital health care. *Lancet Digit Health* 2021 Apr;3(4):e260-e265 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30317-4](https://doi.org/10.1016/S2589-7500(20)30317-4)] [Medline: [33678589](https://pubmed.ncbi.nlm.nih.gov/33678589/)]
13. Seifert A, Reinwand DA, Schlomann A. Designing and using digital mental health interventions for older adults: being aware of digital inequality. *Front Psychiatry* 2019 Aug 9;10:568 [FREE Full text] [doi: [10.3389/fpsy.2019.00568](https://doi.org/10.3389/fpsy.2019.00568)] [Medline: [31447716](https://pubmed.ncbi.nlm.nih.gov/31447716/)]
14. Andrews JA, Brown LJ, Hawley MS, Astell AJ. Older adults' perspectives on using digital technology to maintain good mental health: interactive group study. *J Med Internet Res* 2019 Feb 13;21(2):e11694 [FREE Full text] [doi: [10.2196/11694](https://doi.org/10.2196/11694)] [Medline: [30758292](https://pubmed.ncbi.nlm.nih.gov/30758292/)]
15. Psychosocial Interventions for Older Adults With Serious Mental Illness. Publication No. PEP21-06-05-001. Substance Abuse and Mental Health Services Administration. Rockville, MD, USA: Substance Abuse and Mental Health Services Administration (US); 2021. URL: https://store.samhsa.gov/sites/default/files/SAMHSA_Digital_Download/PEP21-06-05-001.pdf [accessed 2022-10-04]
16. Borghouts J, Eikley E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *J Med Internet Res* 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: [10.2196/24387](https://doi.org/10.2196/24387)] [Medline: [33759801](https://pubmed.ncbi.nlm.nih.gov/33759801/)]

17. Egede LE, Acierno R, Knapp RG, Lejuez C, Hernandez-Tejada M, Payne EH, et al. Psychotherapy for depression in older veterans via telemedicine: a randomised, open-label, non-inferiority trial. *Lancet Psychiatry* 2015 Aug;2(8):693-701. [doi: [10.1016/S2215-0366\(15\)00122-4](https://doi.org/10.1016/S2215-0366(15)00122-4)] [Medline: [26249300](https://pubmed.ncbi.nlm.nih.gov/26249300/)]
18. Tomasino KN, Lattie EG, Ho J, Palac HL, Kaiser SM, Mohr DC. Harnessing peer support in an online intervention for older adults with depression. *Am J Geriatr Psychiatry* 2017 Oct;25(10):1109-1119 [FREE Full text] [doi: [10.1016/j.jagp.2017.04.015](https://doi.org/10.1016/j.jagp.2017.04.015)] [Medline: [28571785](https://pubmed.ncbi.nlm.nih.gov/28571785/)]
19. Knaevelsrud C, Böttche M, Pietrzak RH, Freyberger HJ, Kuwert P. Efficacy and feasibility of a therapist-guided internet-based intervention for older persons with childhood traumatization: a randomized controlled trial. *Am J Geriatr Psychiatry* 2017 Aug;25(8):878-888. [doi: [10.1016/j.jagp.2017.02.024](https://doi.org/10.1016/j.jagp.2017.02.024)] [Medline: [28365000](https://pubmed.ncbi.nlm.nih.gov/28365000/)]
20. Chikersal P, Belgrave D, Doherty G, Enrique A, Palacios JE, Richards D, et al. Understanding client support strategies to improve clinical outcomes in an online mental health intervention. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Apr Presented at: CHI '20; April 25-30, 2020; Honolulu, HI, USA p. 1-16. [doi: [10.1145/3313831.3376341](https://doi.org/10.1145/3313831.3376341)]
21. Buck B, Kopelovich SL, Tauscher JS, Chwastiak L, Ben-Zeev D. Developing the workforce of the digital future: leveraging technology to train community-based mobile mental health specialists. *J Technol Behav Sci* (forthcoming) 2022 Aug 03:1-7 [FREE Full text] [doi: [10.1007/s41347-022-00270-6](https://doi.org/10.1007/s41347-022-00270-6)] [Medline: [35967965](https://pubmed.ncbi.nlm.nih.gov/35967965/)]
22. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. *Clin Psychol Sci Pract* 2016 Nov 17;24(1):27-45. [doi: [10.1111/cpsp.12173](https://doi.org/10.1111/cpsp.12173)]
23. myStrength. 2022. URL: <https://mystrength.com/> [accessed 2022-11-08]
24. Marin County, CA, USA. Data USA. URL: <https://datausa.io/profile/geo/marin-county-ca> [accessed 2022-10-04]
25. Community Health Promotores Program. North Marin Community Services. URL: <https://www.northmarinc.org/community-health/promotores-program/> [accessed 2022-10-04]
26. Older Persons. United Nations High Commissioner for Refugees. URL: <https://www.unhcr.org/older-persons.html> [accessed 2022-11-08]
27. O'Connor M, Casey L. The Mental Health Literacy Scale (MHLS): a new scale-based measure of mental health literacy. *Psychiatry Res* 2015 Sep 30;229(1-2):511-516. [doi: [10.1016/j.psychres.2015.05.064](https://doi.org/10.1016/j.psychres.2015.05.064)] [Medline: [26228163](https://pubmed.ncbi.nlm.nih.gov/26228163/)]
28. Hughes ME, Waite LJ, Hawkey LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. *Res Aging* 2004;26(6):655-672 [FREE Full text] [doi: [10.1177/0164027504268574](https://doi.org/10.1177/0164027504268574)] [Medline: [18504506](https://pubmed.ncbi.nlm.nih.gov/18504506/)]
29. Russell D, Peplau LA, Ferguson ML. Developing a measure of loneliness. *J Pers Assess* 1978 Jun;42(3):290-294. [doi: [10.1207/s15327752jpa4203_11](https://doi.org/10.1207/s15327752jpa4203_11)] [Medline: [660402](https://pubmed.ncbi.nlm.nih.gov/660402/)]
30. Steptoe A, Shankar A, Demakakos P, Wardle J. Social isolation, loneliness, and all-cause mortality in older men and women. *Proc Natl Acad Sci U S A* 2013 Apr 09;110(15):5797-5801 [FREE Full text] [doi: [10.1073/pnas.1219686110](https://doi.org/10.1073/pnas.1219686110)] [Medline: [23530191](https://pubmed.ncbi.nlm.nih.gov/23530191/)]
31. Pachana NA, Smith N, Watson M, McLaughlin D, Dobson A. Responsiveness of the Duke Social Support sub-scales in older women. *Age Ageing* 2008 Nov;37(6):666-672. [doi: [10.1093/ageing/afn205](https://doi.org/10.1093/ageing/afn205)] [Medline: [18829690](https://pubmed.ncbi.nlm.nih.gov/18829690/)]
32. Powers JR, Goodger B, Byles JE. Assessment of the abbreviated Duke Social Support Index in a cohort of older Australian women. *Australas J Ageing* 2004 Jun;23(2):71-76. [doi: [10.1111/j.1741-6612.2004.00008.x](https://doi.org/10.1111/j.1741-6612.2004.00008.x)]
33. Kotwal AA, Cenzer IS, Waite LJ, Covinsky KE, Perissinotto CM, Boscardin WJ, et al. The epidemiology of social isolation and loneliness among older adults during the last years of life. *J Am Geriatr Soc* 2021 Nov;69(11):3081-3091 [FREE Full text] [doi: [10.1111/jgs.17366](https://doi.org/10.1111/jgs.17366)] [Medline: [34247388](https://pubmed.ncbi.nlm.nih.gov/34247388/)]
34. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q* 2012 Mar;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
35. SAS® 9.4 DATA Step Statements: Reference. SAS® 9.4 and SAS® Viya® 3.5 Programming Documentation. 2013. URL: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/lestmtsref/titlepage.htm [accessed 2022-10-04]
36. Dedoose Version 9.0.17, web application for managing, analyzing, and presenting qualitative and mixed method research data. dedoose. Los Angeles, CA, USA: SocioCultural Research Consultants, LLC; 2021. URL: <https://www.dedoose.com/> [accessed 2022-10-04]
37. O'Brien HL, Morton E, Kampen A, Barnes SJ, Michalak EE. Beyond clicks and downloads: a call for a more comprehensive approach to measuring mobile-health app engagement. *BJPsych Open* 2020 Aug 11;6(5):e86 [FREE Full text] [doi: [10.1192/bjo.2020.72](https://doi.org/10.1192/bjo.2020.72)] [Medline: [32778200](https://pubmed.ncbi.nlm.nih.gov/32778200/)]
38. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010 Jun 08;5(6):e10939 [FREE Full text] [doi: [10.1371/journal.pone.0010939](https://doi.org/10.1371/journal.pone.0010939)] [Medline: [20544030](https://pubmed.ncbi.nlm.nih.gov/20544030/)]
39. Robinson E, Titov N, Andrews G, McIntyre K, Schwencke G, Solley K. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010 Jun 03;5(6):e10942 [FREE Full text] [doi: [10.1371/journal.pone.0010942](https://doi.org/10.1371/journal.pone.0010942)] [Medline: [20532167](https://pubmed.ncbi.nlm.nih.gov/20532167/)]
40. Hernandez-Ramos R, Palomares K, Ceballos-Corro E, Schueller S, Stadnick N, Eikev EV, et al. Evaluation of a pilot implementation of a digital cognitive-behavioral therapy platform for county mental health services. In: Society for

- Implementation Research Collaboration 6th Annual Conference. 2022 Sep Presented at: SIRC '22; September 8-10, 2022; San Diego, CA, USA.
41. Harjumaa M, Halttu K, Koistinen K, Oinas-Kukkonen H. User experience of mobile coaching for stress-management to tackle prevalent health complaints. In: Proceedings of the 6th Scandinavian Conference on Information Systems - Nordic Contributions in IS Research. 2015 Presented at: SCIS '15; August 9-12, 2015; Oulu, Finland p. 152-164. [doi: [10.1007/978-3-319-21783-3_11](https://doi.org/10.1007/978-3-319-21783-3_11)]
 42. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. *J Med Internet Res* 2012 Jun 22;14(3):e91 [FREE Full text] [doi: [10.2196/jmir.2100](https://doi.org/10.2196/jmir.2100)] [Medline: [22743581](https://pubmed.ncbi.nlm.nih.gov/22743581/)]
 43. Bucci S, Berry N, Morris R, Berry K, Haddock G, Lewis S, et al. "They are not hard-to-reach clients. We have just got hard-to-reach services." Staff views of digital health tools in specialist mental health services. *Front Psychiatry* 2019 May 10;10:344 [FREE Full text] [doi: [10.3389/fpsyt.2019.00344](https://doi.org/10.3389/fpsyt.2019.00344)] [Medline: [31133906](https://pubmed.ncbi.nlm.nih.gov/31133906/)]
 44. Carolan S, de Visser RO. Employees' perspectives on the facilitators and barriers to engaging with digital mental health interventions in the workplace: qualitative study. *JMIR Ment Health* 2018 Jan 19;5(1):e8 [FREE Full text] [doi: [10.2196/mental.9146](https://doi.org/10.2196/mental.9146)] [Medline: [29351900](https://pubmed.ncbi.nlm.nih.gov/29351900/)]
 45. Walsh S, Kaselionyte J, Taylor SJ, Priebe S. What might affect acceptability of online positive psychology interventions for depression: a qualitative study on patient expectations'. *BMC Psychiatry* 2018 Jul 27;18(1):240 [FREE Full text] [doi: [10.1186/s12888-018-1812-x](https://doi.org/10.1186/s12888-018-1812-x)] [Medline: [30053854](https://pubmed.ncbi.nlm.nih.gov/30053854/)]
 46. Pagliari C, Burton C, McKinstry B, Szentatotai A, David D, Serrano Blanco A, et al. Psychosocial implications of avatar use in supporting therapy for depression. *Stud Health Technol Inform* 2012;181:329-333. [Medline: [22954882](https://pubmed.ncbi.nlm.nih.gov/22954882/)]
 47. Urech A, Krieger T, Möseneder L, Biaggi A, Vincent A, Poppe C, et al. A patient post hoc perspective on advantages and disadvantages of blended cognitive behaviour therapy for depression: a qualitative content analysis. *Psychother Res* 2019 Nov;29(8):986-998. [doi: [10.1080/10503307.2018.1430910](https://doi.org/10.1080/10503307.2018.1430910)] [Medline: [29385964](https://pubmed.ncbi.nlm.nih.gov/29385964/)]
 48. Kotov SV, Isakova EV, Sheregeshev VI. Possibilities for correcting emotional and behavioral impairments in stroke patients during rehabilitation therapy. *Neurosci Behav Physi* 2019 Dec 28;50(2):156-161. [doi: [10.1007/s11055-019-00882-1](https://doi.org/10.1007/s11055-019-00882-1)]
 49. Townsend L, Zippay A, Caler K, Forenza B. Technology and opportunity: people with serious mental illness and social connection. *J Soc Social Work Res* 2016 Jun 01;7(2):371-393. [doi: [10.1086/686882](https://doi.org/10.1086/686882)]
 50. Syeda MM, Fournie M, Ibanez MC, Crooks CV. Cocreating an evaluation approach for a healthy relationships program with community partners: lessons learned and recommendations. *Am J Eval* 2021 Oct 04;42(4):466-482. [doi: [10.1177/1098214021997574](https://doi.org/10.1177/1098214021997574)]
 51. Bolda EJ, Lowe JI, Maddox GL, Patnaik BS. Community partnerships for older adults: a case study. *Fam Soc* 2018 May 03;86(3):411-418. [doi: [10.1606/1044-3894.3439](https://doi.org/10.1606/1044-3894.3439)]
 52. Wang C, López-Núñez MI, Pan R, Wan X, Tan Y, Xu L, et al. The impact of the COVID-19 pandemic on physical and mental health in China and Spain: cross-sectional study. *JMIR Form Res* 2021 May 21;5(5):e27818 [FREE Full text] [doi: [10.2196/27818](https://doi.org/10.2196/27818)] [Medline: [33900933](https://pubmed.ncbi.nlm.nih.gov/33900933/)]
 53. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag Sci* 2000 Feb;46(2):186-204 [FREE Full text] [doi: [10.1287/mnsc.46.2.186.11926](https://doi.org/10.1287/mnsc.46.2.186.11926)]
 54. Lund A, Holthe T, Halvorsrud L, Karterud D, Johannessen AF, Lovett HM, et al. Involving older adults in technology research and development discussions through dialogue cafés. *Res Involv Engagem* 2021 May 10;7(1):26 [FREE Full text] [doi: [10.1186/s40900-021-00274-1](https://doi.org/10.1186/s40900-021-00274-1)] [Medline: [33971975](https://pubmed.ncbi.nlm.nih.gov/33971975/)]
 55. Anderson C, Seff LR, Batra A, Bhatt C, Palmer RC. Recruiting and engaging older men in evidence-based health promotion programs: perspectives on barriers and strategies. *J Aging Res* 2016;2016:8981435 [FREE Full text] [doi: [10.1155/2016/8981435](https://doi.org/10.1155/2016/8981435)] [Medline: [27366330](https://pubmed.ncbi.nlm.nih.gov/27366330/)]

Abbreviations

DMHI: digital mental health intervention

Edited by A Mavragani; submitted 03.10.22; peer-reviewed by E Lattie, J Andrews; comments to author 18.10.22; revised version received 08.11.22; accepted 18.11.22; published 13.12.22.

Please cite as:

Borghouts J, Eikay EV, De Leon C, Schueller SM, Schneider M, Stadnick NA, Zheng K, Wilson L, Caro D, Mukamel DB, Sorkin DH
Understanding the Role of Support in Digital Mental Health Programs With Older Adults: Users' Perspective and Mixed Methods Study

JMIR Form Res 2022;6(12):e43192

URL: <https://formative.jmir.org/2022/12/e43192>

doi: [10.2196/43192](https://doi.org/10.2196/43192)

PMID: [36512387](https://pubmed.ncbi.nlm.nih.gov/36512387/)

©Judith Borghouts, Elizabeth V Eikey, Cinthia De Leon, Stephen M Schueller, Margaret Schneider, Nicole A Stadnick, Kai Zheng, Lorraine Wilson, Damaris Caro, Dana B Mukamel, Dara H Sorkin. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Efficacy of the Mental Health App Intellect to Reduce Stress: Randomized Controlled Trial With a 1-Month Follow-up

Sean Han Yang Toh¹, BA; Jessalin Hui Yan Tan², BA; Feodora Roxanne Kosasih¹, BA; Oliver Sündermann¹, PhD

¹Intellect Pte Ltd, Singapore, Singapore

²Department of Psychology, National University of Singapore, Singapore, Singapore

Corresponding Author:

Oliver Sündermann, PhD

Intellect Pte Ltd

171 Tras St, #02-179 Union Building

Singapore, S079025

Singapore

Phone: 65 93571995

Email: oliver@intellect.co

Abstract

Background: Excessive stress is a major global health concern, particularly in young adults. Short skills-focused self-guided interventions (SGIs) on smartphones are a scalable way to improve stress-coping skills at the population level.

Objective: In this randomized controlled trial, we aimed to examine the possible efficacy of a recently developed stress-coping SGI (*Intellect*) in improving psychological distress, relative to an active control group and 2 potential moderators of this predicted relationship (ie, psychological mindedness [PM] and coping self-efficacy [CSE]).

Methods: University students (N=321) were randomly assigned to either an 8-day SGI on stress-coping or an active control group. Self-reported measures were obtained at baseline, after the intervention, and at the 1-month follow-up. The primary outcome was psychological stress (Psychological Stress Measure-9). Secondary outcomes were anxiety (Generalized Anxiety Disorder-7) and depressive symptoms (Patient Health Questionnaire-9). PM and CSE were assessed as potential moderators at baseline.

Results: The final sample (n=264) included 188 (71.2%) female, 66 (25%) male, 7 (2.7%) nonbinary, and 3 (1.1%) others participants with a mean age of 22.5 (SD 5.41) years. The intervention group reported significantly lower perceived stress (partial eta-squared [η^2]=0.018; $P=.03$) and anxiety ($\eta^2=0.019$; $P=.03$) levels after intervention relative to the active control group. The effects on perceived stress levels remained statistically significant at the 1-month follow-up ($\eta^2=0.015$; $P=.05$). Students with the lowest CSE and highest PM experienced the fastest decline in perceived stress levels ($\beta=6.37$, 95% CI 2.98-9.75). Improvements in anxiety levels were not observed at 1-month follow-up. Similarly, no intervention effects were found for depression levels at postintervention and follow-up periods.

Conclusions: This study provides evidence that the *Intellect* stress-coping SGI is effective in reducing perceived stress and anxiety levels among university students. Mobile health apps are brief, scalable, and can make important contributions to public mental health.

Trial Registration: ClinicalTrials.gov NCT04978896; <https://www.clinicaltrials.gov/ct2/show/NCT04978896>

(*JMIR Form Res* 2022;6(12):e40723) doi:[10.2196/40723](https://doi.org/10.2196/40723)

KEYWORDS

mobile health; mHealth; randomized controlled trial; RCT; self-guided interventions; cognitive behavioral therapy; CBT; stress coping; stress management; university students; psychological mindedness; coping self-efficacy; mobile phone

Introduction

Background

Psychological stress occurs when one's perceived demands exceed one's perceived capacity to cope [1,2]. Excessive and

prolonged stress can have multiple negative health effects, including heightened risk of obesity [3], impaired working memory [4,5], and cardiac arrest [6]. Adverse stress effects on mental health are also common, including an increased risk of depression and anxiety [7], insomnia [8], and psychosis [9]. Excessive stress is widely recognized as a major health burden

[10]. Young adults and late adolescents are at particular risk of experiencing heightened stress levels owing to life transitions from childhood to adulthood and taking on new responsibilities [11,12] or uncertainties relating to career prospects [13]. Equipping young adults with effective stress-coping skills is a viable attempt to lower the public burden of excessive stress [9]. Digital tools such as mobile health (mHealth) apps are likely going to play an increasingly important role in public mental health [14].

Self-guided interventions (SGIs) on smartphones are easily accessible, affordable, flexible, and convenient to use [15]. SGIs can also promote well-being and prevent illness by equipping individuals with the skills to address minor stressors before they evolve into larger problems [16]. Various randomized controlled trials (RCTs) have examined the efficacy of SGIs on mental health outcomes [17], and meta-analyses have confirmed that SGIs can improve symptoms of anxiety and depression across age groups [18,19]. A recent meta-analysis by Linardon et al [20] specifically found that SGIs were efficacious in reducing stress levels. However, the effect sizes for most interventions on stress (Hedges $g=0.35$), anxiety, and depression were rather small (Hedges $g=0.19-0.21$) and inconsistent across studies [18,19]. Although many SGIs are freely available on smartphones, only a few are supported by theoretical and empirical data [21]. In a review of 62 stress management SGIs [22], a quarter of them did not use any evidence-based stress reduction strategies. These results correspond to a separate review of 60 stress-coping SGIs whereby a third of them either failed to provide any empirical strategies or delivered a different strategy than that stated in its description [23]. In both reviews, most of the remaining SGIs focused on momentary breathing and mindfulness exercises, which produced short-term modest effects on stress and anxiety levels. Only a minority of SGIs involved evidence-based interventions such as cognitive behavioral therapy (CBT) [21,23]. Indeed, the meta-analysis by Linardon et al [20] found substantially larger effect sizes for SGIs that involved CBT-based practices compared with waitlist control group. These CBT-based effects on stress, anxiety, and mood were also superior to alternative interventions and could be strengthened when combined with brief professional guidance within the mHealth app. The same authors also concluded that research on the overall efficacy and dissemination of evidence-based SGIs is in its infancy because of the small number of studies, especially those involving younger adults (ie, 5 studies). In addition, the meta-analysis by Linardon et al [20] highlighted methodological problems with RCTs involving stress-coping SGIs, such as high risk of bias and lack of follow-up data.

Furthermore, very little is known about who benefits the most from stress-coping SGIs [24]. Previous authors were unable to analyze any moderators of web- and computer-based stress interventions owing to the “lack and inconsistency of information provided by the [reviewed] studies” [25]. To our knowledge, only Coudray et al [26] evaluated moderators in a sample of 920 college students and found that those with lower perceived present control (ie, aspects of stressors perceived to be controllable in the present) experienced significantly greater reduction in stress after 1 to 3 weeks of web-based stress

management interventions compared with those with higher perceived present control. The authors concluded their findings by encouraging similar research on moderators using diverse samples. A potential moderator that may influence the efficacy of CBT-based SGIs on mental health outcomes is coping self-efficacy (CSE). CSE refers to one’s confidence in coping strategies [27], particularly in times of hardship [28,29] and threat [30]. Slightly different from perceived control, which refers to one’s perception of the availability of any effective response, CSE describes one’s confidence in their ability to actually effect that response [31]. Some RCTs have found CSE to be a mediator of the relationship between CBT-based SGIs and improved stress, anxiety, and depression outcomes [32-34], but researchers are less clear about CSE as a moderator. For instance, individuals with higher perceived confidence in coping with stressors were expected to maximize their gains from CBT-based treatments [35], or similar to the findings on perceived present control, it is also possible for individuals who perceive that they need more help with coping have higher motivation to benefit from using a CBT-based SGI, relative to those with higher CSE. Another potential moderator of the relationship between CBT-based SGI and outcomes is psychological mindedness (PM). PM refers to one’s predisposition to be aware of, assess, and reflect on one’s mental states and behaviors cognitively and emotionally [36,37]. Self-reflective practices within CBT were involved in improving depression symptomatology in depressive disorders [38], anxiety disorders [39], and social phobia [40], as it is thought to enhance individuals’ abilities and frequency in self-monitoring and self-evaluating their cognitions, emotions, and behaviors. Despite PM being central to the successful practice of CBT [41], very little is known about the differences in an individual’s capacity for self-reflection and insight that could moderate the efficacy of the modality [36,42]. Wiles et al [43], for example, did not find PM to be a treatment moderator in CBT, although they acknowledged that their small sample size was inadequate for testing interaction effects. Some studies have recommended higher initial levels of self-reflection and insight that may predict faster responses to CBT as potential avenues for future research [36,42]. These authors speculated that metacognitive processes such as higher PM might predispose an individual’s capacity to appreciate and understand the cause-and-effect concepts of CBT-based techniques. As CBT principles are often adopted by evidence-based SGIs, it would be interesting to test how initial levels of PM can influence CBT-based therapy outcomes, as it would allow personalizing CBT-based SGI and optimizing treatment outcomes.

Objectives

Using a randomized controlled design, we examined the possible efficacy of a recently developed CBT-based stress-coping SGI in the mHealth app (*Intellect*) on a sample of Singaporean university students compared with an active control group. We first hypothesized that participants in the intervention group (stress SGI) would report lower perceived stress than the active control group after the intervention. Although the effects of self-guided CBT interventions for stress coping are typically small, we predicted possible gains for the intervention group to be maintained at 1-month follow-up, in line with related RCTs

that also found small effects present at follow-up assessments [44,45].

Second, we predicted that participants in the intervention group would report lower anxiety and depressive symptoms than the active control group after the intervention. Third, we hypothesized that users with higher levels of PM and CSE would experience the greatest reduction in perceived stress levels. As studies have shown that individuals who benefited most from stress interventions typically also experienced greater improvement in anxiety and depressive symptoms [46], we also predicted that this moderation effect by CSE and PM also applies to secondary outcomes.

Methods

Study Design

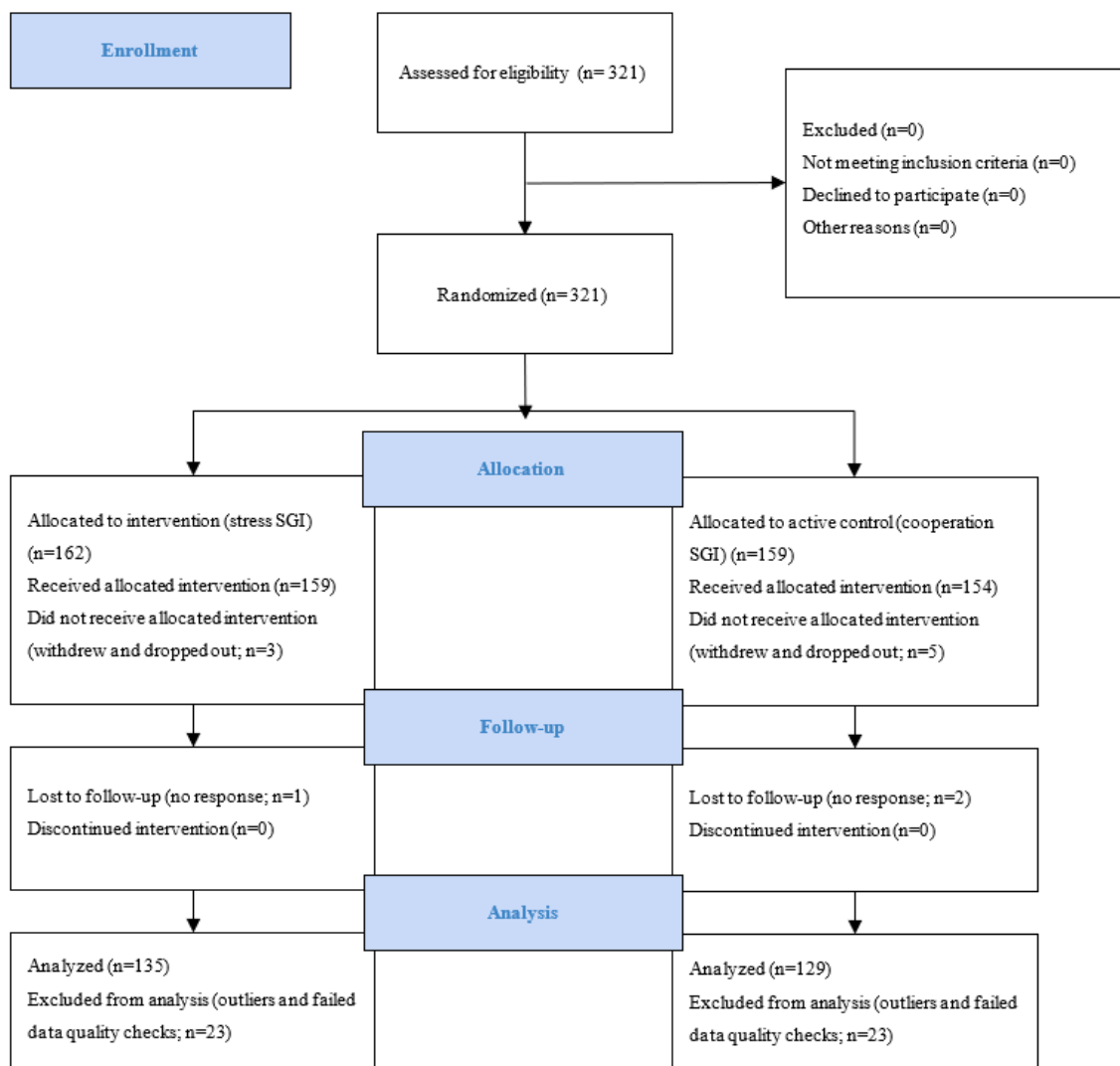
This study was an RCT with two groups: (1) an 8-day stress-coping self-guided program (intervention group) and (2) an 8-day cooperation self-guided program (active control group). A 2×3 mixed factorial experimental design was used with *condition* (intervention vs active control) as a between-group factor and *time of assessment* (baseline vs postintervention vs

1-month follow-up) as a within-group factor. The primary dependent variable measured was perceived stress levels. The secondary dependent variables were depression and anxiety levels. The CSE and Psychological Mindedness Scale (PMS) were also used for independent moderation analyses.

Recruitment and Study Participants

A total of 321 participants were recruited through the Psychology Department's Research Participation Programme and the university's research recruitment platform. Recruitment posters comprising the study procedures, inclusion criteria, and reimbursement for participation were distributed on the web and among the faculty. Of the 321 participants, 46 (14.3%) participants were excluded owing to withdrawal and failing data quality checks, whereas 3 (0.9%) participants were lost to follow-up (Figure 1). Of the 321 participants, the final sample of 264 (82.2%) participants was predominantly female (188/264, 71.2%), with a mean age of 22.45 (SD 5.41; range 18-59) years. All the participants were undergraduate students who were able to read and understand English. Elevated perceived stress was not a requirement for inclusion. Participants also received either course credits or monetary reimbursement for their participation.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart. SGI: self-guided intervention.



Power Calculation

Most studies that investigated app-based stress-coping programs found small to moderate effect sizes for the primary outcome measures of stress, mood, and anxiety [47]. Subjecting a small effect size of Cohen $f=0.1$, an α level of .05, and a power of 0.9 to G*Power revealed a minimum sample size of 214. Web-based studies are prone to attrition, and in line with similar studies, we expected 30% to 50% attrition rate [47,48] and considered that 10% of the questionnaire data may be inconsistent or invalid as commonly found in web-based survey studies [49]. Thus, we aimed to recruit 321 participants.

Procedure

Participants signed up for the study on the university's recruitment sites via a link that directed them to the web-based survey platform hosted on Qualtrics (SAP America Inc). Upon consenting to the web-based Participants Information Sheet, participants completed the primary outcome measure on perceived stress (Psychological Stress Measure [PSM]; PSM-9), secondary outcome measures on mental health (Patient Health Questionnaire [PHQ]; PHQ-9 and Generalized Anxiety Disorder [GAD]; GAD-7), CSE (Coping Self-Efficacy Scale [CSES]), PMS, and demographic information.

Next, participants were randomly assigned to either the intervention or active control condition using simple randomization procedures through the Qualtrics software. They were not informed about the conditions they were allocated to and the real nature of the study. Instead, they were informed that the study would examine the efficacy of SGIs in promoting well-being. They were then guided to download the mHealth app *Intellect* on their personal smartphones from the Apple App Store (Apple Inc) or Google Play Store (Google LLC). According to their assigned conditions, a number code was provided to them to unlock the app. Participants in the intervention condition participated in the 8-day stress-coping program, whereas those in the active control condition took part in the 8-day cooperation learning program. Both programs involved fulfilling a series of tasks aimed at improving well-being. This included content education and short daily activities, averaging 5 minutes. To promote adherence to the

study, standardized daily reminders to complete the program were sent via SMS text messages by the researcher to participants. Participants were only allowed to proceed onto the next page of the app after they had completed the preceding exercises. It was expected that every participant would complete all the self-guided daily activities. All the participants were instructed to refrain from using any other SGIs that affect well-being other than the given SGI throughout the entire duration of the study, lasting from the beginning until the end of the 1-month follow-up. This minimized the potential confounding effects.

Upon completion of the 8-day program through technical verification, participants received a survey link to complete the outcome measures and the App Engagement Scale (AES). A month after the completion of the SGI, participants were provided with a survey link to complete the outcome measures. Reimbursement was given upon completion of the postintervention measures with either 3 course credits or Singapore \$10 (US \$7.5) and after 1-month follow-up measures with an additional 1 course credit or Singapore \$5 (US \$3.75). University students intending to major in psychology were required to collect a minimum of 28 course credits over the course of an academic year. After the follow-up assessment, participants in the active control group were also given access to the stress-coping SGI.

Interventions

Stress-Coping Program

This was an 8-day program that provided psychoeducation on the negative effects of stress and effective stress management skills. Table 1 provides an overview of the 6 topics and content covered. Guided by the principles of CBT, the program helped users to identify and change unhelpful thought patterns and behaviors related to stress. Participants engaged in a series of daily exercises involving reflection and mindfulness, and they were guided to identify and type down their stressors, negative thoughts associated with the stressors, and positive affirmations. Participants were also taught breathing exercises and encouraged to practice them (Table 1 presents an overview of the program).

Table 1. Overview of stress-coping program.

Topic	Content
Topic 1: How stress affects the body	<ul style="list-style-type: none"> • Be able to identify personal stressors • Understand the difference between internal and external stressors
Topic 2: Saying “no” to burnout	<ul style="list-style-type: none"> • Understand the impact of stress on well-being • Understand and identify emotional, physiological, and behavioral signs and symptoms of stress
Topic 3: Learning to manage stress	<ul style="list-style-type: none"> • In-depth introduction of stress • Understand that stress can be healthy if kept at manageable levels
Topic 4: Self-care of the mind	<ul style="list-style-type: none"> • Understand that internal stressors are self-induced feelings and thoughts • Be able to identify negative thoughts that lead to stress • Develop skills to challenge negative thoughts • Develop skills to engage in positive affirmation
Topic 5: Self-care of the body	<ul style="list-style-type: none"> • Understand that external stressors are events or situations caused by the environment • Understand the importance of prioritizing one’s own well-being in stressful situations that cannot be controlled • Develop skills in communicating negative feelings in stressful situations • Be able to manage stress with the help of deep breathing exercises
Topic 6: Continuing daily self-care	<ul style="list-style-type: none"> • Recap of the entire learning path and key information in each learning session

Cooperation Learning Program

The 8-day program on cooperation functioned as the active control group. This program provided psychoeducation through 5 topics for participants to understand and improve collaboration

and interpersonal relationships. Short quizzes and feedback-giving exercises were included (Table 2 provides an overview of the program). The duration of the cooperation SGI matched the stress SGI in terms of time and effort required to complete the program.

Table 2. Overview of cooperation program.

Topic	Content
Topic 1: What is “cooperation”?	<ul style="list-style-type: none"> • Be able to identify personal preferences to cooperation • Understand the differences between cooperation and conformity
Topic 2: Focusing on the bigger picture	<ul style="list-style-type: none"> • Cultivate awareness of collective goals • Be able to identify possible strengths in a team member that can propel toward the collective goal • Be able to identify possible weaknesses in a team member that can hinder team efforts
Topic 3: Understanding group dynamics	<ul style="list-style-type: none"> • In-depth introduction of group dynamics • Be able to identify the correct scenarios that support or threaten group dynamics
Topic 4: Building positive relationships	<ul style="list-style-type: none"> • Understand that healthy relationships, which strengthen group dynamics, are important drivers toward attaining team’s objectives • Learn 2 methods that can build empathy and trust to improve group relationships: support and feedback • Offering support: develop skills to manage tone, body language, and learn a set of encouraging words • Offering support: identify ways to empower team members’ strengths and to cover each other’s weaknesses or blind spots • Providing feedback: in-depth introduction of the many ways one can provide constructive feedback to team members
Topic 5: Learning to work together	<ul style="list-style-type: none"> • In-depth recap of the entire learning path and key information in each learning session

Measures

Primary Outcome Measure

PSM-9 [50,51] measures the affective, cognitive, behavioral, and somatic components of psychological stress. This is a 9-item self-report measure using an 8-point scale (ranging from 1=“Not at all” to 8=“Extremely”). Some examples of the items include “I feel calm” and “I feel a great weight on my shoulders.” As

the midpoint of this scale is 4=“A bit,” an overall score of ≥ 36 on this measure would indicate that the overall sample is at elevated stress. Higher total scores reflect more stress symptoms. The PSM-9 has an acceptable internal consistency with Cronbach α ranging from .74 to .78 in this study.

Secondary Outcome Measures

Patient Health Questionnaire

PHQ-9 [52,53] is a widely used 9-item measure of depression symptoms. The PHQ-9 was included in this study as stress symptoms often include low mood and depression. Items are scored on a 4-point scale (ranging from 0="Not at all" to 3="Nearly every day"), with higher scores indicating more depressive symptoms. The total score of the PHQ-9 ranges from 0 to 27, with scores of 5, 15, and 20 indicating the cutoff points for mild, moderate, and severe depression, respectively [53]. The internal consistency of PHQ-9 in this study was very good with Cronbach α ranging from .86 to .88.

Generalized Anxiety Disorder

GAD-7 [54] is a 7-item self-report instrument that measures symptoms of generalized anxiety. We included the GAD-7 because general anxiety is associated with heightened stress levels. Similar to the PHQ-9, it uses a 4-point scale (ranging from 0="Not at all" to 3="Nearly every day"), with higher scores indicating more severe symptoms. Mild, moderate, and severe anxiety were indicated by scores ranging from ≤ 9 , 10-14, and 15-21, respectively [55]. The internal reliability of the GAD-7 in this study was excellent with Cronbach α ranging from .90 to .91.

Coping Self-Efficacy Scale

CSES [28] is a 26-item scale that assesses perceived self-efficacy for coping with threats and challenges. Items include "Keep from feeling sad" and "Resist the impulse to act hastily when under pressure." An 11-point scale was used (0="Cannot do at all"; 5="Moderately certain can do"; 10="Certain can do"), with higher scores reflecting a stronger belief in one's coping abilities. In this study, CSES possessed an excellent internal consistency of Cronbach α =.90.

Psychological Mindfulness Scale

PMS [37] measures an individual's ability to reflect on psychological processes, emotional processing, and interpersonal relationships. The PMS is a 45-item self-report instrument consisting of items such as "I often find myself thinking about what made me act in a certain way" and "I am sensitive to the changes in my own feelings." It uses a 4-point scale to score items (1="Strongly disagree"; 4="Strongly agree"). Higher total scores indicated higher levels of PM. The PMS produced a good internal consistency of Cronbach α =.84 in this study.

App Engagement Scale

The AES [56] assesses the degree of app engagement, which was included to assess whether both groups were equally engaged in the SGIs. This 7-item scale uses a 5-point Likert scale (ranging from 1="Strongly disagree" to 5="Strongly agree"). The scale had good internal reliability in this study, with a Cronbach α =.85. App engagement was predictive of positive outcomes on measures of mood and anxiety [57].

Statistical Analyses

Data Screening

Incomplete responses were excluded from the study. Following the multiple hurdles approach by Curran [58], the most likely

invalid data were sequentially identified in 2 steps and removed. First, submissions with an overall response time of <800 seconds were flagged, as it indicated that participants sped through the questionnaire and downloading the app, suggesting that they may not have participated properly. The subsequent checks confirmed whether the flagged responses were invalid. Next, data with strings of identical responses (eg, selecting "agree" to all items) for the entire scale of any of the self-report measures were excluded. Subsequently, several attention checks were used in the self-report scales [58]. For example, in the PSM-9, items 1 and 6 were reverse-coded. A contradiction was detected in the participants' self-report if responses to these 2 items were congruent with the other items on the scale [51]. Similarly, in PMS, responses to items 5 and 23 should be similar to each other but opposite to that of item 35 [37]. Contradictory responses were considered invalid and were subsequently removed. Importantly, removing these responses did not change any outcomes significantly. All statistical analyses were performed using IBM SPSS Statistics (version 25.0, IBM Corp). Data were first visually inspected using scatter plots and histograms to examine the distribution of the data and identify significant outliers. Visual inspection of these data revealed a normal distribution within an acceptable range of skewness and kurtosis (all values between -1 and 1), in accordance with the guidelines of Kline [59]. In addition, 5 outliers (ie, participants who reported data further than 2.5 SDs from the mean) were removed before the analyses [60,61]. Independent 2-tailed *t* tests and chi-square tests were used to examine any baseline differences of all demographic and dependent variables between the intervention and control groups. Should group differences of these variables emerge, they would be controlled for in the statistical comparisons.

Main Analyses

Analysis of covariance (ANCOVA) examined changes at postintervention and follow-up periods between the intervention and active control groups. ANCOVA is the recommended analysis for inferential testing of intervention effects [57], as it controls for baseline scores to ensure that group differences are due to intervention effects [62]. The α level was set at $P<.05$. Partial eta-squared (η^2) was the effect size reported for ANCOVA, whereas eta-squared (η^2) was the effect size reported for 2-tailed *t* tests and ANOVA. Guidelines by Cohen [63] for eta-squared were used, whereby 0.01 to 0.05 indicates small effect, 0.06 to 0.13 indicates moderate effect, and ≥ 0.14 indicates large effect.

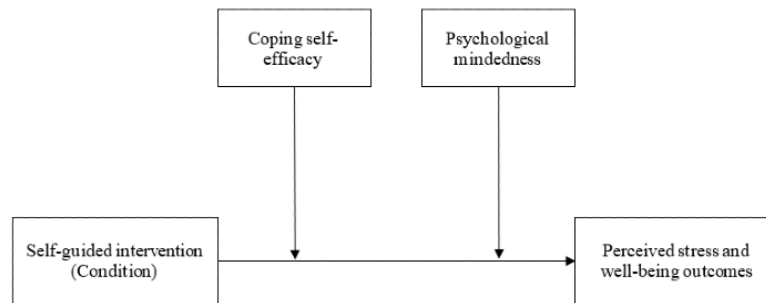
Moderation Analyses

Finally, double moderation analyses were conducted using Hayes PROCESS (version 4.0, IBM Corp, macro Model 2; Figure 2) [64]. A total of 6 separate models were conducted. The first 3 models used the primary (ie, PSM-9) and secondary outcome measures (ie, GAD-7 and PHQ-9) as dependent variables after the intervention. The remaining 3 models used the same outcome measures as dependent variables at follow-up. Each model ran a multiple linear regression double moderation analysis to examine the moderating effects of CSE and PM on the relationship between condition (intervention and control) and PSM-9 or PHQ-9 or GAD-7 at postintervention or follow-up

period. All analyses were conducted with the respective outcome measures before the intervention as covariates. Each moderator level was determined by the SD value of 1 from the mean. CIs

were set at 95%, with 5000 bootstrap iterations to assume normality [65].

Figure 2. Hypothesized relationships with coping self-efficacy and psychological mindedness as independent moderators of the direct effect of self-guided intervention on mental health outcomes.



Ethics Approval

Ethics approval to conduct the study was obtained from the Institutional Review Board of the National University of Singapore (NUS-IRB-2021-85). The study was preregistered with ClinicalTrials.gov (NCT04978896). All the participants provided electronic consent before participating in the study. Data collection took place in an entirely web-based setting, and all user data were deidentified before any analyses.

Results

Participants

The descriptive statistics for both groups are presented in Table 3. Independent 2-tailed *t* tests and chi-square tests did not find

any significant differences in age, sex, and baseline scores for PSM-9, PHQ-9, GAD-7, CSES, and PMS (all $P > .05$) between groups. Participants in both conditions did not show any significant differences in their degree of engagement with their SGIs ($P > .05$) as measured with the AES. The mean scores for the whole sample's characteristics were as follows (PSM-9: mean 37.78, SD 9.58; PHQ-9: mean 7.07, SD 5.27; GAD-7: mean 6.60, SD 4.57; CSES: mean 142.9, SD 42.2; and PMS: mean 126.2, SD 10.5), indicating that the average student was mildly distressed, depressed, and anxious.

Table 3. Descriptive statistics for demographics and outcome variables by condition (n=264).

Variables	Intervention condition (n=135)	Active control condition (n=129)	<i>P</i> value
Age (years), mean (SD)	22.88 (6.26)	22.00 (4.34)	.19
Sex, n (%)			.50
Female	94 (70.1)	94 (72.9)	
Male	38 (27.7)	28 (21.7)	
Nonbinary	3 (2.2)	4 (3.1)	
Others	0 (0)	3 (2.3)	
PSM-9 ^a , mean (SD)	38.5 (9.72)	37.0 (9.40)	.19
GAD-7 ^b , mean (SD)	6.81 (4.62)	6.37 (4.52)	.43
PHQ-9 ^c , mean (SD)	7.35 (5.11)	6.78 (5.44)	.38
CSES ^d , mean (SD)	139.5 (45.1)	146.5 (38.9)	.18
PMS ^e , mean (SD)	125.6 (10.4)	126.9 (10.7)	.31
AES ^f , mean (SD)	27.5 (3.30)	27.2 (3.68)	.56

^aPSM-9: Psychological Stress Measure-9.

^bGAD-7: Generalized Anxiety Disorder-7.

^cPHQ-9: Patient Health Questionnaire-9.

^dCSES: Coping Self-Efficacy Scale.

^ePMS: Psychological Mindedness Scale.

^fAES: App Engagement Scale (collected after intervention).

Outcome Evaluations

The mean values of the outcome measures are listed in Table 4. The intervention group reported significantly lower perceived stress levels at postintervention and follow-up periods compared with the active control group. The small effect sizes were between 0.015 and 0.019. This indicated that the stress-coping SGI was moderately more effective than the cooperation SGI (control) in reducing perceived stress over time. The intervention group also reported significantly lower anxiety levels than the active control group at postintervention period but not at follow-up. No significant differences were observed between

the intervention and active control groups for depressive symptoms at the 2 time points.

For the ANCOVA results (controlling for all baseline measurements) at postintervention period, the time \times group interaction exhibited significant differences for PSM-9 ($F_{1,261}=7.16$; $P=.008$) and GAD-7 ($F_{1,261}=5.86$; $P=.02$) but not for PHQ-9 ($F_{1,261}=1.65$; $P=.20$). Only the interaction effect for PSM-9 ($F_{1,261}=6.08$; $P=.01$) remained significant at follow-up (GAD-7: $F_{1,261}=.796$; $P=.37$; and PHQ-9: $F_{1,261}=3.24$; $P=.07$). Therefore, the main effects on psychological stress scores indicated a significant intervention effect over time.

Table 4. Means (SDs), univariate F values, and effect sizes (ESs) for outcome variables at postintervention and 1-month follow-up.

Variable	Baseline		Postintervention		F test (df)	P value	ES ^a	Follow-up		F test (df)	P value	ES
	Intervention, mean (SD)	Control, mean (SD)	Intervention, mean (SD)	Control, mean (SD)				Intervention, mean (SD)	Control, mean (SD)			
PSM-9 ^b	38.5 (9.72)	37.0 (9.40)	32.5 (9.41)	33.7 (9.16)	4.82 (1,261)	.03 ^c	.018	33.7 (8.89)	34.8 (9.70)	3.89 (1,261)	.05	.015
GAD-7 ^d	6.81 (4.62)	6.37 (4.52)	5.47 (3.94)	5.95 (4.47)	4.99 (1,261)	.03	.019	5.92 (4.33)	5.90 (4.22)	0.094 (1,261)	.76	.000
PHQ-9 ^e	7.35 (5.11)	6.78 (5.44)	5.76 (4.29)	5.78 (4.36)	0.613 (1,261)	.43	.002	5.73 (4.23)	6.06 (4.54)	1.80 (1,261)	.18	.007

^aESs of 0.01=small, 0.06=moderate, 0.14=large [63].

^bPSM-9: Psychological Stress Measure-9.

^cItalicized values indicate significant P values at .05.

^dGAD-7: Generalized Anxiety Disorder-7.

^ePHQ-9: Patient Health Questionnaire-9.

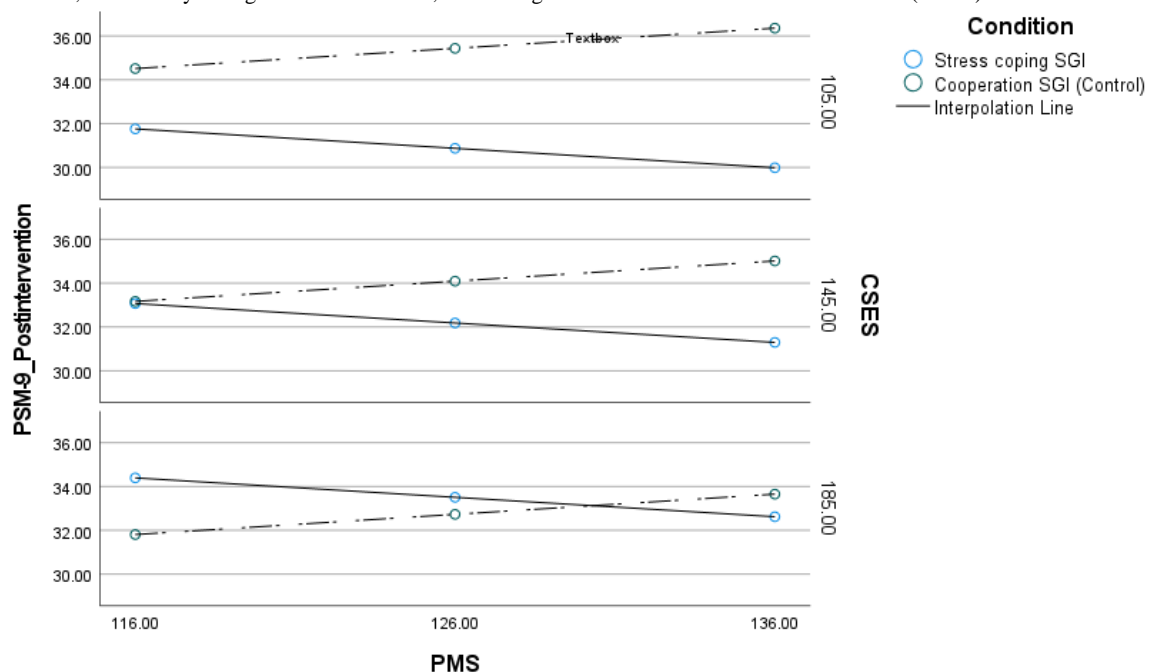
At the postintervention period, the results revealed a significant effect in the overall moderation model with PSM-9 as the dependent variable and CSES and PMS as moderators ($R^2=0.39$; $F_{9,253}=18.3$; $P<.001$). Both interaction terms were significant at postintervention period (CSES \times condition: $b=-0.0655$, $SE=0.0231$; $t_{263}=-2.83$; $P=.005$; 95% CI -0.111 to -0.020 and PMS \times condition: $b=0.181$, $SE=0.0917$; $t_{263}=1.97$; $P=.05$; 95% CI 0.0002 - 0.362), indicating that these specific conditions must be met for the intervention condition to predict lower psychological stress [66]. The stress-coping SGI significantly predicted lower scores on the PSM-9 at postintervention period in participants who (1) experienced lower CSE and (2) had moderate to high baseline PM (low CSE and moderate PM: $b=4.56$, $SE=1.30$; $t_{263}=3.52$; $P<.001$; 95% CI 2.01 - 7.11 and low CSE and high PM: $b=6.37$, $SE=1.72$; $t_{263}=3.71$; $P<.001$; 95% CI 2.98 - 9.75). Similarly, the intervention condition significantly predicted lower psychological stress at postintervention period

for participants with (1) moderate CSE and (2) moderate to high baseline PM (moderate CSE and moderate PM: $b=1.91$, $SE=0.920$; $t_{263}=2.07$; $P=.04$; 95% CI 0.0953 - 3.72 and moderate CSE and high PM: $b=3.72$, $SE=1.27$; $t_{263}=2.92$; $P=.004$; 95% CI 1.21 - 6.22). The results are presented in Table 5. Finally, regardless of the level of PMS, participants who began the stress-coping SGI with high SCE did not experience significantly lower psychological stress at postintervention period. These moderation effects were not observed at follow-up, even as the stress-coping SGI group reported significantly lower perceived stress levels at postintervention period (CSE \times condition: $b=-0.0415$, $SE=0.0241$; $t_{263}=-1.72$; $P=.09$; 95% CI -0.0891 to 0.0060 and PMS \times condition: $b=-0.0071$, $SE=0.0958$; $t_{263}=-.0740$; $P=.94$; 95% CI -0.196 to 0.182). A visualization of these interactions is shown in Figure 3.

Table 5. Intervention condition predicts lower psychological stress at each level of the moderators.

Coping self-efficacy	Psychological mindedness	<i>t</i> test (<i>df</i>)	β (SE; 95% CI)	<i>P</i> value
Low (-1 SD)	Low (-1 SD)	1.91 (262)	2.75 (1.44; -0.090 to 5.60)	.06
Low (-1 SD)	Moderate	3.52 (262)	4.56 (1.30; 2.01 to 7.11)	<.001 ^a
Low (-1 SD)	High (+1 SD)	3.71 (262)	6.37 (1.72; 2.98 to 9.75)	<.001
Moderate	Low (-1 SD)	0.075 (262)	.099 (1.32; -2.51 to 2.71)	.94
Moderate	Moderate	2.07 (262)	1.91 (0.920; 0.095 to 3.72)	.04
Moderate	High (+1 SD)	2.92 (262)	3.72 (1.27; 1.21 to 6.22)	.004
High (+1 SD)	Low (-1 SD)	-1.44 (262)	-2.59 (1.79; -6.11 to 0.940)	.15
High (+1 SD)	Moderate	-0.581 (262)	-0.778 (1.34; -3.41 to 1.86)	.56
High (+1 SD)	High (+1 SD)	0.719 (262)	1.03 (1.43; -1.79 to 3.86)	.47

^aItalicized values indicate significance.

Figure 3. Interactions between SGIs and moderators on psychological stress at postintervention. CSES: Coping Self-Efficacy Scale; PMS: Psychological Mindedness Scale; PSM-9: Psychological Stress Measure; SGI: self-guided intervention. Mean of PMS is 126 (SD 10). Mean of CSES is 145 (SD 40).

Moderation analyses for the secondary measures were not significant. For postintervention and follow-up periods, the interaction terms for both CSE and PM were not significant for both anxiety and depressive outcomes. This means that even as participants in the stress-coping intervention group experienced significantly lower anxiety levels than active control participants at postintervention period, no subgroup of individuals benefited more than the others. Similarly, as there were no significant main effects on anxiety and depressive symptoms at follow-up; there was no moderation effect.

Discussion

Principal Findings

Overview

This study aimed to evaluate the possible efficacy of an 8-day stress-coping SGI compared with an active control in improving

perceived stress and well-being levels in a sample of Asian university students. We also examined PM and CSE as factors that may identify people for whom the SGI worked best. Our hypotheses were largely supported, with the intervention group showing significantly greater improvement in perceived stress levels at postintervention and 1-month follow-up periods. In particular, students with low to medium CSE or medium to high PM benefited from significantly lower perceived stress levels after the intervention than other participants. These moderation effects were no longer observed at follow-up, indicating that these subgroups reaped the benefits of the stress-coping SGI faster than others. There were no additional benefits for participants who began the stress SGI with high CSE. In addition, the intervention group reported significantly lower anxiety levels than the active control group after the intervention without any moderation. The SGIs were also perceived to be satisfactory by both groups of participants. Participants were equally likely to continue using the *Intellect* SGI.

Psychological Stress

Our results extend the previous dearth of research that a CBT-based stress-coping intervention delivered on a mobile-based platform can be effective in reducing psychological stress in Asian university students. Previous studies that have compared CBT-based mobile interventions with waitlist controls have shown similar results. Self-guided CBT programs such as Calm [67], SMART-OP [68], and BioBase [69] previously helped decrease perceived stress in Western college students using an RCT design. Calm, consumer-based mindfulness meditation app that incorporates various CBT techniques, enhanced concentration and present-moment awareness in students using daily mindfulness meditations. This heightened awareness and focus, integrated with an 8-week cognitive training course, and facilitated the development of balanced thinking in the face of stressors [67]. Likewise, student users of the 6-week SMART-OP intervention had access to a variety of cognitive behavioral content, including stress management psychoeducation and cognitive restructuring exercises. These were further complemented with relaxation skills (ie, focused breathing, guided muscle relaxation, and biofeedback challenge) [68]. BioBase delivered a 4-week course with elements of CBT and self-compassion. The course focused on recognizing stressors and increasing one's perception of control to address them [69]. Comparably, *Intellect's* stress intervention comprises very similar CBT-based content. In addition, the stress-coping SGI within *Intellect* demonstrated the efficacy of such content in reducing perceived stress levels despite reduced exposure (ie, 5 minutes per day for 8 days). Our small effect sizes were also consistent with the effect sizes relative to active controls ($g=0.09$) presented in the meta-analyses of Linardon et al [20]. To address the lack of follow-up assessments in mental health mobile interventions [20,25], our study found that reduced perceived stress levels were sustained at the 1-month follow-up. This was in line with the few stress SGI studies that also conducted follow-up assessments for up to 6 weeks [69,70].

Anxiety and Depression

Our second hypothesis was partially supported. The stress-coping SGI was effective in reducing anxiety symptomatology among college students only at postintervention. Similarly, a multitude of stress-coping mobile-based interventions have shown reductions in self-reported anxiety [44,71-73] with similar effect sizes (Hedges $g=0.07-0.29$). However, the length of these interventions was longer, ranging minimally from 2 to 11 weeks [20]. These smartphone interventions used either mindfulness-based practices only [71-73] or both mindfulness and CBT [44]. Regardless of their theoretical orientations, these interventions used relaxation techniques, such as deep breathing and mindful awareness, that are also present in the *Intellect* stress-coping SGI. These methods previously provided short-term relief from anxiety symptoms [74,75]. Despite its brief content, the *Intellect* stress-coping SGI reduced anxiety among university students, possibly because of the use of similar relaxation techniques. In contrast, the stress-coping SGI did not reduce the depressive symptoms. This finding contradicted the few CBT-based smartphone interventions reviewed by Linardon

et al [20,76-78]. There are several possible reasons for the null findings. First, in comparison with our stress-coping SGI, the course content within these CBT-based interventions ranged from 2 weeks [77] to 2 months [78]. Hence, it is plausible that their content was more exhaustive and also more specifically tailored to address the etiology and maintenance factors of depression [78]. The stress-coping SGI was specifically designed to provide relief from perceived stress symptoms and did not incorporate more diverse content addressing depressive symptoms. Consequently, it is plausible that its efficacy against depressive symptoms would be limited. Second, the efficacies of these interventions were demonstrated against waitlist control [77] or across multiple time points and without a control group [76]. Research has shown that treatment effects of CBT interventions on depressive symptoms were significantly more difficult to detect when there is an active control group [79]. It is conceivable that the use of an active control may have obscured the intervention effect in our study, given that participants in both the intervention group and active control group were significantly less depressed after the intervention (intervention: mean 5.74, SD 4.28; $t_{134}=-4.85$; $P<.001$; and active control: mean 5.78, SD 4.47; $t_{128}=-3.07$; $P<.01$) relative to baseline (intervention: mean 7.35, SD 5.11; and active control: mean 6.78, SD 5.44). Considering that poor social dynamics was a potential risk factor for depression in college students [80], it is plausible that the active control group improved group dynamics, which may have lowered depression levels. This has been observed in previous studies [81]. However, this conclusion cannot be drawn, as we did not evaluate the efficacy of the cooperation SGI against a waitlist control group. Future studies may shed further light on this hypothesis using a 3-arm RCT with a waitlist control group, a neutral active control group (eg, attention control), and the stress-coping SGI. Interestingly, a few similar interventions have also detected anxiety effects but not depression [82] or depression effects but no anxiety effects [47]. It remains possible that certain stress-focused strategies (ie, relaxation techniques) were more effective in reducing anxiety than depressive symptoms. There were no significant group differences in anxiety or depressive symptoms at the 1-month follow-up. Sharing similar characteristics as our sample, previous RCTs evaluating the efficacy of SGIs with nonclinical participants experiencing mild baseline symptoms generally found small to no effects of app use after intervention [33,83-85]. Bakker and Rickard [33] postulated that these participants may have experienced lower motivation to seriously engage with the app, thus reducing their chances of maintaining psychological benefits.

Moderators

Our third hypothesis was partially supported. CSE and PM emerged as significant moderators, such that students with the lowest CSE and highest PM faced the greatest reduction in perceived stress levels within the intervention condition at postintervention. There were no benefits for participants with low PM and high CSE. In addition, this moderation effect was no longer observed at follow-up. The finding that perceived stress levels improved relatively quicker within this group of university students was not entirely surprising. Individuals with

a greater propensity for self-reflection and insight were more likely to be aware and thus found it easier to look at their stressors from a distance [36]. Coupled with a deeper motivation to explore the cause of their stressors, these facilitated the use of self-guided CBT strategies to resolve them (eg, reappraisal) [86-88]. McCallum et al [89] also showed that higher levels of PM in psychiatric outpatients were associated with more favorable outcomes in short-term interpretive and supportive therapies. In line with the increasing interest in the nonclinical use of CBT-based self-guided strategies, we found that university students with higher PM reaped faster gains with a CBT-based stress-coping intervention. The finding that individuals with lower CSE also experienced faster gains may be counterintuitive at first. In an RCT that compared the efficacies of 3 different CBT-based SGIs, Bakker et al [47] found that CSE was significantly increased in only 2 of them (ie, MoodKit and MoodMission). Further investigation, however, showed that participants had much lower baseline CSE levels in these 2 samples (152.00 and 154.46, respectively, compared with 169.75 in MoodPrism; [56]). Bakker et al [47] did not find any significant group differences in baseline CSE levels. Beyond CBT-based SGIs, other forms of SGIs have also shown potential ceiling effects for CSE, such that a lack of improvement resulted in no mediation of better stress outcomes [90,91]. In comparison to those with high CSE, individuals with lower perceived confidence to destress successfully would naturally be more motivated and invested in seeking additional help [92]. The belief that one could still learn how to cope with stressors of emerging adulthood was also previously suggested to enhance the effectiveness of stress management interventions [93]. Our results suggest that this may be the case. Future studies may further evaluate CSE and PM as factors to optimize treatment delivery for a nonclinical college student population. The finding that CSE and PM did not moderate the efficacy of the stress-coping SGI on secondary outcomes meant that our final hypothesis was not supported. These results were mostly unsurprising, as there was no main effect of the intervention condition on depression at postintervention and follow-up periods and anxiety at follow-up. Other studies have shown that different forms of coping could be distinctly tailored to different mental health outcomes [94]. For example, lower perceived self-efficacy to problem-solving (problem-focused CSE) has been shown to incline one's need to find ways to destress [92]. Instead, lower perceived self-efficacy to regulate one's own emotions to the problem (emotion-focused CSE) may then incline one's own need to seek emotional help and reduce anxiety. A probable reason that perceived stress levels were moderated, but not anxiety and depression, is likely due to the heavy emphasis on problem-solving internal and external stressors throughout the stress-coping SGI. Individuals with a greater capacity for self-reflection may have mostly reviewed their thought processes and behaviors with regard to the sources of their stressors and adapted with the aim of resolving these root causes. This is similar to those who have lower CSE, as they may channel most of their energy toward equipping new skills to manage stress. As the stress-coping SGI in *Intellect* lacks sufficient emphasis on building emotional resources, we encourage researchers to further evaluate the role of these moderators using alternative CBT-based SGIs.

Strengths and Limitations

The randomized controlled design allowed for causal conclusions [95]. The active waitlist control group also controlled for the effects of attention and other nonspecific factors from the intervention effects, thus strengthening the validity of the results [96]. Finally, significant effects on perceived stress levels at the 1-month follow-up period showed that the stress-coping SGI successfully improved university students' well-being over a short period. These directly addressed the lack of follow-up measurements as stated in the most recent meta-analyses of mental health SGIs [20,25].

This study had several limitations. First, the actual amount of effort put in to complete the SGI was not controlled for. Data on the duration that participants spent thinking about or practicing the skills learned from the SGI beyond the intervention were not collected. It is plausible that the positive effects could also be explained by other variables that were not assessed during the intervention, such as better campus life or a less stressful curriculum. Second, the actual duration spent on the *Intellect* app was not collected or controlled for. This is a potentially serious concern, as low adherence and inconsistent engagement with SGIs were previously found to limit their positive effects [97]. However, efforts were made to ensure that the participants completed the app activities properly. Initially, the participants were instructed to complete every daily activity present in their respective SGIs during the intervention period. *Intellect* SGIs also did not allow the participants to proceed to the next page until they completed the preceding activities. We encouraged maximum adherence through daily text reminders. Importantly, every final participant was technically verified as having completed the activities in their SGI. We acknowledged that it is possible that the participants in the active control group completed the activities in less time and with less care than those in the intervention group, which would threaten the internal validity of our findings. However, both groups did not differ on the subjective AES, which gives some confidence that groups did not differ in their motivation to engage with the SGIs and the time they spent on them [98]. Nonetheless, future studies should control for the time spent on the SGI and evaluate its relationship with SGI outcomes. Third, subjective self-reports are prone to retrospective recall biases [99]. This limits the accuracy of our findings on perceived stress, anxiety, and depressive symptoms given that they vary from time to time. Therefore, our results should be interpreted with caution. Given that the PSM-9 asked for self-reported psychological stress over the last 4 to 5 days, we are less able to infer whether the stress-coping SGI could ecologically reduce perceived stress among university students without the implementation of momentary assessments. To further strengthen the reliability of these findings, future studies can use daily diary methods to evaluate the real-time efficacy of mobile interventions on university students' well-being [100]. In place of self-reports, recent studies have recommended behavioral or physiological measures to objectively assess stress symptoms [101]. This method would also reduce the risk of recall bias and enhance the validity of our findings. Fourth, the study design only compared the intervention group with an active control group. The absence of a waitlist control condition prevented the

exclusion of the possibility of third variables contributing to the findings, which can be addressed in future studies by including a waitlist control alongside an active control. Fifth, the absence of long-term follow-up assessments (eg, 3 or 6 months) makes it difficult to deduce whether improvements can be sustained over time. Therefore, our study could not evaluate the possible long-term benefits on symptoms of chronic stress. Finally, this study has limited external validity for the student population at large, as our sample contained mainly female students from Singapore with slightly elevated stress, anxiety, and depression levels. Therefore, our positive findings on perceived stress and anxiety symptoms may be generalized to the Western student population or clinical samples. Future researchers may consider replicating our findings with a more diverse college sample while administering even longer-term follow-up assessments.

Conclusions

In conclusion, this RCT found evidence for an 8-day stress-coping SGI in improving perceived stress and anxiety levels among Asian university students. The effects on perceived stress levels were sustained at 1-month follow-up, but not for anxiety, thus giving some confidence that a brief, time-limited, and CBT-based SGI can maintain its gains on perceived stress. However, depressive symptoms did not decrease. Students with lower CSE and higher PM experienced reduced stress faster than other students in the CBT-based SGI intervention group. The identification of these moderators can optimize the outcome and treatment delivery of such stress-coping mobile apps. Our findings are useful, given the potential for scaling up such easily accessible, and brief interventions. Several limitations were noted, including the lack of ecological momentary assessments to capture perceived stress more accurately, longer-term follow-up measures to evaluate sustainability of gains, and more diverse samples to evaluate transferability of the findings.

Acknowledgments

This study was partly funded by Intellect Pte Ltd under a Research Collaboration Agreement with the National University of Singapore.

Data Availability

Data supporting the findings of this study are available upon request from the corresponding author.

Conflicts of Interest

OS set up the Research Collaboration Agreement between the National University of Singapore and Intellect in 2020, and in January 2022 joined Intellect Pte Ltd as their Clinical Director. The study design, data management, statistical analysis, interpretation of the data, and reporting of the study are independent of Intellect Pte Ltd.

Multimedia Appendix 1

CONSORT-EHEALTH Checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1243 KB - formative_v6i12e40723_app1.pdf](#)]

References

1. Cohen S, Kessler RC, Gordon LU. *Measuring Stress: A Guide for Health and Social Scientists*. Oxford, UK: Oxford University Press; 1997.
2. Leka S, Griffiths A, Cox T. *Work organisation and stress : systematic problem approaches for employers, managers and trade union representatives*. World Health Organization. 2003 May 23. URL: <https://www.who.int/publications/i/item/9241590475> [accessed 2021-09-30]
3. van der Valk ES, Savas M, van Rossum EF. Stress and obesity: are there more susceptible individuals? *Curr Obes Rep* 2018 Jun;7(2):193-203 [FREE Full text] [doi: [10.1007/s13679-018-0306-y](https://doi.org/10.1007/s13679-018-0306-y)] [Medline: [29663153](https://pubmed.ncbi.nlm.nih.gov/29663153/)]
4. Lupien SJ, Gillin CJ, Hauger RL. Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose-response study in humans. *Behav Neurosci* 1999 Jun;113(3):420-430. [doi: [10.1037//0735-7044.113.3.420](https://doi.org/10.1037//0735-7044.113.3.420)] [Medline: [10443770](https://pubmed.ncbi.nlm.nih.gov/10443770/)]
5. Hsu FC, Garside MJ, Massey AE, McAllister-Williams RH. Effects of a single dose of cortisol on the neural correlates of episodic memory and error processing in healthy volunteers. *Psychopharmacology (Berl)* 2003 Jun;167(4):431-442. [doi: [10.1007/s00213-003-1413-2](https://doi.org/10.1007/s00213-003-1413-2)] [Medline: [12684731](https://pubmed.ncbi.nlm.nih.gov/12684731/)]
6. Gianaros PJ, Wager TD. Brain-body pathways linking psychological stress and physical health. *Curr Dir Psychol Sci* 2015 Aug 01;24(4):313-321 [FREE Full text] [doi: [10.1177/0963721415581476](https://doi.org/10.1177/0963721415581476)] [Medline: [26279608](https://pubmed.ncbi.nlm.nih.gov/26279608/)]
7. Parrish BP, Cohen LH, Laurenceau JP. Prospective relationship between negative affective reactivity to daily stress and depressive symptoms. *J Soc Clin Psychol* 2011 Mar;30(3):270-296. [doi: [10.1521/jscp.2011.30.3.270](https://doi.org/10.1521/jscp.2011.30.3.270)]
8. Morin CM, Rodrigue S, Ivers H. Role of stress, arousal, and coping skills in primary insomnia. *Psychosom Med* 2003;65(2):259-267. [doi: [10.1097/01.psy.0000030391.09558.a3](https://doi.org/10.1097/01.psy.0000030391.09558.a3)] [Medline: [12651993](https://pubmed.ncbi.nlm.nih.gov/12651993/)]

9. Amanvermez Y, Rahmadiana M, Karyotaki E, de Wit L, Ebert DD, Kessler RC, et al. Stress management interventions for college students: a systematic review and meta-analysis. *Clin Psychol Sci Pract* 2020 May 22:e12342. [doi: [10.1111/cpsp.12342](https://doi.org/10.1111/cpsp.12342)]
10. Van Daele T, Hermans D, Van Audenhove C, Van den Bergh O. Stress reduction through psychoeducation: a meta-analytic review. *Health Educ Behav* 2012 Aug;39(4):474-485. [doi: [10.1177/1090198111419202](https://doi.org/10.1177/1090198111419202)] [Medline: [21986242](https://pubmed.ncbi.nlm.nih.gov/21986242/)]
11. Bewick B, Koutsopoulou G, Miles J, Slaa E, Barkham M. Changes in undergraduate students' psychological well-being as they progress through university. *Stud High Educ* 2010 Aug 18;35(6):633-645. [doi: [10.1080/03075070903216643](https://doi.org/10.1080/03075070903216643)]
12. Jerusalem M, Mittag W. Self-efficacy in stressful life transitions. In: Bandura A, editor. *Self-Efficacy in Changing Societies*. Cambridge, UK: Cambridge University Press; 1995:177-201.
13. Millennials, Gen Z and mental health: Managing mental health in the workplace. Deloitte. 2020 Jun. URL: <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/About-Deloitte/gx-millennial-survey-mental-health-whitepaper.pdf> [accessed 2021-11-20]
14. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: rapid meta-review. *J Med Internet Res* 2021 Mar 10;23(3):e23365 [FREE Full text] [doi: [10.2196/23365](https://doi.org/10.2196/23365)] [Medline: [33606657](https://pubmed.ncbi.nlm.nih.gov/33606657/)]
15. Sood M, Chadda RK, Singh P. Mobile health (mHealth) in mental health: scope and applications in low-resource settings. *Natl Med J India* 2016;29(6):341-343 [FREE Full text] [Medline: [28327483](https://pubmed.ncbi.nlm.nih.gov/28327483/)]
16. Min JA, Lee CU, Lee C. Mental health promotion and illness prevention: a challenge for psychiatrists. *Psychiatry Investig* 2013 Dec;10(4):307-316 [FREE Full text] [doi: [10.4306/pi.2013.10.4.307](https://doi.org/10.4306/pi.2013.10.4.307)] [Medline: [24474978](https://pubmed.ncbi.nlm.nih.gov/24474978/)]
17. Reid SC, Kauer SD, Hears SJ, Crooke AH, Khor AS, Sancu LA, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomised controlled trial. *BMC Fam Pract* 2011 Nov 29;12:131 [FREE Full text] [doi: [10.1186/1471-2296-12-131](https://doi.org/10.1186/1471-2296-12-131)] [Medline: [22123031](https://pubmed.ncbi.nlm.nih.gov/22123031/)]
18. 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](https://doi.org/10.1016/j.jad.2017.04.046)] [Medline: [28456072](https://pubmed.ncbi.nlm.nih.gov/28456072/)]
19. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017 Oct;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
20. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 2019 Oct;18(3):325-336 [FREE Full text] [doi: [10.1002/wps.20673](https://doi.org/10.1002/wps.20673)] [Medline: [31496095](https://pubmed.ncbi.nlm.nih.gov/31496095/)]
21. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS One* 2016 May 2;11(5):e0154248 [FREE Full text] [doi: [10.1371/journal.pone.0154248](https://doi.org/10.1371/journal.pone.0154248)] [Medline: [27135410](https://pubmed.ncbi.nlm.nih.gov/27135410/)]
22. Christmann CA, Hoffmann A, Bleser G. Stress management apps with regard to emotion-focused coping and behavior change techniques: a content analysis. *JMIR Mhealth Uhealth* 2017 Feb 23;5(2):e22 [FREE Full text] [doi: [10.2196/mhealth.6471](https://doi.org/10.2196/mhealth.6471)] [Medline: [28232299](https://pubmed.ncbi.nlm.nih.gov/28232299/)]
23. Coulon SM, Monroe CM, West DS. A systematic, multi-domain review of mobile smartphone apps for evidence-based stress management. *Am J Prev Med* 2016 Jul;51(1):95-105. [doi: [10.1016/j.amepre.2016.01.026](https://doi.org/10.1016/j.amepre.2016.01.026)] [Medline: [26993534](https://pubmed.ncbi.nlm.nih.gov/26993534/)]
24. Van Stee S, Yang Q. The effectiveness and moderators of mobile applications for health behavior change. In: Kim J, Song H, editors. *Technology and Health: Promoting Attitude and Behavior Change*. Cambridge, MA, USA: Academic Press; 2020:243-270.
25. Heber E, Ebert D, Lehr D, Cuijpers P, Berking M, Nobs S, et al. The benefit of web- and computer-based interventions for stress: a systematic review and meta-analysis. *J Med Internet Res* 2017 Feb 17;19(2):e32 [FREE Full text] [doi: [10.2196/jmir.5774](https://doi.org/10.2196/jmir.5774)] [Medline: [28213341](https://pubmed.ncbi.nlm.nih.gov/28213341/)]
26. Coudray C, Palmer R, Frazier P. Moderators of the efficacy of a web-based stress management intervention for college students. *J Couns Psychol* 2019 Nov;66(6):747-754. [doi: [10.1037/cou0000340](https://doi.org/10.1037/cou0000340)] [Medline: [30998054](https://pubmed.ncbi.nlm.nih.gov/30998054/)]
27. Thorne KJ, Andrews JJ, Nordstokke D. Relations among children's coping strategies and anxiety: the mediating role of coping efficacy. *J Gen Psychol* 2013;140(3):204-223. [doi: [10.1080/00221309.2013.792235](https://doi.org/10.1080/00221309.2013.792235)] [Medline: [24837655](https://pubmed.ncbi.nlm.nih.gov/24837655/)]
28. Chesney MA, Neilands TB, Chambers DB, Taylor JM, Folkman S. A validity and reliability study of the coping self-efficacy scale. *Br J Health Psychol* 2006 Sep;11(Pt 3):421-437 [FREE Full text] [doi: [10.1348/135910705X53155](https://doi.org/10.1348/135910705X53155)] [Medline: [16870053](https://pubmed.ncbi.nlm.nih.gov/16870053/)]
29. Mennin DS, Ellard KK, Fresco DM, Gross JJ. United we stand: emphasizing commonalities across cognitive-behavioral therapies. *Behav Ther* 2013 Jun;44(2):234-248 [FREE Full text] [doi: [10.1016/j.beth.2013.02.004](https://doi.org/10.1016/j.beth.2013.02.004)] [Medline: [23611074](https://pubmed.ncbi.nlm.nih.gov/23611074/)]
30. Benight CC, Harper ML. Coping self-efficacy perceptions as a mediator between acute stress response and long-term distress following natural disasters. *J Trauma Stress* 2002 Jun;15(3):177-186. [doi: [10.1023/A:1015295025950](https://doi.org/10.1023/A:1015295025950)] [Medline: [12092909](https://pubmed.ncbi.nlm.nih.gov/12092909/)]
31. Litt MD. Self-efficacy and perceived control: cognitive mediators of pain tolerance. *J Pers Soc Psychol* 1988 Jan;54(1):149-160. [doi: [10.1037//0022-3514.54.1.149](https://doi.org/10.1037//0022-3514.54.1.149)] [Medline: [3346804](https://pubmed.ncbi.nlm.nih.gov/3346804/)]

32. Bush NE, Smolenski DJ, Denneson LM, Williams HB, Thomas EK, Dobscha SK. A virtual hope box: randomized controlled trial of a smartphone app for emotional regulation and coping with distress. *Psychiatr Serv* 2017 Apr 01;68(4):330-336. [doi: [10.1176/appi.ps.201600283](https://doi.org/10.1176/appi.ps.201600283)] [Medline: [27842473](https://pubmed.ncbi.nlm.nih.gov/27842473/)]
33. Bakker D, Rickard N. Engagement with a cognitive behavioural therapy mobile phone app predicts changes in mental health and wellbeing: MoodMission. *Aust Psychol* 2019 Aug;54(4):245-260. [doi: [10.1111/ap.12383](https://doi.org/10.1111/ap.12383)]
34. Robbins LB, Chang MW, Ling J, Brown R. Psychosocial factors affecting the association between a healthy lifestyle behavior intervention and depressive symptoms in low-income overweight or obese mothers with young children: a mediational analysis. *J Pediatr Perinatol Child Health* 2022;6(1):54-69 [FREE Full text] [doi: [10.26502/jppch.74050090](https://doi.org/10.26502/jppch.74050090)] [Medline: [35098144](https://pubmed.ncbi.nlm.nih.gov/35098144/)]
35. Keshi AK, Basavarajappa. Effectiveness of cognitive behavior therapy on self-efficacy among high school students. *Asian J Manag Sci Educ* 2013 Oct;2(4):68-79.
36. Grant AM. Rethinking psychological mindedness: metacognition, self-reflection, and insight. *Behav Change* 2001;18(1):8-17. [doi: [10.1375/bech.18.1.8](https://doi.org/10.1375/bech.18.1.8)]
37. Conte HR, Plutchik R, Jung BB, Picard S, Karasu TB, Lotterman A. Psychological mindedness as a predictor of psychotherapy outcome: a preliminary report. *Compr Psychiatry* 1990;31(5):426-431. [doi: [10.1016/0010-440x\(90\)90027-p](https://doi.org/10.1016/0010-440x(90)90027-p)] [Medline: [2225801](https://pubmed.ncbi.nlm.nih.gov/2225801/)]
38. Kishon R, Geronazzo-Alman L, Westphal M, Chen Y, Green D, Sarda A, et al. Psychological mindedness and alexithymia predict symptom reduction in cognitive behavior therapy for major depressive disorder. *Int J Cogn Ther* 2020;13(4):287-302. [doi: [10.1007/s41811-019-00062-6](https://doi.org/10.1007/s41811-019-00062-6)]
39. Butler G, Fennell M, Robson P, Gelder M. Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. *J Consult Clin Psychol* 1991 Feb;59(1):167-175. [doi: [10.1037//0022-006x.59.1.167](https://doi.org/10.1037//0022-006x.59.1.167)] [Medline: [2002134](https://pubmed.ncbi.nlm.nih.gov/2002134/)]
40. Rapee RM. *Overcoming Shyness and Social Phobia: A Step-by-Step Guide*. Lanham, MD, USA: Jason Aronson; 1998.
41. Clark DM, Fairburn CG, Jones JV. The science and practice of cognitive behaviour therapy. *J Cogn Psychother* 1997 Jan 01;11(2):141-144. [doi: [10.1891/0889-8391.11.2.141](https://doi.org/10.1891/0889-8391.11.2.141)]
42. Vaidyanathan S, Menon V, Sarkar S. Identifying patient profiles suitable for cognitive behavior Therapy: the role of psychological mindedness. *J Neurosci Rural Pract* 2018;9(4):654-655 [FREE Full text] [doi: [10.4103/jnrp.jnrp_100_18](https://doi.org/10.4103/jnrp.jnrp_100_18)] [Medline: [30271073](https://pubmed.ncbi.nlm.nih.gov/30271073/)]
43. Wiles N, Thomas L, Abel A, Barnes M, Carroll F, Ridgway N, et al. Clinical effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: the CoBaIT randomised controlled trial. *Health Technol Assess* 2014 May;18(31):1-viii [FREE Full text] [doi: [10.3310/hta18310](https://doi.org/10.3310/hta18310)] [Medline: [24824481](https://pubmed.ncbi.nlm.nih.gov/24824481/)]
44. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. *BMC Psychiatry* 2013 Nov 18;13:312 [FREE Full text] [doi: [10.1186/1471-244X-13-312](https://doi.org/10.1186/1471-244X-13-312)] [Medline: [24237617](https://pubmed.ncbi.nlm.nih.gov/24237617/)]
45. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. *PLoS One* 2015 May 26;10(5):e0126559 [FREE Full text] [doi: [10.1371/journal.pone.0126559](https://doi.org/10.1371/journal.pone.0126559)] [Medline: [26010890](https://pubmed.ncbi.nlm.nih.gov/26010890/)]
46. Borrega-Mouquinho Y, Sánchez-Gómez J, Fuentes-García JP, Collado-Mateo D, Villafaina S. Effects of high-intensity interval training and moderate-intensity training on stress, depression, anxiety, and resilience in healthy adults during coronavirus disease 2019 confinement: a randomized controlled trial. *Front Psychol* 2021 Feb 24;12:643069 [FREE Full text] [doi: [10.3389/fpsyg.2021.643069](https://doi.org/10.3389/fpsyg.2021.643069)] [Medline: [33716913](https://pubmed.ncbi.nlm.nih.gov/33716913/)]
47. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behav Res Ther* 2018 Oct;109:75-83. [doi: [10.1016/j.brat.2018.08.003](https://doi.org/10.1016/j.brat.2018.08.003)] [Medline: [30125790](https://pubmed.ncbi.nlm.nih.gov/30125790/)]
48. Hochheimer CJ, Sabo RT, Krist AH, Day T, Cyrus J, Woolf SH. Methods for evaluating respondent attrition in web-based surveys. *J Med Internet Res* 2016 Nov 22;18(11):e301 [FREE Full text] [doi: [10.2196/jmir.6342](https://doi.org/10.2196/jmir.6342)] [Medline: [27876687](https://pubmed.ncbi.nlm.nih.gov/27876687/)]
49. Krosnick JA, Presser S. Question and questionnaire design. In: Wright JD, Marsden PV, editors. *Handbook of Survey Research*. 2nd edition. San Diego, CA, USA: Elsevier; 2009.
50. Lemyre L, Tessier R. Mesure de stress psychologique (MSP): se sentir stressé-e. *Can J Behav Sci* 1988;20(3):302-321. [doi: [10.1037/h0079945](https://doi.org/10.1037/h0079945)]
51. Lemyre L, Tessier R. Measuring psychological stress. Concept, model, and measurement instrument in primary care research. *Can Fam Physician* 2003 Sep;49:1159-1168 [FREE Full text] [Medline: [14526870](https://pubmed.ncbi.nlm.nih.gov/14526870/)]
52. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
53. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry* 2010;32(4):345-359. [doi: [10.1016/j.genhosppsy.2010.03.006](https://doi.org/10.1016/j.genhosppsy.2010.03.006)] [Medline: [20633738](https://pubmed.ncbi.nlm.nih.gov/20633738/)]

54. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
55. 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](https://doi.org/10.1097/MLR.0b013e318160d093)] [Medline: [18388841](https://pubmed.ncbi.nlm.nih.gov/18388841/)]
56. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J Affect Disord* 2018 Feb;227:432-442. [doi: [10.1016/j.jad.2017.11.016](https://doi.org/10.1016/j.jad.2017.11.016)] [Medline: [29154165](https://pubmed.ncbi.nlm.nih.gov/29154165/)]
57. Stice E, Shaw H. Eating disorder prevention programs: a meta-analytic review. *Psychol Bull* 2004 Mar;130(2):206-227. [doi: [10.1037/0033-2909.130.2.206](https://doi.org/10.1037/0033-2909.130.2.206)] [Medline: [14979770](https://pubmed.ncbi.nlm.nih.gov/14979770/)]
58. Curran PG. Methods for the detection of carelessly invalid responses in survey data. *J Exp Soc Psychol* 2016 Sep;66:4-19. [doi: [10.1016/j.jesp.2015.07.006](https://doi.org/10.1016/j.jesp.2015.07.006)]
59. Kline RB. Principles and practice of structural equation modeling. 4th edition. New York, NY, USA: Guilford Press; 2015.
60. Hendriks M, van Ginkel W, Dijkstra T, Piai V. Dropping beans or spilling secrets: how idiomatic context bias affects prediction. *J Cogn Neurosci* 2022 Jan 05;34(2):209-223. [doi: [10.1162/jocn_a_01798](https://doi.org/10.1162/jocn_a_01798)] [Medline: [34813643](https://pubmed.ncbi.nlm.nih.gov/34813643/)]
61. Fisher M, Mormann M. The off by 100% bias: the effects of percentage changes greater than 100% on magnitude judgments and consumer choice. *J Consum Res* 2022 Dec;49(4):561-573. [doi: [10.1093/jcr/ucac006](https://doi.org/10.1093/jcr/ucac006)]
62. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ* 2001 Nov 10;323(7321):1123-1124 [FREE Full text] [doi: [10.1136/bmj.323.7321.1123](https://doi.org/10.1136/bmj.323.7321.1123)] [Medline: [11701584](https://pubmed.ncbi.nlm.nih.gov/11701584/)]
63. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ, USA: Lawrence Erlbaum Associates; 1988.
64. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach. New York, NY, USA: Guilford Press; 2013.
65. DiCiccio TJ, Efron B. Bootstrap confidence intervals. *Statist Sci* 1996 Sep 1;11(3):189-228. [doi: [10.1214/ss/1032280214](https://doi.org/10.1214/ss/1032280214)]
66. Mantua J, Helms SM, Weymann KB, Capaldi 2nd VF, Lim MM. Sleep quality and emotion regulation interact to predict anxiety in veterans with PTSD. *Behav Neurol* 2018 Jun 5;2018:7940832 [FREE Full text] [doi: [10.1155/2018/7940832](https://doi.org/10.1155/2018/7940832)] [Medline: [29971139](https://pubmed.ncbi.nlm.nih.gov/29971139/)]
67. 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 Jun 25;7(6):e14273 [FREE Full text] [doi: [10.2196/14273](https://doi.org/10.2196/14273)] [Medline: [31237569](https://pubmed.ncbi.nlm.nih.gov/31237569/)]
68. Rose RD, Buckley Jr JC, Zbozinek TD, Motivala SJ, Glenn DE, Cartreine JA, et al. A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behav Res Ther* 2013 Feb;51(2):106-112. [doi: [10.1016/j.brat.2012.11.003](https://doi.org/10.1016/j.brat.2012.11.003)] [Medline: [23262118](https://pubmed.ncbi.nlm.nih.gov/23262118/)]
69. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: [10.2196/17767](https://doi.org/10.2196/17767)] [Medline: [31926063](https://pubmed.ncbi.nlm.nih.gov/31926063/)]
70. McCloud T, Jones R, Lewis G, Bell V, Tsakanikos E. Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. *JMIR Mhealth Uhealth* 2020 Jul 31;8(7):e15418 [FREE Full text] [doi: [10.2196/15418](https://doi.org/10.2196/15418)] [Medline: [32735221](https://pubmed.ncbi.nlm.nih.gov/32735221/)]
71. Villani D, Grassi A, Cognetta C, Toniolo D, Cipresso P, Riva G. Self-help stress management training through mobile phones: an experience with oncology nurses. *Psychol Serv* 2013 Aug;10(3):315-322. [doi: [10.1037/a0026459](https://doi.org/10.1037/a0026459)] [Medline: [23937091](https://pubmed.ncbi.nlm.nih.gov/23937091/)]
72. 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](https://doi.org/10.2196/mental.8324)] [Medline: [29362209](https://pubmed.ncbi.nlm.nih.gov/29362209/)]
73. Bostock S, Crosswell AD, Prather AA, Steptoe A. Mindfulness on-the-go: effects of a mindfulness meditation app on work stress and well-being. *J Occup Health Psychol* 2019 Feb;24(1):127-138 [FREE Full text] [doi: [10.1037/ocp0000118](https://doi.org/10.1037/ocp0000118)] [Medline: [29723001](https://pubmed.ncbi.nlm.nih.gov/29723001/)]
74. Jorm AF, Christensen H, Griffiths KM, Parslow RA, Rodgers B, Blewitt KA. Effectiveness of complementary and self-help treatments for anxiety disorders. *Med J Aust* 2004 Oct 04;181(S7):S29-S46. [doi: [10.5694/j.1326-5377.2004.tb06352.x](https://doi.org/10.5694/j.1326-5377.2004.tb06352.x)] [Medline: [15462640](https://pubmed.ncbi.nlm.nih.gov/15462640/)]
75. Wilczyńska D, Łysak-Radomska A, Podczarska-Głowacka M, Zajt J, Dornowski M, Skonieczny P. Evaluation of the effectiveness of relaxation in lowering the level of anxiety in young adults - a pilot study. *Int J Occup Med Environ Health* 2019 Nov 15;32(6):817-824 [FREE Full text] [doi: [10.13075/ijomeh.1896.01457](https://doi.org/10.13075/ijomeh.1896.01457)] [Medline: [31608900](https://pubmed.ncbi.nlm.nih.gov/31608900/)]
76. Mak WW, Tong AC, Yip SY, Lui WW, Chio FH, Chan AT, et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: randomized controlled noninferiority trial. *JMIR Ment Health* 2018 Oct 11;5(4):e60 [FREE Full text] [doi: [10.2196/mental.8597](https://doi.org/10.2196/mental.8597)] [Medline: [30309837](https://pubmed.ncbi.nlm.nih.gov/30309837/)]
77. Levin ME, Haeger J, An W, Twohig MP. Comparing cognitive defusion and cognitive restructuring delivered through a mobile app for individuals high in self-criticism. *Cogn Ther Res* 2018 Jul 13;42(6):844-855. [doi: [10.1007/s10608-018-9944-3](https://doi.org/10.1007/s10608-018-9944-3)]

78. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. *BMJ Open* 2014 Jan 09;4(1):e003440 [FREE Full text] [doi: [10.1136/bmjopen-2013-003440](https://doi.org/10.1136/bmjopen-2013-003440)] [Medline: [24413342](https://pubmed.ncbi.nlm.nih.gov/24413342/)]
79. Keles S, Idsoe T. Six- and twelve-month follow-up results of a cluster randomized controlled trial of a CBT-based group course. *Prev Sci* 2021 May;22(4):409-418. [doi: [10.1007/s11121-020-01160-0](https://doi.org/10.1007/s11121-020-01160-0)] [Medline: [32889703](https://pubmed.ncbi.nlm.nih.gov/32889703/)]
80. Thorley C. Not by degrees: Improving student mental health in the UK's universities. Institute for Public Policy Research. 2017 Sep 4. URL: <https://www.ippr.org/publications/not-by-degrees> [accessed 2022-03-07]
81. Spray J, Floyd B, Littleton J, Trnka S, Mattison S. Social group dynamics predict stress variability among children in a New Zealand classroom. *Homo* 2018 Mar;69(1-2):50-61. [doi: [10.1016/j.jchb.2018.03.005](https://doi.org/10.1016/j.jchb.2018.03.005)] [Medline: [29705502](https://pubmed.ncbi.nlm.nih.gov/29705502/)]
82. Trzcieniecka-Green A, Steptoe A. The effects of stress management on the quality of life of patients following acute myocardial infarction or coronary bypass surgery. *Eur Heart J* 1996 Nov;17(11):1663-1670. [doi: [10.1093/oxfordjournals.eurheartj.a014749](https://doi.org/10.1093/oxfordjournals.eurheartj.a014749)] [Medline: [8922914](https://pubmed.ncbi.nlm.nih.gov/8922914/)]
83. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *J Happiness Stud* 2014 Oct 29;17(1):163-185. [doi: [10.1007/s10902-014-9589-1](https://doi.org/10.1007/s10902-014-9589-1)]
84. Bidargaddi N, Musiat P, Winsall M, Vogl G, Blake V, Quinn S, et al. Efficacy of a web-based guided recommendation service for a curated list of readily available mental health and well-being mobile apps for young people: randomized controlled trial. *J Med Internet Res* 2017 May 12;19(5):e141 [FREE Full text] [doi: [10.2196/jmir.6775](https://doi.org/10.2196/jmir.6775)] [Medline: [28500020](https://pubmed.ncbi.nlm.nih.gov/28500020/)]
85. Whittaker R, Stasiak K, McDowell H, Doherty I, Shepherd M, Chua S, et al. MEMO: an mHealth intervention to prevent the onset of depression in adolescents: a double-blind, randomised, placebo-controlled trial. *J Child Psychol Psychiatry* 2017 Sep;58(9):1014-1022. [doi: [10.1111/jcpp.12753](https://doi.org/10.1111/jcpp.12753)] [Medline: [28573672](https://pubmed.ncbi.nlm.nih.gov/28573672/)]
86. 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](https://doi.org/10.1016/j.adolescence.2016.09.009)] [Medline: [27814499](https://pubmed.ncbi.nlm.nih.gov/27814499/)]
87. Hur JW, Kim B, Park D, Choi SW. A scenario-based cognitive behavioral therapy mobile app to reduce dysfunctional beliefs in individuals with depression: a randomized controlled trial. *Telemed J E Health* 2018 Sep;24(9):710-716. [doi: [10.1089/tmj.2017.0214](https://doi.org/10.1089/tmj.2017.0214)] [Medline: [29323626](https://pubmed.ncbi.nlm.nih.gov/29323626/)]
88. Marciniak MA, Shanahan L, Rohde J, Schulz A, Wackerhagen C, Kobylińska D, et al. Standalone smartphone cognitive behavioral therapy-based ecological momentary interventions to increase mental health: narrative review. *JMIR Mhealth Uhealth* 2020 Nov 12;8(11):e19836 [FREE Full text] [doi: [10.2196/19836](https://doi.org/10.2196/19836)] [Medline: [33180027](https://pubmed.ncbi.nlm.nih.gov/33180027/)]
89. McCallum M, Piper WE, Ogrodniczuk JS, Joyce AS. Relationships among psychological mindedness, alexithymia and outcome in four forms of short-term psychotherapy. *Psychol Psychother* 2003 Jun;76(Pt 2):133-144. [doi: [10.1348/147608303765951177](https://doi.org/10.1348/147608303765951177)] [Medline: [12855060](https://pubmed.ncbi.nlm.nih.gov/12855060/)]
90. Rumbold JL. The design and delivery of stress management in professional sport. Loughborough University. 2014. URL: https://repository.lboro.ac.uk/articles/thesis/The_design_and_delivery_of_stress_management_in_professional_sport/9610061 [accessed 2021-06-21]
91. Dworkin MS, Lee S, Chakraborty A, Monahan C, Hightow-Weidman L, Garofalo R, et al. Acceptability, feasibility, and preliminary efficacy of a theory-based relational embodied conversational agent mobile phone intervention to promote HIV medication adherence in young HIV-positive African American MSM. *AIDS Educ Prev* 2019 Feb;31(1):17-37. [doi: [10.1521/aeap.2019.31.1.17](https://doi.org/10.1521/aeap.2019.31.1.17)] [Medline: [30742481](https://pubmed.ncbi.nlm.nih.gov/30742481/)]
92. Liu SR, Kia-Keating M. Improving coping self-efficacy among distressed students after exposure to university mass violence: a pilot online intervention. *J Coll Stud Psychother* 2018;32(3):199-219. [doi: [10.1080/87568225.2017.1388754](https://doi.org/10.1080/87568225.2017.1388754)]
93. McGee RE, Windle M, Cooper HL, Thompson NJ. Resilience among unemployed and underemployed emerging adults: a cross-sectional study of how stress and coping relate to depressive symptoms. *Emerg Adulthood* 2020 Aug 28;9(4):372-383. [doi: [10.1177/2167696820949401](https://doi.org/10.1177/2167696820949401)]
94. Vedhara K, Shanks N, Anderson S, Lightman S. The role of stressors and psychosocial variables in the stress process: a study of chronic caregiver stress. *Psychosom Med* 2000;62(3):374-385. [doi: [10.1097/00006842-200005000-00011](https://doi.org/10.1097/00006842-200005000-00011)] [Medline: [10845351](https://pubmed.ncbi.nlm.nih.gov/10845351/)]
95. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: study design: randomised controlled trials. *BJOG* 2018 Dec;125(13):1716 [FREE Full text] [doi: [10.1111/1471-0528.15199](https://doi.org/10.1111/1471-0528.15199)] [Medline: [29916205](https://pubmed.ncbi.nlm.nih.gov/29916205/)]
96. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. *Res Nurs Health* 2018 Jun;41(3):320-328. [doi: [10.1002/nur.21870](https://doi.org/10.1002/nur.21870)] [Medline: [29906317](https://pubmed.ncbi.nlm.nih.gov/29906317/)]
97. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
98. Kruglova K, O'Connell SB, Dawadi S, Gelgoot EN, Miner SA, Robins S, et al. An mHealth app to support fertility patients navigating the world of infertility (Infotility): development and usability study. *JMIR Form Res* 2021 Oct 12;5(10):e28136 [FREE Full text] [doi: [10.2196/28136](https://doi.org/10.2196/28136)] [Medline: [34636741](https://pubmed.ncbi.nlm.nih.gov/34636741/)]

99. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
100. Iyer L, Iyer RB, Kumar V. A relaxation app (HeartBot) for stress and emotional well-being over a 21-day challenge: randomized survey study. *JMIR Form Res* 2021 Jan 29;5(1):e22041 [FREE Full text] [doi: [10.2196/22041](https://doi.org/10.2196/22041)] [Medline: [33455902](https://pubmed.ncbi.nlm.nih.gov/33455902/)]
101. Alzahem AM, Van der Molen HT, De Boer BJ. Effectiveness of a dental students stress management program. *Health Prof Educ* 2015 Dec;1(1):34-42. [doi: [10.1016/j.hpe.2015.11.003](https://doi.org/10.1016/j.hpe.2015.11.003)]

Abbreviations

AES: App Engagement Scale
ANCOVA: analysis of covariance
CBT: cognitive behavioral therapy
CSE: coping self-efficacy
CSES: Coping Self-Efficacy Scale
GAD: Generalized Anxiety Disorder
mHealth: mobile health
PHQ: Patient Health Questionnaire
PM: psychological mindedness
PMS: Psychological Mindedness Scale
PSM: Psychological Stress Measure
RCT: randomized controlled trial
SGI: self-guided intervention

Edited by A Mavragani; submitted 04.07.22; peer-reviewed by D Hasson, E Kuhn; comments to author 08.08.22; revised version received 20.08.22; accepted 24.11.22; published 14.12.22.

Please cite as:

Toh SHY, Tan JHY, Kosasih FR, Sündermann O

Efficacy of the Mental Health App Intellect to Reduce Stress: Randomized Controlled Trial With a 1-Month Follow-up

JMIR Form Res 2022;6(12):e40723

URL: <https://formative.jmir.org/2022/12/e40723>

doi: [10.2196/40723](https://doi.org/10.2196/40723)

PMID: [36515984](https://pubmed.ncbi.nlm.nih.gov/36515984/)

©Sean Han Yang Toh, Jessalin Hui Yan Tan, Feodora Roxanne Kosasih, Oliver Sündermann. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

End User Participation in the Development of an Ecological Momentary Intervention to Improve Coping With Cannabis Cravings: Formative Study

Molly A Anderson¹, PhD; Alan J Budney¹, PhD; Nicholas C Jacobson¹, PhD; Inbal Nahum-Shani², PhD; Catherine Stanger¹, PhD

¹Center for Technology and Behavioral Health, Geisel School of Medicine, Dartmouth College, Lebanon, NH, United States

²Institute for Social Research, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:

Catherine Stanger, PhD

Center for Technology and Behavioral Health

Geisel School of Medicine

Dartmouth College

46 Centerra Parkway

EverGreen Center, Suite 315

Lebanon, NH, 03766

United States

Phone: 1 603 646 7023

Email: catherine.stanger@dartmouth.edu

Abstract

Background: Cannabis misuse in young adults is a major public health concern. An important predictor of continued use is cannabis craving. Due to the time-varying nature of cravings, brief momentary interventions delivered while cravings are elevated may improve the use of strategies to cope with cravings and reduce cannabis use.

Objective: The goal of this manuscript is to describe a formative study to develop coping strategy messages for use in a subsequent intervention.

Methods: Young adults (aged 19-25 years; n=20) who reported using cannabis >10 of the past 30 days recruited via social media participated in this formative study. Participants rated an initial set of 15 mindfulness and 15 distraction coping strategies on a scale from 1 to 4 (very low degree to very high degree) for clarity, usefulness, and tone. They also provided comments about the content.

Results: Participants found the initial distraction messages slightly clearer than mindfulness (mean 3.5, SD 0.4 and mean 3.4, SD 0.4, respectively), both were comparable in tone (mean 3.2, SD 0.5 and mean 3.2, SD 0.4, respectively), and mindfulness messages were more useful than distraction (mean 3.0, SD 0.5 and mean 2.8, SD 0.6, respectively). Of the 30 messages, 29 received a rating of very low or low (<2) on any domain by >3 participants or received a comment suggesting a change. We revised all these messages based on this feedback, and the participants rated the revised messages approximately 2 weeks later. Participants earned US \$10 for completing the first and US \$20 for the second survey. The ratings improved on usefulness (especially the distraction items) with very little change in clarity and tone. The top 10 messages of each coping type (mindfulness and distraction) were identified by overall average rating (collapsed across all 3 dimensions: all rated >3.0). The final items were comparable in clarity (distraction mean 3.6, SD 0.4; mindfulness mean 3.6, SD 0.4), tone (distraction mean 3.4, SD 0.4; mindfulness mean 3.4, SD 0.4), and usefulness (distraction mean 3.1, SD 0.5; mindfulness mean 3.2, SD 0.5).

Conclusions: The inclusion of end users in the formative process of developing these messages was valuable and resulted in improvements to the content of the messages. The majority of the messages were changed in some way including the removal of potentially triggering language. These messages were subsequently used in an ecological momentary intervention.

(JMIR Form Res 2022;6(12):e40139) doi:[10.2196/40139](https://doi.org/10.2196/40139)

KEYWORDS

cannabis; formative; distraction; mindfulness; coping; youth; public health; mental health; health intervention; ecological momentary intervention

Introduction

Young adults between the ages of 19 and 25 years are in a time of major life changes during which they are developing new social relationships, experiencing increased independence, and developing the skills necessary to regularly make healthy choices [1]. This time of transition is a period during which additional support for healthy choices is critical, including refraining from substance misuse. Cannabis use among young adults in the United States has increased considerably over the past few decades, while the perceived risks of using cannabis have decreased [2]. In 2020, an estimated 13.5% of young adults aged 18-25 years met the criteria for cannabis use disorder [3]. Greater quantity and frequency of cannabis use are associated with increased cannabis-related problems, such as increased psychological distress, loneliness, and detrimental effects on memory [4-9].

Craving is one predictor of subsequent cannabis use and may be an important target for intervention when attempting to reduce or quit use [10,11]. The term “craving” has a history of slightly different definitions [12]; however, most definitions have in common that they refer to the desire for the drug or desire to experience the resulting effects of using the drug. Some definitions distinguish between “cravings” and “urges,” suggesting that the term “craving” be restricted to referencing a desire for the effects of using a drug, whereas the term “urge” be used when referencing intent to use a drug [13]. Some definitions conceptualize “cravings” and “urges” as the same phenomenon at different points on a spectrum with “craving” used to indicate an extreme desire and “urge,” indicating a lesser desire [14,15]. Although there have been arguments suggesting distinct definitions, individuals respond similarly to items assessing “cravings” and “urges” [16]. Here, we do not distinguish between cannabis cravings and the urge to use cannabis.

Mindfulness and distraction are two strategies to reduce cravings [17]. These two strategies have distinct theoretical bases: mindfulness involves maintaining attention on an immediate experience while adopting an accepting and curious perspective, whereas distraction involves active engagement with an alternative activity to direct focus away from the craving experience [17,18]. Mindfulness has been shown to be a useful strategy to cope with cravings and prevent relapse following periods of abstinence from cigarette and alcohol use [19-21]. More relevantly, the implementation of a mindfulness practice reduces the relationship between cannabis cravings and subsequent use [22,23]. Support for distraction as a coping strategy is mixed; some studies suggest distraction may be maladaptive [24], and others suggest there may be no relationship between distraction as a coping mechanism and craving [10]. However, other research shows distraction to be an effective coping mechanism, even outperforming mindfulness as a strategy to cope with cravings [25].

Direct comparisons of mindfulness and distraction coping strategies have largely been limited to controlled laboratory settings, limiting the generalizability of these findings. The extent to which mindfulness or distraction are effective as coping mechanisms likely depends on the environmental context at a given moment [26,27]. Additionally, though many therapeutic programs teach strategies for coping with cravings [18,28], these require individuals to learn a strategy at a time when cravings may not be present or distressing, then implement the strategy later when they are experiencing uncomfortable or distressing levels of craving. Because craving levels vary throughout the day, it may be beneficial to provide support when craving levels are high. Digital interventions can provide such time-varying support.

Several app-based interventions have shown success in helping people who use cannabis reduce their use [29,30]. Participants generally found these apps to be acceptable as an intervention for cannabis use. Although these mobile interventions provided support to individuals attempting to reduce their cannabis use—including strategies for coping with cannabis cravings—they did not necessarily provide support at the moment when the need is high (eg, when the urge to use is high). Rather, these apps mimic the structure of traditional therapy wherein participants engage with psychoeducational content and develop important skills but do not assess momentary craving or push momentary support when the need is high. However, digital interventions have the capability of providing momentary support in the form of advice, information, or coping strategies recommended to the participant via text (eg, using SMS text messages or within an app). Such interventions have shown promising results in providing support to individuals who are trying to quit smoking [31,32], and in reducing risky alcohol use [33]. The inexpensive and brief nature of interventions that rely on communication via text or direct messages allows for greater flexibility in intervention timing. This can allow for the delivery of a message when a participant is in great need of support, such as when a participant is in proximity to trigger locations detected via passive sensing [34], when the participant texts requesting support [32], or when elevated need is identified by periodic ecological momentary assessment (EMA). Currently, there are no such interventions to help young adults cope with cravings as they reduce their cannabis use.

An ongoing concern with digital interventions is the promotion of engagement with the intervention [35]. Although digital interventions may reduce barriers to accessing treatment [36], the target individuals still must engage with (ie, invest physical, emotional, and cognitive energy in [37]) the intervention to see positive effects. One way to improve the success of a digital intervention is through user-centered design to include the target population in the formative stages of the intervention [38]. The purpose of this study was to develop messages containing mindfulness- or distraction-based coping strategies for use in

an ecological momentary intervention to help young adults cope with cannabis cravings as they try to reduce their cannabis use.

Methods

Participants

Participants included 20 young adults (19-25 years; mean 21.65, SD 1.79 years; n=9, 45% male; n=9, 40% non-White or Hispanic; see [Table 1](#) for demographic details) who responded

to a Facebook ad recruiting people who use cannabis for a research study and were interested in reducing their use and reported using cannabis ≥ 10 out of the past 30 days. Individuals who were pregnant or breastfeeding or who reported being in treatment for problems related to substance use were excluded from participation. These inclusion and exclusion criteria were selected to ensure consistency with the eligibility criteria for the later trial of the ecological momentary intervention for which the messages were being developed.

Table 1. Demographic characteristics of participants.

Participant characteristics	Values, n (%)
Gender	
Female	9 (45)
Male	9 (45)
Nonbinary	2 (10)
Race	
White	14 (70)
Black or African American	1 (5)
Asian	2 (10)
American Indian or Alaska Native	1 (5)
Other (not specified)	2 (10)
Ethnicity	
Hispanic or Latino	3 (15)
Not Hispanic or Latino	17 (85)

Ethical Considerations

All procedures were reviewed and approved by the institutional review board at Dartmouth College (#32248). Participants clicked on a link in a Facebook ad and were brought to an online survey where they received general information about the study and were asked if they consented to be screened for eligibility to participate. All consent procedures were embedded in questions within the survey. Prospective participants who affirmatively consented to be screened were asked questions to determine their eligibility for participation. Those who met the inclusion criteria were presented with more detailed information about the study and asked if they consented to participate in the study. Affirmative consent was required before progressing to participate in the study survey.

Procedures

After completing the eligibility screening and providing informed consent, the participants completed a survey to evaluate a bank of 30 messages consisting of mindfulness-based or distraction-based suggestions for how to cope with cannabis cravings. The initial message items are in [Multimedia Appendices 1 and 2](#). The initial bank of 15 mindfulness messages were adapted from Witkiewitz et al [33] and Spears et al [39]. These studies used mindfulness messages in mobile interventions with the aim of reducing alcohol use and smoking. The initial bank of 15 distraction messages was adapted from Guarino et al [40], which tested a web-based intervention for

self-management of pain in individuals who have problems managing their opioid medications. We adapted the messages used in these sources to apply to coping with cannabis cravings, primarily by replacing any references to alcohol use, smoking, or opioid use with references to cannabis use. These strategies and messages could be adapted to be relevant to behaviors in addition to cannabis use by adjusting the wording such that it applies to other substances.

The goal of this study was to select a final bank of 20 total messages (10 mindfulness and 10 distraction) for use in an ecological momentary intervention that would present 1 randomly selected message from the bank of 20 messages when participants reported an urge to use cannabis ≥ 4 on a scale of 0-10. The number of messages was selected to prevent habituation and boredom by ensuring that a variety of coping strategies are delivered to the participants across a 4-week intervention. All participants in this study were given the same 30 messages at the same time to provide their ratings. Beginning with a bank of 30 possible messages allowed for the 10 lowest-rated messages to be removed from the final message bank.

We implemented a rating scale used in other formative research that aimed to develop a text-based intervention to reduce alcohol use among college students [41]. Participants rated messages on a scale from 1 to 4 (very low degree to very high degree) for understanding (this message is easy to understand), usability (this message is useful), and tone (this message has a good

overall tone [41]). We also asked participants to provide any comments they had about how to improve each message in a free response box. Following the first round of feedback, we reviewed all messages and revised any messages that received a rating of very low or low on any of the domains by at least 3 participants. After revising the messages, we sent a second survey (approximately 2 weeks after completing the first survey) asking participants to reevaluate the messages using the same scale and criteria. Participants were compensated US \$10 for completing the first survey and US \$20 for completing the second survey. The top 10 messages in each category (mindfulness and distraction) were selected for use in the ecological momentary intervention.

Results

Initial Quantitative Results

Figure 1 shows the mean ratings for each category and message type following the first round of ratings. Table 2 shows the mean ratings for each category and message type at both timepoints. The participants found the initial distraction messages more clear than the mindfulness messages ($t_{19}=2.64, P=.02$), both were comparable in tone ($t_{19}=0.28, P=.78$), and mindfulness messages more useful than distraction ($t_{19}=2.36, P=.03$).

Figure 1. Initial participant ratings (n=20) of the distraction and mindfulness messages. Each data point shows one participant’s average rating of the 15 distraction or 15 mindfulness messages on each of the three domains (clarity, tone, and usefulness). Horizontal bars show average ratings and error bars are 95% CIs.

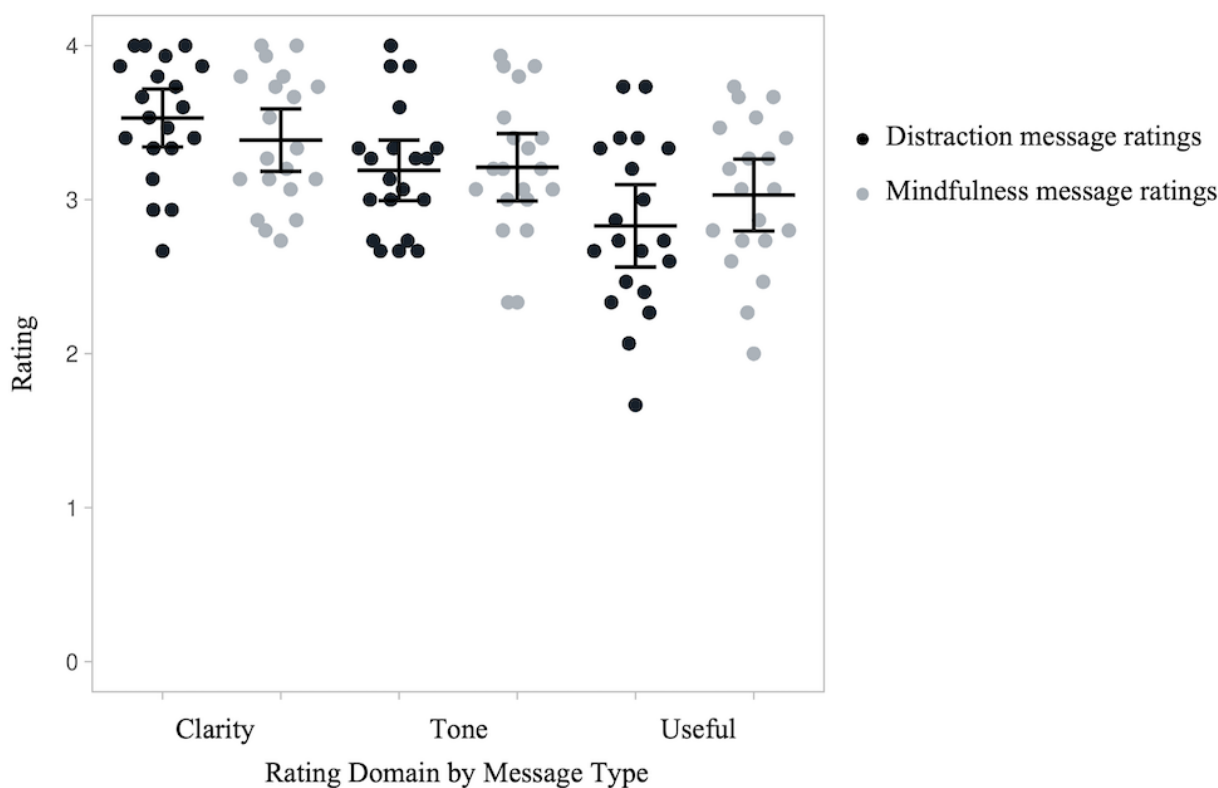


Table 2. Mean (SD) and range of message ratings.

Timepoint	Distraction			Mindfulness		
	Clear	Tone	Useful	Clear	Tone	Useful
1 (N=20)						
Mean (SD)	3.53 (0.39)	3.21 (0.46)	2.83 (0.56)	3.39 (0.42)	3.19 (0.41)	3.03 (0.49)
Range	2.67-4.00	2.33-3.93	1.67-3.73	2.73-4.00	2.67-4.00	2.00-3.73
2 (n=18)						
Mean (SD)	3.50 (0.44)	3.27 (0.51)	2.96 (0.48)	3.49 (0.44)	3.32 (0.46)	3.07 (0.54)
Range	2.67-4.00	2.47-4.00	2.20-4.00	2.67-4.00	2.20-3.93	2.00-3.93

Initial Qualitative Results

Major feedback themes for the mindfulness messages included concerns about the clarity of the messages, suggestions for rewording the message content, concerns about the suggestion being too difficult to implement at the time cravings were high, and positive responses to the nonjudgmental nature of the messages. Major feedback themes for the distraction messages included confusion about the rationale of the message and how it relates to coping with cannabis cravings, indications that the strategy would not be helpful to them, and concerns that the strategy could be triggering (eg, use of social media as a distraction technique could bring up images that are related to cannabis use). [Textbox 1](#) shows feedback categories and example quotes from participants. The first author identified the feedback categories by finding common themes across participant comments. The first author and a research assistant each independently categorized all feedback statements with

82% agreement. Out of the 30 messages we asked the participants to rate, 29 messages received a rating of very low or low (≤ 2) on at least 1 of the 3 domains by ≥ 3 participants or received a comment suggesting a change to the message. Although the sample size is relatively small for this study, we kept the threshold for revising items low with the intention of creating the best possible messages for our subsequent intervention. The lead and senior authors (MAA and CS) reviewed these messages, the quantitative ratings, and the qualitative feedback and revised the messages based on both the domains that were rated low and the open-ended feedback the participants provided. For example, if a message received a low rating on “tone,” we revised the language of the message to make the message friendlier, more empathetic, and more encouraging while maintaining the general strategy the message provided (see [Textbox 2](#) for an example). We implemented participants’ suggestions about how to reword messages whenever possible.

Textbox 1. Examples of participant feedback. The first author found common themes across participant feedback comments which resulted in these feedback categories. This textbox shows the example of participant feedback for each category.

<p>Reword</p> <ul style="list-style-type: none"> Last sentence could reverse: “you don’t need to act on any urges you may feel to use cannabis” <p>Clarity</p> <ul style="list-style-type: none"> Had to read over a second time to understand the message <p>Too difficult to implement</p> <ul style="list-style-type: none"> This may be true, but it’s very hard to implement in real life. Message doesn’t provide a good strategy to use this method IMO <p>Nonjudgmental</p> <ul style="list-style-type: none"> This perspective feels empathic and helpful <p>Rationale</p> <ul style="list-style-type: none"> But why? What’s the benefit <p>Triggering</p> <ul style="list-style-type: none"> As I said before, music tends to be more enjoyable when high. I think it is still effective, but it also may make the person want to use cannabis before zeroing in on his or her favorite song. <p>Not helpful</p> <ul style="list-style-type: none"> Vague and not particularly helpful.
--

Textbox 2. Example of message revision. The lead and senior authors used the quantitative ratings and qualitative feedback to revise the messages. This textbox shows an example of how a message was revised using qualitative feedback.

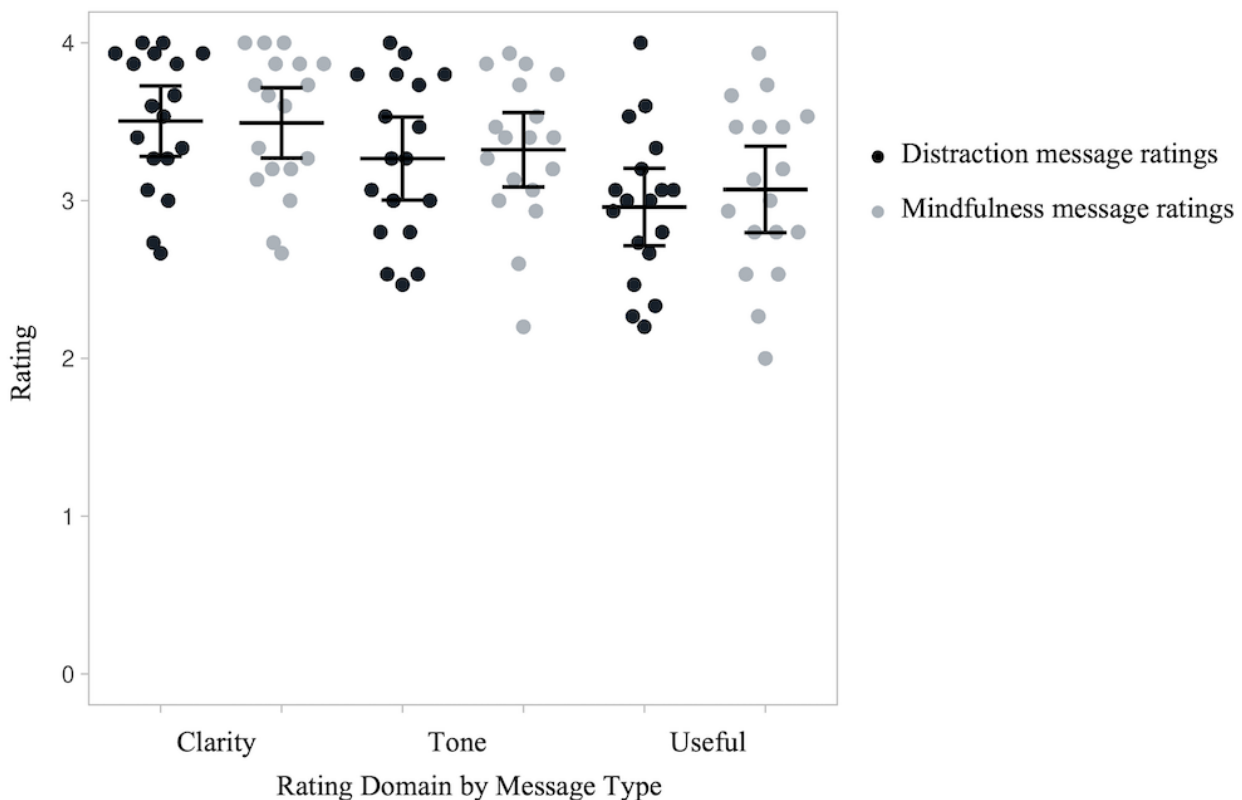
<p>Original message</p> <ul style="list-style-type: none"> Focus on something new by doing something hard. Try counting backward from 100 by sevens. <p>Feedback</p> <ul style="list-style-type: none"> This feels silly and I wouldn’t do it. <p>Revised message</p> <ul style="list-style-type: none"> Reduce your urge to use by focusing on something new and challenging. Try saying the alphabet backwards. Or, try coming up with as many words as you can that rhyme with “think.”

Quantitative Results Following Revision

We were unable to reach 2 participants to complete the second round of revisions. Therefore, 18 participants completed the second round of message evaluations. The 2 participants whom we could not reach were excluded from these analyses. [Figure 2](#) shows the mean ratings for each category and message type following the second round of ratings. The ratings for the usefulness of the distraction messages improved significantly ($t_{17}=2.52, P=.02$) with little change in clarity ($t_{17}=0.33, P=.75$) or tone ($t_{17}=1.55, P=.14$). Ratings for the tone of the mindfulness messages improved significantly ($t_{17}=2.64, P=.02$) with very

little change on clarity ($t_{17}=1.97, P=.07$) and usefulness ($t_{17}=1.25, P=.23$). All participants received the same number of messages to rate in the same order, and participants read and rated 1 message before the next message displayed. It is possible that the number of messages presented to the participants or the order in which the participants read the messages could have influenced the participants' ratings. However, because the number of messages and the order in which they were presented were fixed, we cannot test the impact of order or number of messages on message satisfaction. Mean ratings for all items are in [Table 2](#).

Figure 2. Participant ratings (n=18) of the distraction and mindfulness messages after the first round of message revisions. Each data point shows one participant's average rating of the 15 distraction or 15 mindfulness messages on each of the three domains (clarity, tone, and usefulness). Horizontal bars show average ratings and error bars are 95% CIs.



Qualitative Results Following Revision

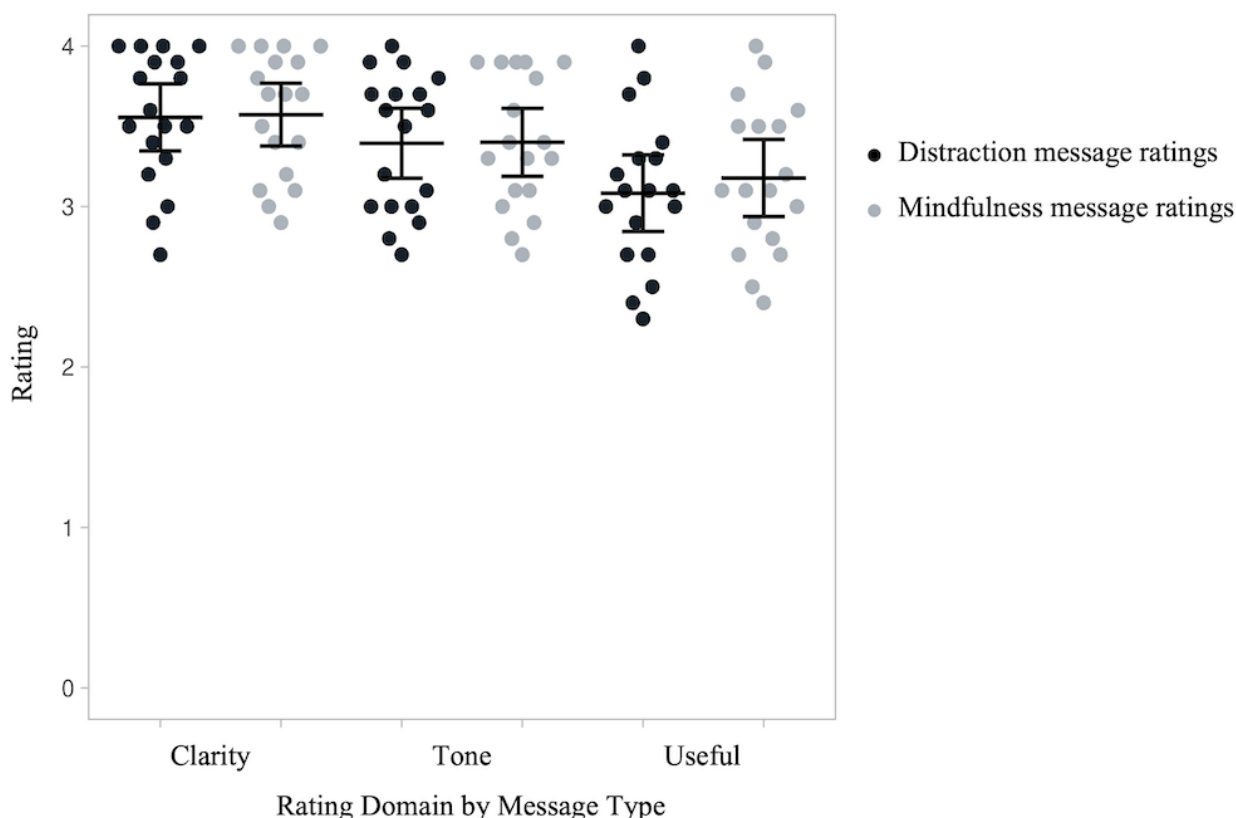
A number of themes emerged from the second round of feedback. For the mindfulness messages, the remaining concerns were primarily about some messages not being helpful (eg, 4 comments with concerns about the strategy not working or not being relatable), and there were some suggestions about how to reword messages to improve their clarity and usefulness (4 comments). For the distraction messages, the remaining concerns were primarily about the strategies not being helpful or practical (12 comments). We made final revisions to the messages based on the participants' comments.

Final Messages

The top 10 messages of each coping type (mindfulness and distraction) were identified by overall (collapsed across all 3

dimensions) average rating. [Figure 3](#) shows the mean ratings for each category and message type for the final messages selected. The final message items are in [Multimedia Appendices 3 and 4](#). All final messages had an overall average rating ≥ 3.0 but selecting the top 10 messages allowed us to exclude messages that participants indicated may be triggering, suboptimal in tone, or particularly unhelpful from the pilot intervention. The second round of ratings of the final 20 distraction and mindfulness messages were not significantly different from each other in clarity (distraction mean 3.56, SD 0.41; mindfulness mean 3.57, SD 0.38; $t_{17}=0.25, P=.81$), tone (distraction mean 3.39, SD 0.43; mindfulness mean 3.40, SD 0.41; $t_{17}=0.08, P=.94$), and usefulness (distraction mean 3.08, SD 0.47; mindfulness mean 3.19, SD 0.47; $t_{17}=0.98, P=.34$). These messages were used in a subsequent ecological momentary intervention study.

Figure 3. Participant ratings (n=18) of the final distraction and mindfulness messages after removing the unused messages. Each data point shows one participant's average rating of the final 10 distraction or 10 mindfulness messages on each of the three domains (clarity, tone, and usefulness). Horizontal bars show average ratings and error bars are 95% CIs.



Discussion

Principal Findings

We set out to develop messages to be used in a pilot digital intervention to help young adults cope with cannabis cravings, including young adults who use cannabis in the message development process. This formative process resulted in changes to the original messages, and the selection of the highest-rated messages to be used in the subsequent intervention. Items improved significantly in terms of usefulness (distraction messages) and tone (mindfulness messages) following the first round of revisions. Although there were no significant improvements on the other domains following revisions, this was likely due to a ceiling effect, given the high initial ratings (mean >3.0 out of the maximum 4.0) on these domains. Initially, the participants found the distraction items to be clearer than the mindfulness messages, and the mindfulness messages to be more useful than the distraction messages. Following the revisions and final selection of 10 messages each from the mindfulness and distraction categories, the ratings of each message type were high and comparable across all domains.

Comparison With Prior Work

Prior work on SMS text message intervention development has highlighted the importance of including end user populations in the development and design of the intervention [38,42]. Formative research to develop mobile health interventions has used focus groups and surveys to gain feedback from the target audience about elements of apps that participants find useful,

to help prioritize various app features, and to identify barriers to implementation [43,44]. Additionally, previous studies developing text-based interventions have involved members of the target population in focus groups to develop message content and offered input on the timing and frequency of the SMS text messages [32,41]. Although focus groups and semistructured interviews may be particularly informative when there are multiple developmental components to consider such as design features and intervention content, our mobile intervention was developed for delivery via a commercial platform thus limiting the flexibility of app design.

Unlike other preliminary studies focused on text-based interventions, we did not conduct focus groups. Instead, we developed our initial messages based on existing evidence-based interventions, sought input from users on the content we developed, and modified our content based on participants' feedback. Future research could implement a more comprehensive participatory approach in developing the various components of the intervention such as the timing and frequency of message delivery. One major finding of previous studies has been the importance of message tone on user engagement. The tone of the messages we developed is one domain on which our messages improved as a result of participant feedback, again highlighting the importance and benefits of including end users in the development of interventions.

Limitations

The findings of this study should be considered in the context of a few limitations. We limited the involvement of our target

population in developing the intervention to providing feedback on messages adapted from established interventions. This limitation is partially due to the constraints of the platform we used for our intervention delivery. The platform used allowed for flexible and diverse experimental designs and made the intervention possible at a low cost without the need for outsourcing programming but was limited in terms of app customization. After appropriate efficacy testing, it may be appropriate to distribute the app using a self-pay subscription model or by making the app freely available to consumers similar to other apps designed for those who are seeking assistance to moderate or abstain from substance use such as SoberTool developed by Blitzen, LLC. The costs of such apps can be sustained using in-app advertisements or possibly reimbursement through medical insurance [45].

A second limitation is regarding our sampling strategy. We limited our recruitment to social media in an attempt to reach a diverse population and allow users across the country to participate in the development of our intervention messages. This strategy allowed us to recruit nationwide and is the same planned strategy for recruiting our intervention participants. However, this method of recruitment may have resulted in bias due to self-selection. Additionally, due to time and monetary constraints, we solicited feedback from the participants using an asynchronous survey instead of more involved focus groups.

One additional limitation is that the user-rated usefulness of messages when craving levels are not currently elevated may be a poor proxy for clinical utility in the context of high craving; however, the usefulness of the messages will be evaluated in

the subsequent intervention. We also did not distinguish between craving as a desire to use cannabis and urge as the intent to use cannabis. Although this is beyond the scope of this study, future studies may make this distinction and test whether 1 strategy (mindfulness or distraction) is better suited for managing the desire for the effects (ie, cravings) versus managing intent to use cannabis (ie, urges).

Conclusions

The findings of this study support the importance and highlight the value of including the target intervention population in the formative process of intervention development. The content of the messages was significantly improved over the course of this formative process.

Additionally, the community identified possible triggers embedded in the messages that may have been counterproductive to our intervention—triggers that would not have been identified without their lived experience and inclusion in the message development. Focus groups may be more useful in developing and revising messages, allowing for additional conversation between the participants and the researchers, thus giving space for clarification and conversation. Sometimes, open-ended participant feedback comments were somewhat unclear, and having the opportunity to have an ongoing discourse could help develop messages further. The final messages developed in this study were subsequently used in a pilot intervention aiming to provide young adults who use cannabis with support for coping with their cannabis cravings as they attempt to reduce their use.

Acknowledgments

This research was supported by National Institute on Drug Abuse grants P30 DA029926 and T32 DA037202. In addition, National Institute on Drug Abuse grants P50 DA054039 and R01 DA039901 supported the effort of IN-S.

Conflicts of Interest

AJB is a Science Advisor at Canopy Growth Inc.

Multimedia Appendix 1

Initial bank of mindfulness messages.

[[DOCX File , 15 KB - formative_v6i12e40139_app1.docx](#)]

Multimedia Appendix 2

Initial bank of distraction messages.

[[DOCX File , 15 KB - formative_v6i12e40139_app2.docx](#)]

Multimedia Appendix 3

Final mindfulness messages.

[[DOCX File , 14 KB - formative_v6i12e40139_app3.docx](#)]

Multimedia Appendix 4

Final distraction messages.

[[DOCX File , 14 KB - formative_v6i12e40139_app4.docx](#)]

References

1. Scales PC, Benson PL, Oesterle S, Hill KG, Hawkins JD, Pashak TJ. The dimensions of successful young adult development: a conceptual and measurement framework. *Appl Dev Sci* 2016;20(3):150-174 [FREE Full text] [doi: [10.1080/10888691.2015.1082429](https://doi.org/10.1080/10888691.2015.1082429)] [Medline: [30344455](https://pubmed.ncbi.nlm.nih.gov/30344455/)]
2. Schulenberg JE, Johnston LD, O'Malley PM, Bachman JG, Miech RA, Patrick ME. *Monitoring the Future National Survey Results on Drug Use, 1975-2019: Volume II, College Students and Adults Ages 19-60*. Ann Arbor: Institute for Social Research, The University of Michigan; 2020.
3. Quality CfbHSA. *Results from the 2020 National Survey on Drug Use and Health: Detailed Tables*. Rockville: National Survey on Drug Use and Health; 2021.
4. Volkow ND, Swanson JM, Evins AE, DeLisi LE, Meier MH, Gonzalez R, et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. *JAMA Psych* 2016 Mar;73(3):292-297. [doi: [10.1001/jamapsychiatry.2015.3278](https://doi.org/10.1001/jamapsychiatry.2015.3278)] [Medline: [26842658](https://pubmed.ncbi.nlm.nih.gov/26842658/)]
5. Budney AJ, Sofis MJ, Borodovsky JT. An update on cannabis use disorder with comment on the impact of policy related to therapeutic and recreational cannabis use. *Eur Arch Psychiatry Clin Neurosci* 2019 Feb;269(1):73-86 [FREE Full text] [doi: [10.1007/s00406-018-0976-1](https://doi.org/10.1007/s00406-018-0976-1)] [Medline: [30604051](https://pubmed.ncbi.nlm.nih.gov/30604051/)]
6. Battisti RA, Roodenrys S, Johnstone SJ, Respondek C, Hermens DF, Solowij N. Chronic use of cannabis and poor neural efficiency in verbal memory ability. *Psychopharmacology (Berl)* 2010 May;209(4):319-330. [doi: [10.1007/s00213-010-1800-4](https://doi.org/10.1007/s00213-010-1800-4)] [Medline: [20217055](https://pubmed.ncbi.nlm.nih.gov/20217055/)]
7. Rhew IC, Cadigan JM, Lee CM. Marijuana, but not alcohol, use frequency associated with greater loneliness, psychological distress, and less flourishing among young adults. *Drug Alcohol Depend* 2021 Jan 01;218:108404 [FREE Full text] [doi: [10.1016/j.drugalcdep.2020.108404](https://doi.org/10.1016/j.drugalcdep.2020.108404)] [Medline: [33250378](https://pubmed.ncbi.nlm.nih.gov/33250378/)]
8. Gilman JM, Lee S, Kuster JK, Lee MJ, Kim BW, van der Kouwe A, et al. Variable activation in striatal subregions across components of a social influence task in young adult cannabis users. *Brain Behav* 2016 May;6(5):e00459 [FREE Full text] [doi: [10.1002/brb3.459](https://doi.org/10.1002/brb3.459)] [Medline: [27257518](https://pubmed.ncbi.nlm.nih.gov/27257518/)]
9. Dumont GJH, van Hasselt JGC, de Kam M, van Gerven JMA, Touw DJ, Buitelaar JK, et al. Acute psychomotor, memory and subjective effects of MDMA and THC co-administration over time in healthy volunteers. *J Psychopharmacol* 2011 Apr;25(4):478-489. [doi: [10.1177/0269881110376687](https://doi.org/10.1177/0269881110376687)] [Medline: [20817749](https://pubmed.ncbi.nlm.nih.gov/20817749/)]
10. Enkema MC, Hallgren KA, Larimer ME. Craving is impermanent and it matters: investigating craving and cannabis use among young adults with problematic use interested in reducing use. *Drug Alcohol Depend* 2020 May 01;210:107957 [FREE Full text] [doi: [10.1016/j.drugalcdep.2020.107957](https://doi.org/10.1016/j.drugalcdep.2020.107957)] [Medline: [32200158](https://pubmed.ncbi.nlm.nih.gov/32200158/)]
11. Norberg M, Barnier E, Weidemann G, Chakerian K, Cornish J, Rapee R. Randomised pilot study of cannabis cue exposure: reducing cue reactivity while building tolerance. *Clin Psychol* 2018 Nov 10;22(2):126-136 [FREE Full text] [doi: [10.1111/cp.12157](https://doi.org/10.1111/cp.12157)]
12. Sayette M, Shiffman S, Tiffany S, Niaura R, Martin C, Shadel W. The measurement of craving. *Addiction* 2000;95:S189-S210. [doi: [10.1046/j.1360-0443.95.8s2.8.x](https://doi.org/10.1046/j.1360-0443.95.8s2.8.x)]
13. Larimer ME, Palmer RS, Marlatt GA. Relapse prevention. An overview of Marlatt's cognitive-behavioral model. *Alcohol Res Health* 1999;23(2):151-160 [FREE Full text] [Medline: [10890810](https://pubmed.ncbi.nlm.nih.gov/10890810/)]
14. Shiffman S. Craving: don't let us throw the baby out with the bathwater. *Addiction* 1987 Jan;82(1):37-38. [doi: [10.1111/j.1360-0443.1987.tb01431.x](https://doi.org/10.1111/j.1360-0443.1987.tb01431.x)]
15. West R, Schneider N. Craving for cigarettes. *Br J Addict* 1987 Apr;82(4):407-415. [doi: [10.1111/j.1360-0443.1987.tb01496.x](https://doi.org/10.1111/j.1360-0443.1987.tb01496.x)] [Medline: [3555575](https://pubmed.ncbi.nlm.nih.gov/3555575/)]
16. Tiffany ST, Singleton E, Haertzen CA, Henningfield JE. The development of a cocaine craving questionnaire. *Drug Alcohol Depend* 1993 Dec;34(1):19-28. [doi: [10.1016/0376-8716\(93\)90042-o](https://doi.org/10.1016/0376-8716(93)90042-o)]
17. Ashe ML, Newman MG, Wilson SJ. Delay discounting and the use of mindful attention versus distraction in the treatment of drug addiction: a conceptual review. *J Exp Anal Behav* 2015 Jan;103(1):234-248 [FREE Full text] [doi: [10.1002/jeab.122](https://doi.org/10.1002/jeab.122)] [Medline: [25545725](https://pubmed.ncbi.nlm.nih.gov/25545725/)]
18. Marlatt G, Witkiewitz K. *Relapse Prevention for Alcohol and Drug Problems*. Relapse prevention: Maintenance Strategies in the Treatment of Addictive Behaviors, 2nd ed. New York: The Guilford Press; 2005:1-44.
19. Bowen S, Marlatt A. Surfing the urge: brief mindfulness-based intervention for college student smokers. *Psychol Addict Behav* 2009 Dec;23(4):666-671. [doi: [10.1037/a0017127](https://doi.org/10.1037/a0017127)] [Medline: [20025372](https://pubmed.ncbi.nlm.nih.gov/20025372/)]
20. Witkiewitz K, Bowen S. Depression, craving, and substance use following a randomized trial of mindfulness-based relapse prevention. *J Consult Clin Psychol* 2010 Jun;78(3):362-374 [FREE Full text] [doi: [10.1037/a0019172](https://doi.org/10.1037/a0019172)] [Medline: [20515211](https://pubmed.ncbi.nlm.nih.gov/20515211/)]
21. Korecki JR, Schwebel FJ, Votaw VR, Witkiewitz K. Mindfulness-based programs for substance use disorders: a systematic review of manualized treatments. *Subst Abuse Treat Prev Policy* 2020 Jul 29;15(1):51 [FREE Full text] [doi: [10.1186/s13011-020-00293-3](https://doi.org/10.1186/s13011-020-00293-3)] [Medline: [32727559](https://pubmed.ncbi.nlm.nih.gov/32727559/)]
22. Enkema MC, Bowen S. Mindfulness practice moderates the relationship between craving and substance use in a clinical sample. *Drug Alcohol Depend* 2017 Oct 01;179:1-7. [doi: [10.1016/j.drugalcdep.2017.05.036](https://doi.org/10.1016/j.drugalcdep.2017.05.036)] [Medline: [28734167](https://pubmed.ncbi.nlm.nih.gov/28734167/)]
23. Enkema MC, Hallgren KA, Bowen S, Lee CM, Larimer ME. Craving management: exploring factors that influence momentary craving-related risk of cannabis use among young adults. *Addict Behav* 2021 Apr;115:106750. [doi: [10.1016/j.addbeh.2020.106750](https://doi.org/10.1016/j.addbeh.2020.106750)] [Medline: [33383565](https://pubmed.ncbi.nlm.nih.gov/33383565/)]

24. Garland EL, Brown SM, Howard MO. Thought suppression as a mediator of the association between depressed mood and prescription opioid craving among chronic pain patients. *J Behav Med* 2016 Feb;39(1):128-138. [doi: [10.1007/s10865-015-9675-9](https://doi.org/10.1007/s10865-015-9675-9)] [Medline: [26345263](https://pubmed.ncbi.nlm.nih.gov/26345263/)]
25. Murphy CM, MacKillop J. Mindfulness as a strategy for coping with cue-elicited cravings for alcohol: an experimental examination. *Alcohol Clin Exp Res* 2014 Apr;38(4):1134-1142 [FREE Full text] [doi: [10.1111/acer.12322](https://doi.org/10.1111/acer.12322)] [Medline: [24428808](https://pubmed.ncbi.nlm.nih.gov/24428808/)]
26. Murphy MJ, Mermelstein LC, Edwards KM, 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](https://doi.org/10.1080/07448481.2011.629260)] [Medline: [22686356](https://pubmed.ncbi.nlm.nih.gov/22686356/)]
27. Cheavens JS, Strunk DR, Lazarus SA, Goldstein LA. The compensation and capitalization models: a test of two approaches to individualizing the treatment of depression. *Behav Res Ther* 2012 Nov;50(11):699-706. [doi: [10.1016/j.brat.2012.08.002](https://doi.org/10.1016/j.brat.2012.08.002)]
28. Budney AJ, Fearer S, Walker DD, Stanger C, Thostenson J, Grabinski M, et al. An initial trial of a computerized behavioral intervention for cannabis use disorder. *Drug Alcohol Depend* 2011 May 01;115(1-2):74-79 [FREE Full text] [doi: [10.1016/j.drugalcdep.2010.10.014](https://doi.org/10.1016/j.drugalcdep.2010.10.014)] [Medline: [21131143](https://pubmed.ncbi.nlm.nih.gov/21131143/)]
29. Coughlin LN, Nahum-Shani I, Philyaw-Kotov ML, Bonar EE, Rabbi M, Klasnja P, et al. Developing an adaptive mobile intervention to address risky substance use among adolescents and emerging adults: usability study. *JMIR Mhealth Uhealth* 2021 Jan 15;9(1):e24424. [doi: [10.2196/24424](https://doi.org/10.2196/24424)]
30. Albertella L, Gibson L, Rooke S, Norberg MM, Copeland J. A smartphone app intervention for adult cannabis users wanting to quit or reduce their use: a pilot evaluation. *J Cannabis Res* 2019 Aug 16;1(1):9 [FREE Full text] [doi: [10.1186/s42238-019-0009-6](https://doi.org/10.1186/s42238-019-0009-6)] [Medline: [33526112](https://pubmed.ncbi.nlm.nih.gov/33526112/)]
31. Keoleian V, Polcin D, Galloway GP. Text messaging for addiction: a review. *J Psychoactive Drugs* 2015;47(2):158-176 [FREE Full text] [doi: [10.1080/02791072.2015.1009200](https://doi.org/10.1080/02791072.2015.1009200)] [Medline: [25950596](https://pubmed.ncbi.nlm.nih.gov/25950596/)]
32. 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 07;7(1):e11246 [FREE Full text] [doi: [10.2196/11246](https://doi.org/10.2196/11246)] [Medline: [30617043](https://pubmed.ncbi.nlm.nih.gov/30617043/)]
33. Witkiewitz K, Desai SA, Bowen S, Leigh BC, Kirouac M, Larimer ME. Development and evaluation of a mobile intervention for heavy drinking and smoking among college students. *Psychol Addict Behav* 2014 Sep;28(3):639-650 [FREE Full text] [doi: [10.1037/a0034747](https://doi.org/10.1037/a0034747)] [Medline: [25000269](https://pubmed.ncbi.nlm.nih.gov/25000269/)]
34. Gustafson DH, Boyle MG, Shaw BR, Isham A, McTavish F, Richards S, et al. An e-health solution for people with alcohol problems. *Alcohol Res Health* 2011;33(4):327-337 [FREE Full text] [Medline: [23293549](https://pubmed.ncbi.nlm.nih.gov/23293549/)]
35. Milward J, Drummond C, Fincham-Campbell S, Deluca P. What makes online substance-use interventions engaging? A systematic review and narrative synthesis. *Digit Health* 2018;4:2055207617743354 [FREE Full text] [doi: [10.1177/2055207617743354](https://doi.org/10.1177/2055207617743354)] [Medline: [29942622](https://pubmed.ncbi.nlm.nih.gov/29942622/)]
36. Hoch E, Preuss UW, Ferri M, Simon R. Digital interventions for problematic cannabis users in non-clinical settings: findings from a systematic review and meta-analysis. *Eur Addict Res* 2016;22(5):233-242 [FREE Full text] [doi: [10.1159/000445716](https://doi.org/10.1159/000445716)] [Medline: [27160333](https://pubmed.ncbi.nlm.nih.gov/27160333/)]
37. Nahum-Shani I, Shaw SD, Carpenter SM, Murphy SA, Yoon C. Engagement in digital interventions. *Am Psychol* 2022 Oct;77(7):836-852 [FREE Full text] [doi: [10.1037/amp0000983](https://doi.org/10.1037/amp0000983)] [Medline: [35298199](https://pubmed.ncbi.nlm.nih.gov/35298199/)]
38. 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](https://doi.org/10.2196/mhealth.3779)] [Medline: [25714907](https://pubmed.ncbi.nlm.nih.gov/25714907/)]
39. Spears CA, Abrams LC, Glass CR, Hedeker D, Eriksen MP, Cottrell-Daniels C, et al. Mindfulness-based smoking cessation enhanced with mobile technology (iQuit Mindfully): pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Jun 24;7(6):e13059 [FREE Full text] [doi: [10.2196/13059](https://doi.org/10.2196/13059)] [Medline: [31237242](https://pubmed.ncbi.nlm.nih.gov/31237242/)]
40. Guarino H, Fong C, Marsch LA, Acosta MC, Syckes C, Moore SK, et al. Web-based cognitive behavior therapy for chronic pain patients with aberrant drug-related behavior: outcomes from a randomized controlled trial. *Pain Med* 2018 Dec 01;19(12):2423-2437 [FREE Full text] [doi: [10.1093/pm/pnx334](https://doi.org/10.1093/pm/pnx334)] [Medline: [29346579](https://pubmed.ncbi.nlm.nih.gov/29346579/)]
41. Thomas K, Linderoth C, Bendtsen M, Bendtsen P, Müssener U. Text message-based intervention targeting alcohol consumption among university students: findings from a formative development study. *JMIR Mhealth Uhealth* 2016 Oct 20;4(4):e119 [FREE Full text] [doi: [10.2196/mhealth.5863](https://doi.org/10.2196/mhealth.5863)] [Medline: [27765732](https://pubmed.ncbi.nlm.nih.gov/27765732/)]
42. Thomas K, Müssener U, Linderoth C, Karlsson N, Bendtsen P, Bendtsen M. Effectiveness of a text messaging-based intervention targeting alcohol consumption among university students: randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Jun 25;6(6):e146 [FREE Full text] [doi: [10.2196/mhealth.9642](https://doi.org/10.2196/mhealth.9642)] [Medline: [29941417](https://pubmed.ncbi.nlm.nih.gov/29941417/)]
43. Bagot K, Hodgdon E, Sidhu N, Patrick K, Kelly M, Lu Y, et al. End user-informed mobile health intervention development for adolescent cannabis use disorder: qualitative study. *JMIR Mhealth Uhealth* 2019 Oct 04;7(10):e13691 [FREE Full text] [doi: [10.2196/13691](https://doi.org/10.2196/13691)] [Medline: [31588909](https://pubmed.ncbi.nlm.nih.gov/31588909/)]
44. Helseth SA, Guigayoma J, Price D, Spirito A, Clark MA, Barnett NP, et al. Developing a smartphone-based adjunct intervention to reduce cannabis use among juvenile justice-involved adolescents: protocol for a multiphase study. *JMIR Res Protoc* 2022 Mar 11;11(3):e35402 [FREE Full text] [doi: [10.2196/35402](https://doi.org/10.2196/35402)] [Medline: [35275086](https://pubmed.ncbi.nlm.nih.gov/35275086/)]

45. Powell AC, Bowman MB, Harbin HT. Reimbursement of apps for mental health: findings from interviews. *JMIR Ment Health* 2019 Aug 06;6(8):e14724 [FREE Full text] [doi: [10.2196/14724](https://doi.org/10.2196/14724)] [Medline: [31389336](https://pubmed.ncbi.nlm.nih.gov/31389336/)]

Abbreviations

EMA: ecological momentary assessment

Edited by A Mavragani; submitted 07.06.22; peer-reviewed by C Drummond, S Rodda; comments to author 27.07.22; revised version received 28.10.22; accepted 01.11.22; published 15.12.22.

Please cite as:

Anderson MA, Budney AJ, Jacobson NC, Nahum-Shani I, Stanger C

End User Participation in the Development of an Ecological Momentary Intervention to Improve Coping With Cannabis Cravings: Formative Study

JMIR Form Res 2022;6(12):e40139

URL: <https://formative.jmir.org/2022/12/e40139>

doi: [10.2196/40139](https://doi.org/10.2196/40139)

PMID: [36520509](https://pubmed.ncbi.nlm.nih.gov/36520509/)

©Molly A Anderson, Alan J Budney, Nicholas C Jacobson, Inbal Nahum-Shani, Catherine Stanger. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 15.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Electronic Source Data Transcription for Electronic Case Report Forms in China: Validation of the Electronic Source Record Tool in a Real-world Ophthalmology Study

Bin Wang^{1*}, MD; Junkai Lai^{2*}, MD; Mimi Liu^{3*}, MM; Feifei Jin^{4,5,6}, MD; Yifei Peng⁷, MS; Chen Yao^{1,8}, MPH

¹Peking University Clinical Research Institute, Peking University First Hospital, Beijing, China

²Institute of Automation, Chinese Academy of Sciences, Beijing, China

³School of Ophthalmology and Optometry, Eye Hospital of Wenzhou Medical University, Wenzhou, China

⁴Trauma Medicine Center, Peking University People's Hospital, Beijing, China

⁵Key Laboratory of Trauma treatment and Neural Regeneration, Peking University, Ministry of Education, Beijing, China

⁶National Center for Trauma Medicine of China, Beijing, China

⁷Hangzhou Tigermed Consulting Co, Ltd, Hangzhou, China

⁸Hainan Institute of Real World Data, Qionghai, China

*these authors contributed equally

Corresponding Author:

Chen Yao, MPH

Peking University Clinical Research Institute

Peking University First Hospital

No 8, Xishiku Street

Xicheng District

Beijing, 100034

China

Phone: 86 01083325822

Email: yaochen@hsc.pku.edu.cn

Abstract

Background: As researchers are increasingly interested in real-world studies (RWSs), improving data collection efficiency and data quality has become an important challenge. An electronic source (eSource) generally includes direct capture, collection, and storage of electronic data to simplify clinical research. It can improve data quality and patient safety and reduce clinical trial costs. Although there are already large projects on eSource technology, there is a lack of experience in using eSource technology to implement RWSs. Our team designed and developed an eSource record (ESR) system in China. In a preliminary prospective study, we selected a cosmetic medical device project to evaluate ESR software's effect on data collection and transcription. As the previous case verification was simple, we plan to choose more complicated ophthalmology projects to further evaluate the ESR.

Objective: We aimed to evaluate the data transcription efficiency and quality of ESR software in retrospective studies to verify the feasibility of using eSource as an alternative to traditional manual transcription of data in RWS projects.

Methods: The approved ophthalmic femtosecond laser project was used for ESR case validation. This study compared the efficiency and quality of data transcription between the eSource method using ESR software and the traditional clinical research model of manually transcribing the data. Usability refers to the quality of a user's experience when interacting with products or systems including websites, software, devices, or applications. To evaluate the system availability of ESR, we used the System Usability Scale (SUS). The questionnaire consisted of the following 2 parts: participant information and SUS evaluation of the electronic medical record (EMR), electronic data capture (EDC), and ESR systems. By accessing log data from the EDC system previously used by the research project, all the time spent from the beginning to the end of the study could be counted.

Results: In terms of transcription time cost per field, the eSource method can reduce the time cost by 81.8% (11.2/13.7). Compared with traditional manual data transcription, the eSource method has higher data transcription quality (correct entry rate of 2356/2400, 98.17% vs 47,991/51,424, 93.32%). A total of 15 questionnaires were received with a response rate of 100%. In terms of usability, the average overall SUS scores of the EMR, EDC, and ESR systems were 50.3 (SD 21.9), 51.5 (SD 14.2), and 63.0 (SD 11.3; contract research organization experts: 69.5, SD 11.5; clinicians: 59.8, SD 10.2), respectively. The Cronbach α

for the SUS items of the EMR, EDC, and ESR systems were 0.591 (95% CI -0.012 to 0.903), 0.588 (95% CI -0.288 to 0.951), and 0.785 (95% CI 0.576-0.916), respectively.

Conclusions: In real-world ophthalmology studies, the eSource approach based on the ESR system can replace the traditional clinical research model that relies on the manual transcription of data.

(*JMIR Form Res* 2022;6(12):e43229) doi:[10.2196/43229](https://doi.org/10.2196/43229)

KEYWORDS

electronic medical record; electronic health record; electronic source; eSource; eSource record tool; real-world data; data transcription; data quality; System Usability Scale; ophthalmology

Introduction

Background

An electronic source (eSource) generally includes the direct capture, collection, and storage of electronic data (eg, electronic medical records [EMRs], electronic health records [EHRs], or wearable devices) to simplify clinical research [1]. It can improve data quality and patient safety and reduce clinical trial costs. Despite the existence of several United States Food and Drug Administration guidelines [1,2] and European Medicines Agency guidelines [3], the development, implementation, and evaluation of EMR-specific electronic resource solutions are limited. Owing to known challenges such as the limited interoperability of EMRs and electronic data capture (EDC) systems, unstructured data (eg, researcher notes or comments), and the need for some data (eg, research-specific data not included in the EMR) to be manually transcribed and treated, accessing and correcting the source data in real time during data collection can be slow [4]. The direct use of EHR data in clinical research helps to improve data quality and reduce costs. Some research progress has already been achieved in eSource technology [4-6] in relatively large projects [7-9]. The review by Garza et al [5] included 14 studies detailing recent advances in eSource technology in clinical research. In total, 57% (8/14) of studies described single-site, single-EHR system implementation; 67% (4/6) of multisite studies were part of the same pilot study (EHR4CR European Pilot), a collaborative initiative across multiple European countries. Owing to the sensitivity of medical data, the variety of suppliers of medical information systems, and the low interoperability between medical systems, there is no similar eSource-related project in China. Therefore, it is common to use manual transcription data to conduct clinical research in China.

Real-world data (RWD) are data related to the patient health status and the delivery of health care that are routinely collected from a variety of sources [10]. A real-world study (RWS) collects RWD in a real-world environment and obtains real-world evidence (RWE) for the use value and potential benefits or risks of medical products through analysis [11]. Global regulatory agencies have issued a series of RWE-related guidelines, and researchers in different fields have shown interest in using RWD to conduct clinical research. Despite the availability of many relevant guidelines, various challenges in the use of RWD persist, such as inefficient data collection, lack of data quality control, and diversification of data standards and data compliance [12]. In China, data in the medical system are limited to the local area network, and external access and data

sharing cannot be performed. Owing to the inability to coordinate hospitals' concerns about the privacy of patients' medical data and the needs of researchers for data transparency, the transformation and upgrading of existing medical system suppliers cannot meet the requirements of clinical research. The use of eSource technology in RWS is expected to solve the challenges of data collection efficiency and data quality, thereby reducing the cost of relying on manual transcription data to conduct research.

eSource Record Project

In 2019, the China National Medical Products Administration established the Hainan Boao Lecheng Medical Tourism Pilot Zone as a pilot base for RWS, allowing domestic citizens to use global innovative products in China without first obtaining domestic market approval. Collecting and properly analyzing the RWD generated from patient visit data in Boao after using innovative medical products can help generate RWE that can be used for further domestic market approval. In 2020, the Hainan Real World Research Institute launched an eSource record (ESR) project to form an integrated solution and tool for hospital RWD collection, governance, and management. In this project, we completed the design and development of ESR software. The original intention of ESR was to provide a general tool that can implement RWS, thereby improving the implementation efficiency and quality of RWS research. Considering the need to develop many functions for old and underdeveloped medical supply systems to achieve automatic data capture and traceability of research data and to address the role of regulatory agencies and contract research organizations (CROs), we designed the ESR tool to meet the functional needs of clinical research. ESR tools can act as a bridge to connect EMR and EDC systems to achieve eSource technology. ESR was developed by a vendor to be used as a cost-saving method for conducting an RWS for sponsors.

After completing the development of the ESR tool, we tested and evaluated it in an RWS project in the Hainan Boao Lecheng Medical Tourism Pilot Zone. In a preliminary prospective study [13], we selected a cosmetic medical device project to evaluate the effect of ESR software on data collection and transcription. After completing the first evaluation of the ESR tool in the actual RWS project, we decided to expand the evaluation environment from simple RWS projects to more types of projects.

This Study

As previously used medical esthetics research projects are relatively simple [13], we selected a completed ophthalmic

femtosecond laser device project for retrospective data extraction and evaluation. Our team recently published a paper in the PharmaSUG China 2022 conference that introduced and evaluated a standardized method to convert source data to electronic case report form (eCRF) data in a real-world ophthalmology study [14]. This conference paper focused on the performance of natural language processing (NLP) in different types of data extraction. The accuracy of the data standardization method used in this study was 98.6%, but eCRF data completeness remained at 23.9% [14]. Through the case study, we summarized 2 key problems during the process of data transformation from real-world source data to eCRF data, namely, the lack of research-relevant source data and the complexity associated with the standardization of unstructured source data [14].

The purpose of this study was to evaluate the data transcription efficiency and quality of ESR software in retrospective studies to verify the feasibility of using eSource software as an alternative to traditional manual transcription of data in RWS projects. The ESR tool is currently in the exploration and evaluation stage; it is not yet fully mature and has not been pushed to the market for application. The contribution of this study was to provide more cases that use ESR tools in RWS projects to improve the design of tools and provide practical experience in exploring the application of eSource technology in China.

Methods

ESR Application Scenarios in the Hospital

ESR software links the EDC and EMR systems. The ESR system needs to receive the document format used in the EMR system and the reporting table field of the case in the EDC system and then convert the research field in the EDC system into writing suggestions in the EMR system document and send it to the EMR system. This strategy conforms to clinical doctors' routine writing habits and data needs in research and increases the interoperability between systems. More details on the ESR software are available in a previous study [13].

The deployment of the ESR system in the hospital ensures the safety of the medical data. The deployment plan is described below and mainly involves two aspects:

1. The source data of various paths are required for integrated research to form a copy of the certificate. The overall theoretical framework is written back to the EMR system, according to whether the source data supplemented by the ESR are divided into 2 schemes:
 - Plan A: ESR supplementary source data do not write back to the EMR system. The data of the health information system of the hospital move in one direction to the ESR system, and no other interaction is generated. In this mode, only data transmission interfaces from the clinical data repository of the hospital (EMR, laboratory information systems, and picture archiving and communication systems) to the ESR need to be established. The research source data required by clinicians only selected the research source

data for ESR (excluding the full data of the patient). ESR forms a copy of the hospital source to copy the database for backup.

- Plan B: ESR supplementary source data are written back to the EMR system. This scheme was applied after the clinician used the ESR system to collect research data according to the research plan. This part of the supplementary data was selected to synchronously write back to the EMR system. Therefore, based on the content of regular EMRs for medical care, some ESR supplementary data are posted to realize the transformation of medical records into a more detailed scientific research medical record process and meet the data collection requirements of clinical research. This model can be used to develop more interfaces to interact with the EMR system.
2. The copy of the database is used to form a clinical research database (ie, ESR and EDC system docking).

When the data required by the institute are not suitable for the conventional diagnosis and treatment process, ESR can be used as a supplementary source data recording tool to prepare patients' medical records through voice recognition technology and optical character recognition technology and to report adverse events. This function makes it convenient for doctors to efficiently complete supplementary collection of clinical research source data. When ESR is collected as a supplementary source, EMRs can be completed in a manner similar to conventional EMR systems in ESR. ESR can also follow the habits of doctors, configure structured meters for efficient data collection, and generate and record supplementary data required by clinical research.

ESR-Based RWS Implementation Process

The research mode for implementing RWS on the basis of ESR software can be summarized in the following five steps:

1. Determining the research plan: first, as the premise of clinical trials, researchers must provide RWS research solutions and eCRF to collect data.
2. Configuring the traceability path of the eCRF in the EMR or other source files: eCRF topics can be associated with the EMR form to configure the traceability paths of the different eCRF topics. For example, demographic data in the eCRF can be traced back to the admission record form in the EMR.
3. Clinicians collect medical records and source data according to the prompts of medical record writing: routine medical records do not record certain necessary research-specific data such as scale scores. Therefore, after completing the eCRF traceability configuration, clinicians can design medical record writing prompts and rules for the eCRF that conform to clinical habits and meet their data collection requirements to cover the elements required for research and standardize the EMR recording process among different clinicians.
4. NLP technology intelligently extracts structured research data from the certified copy database into an eCRF: by connecting different medical systems that connect hospitals, the ESR system summarizes the source data of the hospital

to form a copy of the database. The data outside the hospital and the source data of the EMRs recorded in ESR were entered into the certification copy database simultaneously. Using the NLP model, the ESR tool can capture data from free-text medical records. For structured data, ESR directly extracts data.

5. Data traceability and correction: ESR uses NLP technology to automatically extract data from the eCRF of the certification copy database in real time. This also supports the traceability of the source data for viewing these data. The clinical research coordinator (CRC) does not need to manually fill in the eCRF but can trace the source verification of the eCRF in the ESR. Through the traceability interface developed in the EDC system, the clinical research associate performs conventional source data verification and questioning work and sends queries to the ESR to remind clinicians to correct medical records.

Pilot Case Selection

The femtosecond laser project is a prospective, single-group, and observational RWS of a femtosecond laser eye therapy system (CATALYS Precision Laser System [CATALYST]) for actual clinical diagnosis and treatment. The project was approved for marketing in 2021 and can be used in a typical case of ESR system performance evaluation in retrospective studies. CATALYST clinical research data were initially recorded in hospital information systems including EMRs and then manually entered into the eCRF rather than filling in the

eCRF directly from eSource data. We selected 2 medical record forms (admission and surgical records) and the corresponding eCRFs for the research. The collected research data included the time spent on ESR system data transcription, correct rate of eCRF filling of the ESR system, and overall performance scores of different systems. The correct rate of eCRF filling referred to judging whether the filling value of the eCRF question was consistent with the source data according to the source data of the EMRs. After source data errors were excluded, if they were consistent, the question was completed correctly. By accessing log data from the EDC system previously used by the research project, all time spent from the start of the study to the end of the study could be counted. The time required for traditional manual transcription included the time required for data entry and correction by the CRC. The ESR software used NLP to automatically extract data from the EMRs in text form and used them to fill in the eCRF. When checking the source, the CRC needed to check the correctness of the fields that were filled in by the NLP system and needed to manually correct the incorrectly entered fields. The eSource transcription time included the CRC checking the correctness of the fields filled in by the NLP system and the time spent manually correcting incorrectly entered fields.

Times transcribed by the eSource software were timed using a stopwatch and manually recorded in Microsoft Excel. The data sources of the eCRF data variables and extraction method using the ESR system are listed in [Table 1](#).

Table 1. Data sources for electronic case report form data variables and extraction methods using the eSource record system.

Research variable	Data sources	Source data record type	Extraction method
Preoperative visit			
Subject information	Admission note (demographics)	Structured	Field mapping
Date of visit	Admission note (medical information)	Structured	Field mapping
Preoperative exam	Admission note (auxiliary examination)	Free text	NLP ^a technology
Ocular history and medications	Admission note (history)	Free text	NLP technology
Uncorrected distant visual acuity	Admission note (specialist examination)	Free text	NLP technology
Best corrected distance visual acuity	Admission note (specialist examination)	Free text	NLP technology
Manifest refraction	Admission note (specialist examination)	Free text	NLP technology
Slit-lamp exam	Admission note (specialist examination)	Free text	NLP technology
Intraocular pressure	Admission note (specialist examination)	Free text	NLP technology
Biometry	Admission note (auxiliary examination)	Free text	NLP technology
Intraocular lens power calculation	Admission note (auxiliary examination)	Free text	NLP technology
Corneal topography	Admission note (auxiliary examination)	Free text	NLP technology
Cataract status	Admission note (history of present illness)	Free text	NLP technology
Dilated fundus exam	Admission note (specialist examination)	Free text	NLP technology
Ocular symptoms	Admission note (history of present illness)	Free text	NLP technology
Intraoperative visit			
Surgery date	Operative report (operation time)	Structured	Field mapping
Corneal incision type	Operative report (procedure)	Free text	NLP technology
Capsulotomy size	Operative report (procedure)	Free text	NLP technology
Lens removal	Operative report (procedure)	Free text	NLP technology
Phacofragmentation	Operative report (procedure)	Free text	NLP technology
Viscoelastic agent	Operative report (procedure)	Free text	NLP technology
Intraocular lens	Operative report (procedure)	Free text	NLP technology
Surgical medication	Operative report (procedure)	Free text	NLP technology
Anesthesia	Operative report (procedure)	Free text	NLP technology
Type of closure	Operative report (procedure)	Free text	NLP technology
Other surgical procedures	Operative report (procedure)	Free text	NLP technology

^aNLP: natural language processing.

Ethics Approval

This study was conducted in accordance with the principles of the Declaration of Helsinki. Ethics approval was obtained from the Peking University Institutional Review Board (number IRB00001052–21081). Patient data were anonymized in accordance with the standards of clinical trials.

Implementation Process

We configured the project in the development environment of the ESR system according to the research protocol and eCRF information from a previous project. Ophthalmic clinicians and CRO Company members who were residents of the department and had participated in numerous RWS projects were invited to participate in our testing assignment. Clinicians screened the information of all previous patient admissions in the hospital's EMR system for the project. After exporting the basic

hospitalization information of 90 patients, the EMRs of 30 patients were randomly selected according to the number of hospitalizations required to test the ESR system. The admission notes and operative report document content exported from the EMR system were copied into the ESR system to generate the same patient records. Using the medical record writing template, clinicians processed and generated 30 text corpora using fictitious information to train the NLP model. During the process of generating test medical records, clinicians could explore and experience different data input methods for the ESR system, such as voice input and optical character recognition. According to the annotation guidelines, 20 text corpora were annotated by the medical specialists. The technicians trained the NLP model using the annotated text, generated predictions for the remaining 10 patient records, and manually corrected the extracted fields to achieve model optimization.

Data Standard Conversion Process

The EMR and EDC systems transmit data to the ESR system through the data standards of Health Level 7 Clinical Document Architecture and Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model, respectively. The core of the data conversion process is to formulate text data labels on the basis of the most simplified data model, improve the efficiency of the NLP algorithm, and optimize the interoperability of clinical data models and the standard term library required by auxiliary extraction research. This process includes the five steps listed as follows [15]. The CDISC model used in ophthalmology is presented in [Multimedia Appendix 1](#):

1. Send an eCRF to the ESR system from the EDC and send medical records from the EMR to the ESR system: the source data collection module of the ESR system is responsible for the EMRs and the collection of source data. The data transcription module of the ESR system is responsible for positioning the eCRF field to capture the text segment of the source data and to fill in the eCRF.
2. Research data collection models and the generation of labels: structural data are first mapped to the Observational Medical Outcomes Partnership model and then mapped to the CDISC model. Nonstructured data do not have a wide range of intermediate layers; they are directly converted to the CDISC model without considering the Observational Medical Outcomes Partnership model. The process of converting nonstructured data into research data is used to comment on and extract related content using the NLP model.
3. Extraction of model training and entity and entity relationships: in terms of physical extraction, a Chinese entity identification model, Bidirectional Encoder Representations from Transformers + Bidirectional Long Short-Term Memory Networks + Conditional Random Field, was adopted.
4. Generating a special research term: the dedicated term database is a mapping library between the actually extracted terms and standard terms in the indicator.
5. Standardization before entity extraction before filling in the eCRF: the output of the NLP model mainly includes the relationship table between all extracted label values and entities.

Data Conversion Metadata

The data conversion from source data to CRF fields includes the conversion of both structured and unstructured text data. The details of this step are provided in previous papers by our team [14]. The classification of the data conversion methods

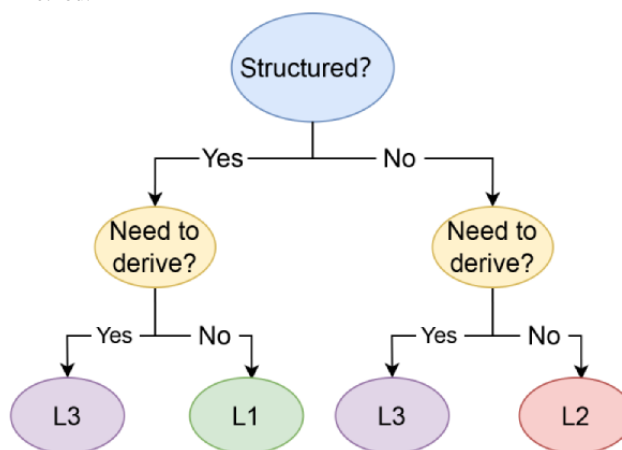
can be seen in decision trees ([Figure 1](#)). Several data conversion examples are provided in [Multimedia Appendix 2](#).

The first condition is whether the source data are structured or based on text. For surgery time fields (L1 category conversion), the structured source data can be directly used and converted into a standard time format. If the first condition is not met, it is necessary to use data transformation of the L2 category. NLP is used to extract the entity from the text. When extracting the entity, it is necessary to build an extraction rule such as using regular terms to extract related texts. To fill in a single research field, multiple entities must be extracted to determine which entity corresponds to the field in the report form. Therefore, after entity extraction, the output rules must be defined to extract field-related content.

After the physical extraction, regular extraction and term matching were performed. However, doctors' use of diagnosis and treatment has not been standardized. They often need to add unrecognized terms to existing standard terms. The dictionary adds costs such as labor costs and employees who need clinical knowledge to identify the term in the text and the relation of the term to the corresponding standard terms. When constructing the output rules, information technology and clinical knowledge experts must capture the habits of the physician and write rules that can be summarized. In retrospective studies, establishing rules is the most cost-effective job, because doctors usually change their descriptions.

The second condition in the decision tree is whether the field must be derived to obtain the relevant content of the field, which is classified as the data conversion method of L3. Derivative data can originate from text or structural data. Structural data can be used in simple derivative algorithms. For example, the age field can use the extracted birthday field to derive the age of a subject. Data derived from text types must be treated as structured and derived fields. Taking the history of the disease as an example, it is necessary to derive fields after the extraction entity, and the output rules are judged. Here, the derivative algorithm is merged with the output rules as follows: (1) if (Past History-Regular) OR (Previous History-Disease-Judgment) has nothing to do with (Negative Words), then output "Yes"; (2) if (Past History-Regular) OR (Previous History-Disease-Judgment) is related to (Negative Words), then output "No." L3 fields sometimes need to use multiple outputs to derive research field content. For example, the field can include whether there is a discovery other than the specified scope of research. It is necessary to first exclude the output within the specified range to address other findings. L3 data transformation should not involve inferred or subjective judgment such as determining whether the event is a significant adverse event.

Figure 1. Data conversion classification method.



Rating Scale

Usability includes effectiveness, efficiency, and overall user satisfaction. The System Usability Scale (SUS) is used to evaluate system usability [16]. The cross-industry average system availability scale score is 68, so this value is considered the threshold of acceptable availability. More details on SUS scores are available in previous studies [13]. The questionnaire consisted of the following 2 parts: participant information and SUS evaluation of the EMR, EDC, and ESR systems. In the questionnaire, the ESR system’s SUS score was the basic required question. For clinicians, the SUS score of the EMR system had to be completed simultaneously; for CRO company members, the SUS score of the EDC system had to be completed simultaneously.

Data Analysis

Descriptive statistical methods were used to analyze the population participating in the questionnaire and their SUS scores. The data analysis software used in this study was Python

(version 3.7.11). The Python *tableone* package (version 0.7.10) was used to generate demographic information for the questionnaires [17]. Cronbach α was used to evaluate the reliability of SUS. The Python *Pingouin* package (version 0.5.2) was used to calculate Cronbach α . A general acceptable range of Cronbach α is a value $\geq .70$ [18].

Results

eCRF Data Transcription Time

Admission notes and operative reports corresponded to the preoperative and intraoperative visit points of the eCRF, respectively. In the traditional method, the total entry time for the 51,424 fields at these 2 visit points was 11,738.85 minutes, that is, the average entry time for each field was 13.7 seconds. The eSource method required 6100 seconds for 2400 fields, which corresponded to an average entry time of 2.5 seconds per field. Therefore, the eSource method can save 11.2 seconds, that is, an overall time savings of 81.8%. The results are presented in Table 2.

Table 2. Data transcription times for the electronic source method (unit: seconds).

	Total number of eCRF ^a fields, n	Total time (s)	Time spent per patient (s), mean (SD)
Admission notes	1890	5173	172.4 (15.0)
Operative reports	510	927	30.9 (4.8)
Total	2400	6100	203.3 (16.9)

^aeCRF: electronic case report form.

eCRF Data Transcription Quality

When using the traditional method, among the 51,424 fields entered, 47,991 fields were entered correctly, and the correct rate of entry was 93.32%. In the eSource method, the total correct entry rate was 98.17% (2356/2400). Using the eSource

method for data extracted by NLP to fill in the wrong fields mainly focused on “slit-lamp examination lens” and “eye symptoms.” The main reason was that some described words were not in the dictionary of the NLP model, so they were not recognized. The results are presented in Table 3.

Table 3. Data transcription quality of the eSource method.

	Total number of eCRF ^a fields	Fields filled in correctly by NLP ^b , n (%)	CRC ^c -corrected fields, n (%)
Admission notes	1890	1849 (97.83)	41 (2.17)
Operative reports	510	507 (99.41)	3 (0.59)
Total	2400	2356 (98.17)	44 (1.83)

^aeCRF: electronic case report form.

^bNLP: natural language processing.

^cCRC: clinical research coordinator.

System Performance Evaluation Questionnaire

Questionnaires were sent to 15 individuals participating in the femtosecond laser project test, and 15 questionnaires were received for a response rate of 100%. The characteristics of the participants of the questionnaire survey are presented in [Table 4](#). In terms of usability, the average overall SUS scores of the

EMR, EDC, and ESR systems were 50.3 (SD 21.9), 51.5 (SD 14.2), and 63.0 (SD 11.3; CRO experts: 69.5, SD 11.5; clinicians: 59.8, SD 10.2), respectively. The Cronbach α for the SUS items of the EMR, EDC, and ESR systems were 0.591 (95% CI -0.012 to 0.903), 0.588 (95% CI -0.288 to 0.951), and 0.785 (95% CI 0.576-0.916), respectively.

Table 4. The characteristics of the population participating in the questionnaire.

Items	Total (N=15)	CRO ^a experts (n=5)	Clinicians (n=10)
Sex, n (%)			
Female	9 (60)	5 (100)	4 (40)
Male	6 (40)	0 (0)	6 (60)
Age (years), mean (SD)	27.3 (4.4)	27.8 (1.5)	27.1 (5.4)
Profession, n (%)			
Clinical research associate	1 (7)	1 (20)	0 (0)
Clinical research coordinator	4 (27)	4 (80)	0 (0)
Clinician	10 (67)	0 (0)	10 (100)
Highest level of education, n (%)			
College degree and below	3 (20)	0 (0)	3 (30)
Undergraduate	4 (27)	3 (60)	1 (10)
Postgraduate	8 (53)	2 (40)	6 (60)
Experience in the medical field, n (%)			
1-3 years	7 (47)	2 (40)	5 (50)
≥10 years	1 (7)	0 (0)	1 (10)
<1 year	3 (20)	0 (0)	3 (30)
4-6 years	4 (27)	3 (60)	1 (10)
Frequency of EMR^b system use, n (%)			
Regular	2 (13)	0 (0)	2 (20)
Not applicable	7 (47)	5 (100)	2 (20)
Occasional use	1 (7)	0 (0)	1 (10)
Use every day	2 (13)	0 (0)	2 (20)
Frequent use	3 (20)	0 (0)	3 (30)
Frequency of EDC^c system use, n (%)			
Regular	1 (7)	1 (20)	0 (0)
Not applicable	10 (67)	0 (0)	10 (100)
Use every day	3 (20)	3 (60)	0 (0)
Frequent use	1 (7)	1 (20)	0 (0)

^aCRO: contract research organization.

^bEMR: electronic medical record.

^cEDC: electronic data capture.

Discussion

Principal Findings

This study provides specific examples of the use of ESR software in ophthalmology equipment to transform RWD into research data. Similar to previous case results from prospective studies [13], we found that the eSource approach was superior to manual data transcription in terms of data transcription quality, indicating that the improvement in data transcription efficiency provided by ESR systems does not sacrifice data quality. In the practice of the project, we found that although data transformation had good accuracy, more effective ways to improve the completeness of research data and implementation

efficiency were lacking. Retrospective research typically does not use data standards before and during collection, resulting in difficulties in the process of data standard conversion in the later period. Therefore, data cannot be appropriately applied to eCRF, and information may be lost. This limitation explains why the data collection method used in prospective research based on research data standards is very important because it may directly affect the quality of the data from the starting point of data collection, thereby achieving standard conversion from RWD to clinical research data.

The main problems of low completeness include the lack of content specified in the description provided by the doctor recording the disease, lack of standard expression methods, and

logic of describing clinical events. The process of development and transformation models must also be optimized, because they can only modularize a small number of fields but do not classify the same type of data, resulting in low development efficiency. In addition, the decrease in development efficiency is mainly because the clinical events described by doctors and the information required by the study are difficult to match. The overall data conversion process is long, and a real-time record of the details of all conversion processes and timely feedback on the quality of the data are unavailable. Therefore, the standardization of the doctor's record, application of the data model to reduce the repeated transformation of the research field, details of the process of data transformation, and quality of the feedback data are necessary to optimize completeness and efficiency.

After summarizing the problems encountered, we proposed some possible solutions. First, the problem of the lack of matching of the research term with doctors' habits may be solved by collecting commonly used terms to expand the coverage of the term and match the term using more automated approaches. Second, regarding the differences in context descriptions between doctors, some clinicians have proposed the use of recommended texts to promote the consistency of clinical event descriptions. Third, the lower development efficiency caused by the differences between fields may be improved by implementing suggestions for using data models. Finally, the problems of timely recording, data standardization, and feedback data quality may be solved by establishing source data management platforms to strengthen the source data and transparency of the data standardization process.

In retrospective case studies, the rate of complete data extraction is affected by the degree of consistency and vocally described by doctors. The solution at the time was to send the recommended text to the hospital EMR through ESR to strengthen the consistency of the physician's description. However, from the perspective of experience, doctors still use different expressions for research data and efficiency cannot be improved. Therefore, technicians should use the NLP algorithm to extend the extraction rules and lists effectively. The improvement in the efficiency of artificial intelligence technology was attributed to the inclusion of each sample in the learning sheet of the model. In addition, the operating threshold during development was low and the efficiency was high. As long as the technical personnel responsible for the medical information can identify the entity and entity relationships in the text content, the model can automatically learn.

In a retrospective research environment, semantics allows researchers to affect the direction of data collection. After linking the hospital EMR documents and research data, clinical researchers and medical experts used research data models to generate text suggestions. When recommended terms are generated, researchers must consider methods that facilitate doctors' use of the standard of research employed by the research terms to enable the collection of research data. According to the needs of field-specific terms for research, auxiliary instructions must be added to the prompt of medical records. However, if a relatively large standard library is needed, such as the Eleventh Revision of the International Classification

of Diseases, we expect that the term query mode will be switched during the writing process such that the original terms and selected standard terms can be recorded. Therefore, this method can collect conventional terminology and match standard terms.

When applying ESR to real-world ophthalmology studies, we believe that the following conditions must be met: (1) first, specific research protocols and eCRF exist, and the research design is prospective research; (2) researchers must design a recommended text template for hospital EMRs according to the requirements of the research data, thereby promoting the standardization of data records; and (3) modules covering surgical medication, eye symptoms, additional surgery, and surgery have many prespecified events, such as whether star-shaped eye symptoms are observed. These events are rarely recorded in the research data. If technicians negotiate with the researchers in advance, the content that is not included may represent the incident or the researchers must clearly describe the content, which will improve the completeness of the eCRF data.

Similar to the EHR (hospital or clinic) SUS scores [19,20], we found that the hospital EMR SUS scores were also <60 points, which is the F level. In terms of system usability, the SUS score of ESR software obtained in this study was 63. Our research found that in hospital EMR and EDC, the SUS Cronbach α value was <.6. According to the classification of Cronbach α [18], a result ≤ 6 indicates poor reliability of the hospital EMR and EDC SUS. The cause of this phenomenon may be related to the lower number of participants in our survey. However, the Cronbach α of the ESR SUS was >.7, which is acceptable. This further increases our confidence in ESR SUS scores.

The Duke University Clinical Research Institute introduced the RADaptor tool [21,22] as a solution to improve the efficiency of clinical research. The RADaptor tool acts as an intermediate plug-in to connect the EHR system (Epic) and the eCRF of REDCap (Research Electronic Data Capture; Vanderbilt University) software. The study by Nordo et al [22] used the RADaptor tool for case validation and evaluation and showed that this tool outperformed the traditional manual data transcription process. Unlike ESR systems that can be used for RWS, the current scope of the RADaptor tool is limited to single-site registration data. Following the eSource initiative of TransCelerate in 2016 [23], the latest work has now transitioned to the Health Level 7 Project Vulcan Fast Healthcare Interoperability Resources Accelerator [24]. The Vulcan project aims to help health care researchers more efficiently acquire, exchange and use, and translate data in clinical research using its widely recognized data exchange standard.

Limitations

In terms of limitations, our study was limited by investigator selection and representativeness, similar to previous studies evaluating RWS projects on medical aesthetics [13]. Furthermore, we were unable to obtain authorization to deploy the ESR system in the hospital intranet at the outset because of the need to demonstrate the value of the ESR software in an ophthalmology-based RWS project to hospital administrators. Therefore, we did not select all cases for evaluation but only

sampled some cases to reduce the time spent transferring medical records from the hospital EMR system to the ESR system. Finally, the premise of this study is that the necessary fields for the study have been recorded in the previous CATALYST study medical records. Therefore, the extraction effect of ESR software largely depends on the method of scientific medical records, which is why we advocate project-based RWS. A homogenized medical record template will facilitate the standardization of terminology records in prospective studies, thereby reducing the burden of late NLP technology extraction.

Conclusions

In ophthalmic RWS, the eSource approach based on the ESR system can replace the traditional clinical research model that relies on manual transcription of data. On the basis of this specific case, we provide experience in applying eSource technology to the transformation of RWD into research data. Follow-up research will focus on the deep functional integration of ESR and EMR systems to cope with complex research projects and optimize the RWS implementation process on the basis of the eSource approach.

Acknowledgments

This study was supported by the China Center for Food and Drug International Exchange.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

All the authors contributed to the study. BW wrote the first draft of the manuscript. CY conceived the idea of this study. BW, JL, and FJ participated in the preliminary design of the method. ML provided guidance from the ophthalmic clinician's perspective. YP provided technical guidance for the clinical research on medical devices. BW conducted questionnaire surveys, data collection, and analysis. JL provided technical support for data standardization and natural language processing technology. CY provided comments and revised the manuscript accordingly. All the authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Clinical Data Interchange Standards Consortium Model used in the ophthalmology field.

[PDF File (Adobe PDF File), 133 KB - [formative_v6i12e43229_app1.pdf](#)]

Multimedia Appendix 2

Several data conversion examples.

[PDF File (Adobe PDF File), 214 KB - [formative_v6i12e43229_app2.pdf](#)]

References

1. Guidance for industry electronic source data in clinical investigations. Food and Drug Administration (FDA). 2013 Sep. URL: <https://www.fda.gov/media/85183/download> [accessed 2021-08-24]
2. Use of electronic health record data in clinical investigations guidance for industry. Food and Drug Administration (FDA). 2018 Jul. URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry> [accessed 2021-08-24]
3. Qualification opinion on eSource Direct Data Capture (DDC). European Medicines Agency (EMA). 2019 Jul 25. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-esource-direct-data-capture-ddc_en.pdf [accessed 2021-10-16]
4. Parab AA, Mehta P, Vattikola A, Denney CK, Cherry M, Maniar RM, et al. Accelerating the adoption of eSource in clinical research: a transclerate point of view. *Ther Innov Regul Sci* 2020 Sep 03;54(5):1141-1151 [FREE Full text] [doi: [10.1007/s43441-020-00138-y](https://doi.org/10.1007/s43441-020-00138-y)] [Medline: [32128701](https://pubmed.ncbi.nlm.nih.gov/32128701/)]
5. Garza M, Myneni S, Nordo A, Eisenstein E, Hammond W, Walden A. Esource for standardized health information exchange in clinical research: a systematic review. *Stud Health Technol Inform* 2019;257. [doi: [10.47912/jscdm.66](https://doi.org/10.47912/jscdm.66)]
6. Ethier J, Curcin V, McGilchrist MM, Choi Keung SN, Zhao L, Andreasson A, et al. eSource for clinical trials: implementation and evaluation of a standards-based approach in a real world trial. *Int J Med Inform* 2017 Oct;106:17-24 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.06.006](https://doi.org/10.1016/j.ijmedinf.2017.06.006)] [Medline: [28870379](https://pubmed.ncbi.nlm.nih.gov/28870379/)]

7. Rocca M, Asare A, Esserman L, Dubman S, Gordon G. Source data capture from EHRs: using standardized clinical research data. U.S. Food & Drug Administration. 2019 Oct 30. URL: <https://www.fda.gov/media/132130/download> [accessed 2021-08-16]
8. Hussain S, Ouagne D, Sadou E, Dart T, Daniel C. EHR4CR: a semantic web based interoperability approach for reusing electronic healthcare records in protocol feasibility studies. CEUR-WS. 2012. URL: http://ceur-ws.org/Vol-952/paper_31.pdf [accessed 2021-10-12]
9. Delaney B, Curcin V, Andreasson A, Arvanitis T, Bastiaens H, Corrigan D, et al. Translational medicine and patient safety in Europe: TRANSFoRM--Architecture for the learning health system in Europe. Biomed Res Int 2015;2015:961526 [FREE Full text] [doi: [10.1155/2015/961526](https://doi.org/10.1155/2015/961526)] [Medline: [26539547](https://pubmed.ncbi.nlm.nih.gov/26539547/)]
10. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-world evidence - what is it and what can it tell us? N Engl J Med 2016 Dec 08;375(23):2293-2297. [doi: [10.1056/NEJMsb1609216](https://doi.org/10.1056/NEJMsb1609216)] [Medline: [27959688](https://pubmed.ncbi.nlm.nih.gov/27959688/)]
11. Real-World Data Used to Generate Real-World Evidences (Interim). Center for Drug Evaluation of NMPA. 2021. URL: <http://zhengbiaoke-com.oss-cn-beijing.aliyuncs.com/2104/2104-773f72d9-7fc4-4f02-ae13-ad4aeafc2f98.pdf> [accessed 2022-02-11]
12. Grimberg F, Asprion PM, Schneider B, Miho E, Babrak L, Habbabeh A. The real-world data challenges radar: a review on the challenges and risks regarding the use of real-world data. Digit Biomark 2021 Jun 24;5(2):148-157. [doi: [10.1159/000516178](https://doi.org/10.1159/000516178)]
13. Wang B, Hao X, Yan X, Lai J, Jin F, Liao X, et al. Evaluation of the clinical application effect of eSource record tools for clinical research. BMC Med Inform Decis Mak 2022 Apr 11;22(1):98 [FREE Full text] [doi: [10.1186/s12911-022-01824-7](https://doi.org/10.1186/s12911-022-01824-7)] [Medline: [35410214](https://pubmed.ncbi.nlm.nih.gov/35410214/)]
14. Lai J, Yao C, Wang B, Wang K, Liao X. Real world data standardization for an ophthalmology real world study in china and exploration of new eSource solutions. PharmaSUG. 2022. URL: <https://www.pharmasug.org/proceedings/china2022/RW/Pharmasug-China-2022-RW146.pdf> [accessed 2022-10-26]
15. Lai J, Wang B, Yao C, Ren Y, Jin F, Wang K. Research on standards transformation from real world data to clinical research data. China Food Drug Administration Magazine 2021(11):39-46.
16. Jordan P, Thomas B, McClelland I, Weerdmeester B. Usability Evaluation In Industry. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 1996.
17. Pollard T, Johnson A, Raffa J, Mark R. tableone: an open source Python package for producing summary statistics for research papers. JAMIA Open 2018 Jul;1(1):26-31 [FREE Full text] [doi: [10.1093/jamiaopen/ooy012](https://doi.org/10.1093/jamiaopen/ooy012)] [Medline: [31984317](https://pubmed.ncbi.nlm.nih.gov/31984317/)]
18. Taber KS. The use of Cronbach's Alpha when developing and reporting research instruments in science education. Res Sci Educ 2017 Jun 7;48(6):1273-1296. [doi: [10.1007/s11165-016-9602-2](https://doi.org/10.1007/s11165-016-9602-2)]
19. Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The association between perceived electronic health record usability and professional burnout among US physicians. Mayo Clin Proc 2020 Mar;95(3):476-487 [FREE Full text] [doi: [10.1016/j.mayocp.2019.09.024](https://doi.org/10.1016/j.mayocp.2019.09.024)] [Medline: [31735343](https://pubmed.ncbi.nlm.nih.gov/31735343/)]
20. Melnick E, West C, Nath B, Cipriano P, Peterson C, Satele D, et al. The association between perceived electronic health record usability and professional burnout among US nurses. J Am Med Inform Assoc 2021 Jul 30;28(8):1632-1641 [FREE Full text] [doi: [10.1093/jamia/ocab059](https://doi.org/10.1093/jamia/ocab059)] [Medline: [33871018](https://pubmed.ncbi.nlm.nih.gov/33871018/)]
21. Evaluation of the effect of RFD functionality on completion of Clinical Research data collection. Duke University School of Medicine. URL: https://dukeinformatics.org/wp-content/uploads/2017/12/FDA-CDER-2017_RADaptor-003.pdf [accessed 2021-12-09]
22. Nordo AH, Eisenstein EL, Hawley J, Vadakkevedu S, Pressley M, Pennock J, et al. A comparative effectiveness study of eSource used for data capture for a clinical research registry. Int J Med Inform 2017 Jul;103:89-94 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.04.015](https://doi.org/10.1016/j.ijmedinf.2017.04.015)] [Medline: [28551007](https://pubmed.ncbi.nlm.nih.gov/28551007/)]
23. Kellar E, Bornstein SM, Caban A, Célingant C, Crouthamel M, Johnson C, et al. Optimizing the use of electronic data sources in clinical trials: the landscape, part 1. Ther Innov Regul Sci 2016 Nov 30;50(6):682-696. [doi: [10.1177/2168479016670689](https://doi.org/10.1177/2168479016670689)] [Medline: [30231749](https://pubmed.ncbi.nlm.nih.gov/30231749/)]
24. HL7 starts project Vulcan FHIR accelerator program. Electronic Health Reporter. 2020 Sep 14. URL: <https://tinyurl.com/8e92ymyh> [accessed 2020-10-08]

Abbreviations

CATALYST: CATALYS Precision Laser System
CDISC: Clinical Data Interchange Standards Consortium
CRC: clinical research coordinator
CRO: contract research organization
eCRF: electronic case report form
EDC: electronic data capture
EHR: electronic health record
EMR: electronic medical record

eSource: electronic source
ESR: electronic source record
NLP: natural language processing
REDCap: Research Electronic Data Capture
RWD: real-world data
RWE: real-world evidence
RWS: real-world study
SUS: System Usability Scale

Edited by A Mavragani; submitted 05.10.22; peer-reviewed by D Chrimes, KM Kuo; comments to author 24.10.22; revised version received 09.11.22; accepted 23.11.22; published 16.12.22.

Please cite as:

Wang B, Lai J, Liu M, Jin F, Peng Y, Yao C

Electronic Source Data Transcription for Electronic Case Report Forms in China: Validation of the Electronic Source Record Tool in a Real-world Ophthalmology Study

JMIR Form Res 2022;6(12):e43229

URL: <https://formative.jmir.org/2022/12/e43229>

doi: [10.2196/43229](https://doi.org/10.2196/43229)

PMID: [36525285](https://pubmed.ncbi.nlm.nih.gov/36525285/)

©Bin Wang, Junkai Lai, Mimi Liu, Feifei Jin, Yifei Peng, Chen Yao. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 16.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Assessing and Promoting Cardiovascular Health for Adolescent Women: User-Centered Design Approach

Kolbi Bradley¹, BA; Santiago J Arconada Alvarez^{2,3*}, MSc; Amanda K Gilmore^{4,5*}, MSc, PhD; Morgan Greenleaf^{2,3*}, MSc; Aayahna Herbert^{6*}, BA; Melissa J Kottke^{7*}, MPH, MBA, MD; Maren Parsell^{2,8*}, MBA; Sierra Patterson^{9*}, MPH; Tymirra Smith^{10*}, BA; Mercedes Sotos-Prieto^{11,12,13*}, MSc, PhD; Elizabeth Zeichner³, BA; Holly C Gooding^{1,14*}, MSc, MD

¹Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, United States

²Georgia Clinical and Translational Science Alliance, Atlanta, GA, United States

³Emory University School of Medicine, Atlanta, GA, United States

⁴Department of Health Policy and Behavioral Sciences, School of Public Health, Georgia State University, Atlanta, GA, United States

⁵National Center for Sexual Violence Prevention, Mark Chaffin Center for Healthy Development, School of Public Health, Georgia State University, Atlanta, GA, United States

⁶College of Computing, School of Interactive Computing, Georgia Tech, Atlanta, GA, United States

⁷Jane Fonda Center, Department of Gynecology and Obstetrics, Emory University School of Medicine, Atlanta, GA, United States

⁸Emory Healthcare, Atlanta, GA, United States

⁹Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

¹⁰College of Design, School of Industrial Design, Georgia Tech, Atlanta, GA, United States

¹¹Department of Preventive Medicine and Public Health, School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

¹²Centro de Investigacion Biomedica en Red Epidemiologica y Salud Publica, Madrid, Spain

¹³Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, United States

¹⁴Children's Healthcare of Atlanta, Atlanta, GA, United States

*these authors contributed equally

Corresponding Author:

Holly C Gooding, MSc, MD

Department of Pediatrics

Emory University School of Medicine

49 Jesse Hill Jr Dr SE

Atlanta, GA, 30303

United States

Phone: 1 4047781429

Email: holly.gooding@emory.edu

Abstract

Background: Cardiovascular disease (CVD) is the leading cause of death among women in the United States. A considerable number of young women already have risk factors for CVD. Awareness of CVD and its risk factors is critical to preventing CVD, yet younger women are less aware of CVD prevalence, its risk factors, and preventative behaviors compared to older women.

Objective: The purpose of this study is to assess CVD awareness among adolescent and young adult women and develop a lifestyle-based cardiovascular risk assessment tool for the promotion of CVD awareness among this population.

Methods: This study used a 3-phase iterative design process with young women and health care practitioners from primary care and reproductive care clinics in Atlanta, Georgia. In phase 1, we administered a modified version of the American Heart Association Women's Health Survey to young women, aged 15-24 years (n=67), to assess their general CVD awareness. In phase 2, we interviewed young women, aged 13-21 years (n=10), and their health care practitioners (n=10), to solicit suggestions for adapting the Healthy Heart Score, an existing adult cardiovascular risk assessment tool, for use with this age group. We also aimed to learn more about the barriers and challenges to health behavior change within this population and the clinical practices that serve them. In phase 3, we used the findings from the first 2 phases to create a prototype of a new online cardiovascular risk assessment tool designed specifically for young women. We then used an iterative user-centered design process to collect feedback from approximately 105 young women, aged 13-21 years, as we adapted the tool.

Results: Only 10.5% (7/67) of the young women surveyed correctly identified CVD as the leading cause of death among women in the United States. Few respondents reported having discussed their personal risk (4/67, 6%) or family history of CVD (8/67, 11.9%) with a health care provider. During the interviews, young women reported better CVD awareness and knowledge after completing the adult risk assessment tool and suggested making the tool more teen-friendly by incorporating relevant foods and activity options. Health care practitioners emphasized shortening the assessment for easier use within practice and discussed other barriers adolescents may face in adopting heart-healthy behaviors. The result of the iterative design process was a youth-friendly prototype of a cardiovascular risk assessment tool.

Conclusions: Adolescent and young adult women demonstrate low awareness of CVD. This study illustrates the potential value of a cardiovascular risk assessment tool adapted for use with young women and showcases the importance of user-centered design when creating digital health interventions.

(*JMIR Form Res* 2022;6(12):e42051) doi:[10.2196/42051](https://doi.org/10.2196/42051)

KEYWORDS

adolescent; heart disease; mHealth; digital health intervention; user-centered design; cardiovascular disease; CVD; women's health; risk assessment; young adults; assessment tool

Introduction

Cardiovascular disease (CVD) remains the leading cause of death for women in the United States, despite decades of progress in risk factor detection and treatment [1,2]. Awareness of CVD and its risk factors is critical to reversing this trend, yet national data from the American Heart Association (AHA) Women's Health Survey reveals that younger women aged 25-34 years are much less likely than older women to be aware of CVD and its risk factors, a trend that is worsening over time [3]. A previous study of young women aged 18-39 years participating in the National Health and Nutrition Examination Survey in 2011-2014 found a considerable number already had hypercholesterolemia, hypertension, and diabetes, yet many were unaware of their diagnosis [4]. Compounding this lack of CVD awareness, people of all ages tend to underestimate the likelihood of experiencing negative events and overestimate the likelihood of experiencing positive events [5]. Adolescents especially have been characterized as "young invincibles" with little consideration for their long-term health [6]. However, adolescents are cognitively capable of understanding short-, medium-, and long-term risks for CVD [7] and have been shown in prior studies to report motivation to act now to prevent future CVD [8].

Previous interventions to improve CVD awareness and mitigate CVD risk in adolescent and young adult (AYA) women have been limited to distinct populations, such as college students [9], or have focused on addressing only a few CVD risk factors at a time [10-13]. Mobile health (mHealth) interventions, defined as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [14], are increasingly used to engage patients in CVD health promotion and behavior change [15-17]. There is a need to further explore the use of mHealth tools in diverse populations [18], including adolescents and women, and to optimize their use within clinical practice [17].

The Healthy Heart Score (HHS) is an existing online assessment tool that estimates one's risk for CVD based on self-reported modifiable health behaviors [19]. The HHS was developed using epidemiologic data from adults of 40-75 years participating in

the Nurses' Health Study and Health Professionals Follow-Up Study, who were then followed for over 24 years [19]. An algorithm consisting of age, smoking status, physical activity, diet, alcohol intake, and body mass index was found to adequately predict their 20-year risk of a CVD event. We subsequently validated the use of this same algorithm in Black and White young adults aged 18-30 years participating in the Coronary Artery Risk Development in Young Adults cohort study and found that it predicted risk of a CVD event before 55 years of age as well [20]. The HHS is especially well suited for use with younger women because it does not require clinical measurements or laboratory assessments, assesses modifiable health behaviors, and performs best in those without established CVD risk factors [21].

In this project, we aimed to create a developmentally appropriate mobile CVD risk assessment tool based on the HHS algorithm for use in wellness, mental health, and reproductive health visits for young women in a diverse community in Atlanta, Georgia. In the first phase of the study, we assessed the baseline understanding of CVD and its risk factors in our participant population using the AHA Women's Health Survey [3] to target the new tool at existing knowledge gaps. In the second phase, we presented the existing adult HHS online risk assessment to young women and their health care providers (HCPs) and solicited their ideas for necessary adaptations to make the tool more teen-friendly. In the third phase, we used an iterative user-centered design process to continuously adapt and test revisions to the tool. The final product is an online tool developed with and for young women to address identified gaps in their CVD awareness and promote positive behavior change.

Methods

Study Population

AYA women, aged 13-24 years, were recruited from a primary care practice and an adolescent reproductive health clinic in southeast Atlanta for all 3 phases of the study. The adolescent reproductive health clinic was selected for recruitment in addition to the primary care clinic because many young women are only present for reproductive health care [22], and in our prior work, young women suggested tying cardiovascular health

promotion to their reproductive health concerns as a viable prevention strategy [23]. Southeast Atlanta is a diverse location of Atlanta, with over 70% of the population identifying as Black or African American [24]. The majority of households, served by these 2 practices, report under US \$50,000 per year in annual income [24]. The primary care practice provides primary care, mental health care, and sexual health care services to over 1700 adolescents aged 13-21 years annually and is staffed by adolescent medicine physicians, nurse practitioners, a psychotherapist, and a supporting clinical team. The reproductive health clinic provides confidential sexual health services to young women aged 13-24 years and is staffed by gynecologists, nurse midwives, a health educator, and a supporting clinical team.

Ethical Considerations

This study was approved by the Emory University Institutional Review Board, the Children's Healthcare of Atlanta Institutional Review Board, and the Grady Health System Research Oversight Committee (approval number 00001418). Informed consent was obtained from all surveyed and interviewed participants and their guardians, except in cases where minors presented alone to the clinic, in which an approved waiver of guardian consent was invoked. All data collected from surveys and interviews were deidentified and stored on Emory University's password-protected servers. A US \$25 gift card was given to those who participated in surveys and recorded interviews. A waiver of written informed consent was approved by the aforementioned review boards in cases of observational feedback. No identifying information was recorded for observational feedback, and no financial compensation was provided.

Study Procedures

Phase 1: Baseline Assessment of Adolescent Understanding of CVD Risk

Young women presenting for annual well visits from March 2021 to December 2021 were eligible for this phase of the study. Trained research coordinators approached young women in the waiting room and asked them to complete a modified version of the AHA Women's Health Study survey [25]. Participants provided electronic informed consent before agreeing to complete the survey. A total of 67 young women completed the survey. We calculated and reported descriptive statistics using SAS 9.4 (SAS Institute).

Phase 2: Adolescent and HCP Feedback on the Existing CVD Screening Tool (HHS)

Young women and HCPs from both practices were eligible for this phase of the study, which occurred from August 2021 to December 2021. All participants completed the existing HHS CVD risk assessment online and answered questions administered by a member of the research team, including feedback on the HHS architecture (layout, word choice, use of visuals), conditions under which they thought adolescents would engage with the assessment (eg, at school or the doctor's office), and general suggestions for improvement. Four initial interviews occurred over 1 hour using Health Insurance Portability and

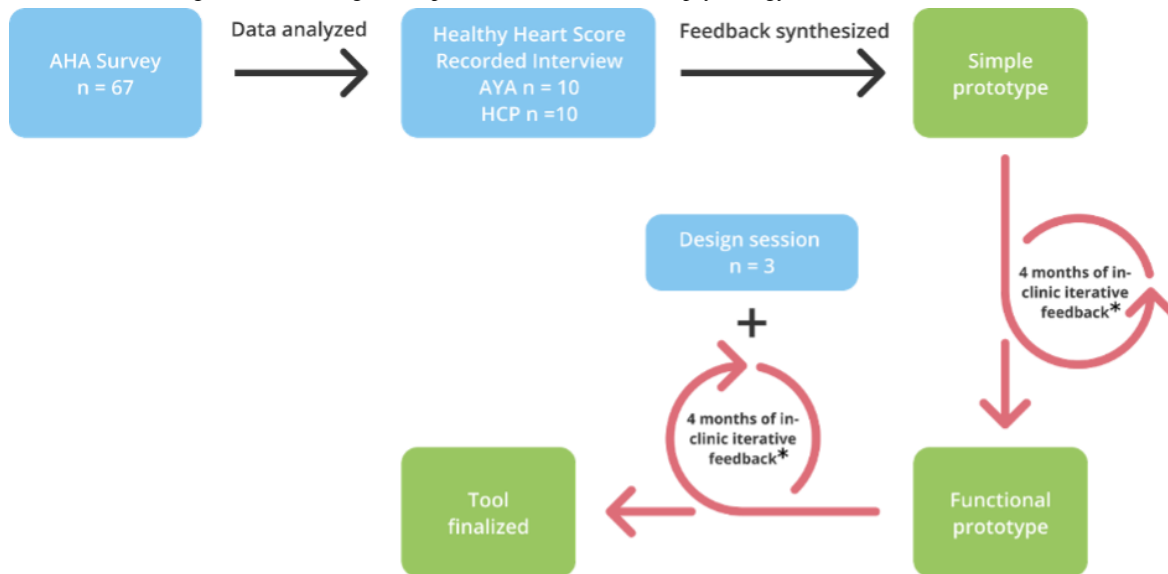
Accountability Act-compliant Zoom during the COVID-19 pandemic. Due to difficulty scheduling the Zoom interviews, we switched to a 15-minute semistructured interview in the clinics in October 2021, addressing the same topics in an abbreviated form. While interviews were conducted in quiet areas, we allowed friends, family members, or guardians to be present during interviews if preferred by the participant. We compared the results from the in-person interviews to those from the Zoom interviews to check for differences; no qualitative differences were found in the results. A total of 11 adolescents and 10 HCPs completed interviews. All interviews were recorded and audio-transcribed verbatim. Due to the failure to record, 1 interview was lost, leaving 10 AYA interviews for textual analysis. All participants in this phase provided written informed consent to be interviewed and recorded. The HCP interview guide is available in the [Multimedia Appendix 1](#) and the AYA interview guide is available in the [Multimedia Appendix 2](#).

Transcriptions were read by the research team members, including an adolescent medicine physician, a clinical research coordinator, and a medical student. Using an inductive coding approach, the research team met to discuss themes and decide on a set of codes and subcodes (see [Table 1](#)). Transcriptions were then coded in teams of 2 using Dedoose, version 8.3.10 (SocioCultural Research Consultants). After double-coding the first 3 transcripts, the team met to review the code application and ensure agreement. Any disagreements were discussed during the team meeting. This process was repeated to ensure consistency in coder agreement.

Phase 3: Adolescent CVD Screening Tool Development

The clinical research team collaborated with the programming and design team to synthesize findings from the AHA survey from phase 1 and HHS interviews from phase 2 into a simple wireframe prototype of a new CVD risk assessment tool for phase 3. A member of the clinical research team then presented the prototype to young women in the waiting rooms of the 2 practices from October 2021 to June 2022. Patients were asked to provide brief feedback on various topics, including layout, content, usability, and areas for improvement. The research team member took notes and presented this feedback to the programming and design teams to inform subsequent iterations of the tool. All data was observational and anonymous, and written informed consent was not required per the institutional review board approvals for this phase. During the fourth month of this user-centered design process, the team facilitated an hour-long design session with 3 participating adolescents to supplement in-clinic feedback using the Ten Design Heuristics [26]. The team consulted with a clinical psychologist and nutritional epidemiologist throughout this iterative design phase to ensure that adapted features and content appropriately used behavioral and health psychology frameworks such as motivational interviewing [27], the health belief model (HBM) [28], the behavioral intervention technology model [29], and social cognitive theory [30], as described further in the Results section. The complete iterative design process, including all 3 phases, is depicted in [Figure 1](#).

Figure 1. Study flow chart. AHA: American Heart Association; AYA: adolescent and young adult; HCP: health care provider. *A nutritional epidemiologist and clinical psychologist were consulted throughout phase 3's iterative design process to ensure the tool was aligned with principles of motivational interviewing, accurate serving size depictions, and several health psychology frameworks.



Results

Phase 1: Baseline Assessment of Young Women's Understanding of CVD Risk

Of the 67 participants who completed the survey, 47 (71.2%) were 15-17 years of age, 18 (27.3%) were 18-21 years of age, 1 was older than 21 years, and 1 did not state her age. Most respondents identified as Black (56/67, 75.4%) and 13 of 67 (22.4%) identified as Hispanic. The majority of the participants (48/67, 71.6%) reported that their primary caregiver had less than a college education. Only 10.5% (7/67) of the participants were aware that CVD is the leading cause of death among women in the United States. Participants were asked which health topics they discussed with their HCP. Very few reported that their HCP discussed their personal risk of CVD (4/67, 6%) or their family history of CVD (8/67, 11.9%) with them. Of the remaining topics surveyed, 46.3% (31/67) reported discussing their exercise habits; 49.3% (33/67) reported discussing their weight; and 26.9% (18/67) reported discussing none of the listed topics with their provider. Participants were also asked to identify major causes of CVD. Less than half of participants were aware that low physical activity (19/67, 28.4%), diabetes (27/67, 40.3%), smoking (29/67, 43.3%), and high cholesterol (33/67, 49.3%) contribute to CVD. High blood pressure (37/67, 55.3%) and being overweight (44/67, 65.7%) were identified as major causes by over half of respondents. Participants were also asked what CVD preventative behaviors they took within the past year. Fewer than half of the participants reported any of the CVD prevention behaviors surveyed, with only 4.5% (3/67) reporting maintaining healthy cholesterol, 10.4% (7/67) reporting maintaining a healthy blood pressure, 11.9% (8/67) reporting reducing sodium/salt intake; 23.9% (16/67) reporting reducing sugar intake; 38.9% (26/67) reporting attempting to lose weight, and 49.3% (33/67) reporting getting physical exercise.

Phase 2: Adolescent Feedback on the Existing Adult CVD Screening Tool

A total of 10 young women aged 13-21 years and 10 HCPs aged 29-65 years contributed interviews for the analysis. All the adolescent participants identified themselves as Black. All HCPs identified as women, with 6 identifying as Black and 4 identifying as White. One HCP also identified as Hispanic. HCPs held various positions and degrees, including clinical assistant (n=1), nurse (n=3), mental health clinician (n=1), physician (n=2), advanced practice provider (n=2), and health educator (n=1).

Table 1 summarizes key themes and subthemes derived from participant interviews. Consistent with the survey data, adolescent participants demonstrated low knowledge of their own risk for CVD prior to completing the HHS, but felt they had increased CVD awareness after completing the HHS assessment. HCPs acknowledged that while they felt comfortable talking about CVD-related topics such as nutrition, having access to more resources would help increase their comfort discussing CVD with patients. Both adolescents and HCPs suggested: (1) adding more teen-friendly food and activity options, (2) using explanations and examples throughout the tool (including visuals), (3) improving the layout and design of the tool, (4) using motivating language more effectively, and (5) minding the reading level of the tool. Both adolescents and HCPs expressed concern about teens providing honest answers in the assessment as well as ongoing engagement with behavior change goals. The barriers and facilitators to use such a tool in practice were also shared, with HCPs speaking directly about limited clinic resources and time. Both HCPs and adolescents brought up the potential for fear of judgement, shame, and embarrassment surrounding answers to certain questions (eg, weight, alcohol intake, and nicotine use), as well as fear of learning about one's CVD risk. After completing the assessment, many adolescents demonstrated both curiosity and shock at their predicted risk. HCPs commented on additional barriers (eg, healthy food affordability) and facilitators of adolescent

health behavior change (eg, working with parents to assist in lifestyle changes). To facilitate the implementation of the assessment in practice, HCPs recommended shortening the assessment as time is often limited in the clinical setting. Both

groups felt that going through the HHS assessment itself could be a cue to action for adolescents to change their behavior, consistent with the HBM [28].

Table 1. Coding themes and representative quotes from 10 adolescent and young adult women (AYA) and 10 health care providers (HCPs) giving feedback on the Healthy Heart Score risk assessment.

Themes	Subthemes	Illustrative quotes
Health knowledge	<ul style="list-style-type: none"> • HCP CVD^a knowledge and comfortability • Teen knowledge of heart attack and stroke • Tool as increasing health awareness • Sources of knowledge 	<ul style="list-style-type: none"> • "...people's lack...of awareness of that risk or seriousness of that risk, too, especially for young women. They...may just feel like that's not a thing they necessarily should be worried about..." HCP • "Because I think if I eat healthy and take care of myself right, I don't think that's going to happen to me, but I don't know" AYA • "I guess things...that surprised me a little, like my instincts with certain foods like, that didn't make me think, oh man, I'm going to get like a heart attack. But made me think OK, like maybe I should try a little more on that." AYA
Content of risk assessment tool	<ul style="list-style-type: none"> • Choices • Explanations and examples • Layout and design • Length • Wording • Motivating language • Reading or content level • Teen-friendly • Visuals 	<ul style="list-style-type: none"> • "I feel like, the reading, like on some of them, like it's helpful, but it's also a lot. Like I didn't read through it." AYA • "Well, [the activity choices] were perfectly fine for my kinds of activities, but I can imagine for kids they need to be different because of organized sports." HCP • "I think the time estimates are maybe going to be hard for some people to make like you have to do a little bit of mental math." HCP • [On what was confusing] "The, like how might categorize [physical activity or food intake] like the one, two, three per week." AYA • "I guess it'd be you wanted to have a Spanish version" HCP
Conditions to tool usage	<ul style="list-style-type: none"> • Honest engagement with tool • Initial assessment • Ongoing work • Barriers to tool usage • Facilitators to tool usage 	<ul style="list-style-type: none"> • "I know a few people my age who do drink. And so, I feel like they wouldn't answer honestly." AYA • "It would be nice to be able to communicate with [teens that] are feeling the same way that actually want to improve their health [...] you can relate to someone in the process." AYA • "Sometimes just due to scheduling and the timing of things, you may not have time enough to complete the survey" HCP • "This tool is easy, and with technology now, patients are tired of paper." – HCP
Emotional reaction	<ul style="list-style-type: none"> • Curiosity about personal results • Fear of judgment • Fear of results • Shame or embarrassment • Sharing personal details • Surprise or shock • Weight 	<ul style="list-style-type: none"> • "I'm scared [...] because it'll probably be like, oh well, you probably need to get better with your health or [you'll have] a heart attack." AYA • "It was kind of shocking to see the result." AYA • "Especially girls. Like they'd be real self-conscious about [their weight]. And so, with the way society and stuff is, and all the pressure to look a certain way and be a certain size or weight this much." AYA • "It's just a form of education or awareness... I don't want to have our teens freaking out once they see that." HCP
Health behaviors	<ul style="list-style-type: none"> • Barriers to behavior change • Facilitators to behavior change • Existing teen behavior 	<ul style="list-style-type: none"> • "High fiber cold cereal? I really don't eat cereal like that. I don't really eat breakfast..." AYA • "Affordability is an issue as far as what's at home, what's available at home or if the teen is even cooking, or does mom cook every day." – HCP
Motivation	<ul style="list-style-type: none"> • Tool as cue to action • External motivation • Internal motivation 	<ul style="list-style-type: none"> • "I definitely need to get serious about my weight...and move my body more." AYA • "I have a lot of people in my family that have diabetes, high blood pressure, and high blood sugar." AYA
HCP role in CVD risk assessment	<ul style="list-style-type: none"> • N/A^b 	<ul style="list-style-type: none"> • "One of the focus points of the well visit is to assess people's risk and to address that either by doing further screenings like checking their cholesterol or based on like the data that you have of what their risk might be kind of talking to them about lifestyle changes that could help prevent cardiovascular disease." HCP

^aCVD: cardiovascular disease.^bN/A: not applicable.

Phase 3: Adolescent CVD Screening Tool Development

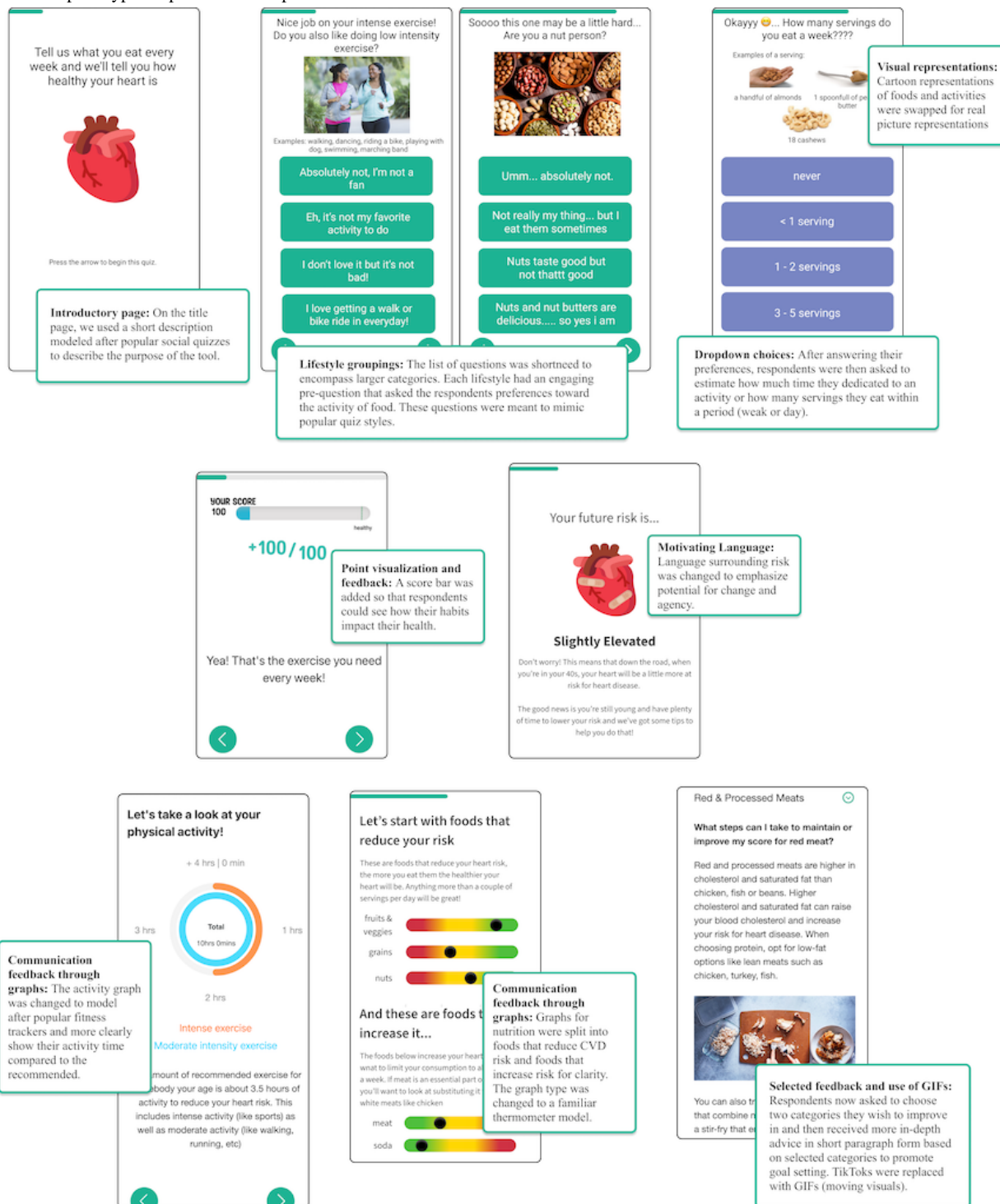
Using data from the phase 1 survey and phase 2 interviews, the programming and design team worked with the clinical research team to develop an initial simple prototype of the tool. In concordance with the behavioral intervention technology model [29], the research team further defined the aims of our tool, focusing on increasing physical activity and the consumption of fruits, vegetables, and grains, and decreasing the consumption of sugar, red meat, and tobacco. With the design team, behavioral strategies were identified, such as education, feedback, agency emphasis, and goal setting. This resulted in an initial prototype that addressed low-CVD prevention awareness activities identified in phase 1, including more relatable teen activities such as team sports (eg, cheerleading, basketball) and food options (eg, oatmeal, popcorn). It also depicted activities and foods with cartoon icons, as the interview findings suggested using visual representations to aid comprehension. Next, the team addressed HCP concerns about assessment length, reducing the tool's length by combining individual foods (eg, cereal, soda, hot dogs, almonds, apples, green beans) into food groups (high-fiber grains, sugary or sweetened drinks, red meats, processed meats, nuts and seeds, fruits, and vegetables), reducing the total number of nutrition questions from 23 to 14. Similarly, we grouped all vigorous activities and moderate activities together into these 2 categories, reducing the total number of physical activity questions from 12 to 4. The total number of questions was reduced from 35 to 18 with these changes. To help avoid feelings of judgment surrounding body size, we deprioritized weight by moving this question to the end and removing feedback on body mass index. We also avoided using any leading or judgmental language in the feedback that may imply a "correct" answer and dissuade teens from answering honestly, opting instead to use more neutral language as suggested by motivational interviewing [27].

This initial storyboard prototype was then shown to teens in the clinic waiting rooms to further refine the examples and language in the questions. Using the first round of in-clinic feedback on

the simple prototype, a functional web-based prototype was then created that fully showed risk, results, and recommendations. Due to the overall low 20-year CVD risk that adolescents have because of their age, the calculated relative scores were translated into point values and shown to the user. These points correspond to the relative importance of the individual elements in the original HHS algorithm. At the end of the assessment, feedback was given in the form of graphs that portrayed results compared to a healthy individual of the same age. TikTok videos demonstrating healthy eating and physical activity tips were initially included based on adolescent suggestions for engaging content. The TikToks were later removed due to the team's inability to find high-quality, evidence-based videos on the platform, and replaced with nutrition and activity advice from the AHA combined with GIFs (moving visuals).

This functional prototype was then shown to additional young women in the clinic waiting rooms and to 3 adolescents in the design session. Further feedback led to refinements, including clearly highlighting the purpose of the tool in bold and addressing both visual inconsistencies (eg, the dropdown for choices sometimes blocked the questions) and semantic issues (eg, some questions did not have the option to select "never"). Points accrued with each health behavior reported were highlighted to improve engagement throughout the assessment. Incorporating motivational interviewing techniques [27], we attempted to emphasize adolescents' autonomy by giving respondents options on which areas they would like to receive additional feedback for. We used the foundations of the HBM [28] as guidance for the feedback, focusing first on raising awareness of an individual's CVD susceptibility and then linking this to specific actions to reduce their CVD risk, along with addressing barriers with specific tips. Language within the feedback also focused on promoting adolescents' self-efficacy and agency to change their scores. Multiple iterative changes were made to the feedback graphs to improve the communication of results. A representation of the final tool can be seen in [Figure 2](#).

Figure 2. Final prototype adaptations and explanations.



Discussion

Principal Results

CVD awareness, knowledge of CVD risk factors, and preventive behaviors were remarkably low among the young women surveyed in this population, which is consistent with other national [3] and regional [25] studies. Few young women reported having never spoken to an HCP about their personal risk for or family history of CVD, despite this study occurring in a region of the United States with some of the highest morbidity and mortality from CVD [31]. Together, these

findings confirmed the urgent need to create a tool to promote CVD awareness, teach preventative measures, and prompt discussion of CVD between young women and their HCPs.

We chose the existing adult HHS as our initial tool for adaptation based on its alignment with the knowledge gaps identified in the survey and existing evidence for its ability to predict early CVD events in both middle-aged and young adults [19,20]. We found that such a tool can indeed prompt adolescents to reflect on their health and habits, as has been shown in a study with middle-aged adults [32]. Similarly, HCPs discussed the low CVD awareness amongst their young female

patients and agreed this risk assessment tool could be a value-added resource to improve awareness. Further discussion with both young women and HCPs revealed the need for such a tool to incorporate teen-friendly foods and activities, clear and concise language, and pictures to improve overall understanding and the accuracy of responses. One issue presented by both HCPs and AYAs was the concern that AYAs would not be honest when answering the survey due to certain topics such as weight, alcohol intake, and nicotine use. It should be noted that the presence of friends, family members, or guardians during interviews may have also affected the candor of participants and therefore the information collected. Consideration of when and where the tool is administered should be taken to ensure that AYA respondents feel they can answer the questions honestly. Additionally, HCPs presented potential clinical barriers, to implementation, including time, and a lack of their own CVD knowledge. Other studies have discussed similar clinical barriers such as practice buy-in, obtaining cooperation and resources, and lack of HCP subject knowledge that have impacted successful intervention implementation [33,34]. Mitigating these challenges is essential to successfully integrating a cardiovascular risk assessment tool into a clinical practice and will be the focus of future studies.

Our iterative design process built upon this feedback and continuously incorporated input from young women through brief interviews in clinic and a design session, resulting in an age-appropriate and culturally tailored cardiovascular risk assessment tool. The input gathered allowed us to make changes to the tool to improve its aesthetic and enjoyability, along with its clarity and functionality. This process highlighted the importance of user-centered design when creating mHealth interventions. Researchers typically use a top-down approach when creating mHealth interventions, preferring to integrate evidence-based care directly into the digital sphere, resulting in mHealth interventions that may not effectively engage their audience [35]. Research has shown that while adolescents are one of the biggest consumers of mHealth interventions, many such tools do not effectively engage them [35]. Studies that have used a user-centered design process have shown great promise in engaging their target population [15]. Involving adolescents in the design process through workshops and direct interviews can help create tools with features that will engage

this population. Combining this user-centered design with evidence-based protocols allows for more effective interventions that not only provide evidence-based guidance but also reach and engage users [36].

Limitations

A major strength of this study is its mixed-methods approach, which leads to a tool that is uniquely tailored to the needs of this specific patient population of young women but is also likely applicable to young women from other backgrounds. Our patient population consisted of young women from minorities and low-income backgrounds, a population whose preferences and experiences are noted as lacking within mHealth intervention development but essential to creating broadly applicable tools [17,36]. The original tool from which our tool is adapted has also been shown to accurately predict the risk of early CVD events in young adults [20]. However, there are limitations to our work. A convenience sample was used to collect survey data and input on the new tool development. As a result, our findings may lack generalizability beyond our local sample and will need to be tested in other populations of adolescents. We also made several adaptations to the validated cardiovascular risk prediction tool to make it usable and acceptable for teens, such as consolidating questions, modifying phrasing, and adjusting the calculation to predict relative risk, and thus, the tool may no longer represent the actual numerical risk of an early CVD event with the same predictive accuracy. Finally, whether the tool prompts true behavior change and correlates with CVD risk reduction will be explored in future studies.

Conclusions

The young women in this study and others consistently demonstrate low awareness of CVD and CVD prevention. More resources are needed to educate young women about CVD and promote CVD discussions with HCPs. An online risk assessment tool is one potential resource that can help increase awareness of CVD if tailored to the needs and preferences of this population. Further research is needed surrounding the feasibility and usability of such a tool in clinical practice, as well as its efficacy in raising CVD awareness and inspiring young women to take action for their future heart health.

Acknowledgments

The authors would like to thank the patients and the clinical and operational teams at Children's Healthcare of Atlanta and the Grady Health System for their time, ideas, and support of this study; the Emory Pipeline Collaborative (EPiC) and Director Allen Lee for assistance with the design session, which is supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR002378; and Brianna Karim for technical assistance with manuscript revisions. This study was funded by NHLBI R03 HL155253 (PI: Gooding) and supported by the Georgia Clinical and Translational Science Alliance through the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR002378. The manuscript preparation was partially funded by a grant to AKG from the National Institute on Drug Abuse (K23DA042935). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

MJK receives research funds from Organon.

Multimedia Appendix 1

Health care provider semistructured interview guide.

[\[PDF File \(Adobe PDF File\), 64 KB - formative_v6i12e42051_app1.pdf \]](#)

Multimedia Appendix 2

Adolescent semistructured interview guide.

[\[PDF File \(Adobe PDF File\), 123 KB - formative_v6i12e42051_app2.pdf \]](#)

References

1. Tsao CW, Aday AW, Almarzooq ZI, Alonso A, Beaton AZ, Bittencourt MS, et al. Heart disease and stroke statistics-2022 update: a report from the American heart association. *Circulation* 2022 Feb 22;145(8):e153-e639 [FREE Full text] [doi: [10.1161/CIR.0000000000001052](https://doi.org/10.1161/CIR.0000000000001052)] [Medline: [35078371](https://pubmed.ncbi.nlm.nih.gov/35078371/)]
2. Women and heart disease. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division for Heart Disease and Stroke Prevention. URL: <https://www.cdc.gov/heartdisease/women.htm> [accessed 2022-11-14]
3. Cushman M, Shay CM, Howard VJ, Jiménez MC, Lewey J, McSweeney JC, American Heart Association. Ten-year differences in women's awareness related to coronary heart disease: results of the 2019 American Heart Association national survey: a special report from the American heart association. *Circulation* 2021 Feb 16;143(7):e239-e248 [FREE Full text] [doi: [10.1161/CIR.0000000000000907](https://doi.org/10.1161/CIR.0000000000000907)] [Medline: [32954796](https://pubmed.ncbi.nlm.nih.gov/32954796/)]
4. Bucholz EM, Gooding HC, de Ferranti SD. Awareness of cardiovascular risk factors in U.S. young adults aged 18-39 years. *Am J Prev Med* 2018 Apr;54(4):e67-e77 [FREE Full text] [doi: [10.1016/j.amepre.2018.01.022](https://doi.org/10.1016/j.amepre.2018.01.022)] [Medline: [29433955](https://pubmed.ncbi.nlm.nih.gov/29433955/)]
5. Sharot T. The optimism bias. *Curr Biol* 2011 Dec 06;21(23):R941-R945 [FREE Full text] [doi: [10.1016/j.cub.2011.10.030](https://doi.org/10.1016/j.cub.2011.10.030)] [Medline: [22153158](https://pubmed.ncbi.nlm.nih.gov/22153158/)]
6. Bibbins-Domingo K, Peña MB. Caring for the "young invincibles". *J Gen Intern Med* 2010 Jul;25(7):642-643 [FREE Full text] [doi: [10.1007/s11606-010-1388-8](https://doi.org/10.1007/s11606-010-1388-8)] [Medline: [20499197](https://pubmed.ncbi.nlm.nih.gov/20499197/)]
7. Steinberg L, Graham S, O'Brien L, Woolard J, Cauffman E, Banich M. Age differences in future orientation and delay discounting. *Child Dev* 2009;80(1):28-44. [doi: [10.1111/j.1467-8624.2008.01244.x](https://doi.org/10.1111/j.1467-8624.2008.01244.x)] [Medline: [19236391](https://pubmed.ncbi.nlm.nih.gov/19236391/)]
8. Gooding HC, Sheldrick RC, Leslie LK, Shah S, de Ferranti SD, Mackie TI. Adolescent perceptions of cholesterol screening results: "young invincibles" or developing adults. *J Adolesc Health* 2016 Aug;59(2):162-170 [FREE Full text] [doi: [10.1016/j.jadohealth.2016.03.027](https://doi.org/10.1016/j.jadohealth.2016.03.027)] [Medline: [27155960](https://pubmed.ncbi.nlm.nih.gov/27155960/)]
9. Warnick JL, Pfammatter A, Champion K, Galluzzi T, Spring B. Perceptions of health behaviors and mobile health applications in an academically elite college population to inform a targeted health promotion program. *Int J Behav Med* 2019 Apr;26(2):165-174. [doi: [10.1007/s12529-018-09767-y](https://doi.org/10.1007/s12529-018-09767-y)] [Medline: [30632092](https://pubmed.ncbi.nlm.nih.gov/30632092/)]
10. Berg CJ, Stratton E, Sokol M, Santamaria A, Bryant L, Rodriguez R. Novel incentives and messaging in an online college smoking intervention. *Am J Health Behav* 2014 Sep;38(5):668-680 [FREE Full text] [doi: [10.5993/AJHB.38.5.4](https://doi.org/10.5993/AJHB.38.5.4)] [Medline: [24933136](https://pubmed.ncbi.nlm.nih.gov/24933136/)]
11. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. *J Adolesc Health* 2017 Dec;61(6):669-677 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.05.024](https://doi.org/10.1016/j.jadohealth.2017.05.024)] [Medline: [28822682](https://pubmed.ncbi.nlm.nih.gov/28822682/)]
12. Cushing CC, Bejarano CM, Ortega A, Sayre N, Fedele DA, Smyth JM. Adaptive mHealth intervention for adolescent physical activity promotion. *J Pediatr Psychol* 2021 Jun 03;46(5):536-546. [doi: [10.1093/jpepsy/jsaa125](https://doi.org/10.1093/jpepsy/jsaa125)] [Medline: [33484137](https://pubmed.ncbi.nlm.nih.gov/33484137/)]
13. Oosterveen E, Tzelepis F, Ashton L, Hutchesson MJ. A systematic review of eHealth behavioral interventions targeting smoking, nutrition, alcohol, physical activity and/or obesity for young adults. *Prev Med* 2017 Jun;99:197-206. [doi: [10.1016/j.ypmed.2017.01.009](https://doi.org/10.1016/j.ypmed.2017.01.009)] [Medline: [28130046](https://pubmed.ncbi.nlm.nih.gov/28130046/)]
14. WHO Global Observatory for eHealth. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: <https://apps.who.int/iris/handle/10665/44607> [accessed 2022-11-23]
15. Brewer LC, Hayes SN, Caron AR, Derby DA, Breutzman NS, Wicks A, et al. Promoting cardiovascular health and wellness among African-Americans: community participatory approach to design an innovative mobile-health intervention. *PLoS One* 2019;14(8):e0218724 [FREE Full text] [doi: [10.1371/journal.pone.0218724](https://doi.org/10.1371/journal.pone.0218724)] [Medline: [31430294](https://pubmed.ncbi.nlm.nih.gov/31430294/)]
16. Klimis H, Thakkar J, Chow CK. Breaking barriers: mobile health interventions for cardiovascular disease. *Can J Cardiol* 2018 Jul;34(7):905-913. [doi: [10.1016/j.cjca.2018.02.012](https://doi.org/10.1016/j.cjca.2018.02.012)] [Medline: [29887218](https://pubmed.ncbi.nlm.nih.gov/29887218/)]
17. Burke LE, Ma J, Azar KM, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on EpidemiologyPrevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on CardiovascularStroke Nursing, Council on Functional GenomicsTranslational Biology, Council on Quality of CareOutcomes Research, Stroke Council. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American heart association. *Circulation* 2015 Sep 22;132(12):1157-1213 [FREE Full text] [doi: [10.1161/CIR.0000000000000232](https://doi.org/10.1161/CIR.0000000000000232)] [Medline: [26271892](https://pubmed.ncbi.nlm.nih.gov/26271892/)]

18. Ahern DK, Kreslake JM, Phalen JM. What is eHealth (6): perspectives on the evolution of eHealth research. *J Med Internet Res* 2006 Mar 31;8(1):e4 [FREE Full text] [doi: [10.2196/jmir.8.1.e4](https://doi.org/10.2196/jmir.8.1.e4)] [Medline: [16585029](https://pubmed.ncbi.nlm.nih.gov/16585029/)]
19. Chiuve SE, Cook NR, Shay CM, Rexrode KM, Albert CM, Manson JE, et al. Lifestyle-based prediction model for the prevention of CVD: the healthy heart score. *J Am Heart Assoc* 2014 Nov 14;3(6):e000954 [FREE Full text] [doi: [10.1161/JAHA.114.000954](https://doi.org/10.1161/JAHA.114.000954)] [Medline: [25398889](https://pubmed.ncbi.nlm.nih.gov/25398889/)]
20. Gooding HC, Ning H, Gillman MW, Shay C, Allen N, Goff DC, et al. Application of a lifestyle-based tool to estimate premature cardiovascular disease events in young adults: the coronary artery risk development in young adults (CARDIA) study. *JAMA Intern Med* 2017 Sep 01;177(9):1354-1360 [FREE Full text] [doi: [10.1001/jamainternmed.2017.2922](https://doi.org/10.1001/jamainternmed.2017.2922)] [Medline: [28715555](https://pubmed.ncbi.nlm.nih.gov/28715555/)]
21. Sotos-Prieto M, Mattei J, Hu FB, Chomistek AK, Rimm EB, Willett WC, et al. Association between a healthy heart score and the development of clinical cardiovascular risk factors among women: potential role for primordial prevention. *Circ Cardiovasc Qual Outcomes* 2016 Feb;9(2 Suppl 1):S77-S85 [FREE Full text] [doi: [10.1161/CIRCOUTCOMES.115.002372](https://doi.org/10.1161/CIRCOUTCOMES.115.002372)] [Medline: [26908864](https://pubmed.ncbi.nlm.nih.gov/26908864/)]
22. Fortuna RJ, Robbins BW, Halterman JS. Ambulatory care among young adults in the United States. *Ann Intern Med* 2009 Sep 15;151(6):379-385. [doi: [10.7326/0003-4819-151-6-200909150-00002](https://doi.org/10.7326/0003-4819-151-6-200909150-00002)] [Medline: [19755363](https://pubmed.ncbi.nlm.nih.gov/19755363/)]
23. Gooding HC, Brown CA, Revette AC, Vaccarino V, Liu J, Patterson S, et al. Young women's perceptions of heart disease risk. *J Adolesc Health* 2020 Nov;67(5):708-713 [FREE Full text] [doi: [10.1016/j.jadohealth.2020.05.010](https://doi.org/10.1016/j.jadohealth.2020.05.010)] [Medline: [32622924](https://pubmed.ncbi.nlm.nih.gov/32622924/)]
24. SE Atlanta Superdistrict fact sheet. Neighborhood Nexus. URL: http://documents.atlantaregional.com/profiles/Superdistrict/Fulton_County_SE_Atlanta_NN.pdf [accessed 2022-07-21]
25. Gooding HC, Brown CA, Liu J, Revette AC, Stamoulis C, de Ferranti SD. Will teens go red? Low cardiovascular disease awareness among young women. *J Am Heart Assoc* 2019 Mar 19;8(6):e011195 [FREE Full text] [doi: [10.1161/JAHA.118.011195](https://doi.org/10.1161/JAHA.118.011195)] [Medline: [30835591](https://pubmed.ncbi.nlm.nih.gov/30835591/)]
26. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Group. URL: <https://www.nngroup.com/articles/ten-usability-heuristics/> [accessed 2022-11-14]
27. Miller WR, Rose GS. Toward a theory of motivational interviewing. *Am Psychol* 2009 Sep;64(6):527-537 [FREE Full text] [doi: [10.1037/a0016830](https://doi.org/10.1037/a0016830)] [Medline: [19739882](https://pubmed.ncbi.nlm.nih.gov/19739882/)]
28. Glanz K, Lewis FM, Rimer BK. Health behavior and health education: theory, research, and practice. *MSSE* 1991;23(12):1404. [doi: [10.1249/00005768-199112000-00016](https://doi.org/10.1249/00005768-199112000-00016)]
29. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. *J Med Internet Res* 2014 Jun 05;16(6):e146 [FREE Full text] [doi: [10.2196/jmir.3077](https://doi.org/10.2196/jmir.3077)] [Medline: [24905070](https://pubmed.ncbi.nlm.nih.gov/24905070/)]
30. Bandura A. Human agency in social cognitive theory. *Am Psychol* 1989 Sep;44(9):1175-1184. [doi: [10.1037/0003-066x.44.9.1175](https://doi.org/10.1037/0003-066x.44.9.1175)] [Medline: [2782727](https://pubmed.ncbi.nlm.nih.gov/2782727/)]
31. Cooper R, Cutler J, Desvigne-Nickens P, Fortmann SP, Friedman L, Havlik R, et al. Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States: findings of the national conference on cardiovascular disease prevention. *Circulation* 2000 Dec 19;102(25):3137-3147. [doi: [10.1161/01.cir.102.25.3137](https://doi.org/10.1161/01.cir.102.25.3137)] [Medline: [11120707](https://pubmed.ncbi.nlm.nih.gov/11120707/)]
32. Buls S, Beverly EA, Berryman DE, Sotos-Prieto M. Assessing utility of a lifestyle-based tool in the clinical setting as a primordial prevention strategy: the healthy heart score. *Chronic Illn* 2022 Mar;18(1):105-118. [doi: [10.1177/1742395319899431](https://doi.org/10.1177/1742395319899431)] [Medline: [31958999](https://pubmed.ncbi.nlm.nih.gov/31958999/)]
33. Powell BJ, Hausmann-Stabile C, McMillen JC. Mental health clinicians' experiences of implementing evidence-based treatments. *J Evid Based Soc Work* 2013 Oct;10(5):396-409 [FREE Full text] [doi: [10.1080/15433714.2012.664062](https://doi.org/10.1080/15433714.2012.664062)] [Medline: [24066630](https://pubmed.ncbi.nlm.nih.gov/24066630/)]
34. Sharp N, Pineros S, Hsu C, Starks H, Sales A. A qualitative study to identify barriers and facilitators to implementation of pilot interventions in the veterans health administration (VHA) northwest network. *Worldviews Evid Based Nurs* 2004;1(2):129-139. [doi: [10.1111/j.1741-6787.2004.04023.x](https://doi.org/10.1111/j.1741-6787.2004.04023.x)] [Medline: [17129326](https://pubmed.ncbi.nlm.nih.gov/17129326/)]
35. Psihogios AM, Lane-Fall MB, Graham AK. Adolescents are still waiting on a digital health revolution: accelerating research-to-practice translation through design for implementation. *JAMA Pediatr* 2022 Jun 01;176(6):545-546. [doi: [10.1001/jamapediatrics.2022.0500](https://doi.org/10.1001/jamapediatrics.2022.0500)] [Medline: [35404381](https://pubmed.ncbi.nlm.nih.gov/35404381/)]
36. Stiles-Shields C, Potthoff LM, Bounds DT, Burns MTS, Draxler JM, Otwell CH, et al. Harnessing phones to target pediatric populations with socially complex needs: systematic review. *JMIR Pediatr Parent* 2020 Aug 26;3(2):e19269 [FREE Full text] [doi: [10.2196/19269](https://doi.org/10.2196/19269)] [Medline: [32845244](https://pubmed.ncbi.nlm.nih.gov/32845244/)]

Abbreviations

- AHA:** American Heart Association
- AYA:** adolescent and young adult
- CVD:** cardiovascular disease

HBM: health belief model
HCP: health care provider
HHS: Healthy Heart Score
mHealth: mobile health

Edited by A Mavragani; submitted 19.08.22; peer-reviewed by M Kapsetaki, JH Rajendran; comments to author 27.10.22; revised version received 16.11.22; accepted 17.11.22; published 19.12.22.

Please cite as:

Bradley K, Arconada Alvarez SJ, Gilmore AK, Greenleaf M, Herbert A, Kottke MJ, Parsell M, Patterson S, Smith T, Sotos-Prieto M, Zeichner E, Gooding HC

Assessing and Promoting Cardiovascular Health for Adolescent Women: User-Centered Design Approach

JMIR Form Res 2022;6(12):e42051

URL: <https://formative.jmir.org/2022/12/e42051>

doi: [10.2196/42051](https://doi.org/10.2196/42051)

PMID: [36534450](https://pubmed.ncbi.nlm.nih.gov/36534450/)

©Kolbi Bradley, Santiago J Arconada Alvarez, Amanda K Gilmore, Morgan Greenleaf, Aayahna Herbert, Melissa J Kottke, Maren Parsell, Sierra Patterson, Tymirra Smith, Mercedes Sotos-Prieto, Elizabeth Zeichner, Holly C Gooding. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 19.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Vocal Characteristics To Classify Psychological Distress in Adult Helpline Callers: Retrospective Observational Study

Ravi Iyer¹, BSc, MSc; Maja Nedeljkovic¹, PhD; Denny Meyer¹, PhD

Centre for Mental Health, Swinburne University of Technology, Hawthorn, Australia

Corresponding Author:

Ravi Iyer, BSc, MSc

Centre for Mental Health

Swinburne University of Technology

34 Wakefield Street

Hawthorn, 3122

Australia

Phone: 61 456 565 575

Email: raviyer@swin.edu.au

Abstract

Background: Elevated psychological distress has demonstrated impacts on individuals' health. Reliable and efficient ways to detect distress are key to early intervention. Artificial intelligence has the potential to detect states of emotional distress in an accurate, efficient, and timely manner.

Objective: The aim of this study was to automatically classify short segments of speech obtained from callers to national suicide prevention helpline services according to high versus low psychological distress and using a range of vocal characteristics in combination with machine learning approaches.

Methods: A total of 120 telephone call recordings were initially converted to 16-bit pulse code modulation format. Short variable-length segments of each call were rated on psychological distress using the distress thermometer by the responding counselor and a second team of psychologists (n=6) blinded to the initial ratings. Following this, 24 vocal characteristics were initially extracted from 40-ms speech frames nested within segments within calls. After highly correlated variables were eliminated, 19 remained. Of 19 vocal characteristics, 7 were identified and validated as predictors of psychological distress using a penalized generalized additive mixed effects regression model, accounting for nonlinearity, autocorrelation, and moderation by sex. Speech frames were then grouped using k-means clustering based on the selected vocal characteristics. Finally, component-wise gradient boosting incorporating these clusters was used to classify each speech frame according to high versus low psychological distress. Classification accuracy was confirmed via leave-one-caller-out cross-validation, ensuring that speech segments from individual callers were not used in both the training and test data.

Results: The sample comprised 87 female and 33 male callers. From an initial pool of 19 characteristics, 7 vocal characteristics were identified. After grouping speech frames into 2 separate clusters (correlation with sex of caller, Cramer's V =0.02), the component-wise gradient boosting algorithm successfully classified psychological distress to a high level of accuracy, with an area under the receiver operating characteristic curve of 97.39% (95% CI 96.20-98.45) and an area under the precision-recall curve of 97.52 (95% CI 95.71-99.12). Thus, 39,282 of 41,883 (93.39%) speech frames nested within 728 of 754 segments (96.6%) were classified as exhibiting low psychological distress, and 71455 of 75503 (94.64%) speech frames nested within 382 of 423 (90.3%) segments were classified as exhibiting high psychological distress. As the probability of high psychological distress increases, male callers spoke louder, with greater vowel articulation but with greater roughness (subharmonic depth). In contrast, female callers exhibited decreased vocal clarity (entropy), greater proportion of signal noise, higher frequencies, increased breathiness (spectral slope), and increased roughness of speech with increasing psychological distress. Individual caller random effects contributed 68% to risk reduction in the classification algorithm, followed by cluster configuration (23.4%), spectral slope (4.4%), and the 50th percentile frequency (4.2%).

Conclusions: The high level of accuracy achieved suggests possibilities for real-time detection of psychological distress in helpline settings and has potential uses in pre-emptive triage and evaluations of counseling outcomes.

Trial Registration: ANZCTR ACTRN12622000486729; <https://www.anzctr.org.au/ACTRN12622000486729.aspx>

(*JMIR Form Res* 2022;6(12):e42249) doi:[10.2196/42249](https://doi.org/10.2196/42249)

KEYWORDS

machine learning; distress; voice; mental distress; psychological stress; artificial intelligence; emotional distress; voice biomarker; biomarker; digital health intervention; mental health; mental health intervention; psychological well being; speech analysis

Introduction

In recent years, the presence of psychological distress in the community has escalated sharply due to COVID-19 [1,2]. Given the demonstrated impacts of high distress on health and functioning, reliable and efficient ways to detect distress are key to early intervention. Despite this growing problem, surprisingly little attention has been paid to methods for the detection of distress in the broader community. We aimed to use artificial intelligence to automatically classify a sample of helpline call recordings according to high and low psychological distress using voice characteristics alone. Thus, we offer an objective and efficient approach to the real-time detection of psychological distress in a large sample of helpline callers.

Psychological distress has been defined as “the unique discomforting emotional state experienced by an individual in response to a specific stressor or demand that results in harm, either temporary or permanent, to the person” [3]. Although commonly associated with diagnoses of depression and anxiety, psychological distress is also associated with other psychiatric diagnoses such as posttraumatic stress disorder (PTSD) and schizophrenia [4] and is a feature of significant life events such as bereavement and employment loss [5,6]. If left undetected, psychological distress can contribute to declines in physical and mental health, longer hospital stays, poor treatment compliance, and increasing cost of care [7].

Helplines internationally have witnessed substantial increases in call volumes and presentations involving psychological distress in recent years [8]. The support provided by helpline staff has been directly linked to sustained reductions in psychological distress [9]. However, the assessment of psychological distress is challenged by the absence of nonverbal cues, time limitations, and reticence of some callers to discuss relevant issues. Helpline staff have also reported difficulties when asking callers directly about psychological distress [10], suggesting a role for alternative approaches to the detection of psychological distress in helpline settings.

Pilot data support the use of voice characteristics to identify psychological distress in varied settings. Scherer and colleagues [11] investigated whether voice recordings could be separated between high and low psychological distress using vowel space as a measure of clarity of articulation and expressiveness of speech (hypothesizing that greater vowel space would signify psychomotor slowing of speech production due to psychological distress). In this study, participants were sourced from 2 separate databases. The first of these was the Distress Assessment Interview Corpus, which contains audio-visual recordings of semistructured interviews conducted between 253 participants and a virtual interviewer, coded for depression using the Patient Health Questionnaire-9, and PTSD using the PTSD-checklist civilian version. A second sample of 68 recordings was also analyzed, sourced from an audio-visual database of interviews with both participants with and without depression. Small to

moderate effect sizes were found in comparisons of depressed and nondepressed recordings (Hedge’s $G=-0.43$) and in those of PTSD with an absence of PTSD diagnosis (Hedge’s $G=-0.34$). This study indicated that psychologically distressed individuals may have poorer articulatory clarity and expressiveness and therefore suggested voice characteristics as a viable marker of psychological distress more broadly.

Kansberger and colleagues [12] alternatively hypothesized that the average vocal pitch would be elevated when utterances were compared between psychologically distressed and emotionally neutral conditions. Average fundamental frequency values (lowest discernible pitch) extracted from follow-up interviews of 16 patients with head and neck cancer were coded for the presence of overt and implied statements ($n=89$) of psychological distress. A mixed effects logistic regression analysis found that fundamental frequency values were indeed elevated in the presence of psychological distress ($P<.05$; 95% CI 1.82-24.31Hz). Furthermore, elevations in pitch accounted for ~70% of the variance in psychological distress outcomes, suggesting that the results obtained were robust against differences between individual participants.

However, these studies reveal a number of limitations. As Kansberger and colleagues [12] note, fundamental frequency (or for that matter, any single measure) is unlikely to be singly prognostic of psychological distress. As Franklin and colleagues [13] note in their review of assessment measures for suicide risk, research would benefit from the use of predictive algorithms capable of considering multiple predictors simultaneously: a research direction that may well also apply to the optimal identification of psychological distress. Furthermore, although Kansberger et al [12] accounted for differences in pitch, nested within individual participants, the authors did not account for nonlinearity or correlation between utterances, suggesting a missed opportunity for accurate modeling of changes in pitch over time. Conversely, while Scherer et al [11] analyzed a larger sample of recordings, these were obtained from recordings of interviews conducted with a virtual avatar and may lack generalizability to more ecologically valid settings.

Artificial intelligence has the potential to detect states of psychological distress in an accurate, efficient, and timely manner using multiple measures. Although there is initial evidence for the efficacy of such an approach [14], current evidence lacks application to real-world ecologies and real-time assessment, both of which are essential if these insights are to move beyond the laboratory. Thus, we aimed to use artificial intelligence approaches to automatically classify a large sample of telephone counseling calls made to Australian suicide prevention helpline services according to the level of psychological distress using a range of vocal characteristics. By classifying short segments of each call to a high level of accuracy, we aim to demonstrate a viable artificial intelligence

support to existing helpline infrastructure that can be deployed in real time.

Methods

[Multimedia Appendix 1](#) illustrates the steps taken in both preprocessing and the final analysis of the vocal characteristics.

Call Recordings

A total of 537 call recordings were initially sourced from On The Line, Australia (the suicide call-back service) and the Australian Federal Police, Canberra (000 emergency services response), as part of a broader study classifying low from imminent risk of suicide using voice characteristics [15]. On The Line recordings were chosen at random from between July 1, 2019, and June 30, 2021, stratified by suicide risk level and disclosed sex of caller. The Australian Federal Police recordings were purposely chosen over the same time period to reflect imminent risk of suicide necessitating emergency services response.

Preprocessing of Calls

All calls were received as mono-channel 8 kHz, 32-bit float format. Each recording was initially transformed to 16-bit pulse code modulation format and normalized (zero mean) with pre-emphasis added to attenuate low signals and emphasize higher frequency signals, thus clarifying the degree of audibility, particularly where this was compromised (eg, mobile phone calls with considerable background noise). Listwise removal of missing data (silence; 381,655 of 722,274 segments, 52.84%) occurred prior to the subsequent analyses.

Selection of Call Recordings Relating to Psychological Distress

A multigated approach informed the designation of psychological distress ratings to each annotated segment of each call recording. A team of associate researchers (n=6) assigned distress levels to segments selected from within each call as described in the next section. The associate researchers were either provisional or fully registered psychologists (henceforth referred to the “psychologists”) with the Psychology Board of Australia, completing postgraduate qualifications in psychology and who had substantial prior experience working with complex presentations, often involving psychological distress of varying magnitude (eg, psychiatric diagnoses, significant life events). A random sample of the 120 call recordings was assigned to each of the psychologists, with 20% crossover (4 of 22 calls assigned to each rater) used for the assessment of interrater reliability (Cohen κ =0.92).

The psychologists were asked to annotate segments of each recording using audio software (Audacity, version 2.4.2, The Audacity Team). Annotated segments were to be free from the counselor’s voice as much as possible and were to feature a diversity of psychological distress ratings; thus, it was common for a range of distress ratings to feature across annotated segments within each call recording, ensuring our analysis considered within-caller variation of distress. Psychological distress was rated by each associate researcher using the distress thermometer, an 11-point discrete Likert-style scale (0=no

distress and 10=extreme distress). The distress thermometer has a critical level of 4 (SD 1.5) indicating clinical levels of psychological distress that require follow-up and referral [16]. Thus, the ratings of psychological distress for each segment were dichotomized for either side of the recommended clinical psychological distress rating (distress rating=4). The psychologists also described each caller’s presentation using standardized psychiatric mental status examination descriptions.

The distress thermometer has been used extensively in areas other than oncology, for which the scale was first developed, with good concordance achieved between clinician and patient ratings (n=364, Lin’s coefficient of concordance=0.79, sensitivity=0.83 and specificity=0.71) [17]. The distress thermometer has been compared with other well-validated measures for identifying psychological distress including the Hospital Anxiety and Depression Scale, the Structured Clinical Interview for DSM (Diagnostic and Statistical Manual of Mental Disorders) disorders and the Depression Anxiety and Stress Scale-21 [16], with a mean area under the receiver operating characteristic curve (AUROC) of 0.82 (SD 0.07) across all these measures [18].

Power Analysis

Following Zou [19], the intraclass correlation coefficient from Kandsberger [12] (intraclass correlation coefficient=0.31) was used with a theoretical number of distress annotations per call (n=4) indicating a desired sample size of 119 calls necessary to identify a significant main effect for psychological distress (power=0.8, α =.05). Thus, 120 calls were chosen at random from the overall sample of callers.

Derivation of Vocal Characteristics

Twenty-four unique vocal characteristics were initially derived from the digital vocal using the analyze function of the Soundgen package in RStudio (version 2022.07.01). Nineteen vocal characteristics remained after removal of highly correlated variables. Annotated segments of each recording were divided into overlapping (50%) 40-ms Blackman-windowed frames, with the segment level of psychological distress assigned to all frames within each segment. The choice of frame size provided a level of focus on important characteristics of the soundwave by magnifying central frequencies of each frame and ensuring that valuable information was not lost in the tails of each window.

Selection and Validation of Candidate Vocal Characteristics

A penalized 2-level generalized additive mixed-effects regression model (GAMM) was used to remove vocal characteristics that were not significant predictors of psychological distress. The choice of GAMM ensured that each predictor could be tested, allowing for nonlinearity with sex of caller as a moderating variable. The 2-level model reflected our approach to data collection: 40-ms frame voice characteristics (level 1) nested within individual calls (level 2), allowing for moderation by sex of caller.

Splines with differing degrees of freedom were added for each predictor to account for different forms of nonlinearity. Random

intercepts were added to account for differences between individual callers. A binomial model with logit link was used to differentiate low from high psychological distress frames within calls.

Clustering of Speech Frames Using Voice Characteristics

k-means clustering of speech frames was used to derive probabilistic clusters from the overall voice characteristics. This approach has been used when analyzing other forms of biologically derived data (eg, DNA gene expression) [20]. The addition of the principal components used to discriminate between clusters provided a supplementary predictor that improved classification accuracy in the next stage.

Classification of Speech Frames Using Machine Learning

We sought to extend upon the work of Kandsberger [12] by using the reduced vocal characteristic predictor set (obtained via GAMM) to classify speech frames according to high and low psychological distress within the cluster configuration derived from the k-means clustering step.

Gradient boosting is a computationally inexpensive approach that is able to achieve a level of transparency unavailable to other powerful machine learning approaches (eg, support vector machines and neural networks). This transparency is desirable when clinical insights are of importance. In its base implementation, gradient boosting assumes linearity among predictors; however, this can be remedied with alternative implementations. Component-wise gradient boosting can analyze nonlinear data by first estimating a GAMM with smooth spline terms added and then by applying each model component (individual predictors and random components) to achieve the best reduction in classification error. Moderation by sex of caller was added to this model for all vocal variables.

Leave-one-caller-out cross-validation was used to validate the model, ensuring that speech frames from any single caller did not feature in both training and test sets. Ten-fold cross-validation was used to confirm the number of boosting trees, and the Youden J index was used to determine the optimal probability cutoff that maximized upon both precision and recall measures [21]. The efficacy of the classification algorithm was measured using AUCROC and the area under the precision-recall curve to ensure robust measures of accuracy in the presence of possible data-class imbalance.

Plain language definitions of the vocal characteristics included in the final model have been included in [Multimedia Appendix 2](#).

Misclassification

Following classification of speech frames using the component-wise gradient boosting algorithm, misclassifications were inspected at the annotated segment level rather than at the individual frame level. This was done by classifying each segment as high or low distress based on the mean probabilities for all the corresponding segment frames.

This approach ensured that the mental status examination descriptions made against each segment by the team of psychologists could be inspected to determine whether any patterns in caller presentation might be evident for the misclassified segments.

Ethics Approval

No contact information for callers was obtained, and thus a waiver of consent was granted by the institutional human research ethics committee (reference #20214340-5805) in compliance with the Declaration of Helsinki. This study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guidelines. A STROBE checklist is included in [Multimedia Appendix 3](#).

Results

Overview

The sample comprised 87 female and 33 male callers (mean caller age 39.47 years, SD 15.24). In total, 117,387 40-ms frames (low psychological distress=75,504; high psychological distress=41,88; mean distress rating 4.92, SD 2.13) were nested within 1177 annotated call segments (low psychological distress=754, high psychological distress=423; mean number of segments per call 7.70, SD 4.22) and within 120 individual callers.

Reduced Predictor Set via GAMM

Penalized GAMM was used to reduce and validate the number of candidate predictors while allowing for nonlinearity and accounting for moderation by sex of caller. The model overall accounted for 15.5% of the variance in the level of psychological distress (adjusted $R^2=0.15$). [Table 1](#) summarizes the coefficients and degree of nonlinearity of each of the 19 vocal characteristics included in the model, and the differences between male and female callers for each of the significant predictors is illustrated in [Figure 1](#).

Table 1. Summary of voice characteristic coefficients in the generalized additive mixed-effects regression model^a (adjusted R²=0.15; n=117,387).

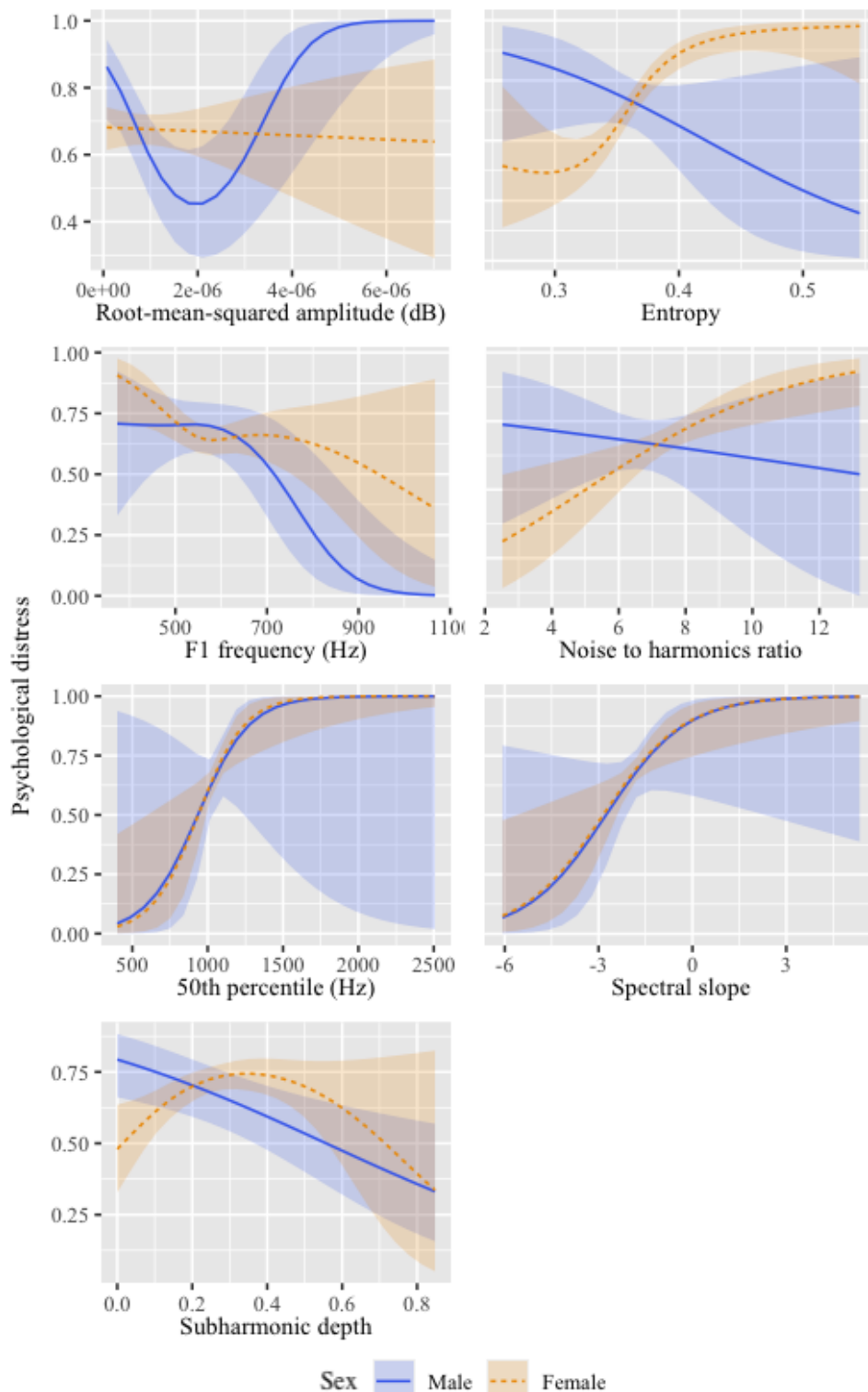
Variable	β	SE	95% CI	edf ^b	F test (df)	P value
Male						
Root-mean-squared amplitude (dB)	8.40	0.08	8.24 to 8.56	2.30	6.21 (2,1455)	<.001
Dominant frequency (Hz)	2.62	0.05	2.54 to 2.71	1.00	2.25 (1,1455)	.13
Entropy	-2.40	0.05	-2.50 to -2.31	1.00	1.70 (1,1455)	.19
First formant frequency (Hz)	-6.05	0.05	-6.15 to -5.95	2.44	5.07 (2,1455)	.01
First formant width (Hz)	1.00	0.03	0.95 to 1.05	1.00	0.98 (1,1455)	.32
Second formant frequency (Hz)	-1.20	0.03	-1.26 to -1.14	1.00	1.14 (1,1455)	.29
Second formant width (Hz)	1.09	0.02	1.05 to 1.13	1.00	1.77 (1,1455)	.18
Third formant frequency (Hz)	-1.11	0.02	-1.16 to -1.07	1.00	1.63 (1,1455)	.20
Third formant width (Hz)	.77	0.04	0.69 to 0.84	1.90	1.91 (2,1455)	.11
Spectral flux	-1.19	0.05	-1.28 to -1.10	1.00	0.44 (1,1455)	.51
Noise to harmonics ratio	-0.58	0.04	-0.65 to -0.51	1.00	0.18 (1,1455)	.67
Loudness (sone)	.19	0.05	0.08 to 0.29	1.00	0.01 (1,1455)	.93
Spectral novelty	1.33	0.03	1.26 to 1.39	1.00	1.03 (1,1455)	.31
Peak frequency (Hz)	6.06	0.09	5.88 to 6.25	1.00	2.72 (1,1455)	.10
25th percentile frequency (Hz)	4.91	0.18	4.55 to 5.27	1.00	0.47 (1,1455)	.49
50th percentile frequency (Hz)	9.38	0.19	9.01 to 9.76	1.00	1.63 (1,1455)	.20
75th percentile frequency (Hz)	5.66	0.13	5.41 to 5.91	1.00	1.29 (1,1455)	.26
Roughness	1.59	0.03	1.54 to 1.65	1.00	2.24 (1,1455)	.14
Spectral centroid (Hz)	-22.60	0.51	-23.60 to -21.59	1.00	1.28 (1,1455)	.26
Spectral slope (Hz)	6.62	0.10	6.41 to 6.82	1.00	2.69 (1,1455)	.10
Depth of the subharmonics (Hz)	-1.56	0.01	-1.59 to -1.53	1.00	8.56 (1,1455)	.004
Female						
Root mean squared amplitude (dB)	-1.17	0.02	-0.20 to -0.13	1.00	0.05 (1,1455)	.82
Dominant frequency (Hz)	1.79	0.04	1.71 to 1.87	1.00	1.39 (1,1455)	.24
Entropy	3.23	0.04	3.16 to 3.29	2.75	13.09 (3,1455)	<.001
First formant frequency (Hz)	-1.79	0.03	-1.86 to -1.73	2.87	1.98 (3,1455)	.08
First formant width (Hz)	-.72	0.02	-0.75 to -0.69	1.00	1.20 (1,1455)	.27
Second formant frequency (Hz)	.07	0.02	0.04 to 0.10	1.00	0.01 (1,1455)	.91
Second formant width (Hz)	3.04	0.04	2.97 to 3.12	2.67	2.28 (3,1455)	.07
Third formant frequency (Hz)	.03	0.02	0.00 to 0.07	1.00	0.03 (1,1455)	.96
Third formant width (Hz)	-0.83	0.02	-0.86 to -0.79	1.00	1.19 (1,1455)	.28
Spectral flux	2.45	0.04	2.38 to 2.52	1.00	3.13 (1,1455)	.08
Noise to harmonics ratio	2.45	0.02	2.41 to 2.49	1.00	9.49 (1,1455)	.002
Loudness (sone)	-.90	0.03	-0.97 to -0.84	1.00	0.46 (1,1455)	.50
Spectral novelty	.07	0.03	0.02 to 0.13	1.00	0.01 (1,1455)	.94
Peak frequency (Hz)	-2.09	0.05	-2.19 to -1.98	1.00	0.98 (1,1455)	.32
25th percentile frequency (Hz)	-3.48	0.11	-3.68 to -3.27	1.00	0.72 (1,1455)	.40
50th percentile frequency (Hz)	10.49	0.10	10.28 to 10.69	1.00	6.77 (1,1455)	.009
75th percentile frequency (Hz)	-4.37	0.08	-4.53 to -4.21	1.00	1.87 (1,1455)	.17
Roughness	-1.11	0.02	-0.14 to -0.07	1.00	0.03 (1,1455)	.87

Variable	β	SE	95% CI	edf ^b	F test (df)	P value
Spectral centroid (Hz)	-13.63	0.29	-14.20 to -13.06	1.00	1.45 (1,1455)	.23
Spectral slope (Hz)	6.53	0.06	6.40 to 6.65	1.00	6.91 (1,1455)	.009
Depth of the subharmonics (Hz)	-.87	0.03	-0.92 to -0.82	2.25	5.25 (2,1455)	.005

^aMale versus female: $\beta=-.06$; SE 0.25, 95% CI -0.44 to 0.56; $t=-0.24$; $P=.81$.

^bedf: estimated degrees of freedom.

Figure 1. Relationship of each voice characteristic to the probability of high psychological distress.



Although sex of caller was not a significant moderator for the model overall ($\beta=-0.06$; SE 0.25, 95% CI -0.44 to 0.56 ; $t=-0.24$; $P=.81$), the profile of significant vocal characteristics did differ between male and female callers. Male callers in our sample appeared to speak louder with increasing psychological distress ($\beta=8.40$, 95% CI 8.24 to 8.56; estimated degrees of freedom [edf]=2.3; $F_{2,1455}=6.21$; $P<.001$). The frequencies of the first formant were also found to fall, suggesting an increase in the articulatory quality of vowel sounds ($\beta=-6.05$; 95% CI -6.15 to -5.95 ; edf=2.44; $F_{2,1455}=5.08$; $P=.01$), while the depth of the subharmonics among male callers was also found to fall, suggesting increasing roughness of speech ($\beta=-1.56$, 95% CI -1.59 to -1.53 ; edf=1.00; $F_{1,1455}=8.56$; $P=.004$).

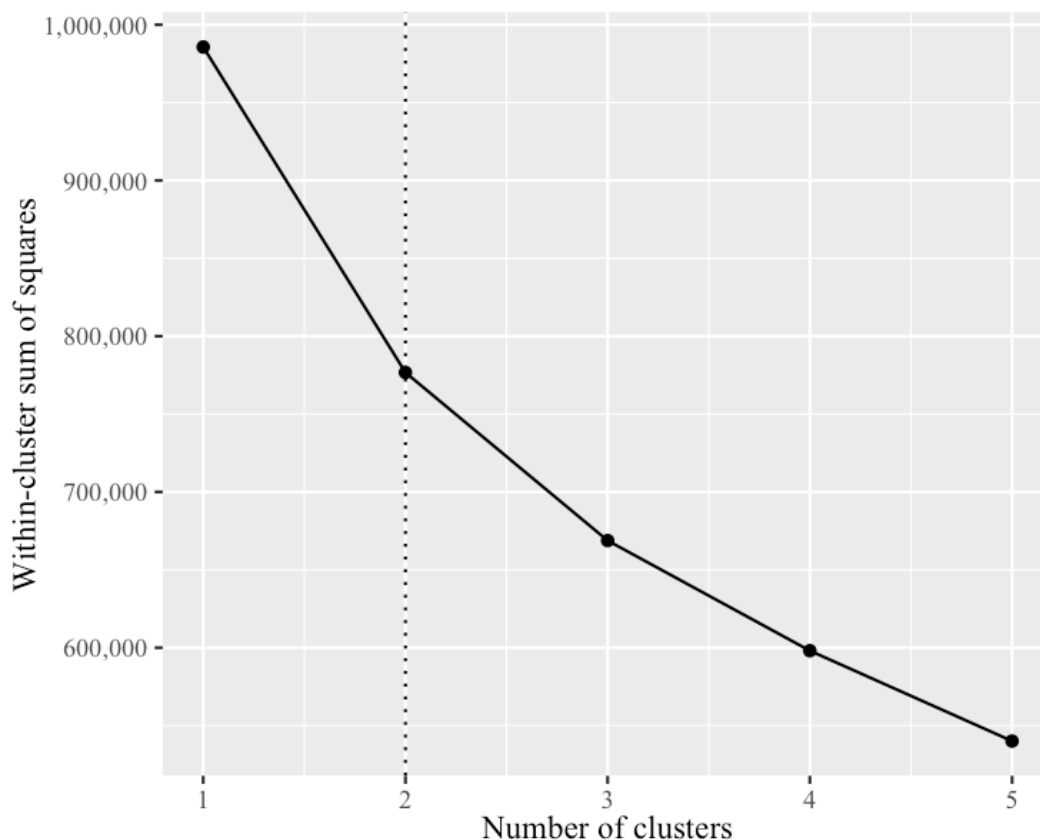
In contrast, the voice quality of female callers was characterized by increasing entropy values, which are commonly associated with a decrease in vocal clarity ($\beta=-3.23$; 95% CI 3.26 to 3.29; edf=2.75; $F_{3,1455}=13.09$; $P<.001$) and an increase in the noise to harmonics ratio, which indicates a greater proportion of noise components within the voice signal ($\beta=-2.45$; 95% CI 2.41 to 2.49; edf=1.00; $F_{1,1455}=9.49$; $P=.002$). There was also an upward shift in the first half of frequencies amongst female callers ($\beta=-10.49$, 95% CI 10.28 to 10.69; edf=1.00; $F_{1,1455}=6.77$; $P=.009$) in conjunction with an increase in spectral slope or breathiness of speech ($\beta=-6.53$; 95% CI 6.40 to 6.65; edf=1.00; $F_{1,1455}=6.91$; $P=.009$). Similar to male callers, the depth of the subharmonics was also seen to fall, suggesting an increase in the roughness of speech ($\beta=-.87$, 95% CI -0.92 to -0.82 ; edf=2.25; $F_{2,1455}=5.25$; $P=.005$).

Clustering of Voice Characteristics

k-means clustering was used to reveal the probabilistic groupings that might be apparent within the voice characteristics data set. The reduced set of 7 vocal characteristics was used in a range of clustering configurations. Figure 2 illustrates a scree plot with different cluster configurations (clusters 1-5). The variance explained by the cluster configurations appears to level off after 2 clusters. The 2-cluster configuration was poorly associated with sex of caller (Cramer's $V=0.02$), suggesting that both clusters were present within the vocal frames of male and female callers.

Binomial logistic regression (logit link) was used to validate the 2-cluster model. The full model accounted for 23.4% of the variance in the 2-cluster response variable. Model coefficients have been summarized in Multimedia Appendix 4. Speech frames in cluster 1 were characterized by higher values of entropy, first formant frequencies, 50th percentile frequency, and spectral slope values, while speech frames in cluster 2 were characterized by higher values in noise to harmonics ratio and the depth of the subharmonics. All call recordings had a mix of cluster 1 and 2 speech frames. The 2 principal component variables used to optimally separate the 2 clusters were added to the component-wise gradient boosting algorithm as additional predictors in the analysis. The subsequent classification model was first trialed without the cluster variables and achieved levels of classification accuracy of 75%, which increased to 94% when the cluster variables were included, validating the inclusion of this clustering step.

Figure 2. Scree plot of variance explained by different cluster configurations.



Classification of Psychological Distress Speech Frames Using Component-Wise Gradient Boosting

Component-wise gradient boosting was used to classify each speech frame according to precoded high versus low psychological distress on the basis of the 2-cluster configuration and the 7 voice characteristics extracted. Leave-one-caller-out cross-validation results are reported below. The gradient boosting model classified each 40-ms speech segment to either high or low psychological distress categories with an AUROC=97.39% (95% CI 96.20-98.45) and an AUCPR=97.52 (95% CI 95.71-99.12). Thus, 39,282 of 41,883 (93.79%) speech frames nested within 728 of 754 (96.6%) segments were correctly classified as exhibiting low psychological distress. Conversely, 71,455 of 75,503 (94.64%) speech frames nested within 382 of 423 (90.3%) segments were correctly classified as exhibiting high psychological distress. In terms of variable importance, the random effect for individual callers contributed 68% to the variation in level of psychological distress in the classification algorithm, followed by cluster configuration (23.4%), spectral slope (4.4%), and the 50th percentile frequency (4.2%)

Misclassification

After accounting for moderation effects by sex of caller, our component-wise gradient boosting classification algorithm incorrectly classified 2601 of 41,883 (6.21%) frames nested within 26 of 754 (3.5%) segments as exhibiting low psychological distress, while 4049 of 75,504 (5.36%) frames nested within 41 of 423 (9.7%) segments were misclassified as exhibiting high psychological distress. A review of the mental status examination descriptions that were made by the reviewing team of psychologists suggested that these call segments were often typified by considerable emotion with high anxiety but were voiced in whispered or soft tones. These particular presentations might have masked the vocal characteristics used to successfully classify the majority of the other recordings.

Discussion

Voice characteristics coupled with artificial intelligence has the potential to yield highly accurate models for the classification of emotional states. Our aim in this novel study was to classify 40-ms speech frames sourced from helpline counseling telephone calls according to high versus low levels of psychological distress. By achieving this aim to a high level of accuracy, we demonstrate an efficient, economical, and scalable approach to the detection of psychological distress in an ecologically valid telehealth setting.

We developed an ensemble approach to achieve an optimal outcome. This approach identified and validated a range of candidate vocal characteristics via penalized GAMM, which together elucidated a set of 7 characteristics (from the initial 19) with a strong predictive relationship to the binary outcome measure (clinical levels of psychological distress) when allowing for sex as a moderator. This analysis also yielded a number of heuristics that illustrate how each voice characteristic changes in response to a shift from high to low psychological distress, providing important clinical insights. In the second phase of the

analysis, k-means clustering combined with component-wise gradient boosting succeeded in accurately classifying the 40-ms speech frames (AUROC=97.39%). Although misclassifications did occur, these were largely confined to a minority of annotated segments within calls (67 of 1177 segments, 5.79%) that often featured anxiety or whispered tones. This suggests that specific caller presentations may not translate as well to vocal-informed classifications of distress, particularly when the level of distress is masked in the caller's voice.

We have achieved comparable results to Kandsberger and colleagues [12] who found significant differences ($P<.05$) between high and low psychological distress when the fundamental frequency of participants' speech was measured. Although we were unable to analyze the fundamental frequency of callers (unavailable within the frequency range of telephone calls), we did obtain significant effects for male and female callers separately on other voice characteristics, including root-mean-squared amplitude, first formant frequencies, and the depth of the harmonics among male callers; and entropy, noise to harmonics ratio, the 50th percentile of frequencies, spectral slope, and the depth of the subharmonics among female callers.

Specifically, with a shift from high to low psychological distress, male callers appeared to speak with increasing average loudness, greater vowel articulatory quality (lower first formant frequencies), and with greater roughness of speech (lower subharmonic depth). In contrast, the voice of female callers was characterized by a decrease in vocal clarity (entropy), an increase in noise in the vocal signal (noise to harmonics ratio), higher frequencies within the first half of the frequency spectrum, increasing breathiness of speech (spectral slope), and increasing roughness of speech (subharmonic depth).

However, we differed substantially from the approaches of previous authors in a number of important ways. First, we measured psychological distress using multiple sources of ratings rather than inferring it via the presence of psychopathology as previous authors have done; second, we employed an ensemble approach to ensure an optimal set of voice characteristics, allowing for nonlinear relationships; third, our use of k-means clustering has revealed probabilistic groupings within each call recording, and finally we have trained and tested a machine learning algorithm that, in conjunction with the 7 critical voice characteristics, has resulted in a more accurate classification of psychological distress. Furthermore, we have demonstrated our approach in an ecologically valid setting of telehealth, replete with background noise and variable call quality.

However, our investigation is not without limitations. We assessed the level of psychological distress using clinical researcher ratings. However, we were unable to ascertain the true level of distress experienced first-hand by the callers themselves due to the practical and ethical issues raised in terms of querying the emotional experience of callers directly when other more pressing issues were at hand (eg, imminent risk of suicide). Instead, we have relied upon objective assessments of psychological distress by trained personnel, which may indeed have advantages over a subjective assessment by callers.

We were able to measure and analyze the quality of caller's voice in terms of frequency across formants and in terms of loudness, breathiness, and clarity. However, we were not able to assess other dimensions of caller's speech, such as timing of speech. This might be important given that other authors have noted a slower speaking style in both depressed and suicidal cohorts [14]. Finally, it might be possible to achieve higher levels of accuracy with other powerful machine learning approaches, such as support vector machines and neural networks. However, this would also sacrifice the level of the transparency we have achieved with our hybrid approach.

Our investigation also has a number of strengths. Similar to Kandsberger and colleagues [12] we analyzed a sample of recordings obtained from ecologically valid settings. In this

way, previous research findings, such as those obtained by Scherer and colleagues [11], have been taken out of clinical settings and trialed within more naturalistic helpline conditions that are often typified by poor call quality and the presence of background noise. We have also classified short 40-ms frames of speech, rather than individual calls at the holistic level, according to high versus low psychological distress. That we have done so to a high level of accuracy suggests the possibility of real-time detection of psychological distress among helpline callers. This is important if such technology is to realize its full potential as a clinical decision aid with the potential for early intervention, such as call triaging. In addition, it provides a method for the measurement of changes in distress over the duration of helpline calls which could be used to evaluate helpline outcomes.

Data Availability

The data sets generated during or analyzed during the current study are not publicly available due to the private and confidential nature of the counseling calls investigated and the possibility of reidentification of the vocal based upon the vocal characteristics reported.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Analysis flowchart.

[[PNG File , 116 KB - formative_v6i12e42249_app1.png](#)]

Multimedia Appendix 2

Plain language definitions of final model vocal characteristics.

[[PDF File \(Adobe PDF File\), 57 KB - formative_v6i12e42249_app2.pdf](#)]

Multimedia Appendix 3

STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.

[[PDF File \(Adobe PDF File\), 47 KB - formative_v6i12e42249_app3.pdf](#)]

Multimedia Appendix 4

Validation of cluster configuration via binomial logistic regression.

[[PDF File \(Adobe PDF File\), 84 KB - formative_v6i12e42249_app4.pdf](#)]

References

1. Luo M, Guo L, Yu M, Jiang W, Wang H. The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public - A systematic review and meta-analysis. *Psychiatry Res* 2020 Sep;291:113190 [[FREE Full text](#)] [doi: [10.1016/j.psychres.2020.113190](https://doi.org/10.1016/j.psychres.2020.113190)] [Medline: [32563745](https://pubmed.ncbi.nlm.nih.gov/32563745/)]
2. Wang Y, Kala M, Jafar T. Factors associated with psychological distress during the coronavirus disease 2019 (COVID-19) pandemic on the predominantly general population: A systematic review and meta-analysis. *PLoS One* 2020;15(12):e0244630 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0244630](https://doi.org/10.1371/journal.pone.0244630)] [Medline: [33370404](https://pubmed.ncbi.nlm.nih.gov/33370404/)]
3. Ridner SH. Psychological distress: concept analysis. *J Adv Nurs* 2004 Mar;45(5):536-545. [doi: [10.1046/j.1365-2648.2003.02938.x](https://doi.org/10.1046/j.1365-2648.2003.02938.x)] [Medline: [15009358](https://pubmed.ncbi.nlm.nih.gov/15009358/)]
4. Kohrt BA, Rasmussen A, Kaiser BN, Haroz EE, Maharjan SM, Mutamba BB, et al. Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. *Int J Epidemiol* 2014 Apr 23;43(2):365-406 [[FREE Full text](#)] [doi: [10.1093/ije/dyt227](https://doi.org/10.1093/ije/dyt227)] [Medline: [24366490](https://pubmed.ncbi.nlm.nih.gov/24366490/)]
5. Drapeau A, Marchand A, Beaulieu-Prévost D. Epidemiology of psychological distress. In: L'Abate L, editor. *Mental Illnesses-Understanding, Prediction and Control*. Rijeka, Croatia: InTech Open; 2012:105-134.

6. Joaquim RM, Pinto AL, Guatimosim RF, de Paula JJ, Souza Costa D, Diaz AP, et al. Bereavement and psychological distress during COVID-19 pandemics: the impact of death experience on mental health. *Current Research in Behavioral Sciences* 2021 Nov;2:100019. [doi: [10.1016/j.crbeha.2021.100019](https://doi.org/10.1016/j.crbeha.2021.100019)]
7. Lee SJ, Katona LJ, De Bono SE, Lewis KL. Routine screening for psychological distress on an Australian inpatient haematology and oncology ward: impact on use of psychosocial services. *Med J Aust* 2010 Sep 06;193(S5):S74-S78. [doi: [10.5694/j.1326-5377.2010.tb03933.x](https://doi.org/10.5694/j.1326-5377.2010.tb03933.x)] [Medline: [21542451](https://pubmed.ncbi.nlm.nih.gov/21542451/)]
8. Scerri J, Sammut A, Cilia Vincenti S, Grech P, Galea M, Scerri C, et al. Reaching out for help: calls to a mental health helpline prior to and during the COVID-19 pandemic. *Int J Environ Res Public Health* 2021 Apr 23;18(9):4505 [FREE Full text] [doi: [10.3390/ijerph18094505](https://doi.org/10.3390/ijerph18094505)] [Medline: [33922749](https://pubmed.ncbi.nlm.nih.gov/33922749/)]
9. McClellan SR, Hunt M, Olsho LEW, Dasgupta A, Chowdhury M, Sparks AC. Satisfaction and mental health outcomes associated with a large regional helpline. *Community Ment Health J* 2022 Aug 11;58(6):1214-1224 [FREE Full text] [doi: [10.1007/s10597-021-00931-5](https://doi.org/10.1007/s10597-021-00931-5)] [Medline: [35015179](https://pubmed.ncbi.nlm.nih.gov/35015179/)]
10. LeCouteur A, Lehmann C, Knott V. Brief psychological distress screening on a cancer helpline: How nurses introduce, and callers respond to, the Distress Thermometer. *Eur J Oncol Nurs* 2021 Aug;53:101986. [doi: [10.1016/j.ejon.2021.101986](https://doi.org/10.1016/j.ejon.2021.101986)] [Medline: [34294575](https://pubmed.ncbi.nlm.nih.gov/34294575/)]
11. Scherer S, Gratch J, Pestian J. Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis. In: ICASSP 2015 - 2015 IEEE International Conference on Acoustics, Speech and Signal Processing.: IEEE; 2015 Presented at: IEEE International Conference on Acoustics, Speech and Signal Processing; Apr 19-24, 2015; Brisbane, Australia. [doi: [10.1109/icassp.2015.7178880](https://doi.org/10.1109/icassp.2015.7178880)]
12. Kandsberger J, Rogers SN, Zhou Y, Humphris G. Using fundamental frequency of cancer survivors' speech to investigate emotional distress in out-patient visits. *Patient Educ Couns* 2016 Dec;99(12):1971-1977. [doi: [10.1016/j.pec.2016.08.003](https://doi.org/10.1016/j.pec.2016.08.003)] [Medline: [27506580](https://pubmed.ncbi.nlm.nih.gov/27506580/)]
13. Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull* 2017 Feb;143(2):187-232. [doi: [10.1037/bul0000084](https://doi.org/10.1037/bul0000084)] [Medline: [27841450](https://pubmed.ncbi.nlm.nih.gov/27841450/)]
14. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. *Speech Communication* 2015 Jul;71:10-49. [doi: [10.1016/j.specom.2015.03.004](https://doi.org/10.1016/j.specom.2015.03.004)]
15. Iyer R, Nedeljkovic M, Meyer D. Using voice biomarkers to classify suicide risk in adult telehealth callers: a retrospective observational study. *JMIR Ment Health* 2022 Aug 15;9(8):e39807 [FREE Full text] [doi: [10.2196/39807](https://doi.org/10.2196/39807)] [Medline: [35969444](https://pubmed.ncbi.nlm.nih.gov/35969444/)]
16. Cuttillo A, O'Hea E, Person S, Lessard D, Harralson T, Boudreaux E. The distress thermometer: cutoff points and clinical use. *Oncol Nurs Forum* 2017 May 01;44(3):329-336 [FREE Full text] [doi: [10.1188/17.ONF.329-336](https://doi.org/10.1188/17.ONF.329-336)] [Medline: [29493167](https://pubmed.ncbi.nlm.nih.gov/29493167/)]
17. Van Lander A, Tarot A, Savanovitch C, Pereira B, Vennat B, Guastella V. Assessing the validity of the clinician-rated distress thermometer in palliative care. *BMC Palliat Care* 2019 Oct 17;18(1):81 [FREE Full text] [doi: [10.1186/s12904-019-0466-x](https://doi.org/10.1186/s12904-019-0466-x)] [Medline: [31623591](https://pubmed.ncbi.nlm.nih.gov/31623591/)]
18. Sousa H, Oliveira J, Figueiredo D, Ribeiro O. The clinical utility of the Distress Thermometer in non-oncological contexts: A scoping review. *J Clin Nurs* 2021 Aug 27;30(15-16):2131-2150. [doi: [10.1111/jocn.15698](https://doi.org/10.1111/jocn.15698)] [Medline: [33555631](https://pubmed.ncbi.nlm.nih.gov/33555631/)]
19. Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med* 2012 Dec 20;31(29):3972-3981. [doi: [10.1002/sim.5466](https://doi.org/10.1002/sim.5466)] [Medline: [22764084](https://pubmed.ncbi.nlm.nih.gov/22764084/)]
20. Botía JA, Vandrovcova J, Forabosco P, Guelfi S, D'Sa K, United Kingdom Brain Expression Consortium, et al. An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst Biol* 2017 Apr 12;11(1):47 [FREE Full text] [doi: [10.1186/s12918-017-0420-6](https://doi.org/10.1186/s12918-017-0420-6)] [Medline: [28403906](https://pubmed.ncbi.nlm.nih.gov/28403906/)]
21. Perkins NJ, Schisterman EF. The Youden Index and the optimal cut-point corrected for measurement error. *Biom J* 2005 Aug;47(4):428-441. [doi: [10.1002/bimj.200410133](https://doi.org/10.1002/bimj.200410133)] [Medline: [16161802](https://pubmed.ncbi.nlm.nih.gov/16161802/)]

Abbreviations

AUROC: area under the receiver operating characteristic

DSM: Diagnostic and Statistical Manual of Mental Disorders

edf: estimated degrees of freedom

GAMM: generalized additive mixed-effects regression model

PTSD: posttraumatic stress disorder

STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 28.08.22; peer-reviewed by R Norel, V Martin; comments to author 28.10.22; revised version received 15.11.22; accepted 05.12.22; published 19.12.22.

Please cite as:

Iyer R, Nedeljkovic M, Meyer D

Using Vocal Characteristics To Classify Psychological Distress in Adult Helpline Callers: Retrospective Observational Study

JMIR Form Res 2022;6(12):e42249

URL: <https://formative.jmir.org/2022/12/e42249>

doi: [10.2196/42249](https://doi.org/10.2196/42249)

PMID: [36534456](https://pubmed.ncbi.nlm.nih.gov/36534456/)

©Ravi Iyer, Maja Nedeljkovic, Denny Meyer. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 19.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Factors Associated With Intention and Use of e–Mental Health by Mental Health Counselors in General Practices: Web-Based Survey

Ann E M De Veirman¹, MA, MSc, PhD; Viviane Thewissen¹, MSc, PhD; Matthijs G Spruijt², MSc; Catherine A W Bolman¹, MPH, PhD

¹Faculty of Psychology, Open University of the Netherlands, Heerlen, Netherlands

²Therapieland, Amsterdam, Netherlands

Corresponding Author:

Catherine A W Bolman, MPH, PhD

Faculty of Psychology

Open University of the Netherlands

valkenburgerweg 177

Heerlen, 6419 AT

Netherlands

Phone: 31 455762626

Email: catherine.bolman@ou.nl

Abstract

Background: Mental health care counselors have a high intention to use e–mental health (EMH), whereas actual use is limited. Facilitating future use requires insight into underlying factors as well as eligibility criteria that mental health care counselors use in their decision to apply EMH.

Objective: The aim of this study was to unfold the intention and underlying reasons for mental health counselors to use EMH and to unveil the criteria they use to estimate patient eligibility for EMH. The theoretical framework was based on the reasoned action approach model, the Unified Theory of Acceptance and Use of Technology, and the Measurement Instrument for Determinants of Innovation model.

Methods: To empirically validate our theoretical model, a web-based survey was conducted among mental health care counselors ($n=132$). To unveil the eligibility criteria, participants were asked to rank their reasons for considering EMH suitable or unsuitable for a patient.

Results: The mean intention to use EMH was positive (mean 4.04, SD 0.64). The mean use of EMH before the COVID-19 pandemic was 38% (mean 0.38, SD 0.22), and it was 49% (mean 0.49, SD 0.25) during the pandemic. In total, 57% of the patient population was considered eligible for EMH. *Usefulness and benefits* ($\beta=.440$; $P<.001$), *Task perception* ($\beta=.306$; $P=.001$), and *Accessibility* ($\beta=.140$; $P=.02$) explained the intention to use EMH ($F_{3,131}=54.151$; $P<.001$; $R^2=0.559$). In turn, intention explained patient eligibility ($F_{1,130}=34.716$; $P<.001$; $R^2=0.211$), whereas intention and patient eligibility explained EMH use ($F_{2,129}=41.047$; $P<.001$; $R^2=0.389$). Patient eligibility partially mediated the relationship between intention to use EMH and EMH use, with a larger direct effect ($c'=0.116$; $P<.001$) than indirect effect ($c=0.065$, 95% CI 0.035-0.099; $P<.001$). Mental health counselors assessed patients' eligibility for EMH mainly through the availability of computers and the internet and patient motivation.

Conclusions: To stimulate the use of EMH, intention and patient eligibility need to be influenced. Intention, in turn, can be enhanced by addressing the perceived usefulness and benefits of EMH, perceived accessibility, and task perception. Access to a computer and patients' motivation to use EMH are important in facilitating patient eligibility. To cause an impact with EMH in general practice, mental health counselors need to be convinced of the benefits of EMH and transfer this enthusiasm to the patient. It is recommended to involve mental health counselors in the development of EMH to increase the (perceived) added value and use.

(JMIR Form Res 2022;6(12):e34754) doi:[10.2196/34754](https://doi.org/10.2196/34754)

KEYWORDS

mental health counselors; general practices; e-mental health; adoption readiness; eligibility for e-mental health; e-mental health use; mental health; eHealth

Introduction

Background

The number of patients who visit their general practitioner (GP) with psychological problems is growing rapidly and, consequently, waiting lists for treatment are increasing [1]. For this reason, GPs and mental health care professionals search for ways to organize this care more efficiently. e-Mental health (EMH) care, often a combined approach with face-to-face care (ie, blended care), could be a solution [2]. This study applied the following definition for EMH: “The use of information and communication technologies for patients with mental health complaints or disorders to inform and/or support them in recovery from their mental health to ultimately improve quality of life” [3]. Interventions involve information and communication technologies, including treatment.

For the treatment of mental problems such as depression, there is convincing evidence of the effectiveness of EMH [4-7]. Furthermore, a growing number of patients are positive about the incorporation of new remote technologies in health care for their convenience and flexibility and the possibility of following treatment at their own pace [8]. In addition, the COVID-19 pandemic required reorganization of care as face-to-face contact was problematic and sometimes even impossible. To illustrate, 64% of Dutch GP practices started with videoconferences with patients during the COVID-19 pandemic [9]. However, the structural implementation of EMH is still limited and faces many difficulties [10-14].

Mental Health Counselors and Adoption of EMH

In the Netherlands, mental health counselors (MHCs) working in general practices operate as gatekeepers in primary care concerning mental health problems [15,16]. MHCs have different educational backgrounds—approximately 50% are sociopsychiatric nurses, 20% are psychologists, and 15% are social workers [16]. These professionals treat patients with mild mental health problems and refer them to specialized care by licensed health care psychologists or psychotherapists in case of severe problems. They use EMH interventions as part of their tools to treat and coach patients. Although technical infrastructures and effective interventions are available, as well as sufficient reimbursement [17,18], actual use is low [10,11]. Facilitating future adoption and use requires insight into the most important underlying factors as well as the eligibility criteria MHCs use in their decision to apply EMH for their (vulnerable) patients. Hence, this study examined the use and nonuse of EMH by MHCs and aimed to unfold the underlying reasons and readiness to adopt EMH.

A study by Lokman et al [19] showed that 80% of GPs used EMH. Half of the GPs used EMH that was available via subscribed commercial eHealth platforms. The other half only referred patients to freely available self-help and psychoeducation websites. However, EMH was applied in <15%

of the patients. According to the MHCs in that study, the currently available EMH was only suitable for one-third of their patients. Van der Vaart et al [15] revealed that MHCs applied EMH more often than psychologists in basic mental health care (49% vs 21%).

According to MHCs, important facilitators of EMH were the following: the perceived benefits, the perceived enhancement of tools it provides to coach and treat patients, the related enrichment for their own work, and its potential to improve the quality of care. In the long term, it can also save time as patients can proactively work through certain assignments and read or reread information at home [10]. Furthermore, MHCs considered themselves sufficiently digitally skilled and capable of providing EMH [11]. However, almost half of MHCs expressed the need for a decision aid and information on the effectiveness of EMH applications [11]. Impediments perceived by MHCs were as follows: the nonadherence of patients, the preference of patients for face-to-face contact, the insufficient possibilities MHCs perceive to be properly equipped to work with the eHealth platforms and the specific EMH applications, the mismatch between the supplied EMH materials and the patients' needs, and the inflexibility of the EMH platform to attune the EMH content to the patients' specific mental health problems and needs [11,20]. The most important reasons for the perceived mismatch were insufficient command of the Dutch language, low health literacy, and lack of a computer or low digital skills [10,11,20]. MHCs often related these reasons to a low level of education. Furthermore, experienced ambiguity in regulations for the reimbursement of EMH also negatively affected the behavior of MHCs, which was still the case in 2019 [21]. More studies have been conducted on facilitators of and barriers to EMH, although they were conducted among licensed psychologists [22] and psychotherapists and often concerned specialized long-lasting psychotherapy [23].

Vulnerable Patients Visiting the MHC and Their Eligibility for EMH

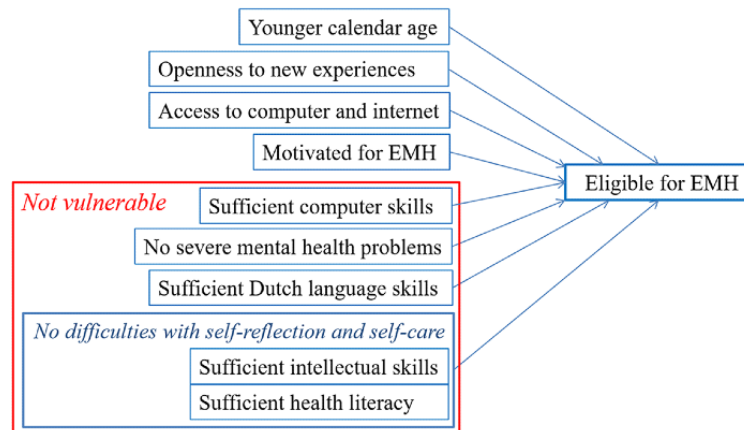
Healthy life expectancy and the prevalence of chronic diseases and mental health problems are strongly socially patterned, disproportionately affecting individuals with a lower socioeconomic position [24]. These underprivileged individuals use EMH less frequently [25,26]. This is also the case for older adults [27-30] and individuals with severe mental health problems [31,32]. This is very undesirable and regrettable as EMH offers great opportunities given that interventions can be fluidly attuned to the needs of these specific groups through the presentation of bite-sized information in plain language accompanied by reading functions, appealing visuals and animations, and speech recognition [33]. As previously mentioned, and of utmost importance, the growing use of EMH may provoke further socioeconomic health inequality [25,34-36].

In the literature, this low use of EMH by patients in lower socioeconomic positions and senior citizens is often associated

with insufficient digital, health-related, and reading skills [20,25,37-39]. It is difficult for patients to find and use information via digital channels to adequately interpret and connect them with behavioral actions. As it is often an individual consideration and decision of the MHC whether EMH elements can be offered to a specific patient, it is important to gain insight into the criteria MHCs use in their consideration of patients' eligibility for EMH. As there was no theoretical framework

available at the time of the study, we developed one, as shown in Figure 1. We combined the aforementioned aspects of vulnerability from the studies by Krijgsman et al [10] and Wouters et al [11], items from the fit-for-blended care checklist [40], and relevant aspects derived from the studies by Titzler et al [12] and Osma et al [41] and from 3 orientational interviews with MHCs before our survey (unpublished).

Figure 1. Eligibility criteria for the use of EMH. EMH: e-mental health.



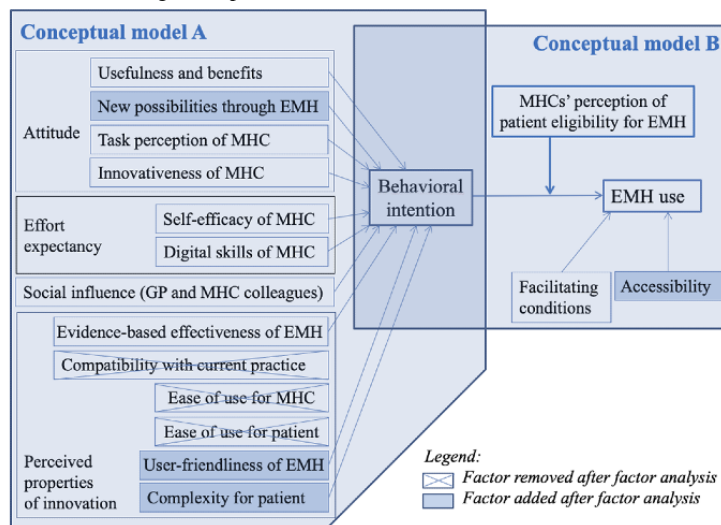
Theoretical Model of Factors Associated With Behavioral Intention and Use of EMH by MHCs

To explain the behavioral intention and use of the MHC, a theoretical model was designed (Figure 2). The model was composed of elements from the reasoned action approach (RAA) model [42], the Unified Theory of Acceptance and Use of Technology (UTAUT) [43,44], the Measurement Instrument for Determinants of Innovation (MIDI) model [45], and the Diffusion of Innovation Theory [46]. In line with the RAA and UTAUT models, use of EMH is explained by behavioral intention to use EMH, whereas behavioral intention, in turn, is explained by the constructs *Attitude* (RAA), *Social Influence* (RAA and UTAUT) and *Self-efficacy* (RAA), *Effort Expectancy* (UTAUT), and *Perceived Usefulness* (UTAUT). With the aim of formulating a universal model, Venkatesh et al [43] integrated in the UTAUT elements from 8 models, including the RAA model, to explain the acceptance and use of IT in organizations. Although the UTAUT has been used extensively also to explain the introduction of eHealth [15,28,47,48], it has been criticized for being too restricted to describe the technology acceptance of individuals [30,49]. As we agreed with this criticism and wanted to develop a model that would show all the key factors influencing the considerations and decisions of MHCs, insights from diffusion and implementation theories were considered to be crucial additions. Hence, elements from the MIDI (*Characteristics of innovations*) [45] and Diffusion of Innovation Theory (*Compatibility with current practice*, *Relative advantage*, and *Complexity*) [46,50] were added, which led to further detailing of the rather general determinant *Attitude* and the selection of the factors of the construct *Perceived properties of innovations*. Figure 2 shows that our theoretical model, in line

with the RAA and UTAUT, consists of 2 parts that will be empirically validated in this study. The first part explains the behavioral intention to use EMH (conceptual model A), and the second part explains the actual use of EMH by the MHC (conceptual model B).

Conceptual model B reflects that it will depend on the MHC's assessment of the eligibility of the patient whether the behavioral intention is transformed into actual use of EMH. This could lead to a situation in which, although an MHC might be willing to use EMH, one could decide not to use it because of certain patient characteristics. This proposition is in line with the observation by Wouters et al [11] that MHCs are generally positive about EMH, although they apply it to a minority of their patients. On the basis of the UTAUT model [43,44] and previous studies [10,20,21,51], facilitating and impeding factors related to the organization of care in the GP practice were also considered as factors that directly influence the actual use of EMH. In line with Venkatesh et al [43], the construct *Facilitating conditions* was defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system. Examples of facilitating or impeding conditions were ambiguity in reimbursement and the availability of time, management support, and information on the innovation. Although our study used the same starting definition for *Facilitating conditions* as Venkatesh et al [43], the scale in the UTAUT questionnaire was operationalized in a different way. It measured not only the availability of resources and technical support (using the definition by Thompson et al [52]), as in the *Facilitating conditions* scale in our theoretical model (model B), but also aspects of the self-efficacy of the user [43].

Figure 2. Theoretical model explaining the behavioral intention of the MHC to use EMH (conceptual model A) and MHCs' actual use of EMH (conceptual model B). EMH: e-mental health; GP: general practitioner; MHC: mental health counselor.



Focus of the Study

This study aimed to explain the behavioral intention and actual use of EMH by MHCs working in general practices. Most previous studies [10,11,18,19,53] described the use of EMH and the reasons for its use by MHCs, but they have not analyzed the correlations between barriers and facilitating factors on the one hand and behavioral intention and use on the other. Others did analyze the predictors of adoption readiness or behavioral intention [22,41,54] but did not assess EMH use itself. Our study addressed both the use of EMH and the factors that are associated with behavioral intention. Moreover, our study focused on MHCs and not (licensed) psychologists. To the best of the authors' knowledge, no such studies have been conducted so far. An additional important aspect of our study is the specific interest in the application of EMH in vulnerable groups as the growing use of EMH might impede their access to health care to a greater extent compared with other groups [25,34-36].

Hypotheses

This study tested 5 hypotheses.

The first hypothesis stated that MHCs with a high behavioral intention to use EMH score significantly higher compared with those with a low behavioral intention on the following 10 factors: (1) perceived usefulness and benefits of EMH, (2) task perception of the MHC, (3) innovativeness of the MHC, (4) social influence experienced by the MHC, (5) self-efficacy of the MHC toward the use of EMH, (6) digital skills of the MHC, (7) evidence-based effectiveness of EMH, (8) compatibility with current practice, (9) perceived ease of use of EMH for the MHC, and (10) perceived ease of use of EMH for the patient [15,22,41,43].

The second hypothesis proposed that perceived usefulness and benefits have the strongest association with intention to use EMH. The construct *Usefulness and benefits* is highly comparable with *Performance expectancy* of the UTAUT model. According to Venkatesh et al [43], this factor is proposed as the strongest predictor of behavioral intention to use technology in all technology acceptance models. Chismar and Wiley-Patton

[55] confirmed the highest importance of perceived usefulness compared with social influence and ease of use in their empirical study.

The third hypothesis was that there is a significant correlation between the behavioral intention to use EMH and actual EMH use and that the strength of this relationship is largely determined by the assessment made by the MHC of the patients' eligibility. Building on the findings that the intention to use EMH is much higher than actual EMH use and that patient characteristics might relate to this disparity, the third hypothesis asserts that the association between intention and use is moderated by the estimated patient eligibility for EMH.

Hypothesis 4 stated that facilitating and inhibiting organizational circumstances in the general practice also have a significant relationship with the use of EMH by the MHC [10,20,21,43,44,51].

According to hypothesis 5, the EMH eligibility assessment that MHCs conduct with regard to their patients includes primarily patient motivation to use EMH and the level of mental health problems. The hypothesized importance of patient motivation and the absence of disease-related contraindications (level of severity, lack of energy, lability, and suicidality) was based on the results of the qualitative study by Titzler et al [12]. Concerning the importance of severity, Osma et al [41] found therapists to have a positive intention to use EMH except in severe cases, such as psychosis, or if basic preconditions for the use of EMH are not met (eg, no internet access or insufficient literacy). Orientational interviews preliminary to our survey (unpublished) confirmed the importance of these factors.

As the validation of the questionnaire used to test the hypotheses was part of this study, the formulated hypotheses and developed model could undergo slight changes before the start of the analysis. The impact is discussed in the Methods section.

Methods

Research Design and Study Population

This was a cross-sectional study among MHCs using a web-based questionnaire (in LimeSurvey; LimeSurvey GmbH). Participants in the study had to be practicing MHCs working for at least 8 hours a week in a general practice.

Recruitment Procedure

Convenience sampling was used. The Dutch eHealth platforms Ksyos, Therapieland, and Minddistrict sent a newsletter and email message to their customer field (total >1000) to invite potential participants. However, as this did not lead to a sufficient number of respondents even after sending a reminder, an advertisement on the web page of the National MHC Association and a LinkedIn message were added. To encourage participation, 5 vouchers worth €10 (US \$10.32) were raffled.

Ethical Considerations

The web-based survey was reviewed and approved by the Ethical Review Committee of the Open University before the start (U/2020/01469/MQF). All participants gave their informed consent before taking part.

Questionnaire and Validation of the Questionnaire

Demographic Questions

The web-based questionnaire registered age, gender, type of general practice (with one or more MHCs), and educational background of the MHC (sociopsychiatric nurse, psychologist, social worker, or other) as well as the number of hours that the MHCs worked per week and their years of experience.

Factors Associated With Behavioral Intention and Use of EMH

In our model (Figure 2, model A), the behavioral intention to use EMH is the dependent variable that is explained by 10 independent variables. In this model (Figure 2, model B), behavioral intention to use EMH, in turn, is one of the 2 variables (together with the factor *Facilitating conditions*) explaining EMH use. To measure the behavioral intention to use EMH, we formulated 4 items in a similar way as for *Behavioral intention to use the system* in the UTAUT questionnaire [43]. The 4 items on this scale were scored by the MHCs on a 5-point Likert scale ranging from strongly disagree (1) to strongly agree (5). After item analysis, 3 of the 4 items remained, resulting in a reliable scale (3 items; Cronbach $\alpha=.85$; Table 1).

In accordance with the theoretical model (conceptual model A; Figure 2), the 10 scales of the factors that are proposed to relate to behavioral intention to use EMH were measured as independent variables (*Usefulness and benefits*, *Task perception of the MHC*, *Innovativeness of the MHC*, *Social influence*, *Self-efficacy of MHC*, *Digital skills of the MHC*, *Evidence-based effectiveness of EMH*, *Compatibility with current practice*, and *Ease of use for the MHC and for the patient*). The factor *Facilitating conditions* was also measured as an independent variable, although it was directly related to EMH use (and part of conceptual model B; Figure 2).

To measure these 11 scales (ie, *Facilitating conditions* and the 10 scales of the factors that relate to behavioral intention), the items of the validated eMental Health Adoption Readiness scale by Feijt et al [56] were used as a starting point and complemented with items based on the MIDI questionnaire [45], the UTAUT questionnaire [43], a digital skill measurement tool, and own insights and literature (eg, related to the COVID-19 pandemic). To refine the items, our own scientific expertise [50,57,58] and input from practice were important. The latter input was collected from one of the authors working at Therapieland and by conducting 3 interviews with MHCs to collect information on important barriers, facilitators, and other considerations for the implementation of EMH and to check the content and completeness of our questionnaire. Before completion, the questionnaire was pretested among 8 MHCs. All items included in the scales were scored on a Likert scale. To validate the subscales of the questionnaire, these 11 scales were subjected to a principal factor analysis [59]. As the factor analysis led to some adjustments of the subscales (item composition, meaning, and sometimes factor name), the theoretical model and hypotheses were slightly modified. The changes in the theoretical model are shown in Figure 2. The factor analysis revealed that items of the scales *Usefulness and benefits of EMH* and *Compatibility with current practice* had to be recategorized in the revised scale of *Usefulness and benefits of EMH* and the new scale *New possibilities through EMH*. In addition, the items of the scales *Ease of use for the MHC* and *Ease of use for the patient* were redistributed, and the scales were renamed (*User-friendliness of EMH for both the MHC and the patient* and *Complexity for the patient*). After these adjustments, the measurement scales of the 10 factors that were assumed to influence behavioral intention showed a reasonable to good internal consistency (Cronbach $\alpha \geq .68$ and $\leq .89$), as shown in Table 1. After validation, the *Facilitating conditions* measurement scale was also adjusted. Its reliability was sufficient (5 items; Cronbach $\alpha=.78$; Table 1). The content referred to organizational infrastructure with items on available time, finances, and administrative workload. In the factor analysis, *External obligation* and *Accessibility* emerged as new scales. The *External obligation* scale comprised 2 items that originally belonged to the *Facilitating conditions* scale. As this scale was not reliable (2 items; Cronbach $\alpha=.36$; Table 1), it was not included in the revised hypothesized model (Figure 2). The *Accessibility* scale had good reliability (2 items; Cronbach $\alpha=.85$; Table 1). *Accessibility* comprised 2 items of the original *Ease of use* scales and was interpreted as complementary to *Facilitating conditions* as it measured whether EMH was easy to obtain and use (because of technical infrastructure). It was presumed to have a direct relation to EMH use as its content came close to the definition of *Facilitating conditions* by Thompson et al [52]: “objective factors, ‘out there’ in the environment, that several judges or observers can agree make an act easy to do.”

For the factors from Table 1, mean scores were used in the regression model. The items that gave a negative evaluation of EMH were recoded. A high score (>3) indicates a positive evaluation by the MHC of that specific factor (eg, the MHC perceives EMH as useful and beneficial, and the MHC does not find that EMH uses complex language).

Table 1. Operationalization of the factors related to e-mental health (EMH) adoption readiness and use after factor analysis.

	Items, N	Cronbach α	Items
Factors explaining intention to use EMH			
Usefulness and benefits of EMH	9	.89	<ul style="list-style-type: none"> • EMH has advantages for the care I give. • EMH does not improve the care I give. • EMH has no added value for my work as MHC^a. • Using EMH treatment allows me to get faster results. • Using EMH has added value for my patients. • EMH is a nice addition to f2f^b contact. • Patients come for f2f treatment, it takes a lot of effort to convince them of benefits of blended care. • EMH fits well with how I am used to working. • A patient who cooperates well in f2f therapy will generally also cooperate in EMH assignments.
New possibilities of EMH	5	.75	<ul style="list-style-type: none"> • Using EMH between f2f sessions makes the sessions more efficient. • In the EMH programs, I can easily give feedback to the patients, and that has a motivating effect for the patient. • EMH ensures that patients can read information about the treatment. • Thanks to EMH, I can provide guidance to the patient even if the circumstances prevent me from making f2f appointments. • If it is temporarily not possible to make f2f appointments, the treatment will continue through EMH.
MHC task perception	3	.78	<ul style="list-style-type: none"> • EMH is an indispensable part of the MHC work. • EMH fits in well with my work as MHC. • EMH does not fit the profession of MHC.
Innovativeness of the MHC	5	.80	<ul style="list-style-type: none"> • I am involved in setting up initiatives for the development of new EMH applications. • Compared to colleagues, I often use EMH. • Compared to colleagues, I take a lot of initiative in the field of EMH. • I have ideas about what more could be developed in EMH applications (eg, virtual reality, gaming, biofeedback) • In my work, I try to encourage colleagues to use EMH.
Social influence	5	.73	<ul style="list-style-type: none"> • My MHC colleagues use EMH. • My GP^c uses the latest eHealth options. • My GP expects me to use EMH when treating patients. • My MHC colleagues expect me to use EMH when treating patients. • Use of EMH is part of the policy of our general practice.
Self-efficacy of the MHC	5	.81	<ul style="list-style-type: none"> • Using EMH applications is easy for me. • I still lack skills to give online therapy (via video calling). • To start using EMH, I need to learn new skills. • I need training in video calling. • I need to practice to give empathetic written feedback.
Digital skills of the MHC	4	.89	<ul style="list-style-type: none"> • I can look up relevant information on the internet. • I can work independently with a computer, and I am able to solve small problems myself. • I can handle email messages well (receiving, sending, adding attachments). • I can work well with digital documents: create, open, close, save in the correct folder, etc.
Evidence-based effectiveness of EMH	2	.77	<ul style="list-style-type: none"> • The EMH programs I work with have been proven to be effective. • The EMH programs I work with are based on correct, scientific knowledge.
User-friendliness of EMH	5	.69	<ul style="list-style-type: none"> • The EMH programs are very user-friendly and invite the patients to use them. • It is quite possible to use parts of the programs, the patient does not have to go through the entire program every time. • I find the range of EMH programs clear. • The EMH programs are very user-friendly and invite the MHC to use them. • During the Covid pandemic, we are helped quickly and well by the suppliers of EMH.

	Items, N	Cronbach α	Items
Complexity for patients	2	.68	<ul style="list-style-type: none"> • Most EMH programs use difficult language. • EMH is only suitable for people with a higher education.
Factors explaining EMH use			
Facilitating conditions	5	.78	<ul style="list-style-type: none"> • I do not get enough time from the GP to study the possibilities of EMH. • The administrative burden prevents me from using EMH. • It is impossible to declare the time it takes to delve into the possibilities of EMH. • There is too little time in the f2f contacts to pay attention to EMH. • Because the agenda is completely filled with f2f appointments, there is too little time to give good feedback to patients in the EMH platform.
Accessibility	2	.85	<ul style="list-style-type: none"> • I use EMH applications that are easily available. • I use EMH applications that are easy to use.
External obligation ^d	3	.36	<ul style="list-style-type: none"> • I am obliged by the organization to use EMH. • The Covid pandemic forces me to use EMH.
Intention to use EMH	3	.85	<ul style="list-style-type: none"> • I plan to use EMH as part of my treatments. • I would like to supplement face-to-face contacts with EMH. • After the Covid pandemic, I plan to continue using EMH.

^aMHC: mental health counselor.

^bf2f: face-to-face.

^cGP: general practitioner.

^dNot added to the theoretical model because of low reliability.

EMH Use and Patient Eligibility

EMH use and MHCs' perception of patient eligibility were both operationalized as estimates by the MHC of the proportion of their patient population (1) who used EMH before the COVID-19 pandemic (ie, EMH use) and (2) whom they considered eligible for EMH (ie, patient eligibility). EMH use was measured using the question "Can you give an estimate of the proportion of patients for whom you have used eMH before the Covid pandemic?" MHCs' perception of patient eligibility was measured using the question "Can you give an estimate of the proportion of your patients that are eligible for EMH?" Answering categories were 90% to 100%, 80% to 90%, 60% to 80%, 40% to 60%, 20% to 40%, 10% to 20%, and <10%. These population estimates were, in fact, alternative assessments as it was impossible in a regression model to relate the eligibility of an individual patient to the decision of the MHC to use EMH, as intended in the draft version of the conceptual model where EMH use was a dichotomous variable. To transform these categorical variables into ratio variables, the response categories were recoded as follows: <10% became 0.05, 10% to 20% became 0.15, 20% to 40% became 0.30, 40% to 60% became 0.50, 60% to 80% became 0.70, 80% to 90% became 0.85, and 90% to 100% became 0.95. The ratio variables were used in the regression model. Subsequently, the main reasons for (not) using EMH were asked using open questions. Respondents with increased use of EMH since the start of the COVID-19 pandemic were asked whether they expected that this increase would become permanent.

This study primarily aimed to explain EMH use before the COVID-19 pandemic. As the study started shortly after the beginning of the COVID-19 pandemic, some questions on the

use of EMH during that period were also included. Social distancing in health care, as enforced by the Dutch government in March 2020, required GPs and MHCs to switch to remote patient contact as much as possible. This significantly increased the use of EMH in general practices [9].

Analysis

All analyses were performed using SPSS software (version 26; IBM Corp). A 95% significance level was used for all tests (Cronbach $\alpha=.05$) except for the 2-tailed *t* tests for hypothesis 1, where a significance criterion of $P<.005$ was used to correct for multiple testing (based on the Bonferroni correction).

To guarantee sufficient power for the regression analysis with 10 independent variables (*m*), the number of respondents (*n*), estimated by the rule of thumb of Tabachnick and Fidell [60], must exceed 130 (ie, 50+8 *m*).

The characteristics of the participants were analyzed using descriptive statistics. Correlation analyses were performed using all variables of the research model. To test hypothesis 1, a dummy variable was first calculated using a median split of EMH adoption readiness. Second, *t* tests were performed to reveal whether there were significant differences in the mean values of the 10 factors between respondents with high and low adoption readiness. To verify the second hypothesis, a multiple backward linear regression analysis was performed with behavioral intention as the dependent variable and the 10 factors from the research model (Figure 2, model A) as predictor variables.

The third hypothesis consisted of 2 parts. The analysis started with the second part: patient eligibility is a moderator variable in the association between behavioral intention and EMH use.

This was investigated with a moderator analysis using the Hayes PROCESS module [61]—EMH use was the dependent variable (Y), behavioral intention was the independent variable (X), patient eligibility was the moderator variable (W), and *Accessibility* and *Facilitating conditions* were the covariates. Although this analysis also tested the first part of hypothesis 3 as well as hypothesis 4, a multiple linear regression analysis was also performed (backward) with EMH use as the dependent variable (Y) and behavioral intention, patient eligibility, and *Accessibility* and *Facilitating conditions* as independent variables as it facilitated the interpretation of the regression coefficients.

For the fifth hypothesis, descriptive statistics were used.

Table 2. Characteristics of the participants (N=132).

Characteristic	Values
Age (years), mean (SD; range)	47.4 (10.7; 27-65)
Age (years; median split), median	49.0
Younger group (<50), n (%)	68 (51.5)
Older group (≥50), n (%)	64 (48.5)
Sex, n (%)	
Male	20 (15.2)
Female	111 (84.1)
Intersex	1 (0.8)
Professional background, n (%)	
Psychologist	30 (22.7)
Sociopsychiatric nurse	26 (19.7)
Social worker	25 (18.9)
Other ^a	51 (38.6)
Work hours per week, mean (SD; range)	23.0 (7.6; 8-40)
Number of patients per week, mean (SD; range)	25.6 (10.5; 4-65)
Work experience (years), mean (SD; range)	4.8 (3.5; 0-20)
≤10, n (%)	128 (97)
≤2, n (%)	45 (34.1)
Number of MHC^b colleagues, n (%)	
None	42 (31.8)
Platform used, n (%)	
Only Therapieland	41 (31.1)
Only Minddistrict	80 (60.6)
Therapieland and Minddistrict	9 (6.8)
Neither Therapieland nor Minddistrict	2 (1.5)

^aThis group was very diverse and comprised, among others, psychiatric nurses, applied psychologists, orthopedagogists, and ergotherapeutists.

^bMHC: mental health counselor.

Variables and Correlation Analysis

In [Multimedia Appendix 1](#), the variables of the research model—mean values and SDs—are listed as well as the Pearson correlation coefficients between them. All Pearson correlation

Results

Description of Participants

The web-based questionnaire was filled out 146 times between April 21, 2020, and August 8, 2020. A total of 4.8% (7/146) of the respondents did not give informed consent and, of the 139 respondents who did give informed consent, 7 (5%) were not active as MHCs. Hence, 132 questionnaires were available for the analysis.

Of the 132 respondents, 111 (84.1%) were female, and the mean age was 47.4 (SD 10.7) years. They worked between 8 and 40 hours per week as MHCs in one or more general practices ([Table 2](#)).

coefficients were below the critical value of 0.80, confirming the independence of the constructs.

The correlation analysis showed that behavioral intention correlated significantly with 8 of the 10 scales. Behavioral

intention had a strong positive correlation with *Usefulness and benefits of EMH* and *Task perception* and a medium correlation with *New possibilities*, *Innovativeness of the MHC*, *Social influence*, and *Evidence-based effectiveness*. The correlations between *User-friendliness* and *Complexity* were low.

Hypothesis 1: Differences Between Groups With Low and High Behavioral Intention

Via a median split (median value of 4.0), the population was divided into a large group (98/132, 74.2%) with high behavioral intention (≥ 4 ; mean 4.32, SD 0.41) and a smaller group (34/132, 25.8%) with lower behavioral intention (< 4 ; mean 3.25, SD 0.48). Using univariate *t* tests, the scores on the 10 factors were

compared for the high- and low-behavioral-intention groups (Table 3). This partly confirmed the first hypothesis: MHCs with high behavioral intention scored significantly higher than MHCs with low behavioral intention on 6 of the 10 factors—*Usefulness and benefits of EMH*, *New possibilities through EMH*, *Task perception of the MHC*, *Innovativeness of the MHC*, *Social influence*, and *Evidence-based effectiveness of the EMH*—but not on perceived *Self-efficacy*, *Digital skills*, *User-friendliness*, or *Complexity of EMH*. These 6 factors were also the factors with a moderate to high correlation with EMH adoption readiness. The mean behavioral intention to use EMH in the total group was positive (mean 4.04, SD 0.64).

Table 3. Mean scale scores and differences for mental health counselors (MHCs) with low and high behavioral intention (BI).

	All MHCs, mean (SD)	MHCs with low BI, mean (SD)	MHCs with high BI, mean (SD)	<i>t</i> test (<i>df</i>) ^a	<i>P</i> value ^b
Usefulness and benefits of EMH ^c	3.80 (0.60)	3.19 (0.69)	4.01 (0.38)	-6.56 (40.2) ^a	<.001 ^a
New possibilities through EMH	3.69 (0.60)	3.27 (0.53)	3.83 (0.56)	-5.12 (130)	<.001
Task perception of the MHC	4.02 (0.79)	3.29 (0.78)	4.27 (0.62)	-7.40 (130)	<.001
Innovativeness of the MHC	2.77 (0.76)	2.38 (0.82)	2.91 (0.69)	-3.67 (130)	<.001
Social influence	3.30 (0.66)	2.97 (0.63)	3.42 (0.63)	-3.55 (130)	.001
Evidence-based effectiveness of EMH	3.73 (0.58)	3.41 (0.72)	3.84 (0.49)	-3.19 (43.8)	.003 ^a
Self-efficacy of the MHC	3.60 (0.73)	3.61 (0.77)	3.60 (0.72)	0.10 (130)	.93
Digital skills of the MHC	4.44 (0.59)	4.57 (0.51)	4.39 (0.61)	1.55 (130)	.12
User-friendliness for the MHC and the patient	3.58 (0.56)	3.48 (0.48)	3.61 (0.59)	-1.23 (130)	.22
Complexity for the patient	3.79 (0.65)	3.69 (0.69)	3.83 (0.64)	-1.04 (130)	.30

^aLevene test for unequal variances; in all other cases, equal variances.

^b $P < .005$ to control for multiple testing.

^cEMH: e-mental health.

Hypothesis 2: Factors Associated With Behavioral Intention to Use EMH

Factors associated with behavioral intention were examined using a multiple regression analysis (backward; Table 4). The first regression model (model 1; Table 4) had 10 independent variables, and these factors explained 56% of the behavioral intention ($F_{10,131}=15.113$; $P < .001$; $R^2=0.555$). However, only

the scales on *Usefulness and benefits* ($b=.401$; $P < .001$) and *Task perception* ($b=.339$; $P=.001$) were significant. The last model (model 9; Table 4) explained 54% of the variance in behavioral intention ($F_{2,131}=76.102$; $P < .001$; $R^2=0.541$) with *Usefulness and benefits* ($b=.471$; $P < .001$) and *Task perception* ($b=.316$; $P=.001$). Hypothesis 2, which stated that the expected usefulness and benefits of EMH are the most important predictors of behavioral intention, was confirmed.

Table 4. Backward regression analysis to explain the behavioral intention to use e-mental health (EMH).^a

	R^2	B (SE)	β	P value
Model 1	0.56			<.001
Constant		0.852 (0.481)	N/A ^b	.08
Usefulness and benefits of EMH		0.425 (0.113)	.401	<.001
New possibilities through EMH		0.012 (0.080)	.011	.88
Task perception of the MHC ^c		0.274 (0.077)	.339	.001
Innovativeness of the MHC		-0.024 (0.065)	-0.028	.72
Social influence		0.034 (0.066)	.035	.61
Self-efficacy of the MHC		-0.057 (0.057)	-0.065	.32
Digital skills of the MHC		-0.012 (0.077)	-0.011	.88
Evidence-based effectiveness of EMH		0.054 (0.076)	.049	.48
User-friendliness for the MHC and the patient		0.089 (0.074)	.079	.23
Complexity for the patient		0.032 (0.065)	.033	.62
Model 9	0.54			<.001
Constant		1.120 (0.244)	N/A	<.001
Usefulness and benefits of EMH		0.499 (0.094)	.471	<.001
Task perception of the MHC		0.255	.316	.001

^aThere was no multicollinearity (all variance inflation factor values were <4 and tolerance values were >0.20).

^bN/A: not applicable.

^cMHC: mental health counselor.

Hypotheses 3 and 4: Factors Explaining EMH Use

EMH use before the COVID-19 pandemic (mean 0.38, SD 0.22) and patient eligibility (mean 0.57, SD 0.23) were determined as percentages of the patient population from the MHCs' estimation.

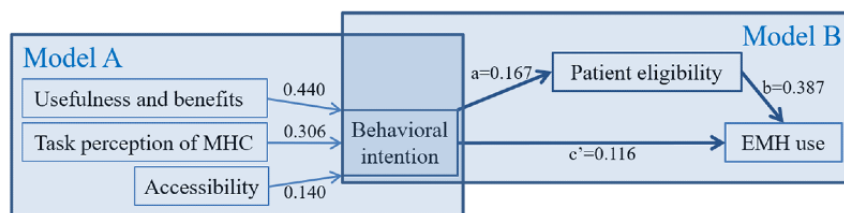
Using the Hayes PROCESS module, a moderator analysis was performed with EMH use as the dependent variable (Y), behavioral intention as the independent variable (X), patient eligibility as the moderator variable (W), and *Accessibility* and *Facilitating conditions* as covariates. The model was significant ($F_{5, 126}=17.246$; $P<.001$; $R^2=0.406$), but the product term $X \times W$ was not ($F_{1, 126}=3.121$; $P=.08$; $DR^2=0.015$). Therefore, patient eligibility was not a moderator and, thus, the second part of the third hypothesis was rejected. However, the first part of the third hypothesis was confirmed—both $r=0.55$ ([Multimedia Appendix 1](#)) and the PROCESS regression analysis ($b=.140$; $P<.001$) showed a significant positive association between behavioral intention and EMH use. *Facilitating conditions* (mean 3.38, SD 0.77) and *Accessibility* (mean 3.77, SD 0.68) were not significantly related to EMH use, nor was the correlation significant ([Multimedia Appendix 1](#)). Hence,

hypothesis 4 (*Facilitating conditions* and *Accessibility* have a significant relationship with EMH use) was rejected.

The backward multiple linear regression analysis that was also performed to facilitate the interpretation of the regression coefficients confirmed that only 2 variables were significant factors explaining EMH use: patient eligibility ($b=.399$; $P<.001$) and behavioral intention ($b=.330$; $P<.001$). Together, they explained 39% of the variance in EMH use ($F_{2,131}=41.047$; $P<.001$; $R^2=0.389$).

As patient eligibility was not a moderator variable but a direct scale of EMH use, the Hayes PROCESS module was used to examine whether it could be a mediator variable instead. The results are shown in [Figure 3](#) (model B). Behavioral intention explained patient eligibility ($F_{1,130}=34.716$; $P<.001$; $R^2=0.211$), whereas behavioral intention and patient eligibility explained EMH use ($F_{2,129}=41.047$; $P<.001$; $R^2=0.389$). Hence, it was concluded that there was a partial mediation between behavioral intention and EMH use through patient eligibility. The direct effect ($c'=0.116$) of behavioral intention was larger than the indirect effect ($c=0.065$, 95% CI 0.035-0.099). All coefficients were significant, with $P<.001$.

Figure 3. Empirically validated theoretical model with non-standardized regression coefficients. Model A shows the three significant factors of behavioral intention. Model B shows the results of the Hayes PROCESS mediation analysis of patient eligibility in the relationship between behavioral intention and EMH use (N=132). EMH: e-mental health; MHC: mental health counselor.



Verification of the Complete Model

To verify the complete theoretical model (ie, models A and B in Figure 2 combined), an additional multiple regression analysis (backward) was performed for behavioral intention. In addition to the remaining factors *Usefulness and benefits of EMH* and *Task perception of the MHC* (model 9; Table 4), the 2 factors for which a direct relationship with EMH use was assumed (*Facilitating conditions* and *Accessibility*) were included in the regression analysis. It turned out that the scales *Usefulness and benefits* ($b=.440$; $P<.001$), *Task perception* ($b=.306$; $P=.001$), and *Accessibility* ($b=.140$; $P=.02$) were significant explanatory factors of behavioral intention ($F_{3,131}=54.151$; $P<.001$; $R^2=0.559$). The empirically validated theoretical model in which only the significant factors related to EMH adoption readiness and use were considered is shown in Figure 3.

Increased Use of EMH Since the Start of the COVID-19 Pandemic

The use of EMH before and since the start of the COVID-19 pandemic was compared. The mean use before the COVID-19 pandemic was 38.1% (mean 0.381, SD 0.219), and it increased to 49.4% after the start of the COVID-19 pandemic (mean 0.494, SD 0.247). There was a significant and strong correlation ($r=0.794$; $P<.001$) between use before and use since the start of the COVID-19 pandemic. More than half (70/132, 53%) of the MHCs had increased the use of EMH since the start of the COVID-19 pandemic, and this increase was also expected to be permanent for slightly more than half (36/70, 51%) of this group.

Hypothesis 5: Factors Taken Into Account by MHCs in Assessing Patient Eligibility for EMH

The answers to the question asking to make a top 5 of the most important characteristics to be perceived as not eligible for EMH are summarized in Figure 4. In this graph, it is shown how often the items were chosen by respondents as top 1, top 2, top 3, top 4, and top 5. The order in which the characteristics are presented is considered the order of importance and is determined by the

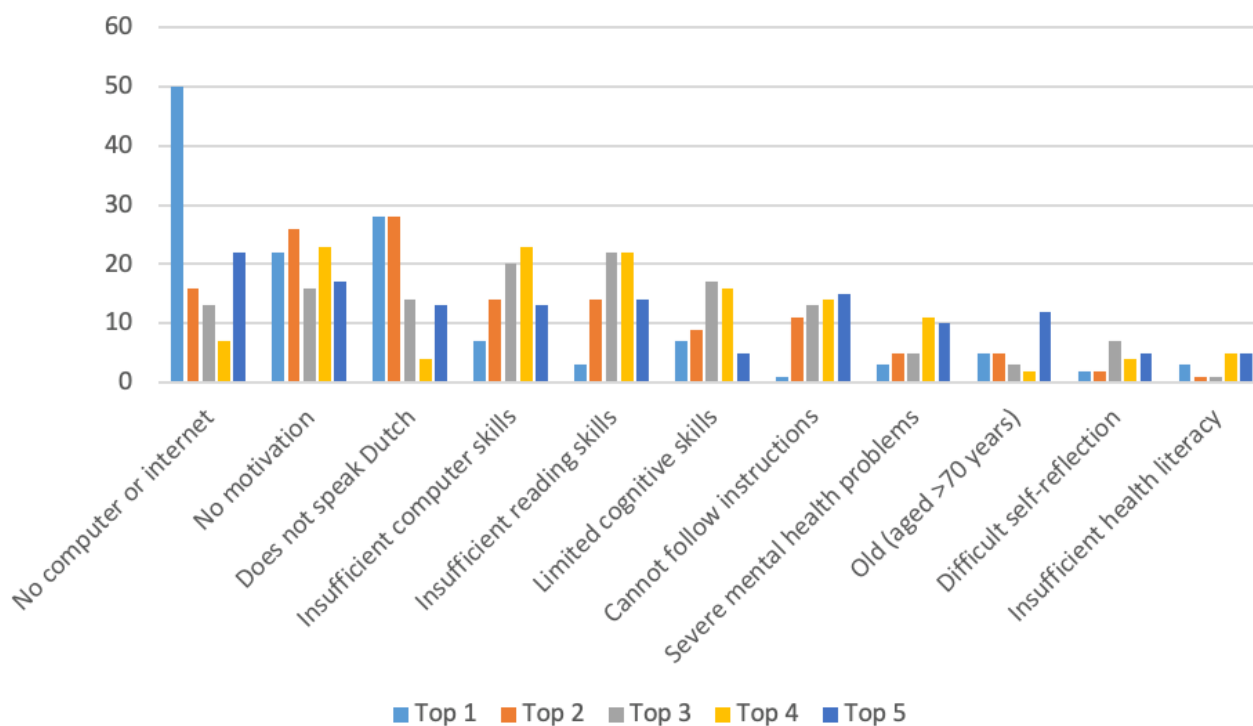
calculated weighted average score (weight factors: top 1×5, top 2×4, top 3×3, top 4×2, and top 5×1). A patient having no access to a computer or the internet was considered to be the most important reason to deem them not eligible for EMH. The total score for this characteristic was 389 (ie, $5 \times 50 + 4 \times 16 + 3 \times 13 + 2 \times 7 + 1 \times 22$, with 50 being the number of times it was chosen as top 1 by the respondents, 16 being the number of times it was chosen as top 2, and so on). *No motivation* came in the second position with a total score of 325, followed by *not speaking Dutch* (whereas EMH programs are offered in Dutch) with a total score of 315.

Regarding the question about the minimum properties or characteristics that a patient needs to have to be eligible for EMH, *access to a computer and the internet* was chosen 114 times, *motivation to work with EMH* was chosen 108 times, *openness to new experiences* was chosen 61 times, *reading and writing skills (in Dutch)* was chosen 57 times, *sufficient cognitive skills* was chosen 38 times, *discipline* was chosen 36 times, *demonstrable computer skills* was chosen 30 times, *other* was chosen 6 times, and *demonstrable health literacy* was chosen once.

The availability of a computer and the internet showed the highest scores in both assessments. When considering the patients' attributes or skills, patient motivation for EMH use emerged as the most important factor. The severity of the mental health problems was given a low score. Therefore, the fifth hypothesis was partly rejected.

Many patient-related answers were given to the open questions about the main reason for (not) using EMH. This not only showed that the patient plays an important role in MHCs' decision to use EMH, as was confirmed by the regression analysis, but also provided information about the factors that determine the assessment of the patients' eligibility. Again, a lack of digital skills, limited language skills, and lack of motivation on the part of the patient were the predominant reasons. Interestingly, 10.6% (14/132) of MHCs mentioned the lack of a suitable EMH module for the patient in question.

Figure 4. Number of times the characteristics were chosen as top-1 to top-5 that make a patient not eligible for e-mental health (EMH; n=131). Weighted total: no access to a computer- or the internet: 389, not motivated to work with EMH: 325, does not speak Dutch (and EMH programs are not available in their language): 315, insufficient computer skills: 210, insufficient reading skills: 195, limited cognitive skills: 159, cannot follow the instructions of the general practitioner or mental health counselor: 131, severe mental health problems: 82, old (aged > 70 years): 70, has difficulties with self-reflection: 52, and insufficient health literacy: 37.



Discussion

Principal Findings

This study aimed to unfold the underlying reasons of MHCs to adopt and use EMH as well as unveil the criteria MHCs use to estimate patient eligibility for EMH. The factor *Perceived usefulness and benefits* was the strongest predictor of behavioral intention to use EMH. In turn, behavioral intention had a direct and indirect effect (via estimated patient eligibility for EMH) on the use of EMH. To estimate patient eligibility for EMH, patients' access to a computer and sufficient digital and language skills and motivation to use EMH were important.

In this study, the intention to use EMH was high, which is important to further disseminate and embed EMH use in practice as intention positively related to use in our study. The fact that intention to use EMH is high among MHCs is in line with previous findings in the Netherlands [10,11] but could be due to the broad definition of EMH and the fact that especially MHCs who are positive toward EMH might have been more likely to participate in this study as other studies have revealed that intentions varied between EMH applications [62] or were higher for the treatment of mild mental problems [17,63].

During the COVID-19 pandemic, EMH use was seen to increase from 38% on average to 49%; both percentages were already considerably higher than in 2016, when it was <15% [19]. More than half (70/132, 53%) of the MHCs reported increased use of EMH since the start of the COVID-19 pandemic, which was often expected to be permanent. It would be interesting to investigate over time whether this expectation also came true.

MHCs with a high behavioral intention scored significantly higher for 6 of the 10 factors (*Usefulness and benefits of EMH*, *New possibilities through EMH*, *Task perception of the MHC*, *Innovativeness of the MHC*, *Social influence*, and *Evidence-based effectiveness of EMH*) compared with those with a low EMH use intention. The results of the regression analysis stress the importance of *Usefulness and benefits of EMH*, *Accessibility*, and *Task perception* to increase MHCs' intention to use EMH. *Usefulness and benefits of EMH* appeared to be the most important predictor of behavioral intention. *Accessibility* and *Task perception* were also predictive, although less strong. The primary importance of *Usefulness and benefits* is in line with relevant literature on the introduction of IT as well as in general [43] as applied to health care environments [15,55]. As a construct, *Usefulness and benefits of EMH* is closely related to *Performance expectancy* from the UTAUT model. *Performance expectancy* was defined as the extent to which a person believes that using the IT system will improve their work performance and is advantageous for patients, and it was also found to be the determinant variable of intention to use [43]. In applying the technology acceptance model, Chismar and Wiley-Patton [55] found that *Perceived usefulness* had a significantly strong effect on the intention of pediatricians to use internet-based applications, but *Subjective norm* and *Perceived ease of use* did not. Van der Vaart et al [15], who used the UTAUT model in their study on the use of web-based self-management interventions, also found *Performance expectancy* to be the most significant predictor of intention to use for both MHCs and primary care psychologists.

The second significant factor of behavioral intention in our study (ie, *Accessibility*) is not 100% equal to any of the scales

of the UTAUT model, although it is related to *Effort expectancy*. *Effort expectancy* was defined by Venkatesh et al [43] as the degree of ease of use and is related to *Accessibility*, *User-friendliness*, and *Perceived self-efficacy* in our models A and B (Figure 2). Apparently, we need 3 factors in our model (Figure 2) to do justice to the different aspects of *Effort expectancy* by Venkatesh et al [43]. In the study by van der Vaart et al [15], *Effort expectancy* was a significant predictor of intention to use for MHCs but not for primary care psychologists. *Perceived ease of use* was not significant in the study by Chismar and Wiley-Patton [55] among pediatricians. These inconsistent results may be indicative of a construct that is formulated rather broadly. Still, there is clear consensus among the studies regarding the dominant role of perceived usability over ease of use. Only if users see added value in the use of EMH will they consider using it, and only then will ease of use become important [64].

The third significant factor of behavioral intention in our study (ie, *Task perception*) does not exist in the UTAUT model and, therefore, is not in the study by van der Vaart et al [15] either. However, *Task perception* relates very well to *Job relevance*, which was the factor determining perceived usefulness in the study by Chismar and Wiley-Patton [55]. In this study, the correlation between *Task perception* and *Usefulness and benefits* emerged as the strongest, confirming a close relationship.

It is interesting to note that *Social influence* was not a significant variable of behavioral intention, which is in line with the studies by van der Vaart et al [15] and Chismar and Wiley-Patton [55]. Seemingly, MHCs, primary care psychologists, and pediatricians make up their minds independent of others' opinions and, therefore, do not need emphasis on influencing intentions to use EMH.

Correlation and regression analyses showed that MHCs with a higher behavioral intention were more likely to use EMH more frequently than their colleagues with a lower behavioral intention. The estimated eligibility of the patient population did not appear to be a moderator variable in the relationship between behavioral intention and EMH use. However, patient eligibility was found to be both a direct factor of EMH use and a mediator variable between behavioral intention and EMH use. The direct effect of patient eligibility on the use of EMH is larger than the direct effect of behavioral intention, which makes it the most determining factor in the decision of the MHC whether to use EMH. It is interesting, though unexpected, that the behavioral intention of MHCs influences their estimation of the eligibility of the patient population. This issue will be discussed later in this section. The question about the inverse relationship (ie, whether the patients' eligibility influences behavioral intention) was not investigated as it is incompatible with the conceptual model. In this model, behavioral intention is a perception of the MHC that already exists before the MHC estimates the eligibility of the patient population involved.

The motivation of the patient emerged as the most important condition to consider a patient eligible for EMH in addition to access to a computer with internet and digital skills. Knowledge of the Dutch language and good reading skills were also high in ranking. The severity of the mental health problems was

given a much lower ranking, which was unexpected considering the literature [12,41]. The fact that MHCs in our study considered the severity of the patients' problems less of a barrier to using EMH may be explained by the fact that their average patient has mild to moderate mental problems compared with the patients with depression in the studies by Titzler et al [12] and Osma et al [41]. In the Netherlands, only mild to moderate mental problems are treated in general practice. Severe cases are treated in a specialized practice. The importance of motivation of the patient was confirmed in the answers to the open questions about the most important reasons for (not) using EMH in our survey and by the study by Wouters et al [11].

In this study, the MHCs estimated 57% of their patient population to be eligible for EMH. This is quite a high estimate compared with the 33% in the study by Lokman et al [19]. It is interesting that the intention of the MHCs to use EMH influences this estimate, as the mediation analysis revealed. This would mean that an MHC with a low intention to use EMH will be inclined to *disqualify* his patients for EMH. In this respect, the MHC can be considered an intermediary who plays a key role in the success of the implementation of EMH. From theory and experiments, it is known that success largely depends on the engagement of intermediaries or champions in the early stages of implementation [50,65,66]. Only an enthusiastic MHC will offer the intervention (EMH) to the patient and will be able to motivate them [12]. Owing to the emphasis that MHCs place on the motivation of the patient, the characteristics of vulnerable patients presumably received a lower ranking. Nevertheless, it became clear that MHCs will not yet use EMH for patients who lack computer and language skills.

Study Limitations

Causal relationships were assumed in this study. However, a causal relationship cannot be established in this type of study. Nevertheless, as the research model is based on theoretical models that have already been empirically verified, this study does provide indications for the assumed causal relationships among EMH use, intention to use EMH, and their determinants.

A second limitation concerns the questionnaire, which, although based on validated questionnaires, was developed specifically for this study, which makes it difficult to compare results with the literature. On the positive side, the reliability of most of the scales proved to be good, and only small adjustments were needed to the predefined scales.

Third, the results of this study might be positively biased as convenience sampling was used and participants were mostly clients of Therapieland or Ksyos and Minddistrict. Obviously, MHCs who were positive about EMH and were already working on the web were more likely to complete the questionnaire than those who were less enthusiastic about web-based activities (*self-selection bias*). In our study, >98% (130/132, 98.5%) of the MHCs had access to a paid EMH platform, whereas, in the study by Lokman et al [19], approximately half of the GPs did not use purchased EMH. Unfortunately, the lack of insight into the response rate impeded a good estimate of the selection bias that might have occurred. A related issue is that the percentage of women in our sample was high (111/132, 84.1%). However, the occupation of MHC is female-dominated in the Netherlands.

A recent (2021) Dutch publication revealed that 71% of MHCs are female [67].

Fourth, there may also have been response bias as self-report questionnaires with Likert scales were used that respondents filled in according to their own interpretation. A statement that is formulated in a fairly general way is perhaps all too easily answered with *agree*. This applies, for example, to the 2 items in the *Accessibility* scale. Feijt et al [56] came to the same conclusion for these items and removed them from the eMental Health Adoption Readiness scale for this reason. Although *Accessibility* is substantively related to *User-friendliness*, it emerged as a separate factor in the factor analysis. In spite of the fact that the reliability was good, there is some doubt regarding its content validity.

The final limitation concerns the difficult interpretation of the intention to use EMH, actual EMH use, and patient eligibility for EMH because of the broad definition of EMH we used (ie, ranging from means of communication to web-based treatment modules). Although the questions were formulated as specifically as possible by adding an example or by speaking of *EMH program* for therapy using web-based programs, the broad concept was used in the measurements of behavioral intention, EMH use, and patient eligibility. The research by van der Vaart et al [15] explicitly focuses on the use of a specific part of EMH (ie, the guided web-based self-management interventions in primary care). Therefore, the conclusions of that study are not entirely comparable with those described in this study.

Recommendations for EMH Practice and Researchers

This study revealed that usefulness and benefits are by far the most important factors influencing the intention to use EMH. Only if potential users see added value in the use of an IT system will they consider using it, and only then will user-friendliness also become important [64]. Therefore, the EMH developer must ensure that EMH has added value for MHCs and patients. This is possible, for example, by developing EMH applications in cocreation with MHCs. Additional qualitative research would be ideally suited to obtain more specific input from MHCs (eg, to understand for which patients MHCs did not find a suitable EMH module, what modules are needed, and what characteristics they need to have). The research by Titzler et al [12] among psychotherapists is a good example of structured interviews that provide useful insights for EMH developers. It would also be good to actively involve the patients as that would help increase patient motivation while at the same time allowing for the determination of whether patient motivation is a real problem and not something that is wrongly estimated or negatively influenced by the MHC or whether the modules themselves are the problem. As the MHCs' personal motivation to support patients with EMH influences their estimation of patients' eligibility, it is important to make MHCs aware of this process and provide skill training on motivating patients

regarding disease self-management and to use EMH. Motivational interviewing skills are posed as a helpful strategy.

This study also showed the importance of task perception of the MHC with regard to EMH. Government policy is necessary to structurally embed EMH in the task perception of the MHC (eg, by promoting digitalization in health care, providing the right reimbursement for EMH activities in general practice, and making EMH part of higher education curricula [68]).

The study results are also useful for the broader population of mental health care professionals, especially the findings with regard to eligible patients and the way in which the intention of health care professionals to use EMH affects their estimation of patient eligibility.

For future use of the questionnaire, specifically in questions on EMH use and patient eligibility, we recommend differentiating between the use of web-based communication and the use of web-based treatment modules. Better ways to determine the factors related to patient eligibility would also be worth investigating.

Recommendations for future research include a longitudinal study to verify the relationships we found in this study and investigate whether EMH use has maintained its upward trend after the COVID-19 pandemic. It would be especially interesting to further explore the influence of behavioral intention on perceived eligibility also among other professionals working in mental health care as it is important for practice to make professionals aware of the fact that personal motivation to support patients with EMH also influences the way in which patients' motivation and skills to use EMH are estimated.

Conclusions

It can be concluded that the intention to use EMH among MHCs was very positive. The most important factors explaining the intention to use were the perceived usefulness and benefits of the use of EMH followed by task perception and accessibility (ie, ease of getting started with it).

The relationship between behavioral intention to use EMH and actual EMH use was partially mediated by the perceived eligibility of the patient population. However, both the behavioral intention to use EMH and the estimated eligibility of the patient population had a significant and direct association with EMH use. The patients' eligibility was most important, which means that an MHC will use little EMH if they consider the patient unsuitable for EMH even if the MHC is positive about the use of EMH.

To determine whether a patient is eligible for EMH, the patients' access to a computer and the internet, digital skills, and Dutch language skills were primarily considered. In addition, patient motivation was found to be of utmost importance.

The study revealed that there will only be a future for blended care if the MHC is convinced of the added value of EMH and can transfer their enthusiasm to the patient.

Acknowledgments

Therapieland, Minddistrict, and the respondents of the questionnaire are acknowledged for their participation in this study. AEMDV thanks Milou Feijt for interesting discussions and the use of the eMental Health Adoption Readiness questionnaire.

Authors' Contributions

The research was initiated by CAWB, VT, and MGS and supervised by CAWB and VT. AEMDV created the web-based survey, gathered the data (with the help of Therapieland and Minddistrict), and performed the statistical analyses. The first draft of the manuscript was written by CAWB (Introduction) and AEMDV (all the other sections). Revisions were mainly made by CAWB and AEMDV. All authors read, modified, and approved the final manuscript.

Conflicts of Interest

MGS is working at Therapieland, in a company that develops and sells e-mental health applications, but he has no personal financial interest related to the subject of this study. The authors have no further conflicts of interest to declare.

Multimedia Appendix 1

Variables of the research model: mean, SD, and Pearson correlation coefficient.

[PDF File (Adobe PDF File), 171 KB - [formative_v6i12e34754_app1.pdf](#)]

References

1. Dutch Healthcare Authority. Factsheet waitinglist mental health care 2021 [wachttijden GGZ 2021]. Government of the Netherlands. 2021 Jul 01. URL: <https://www.rijksoverheid.nl/documenten/publicaties/2021/07/01/factsheet-wachttijden-ggz-2021> [accessed 2021-09-13]
2. van Straten A. E-mental health to bridge the waiting time in secondary mental health care. *Directieve Therapie* 2013;33(3):251-252. [doi: [10.1007/978-90-313-9239-1_7](https://doi.org/10.1007/978-90-313-9239-1_7)]
3. Akwa GGZ (Akwa Mental Health Care). eHealth. Mental Health Care standards [GGZ Standaarden]. 2017. URL: <https://www.ggzstandaarden.nl/generieke-modules/ehealth/introductie> [accessed 2020-03-29]
4. Ebert DD, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. A randomized controlled trial. *J Affect Disord* 2015 May 01;176:9-17. [doi: [10.1016/j.jad.2015.01.056](https://doi.org/10.1016/j.jad.2015.01.056)] [Medline: [25682378](https://pubmed.ncbi.nlm.nih.gov/25682378/)]
5. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. *JAMA Psychiatry* 2017 Apr 01;74(4):351-359 [FREE Full text] [doi: [10.1001/jamapsychiatry.2017.0044](https://doi.org/10.1001/jamapsychiatry.2017.0044)] [Medline: [28241179](https://pubmed.ncbi.nlm.nih.gov/28241179/)]
6. Königbauer J, Letsch J, Doebler P, Ebert D, Baumeister H. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. *J Affect Disord* 2017 Dec 01;223:28-40. [doi: [10.1016/j.jad.2017.07.021](https://doi.org/10.1016/j.jad.2017.07.021)] [Medline: [28715726](https://pubmed.ncbi.nlm.nih.gov/28715726/)]
7. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. *World Psychiatry* 2014 Oct;13(3):288-295 [FREE Full text] [doi: [10.1002/wps.20151](https://doi.org/10.1002/wps.20151)] [Medline: [25273302](https://pubmed.ncbi.nlm.nih.gov/25273302/)]
8. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, et al. Effectiveness of web- and mobile-based treatment of subthreshold depression with adherence-focused guidance: a single-blind randomized controlled trial. *Behav Ther* 2018 Jan;49(1):71-83. [doi: [10.1016/j.beth.2017.05.004](https://doi.org/10.1016/j.beth.2017.05.004)] [Medline: [29405923](https://pubmed.ncbi.nlm.nih.gov/29405923/)]
9. van Tuyll LH, Batenburg R, Keuper JJ, Meurs M, Friele R. Increased use of e-health in general practice during the corona pandemic. Organization of remote care in corona time [Toename gebruik e-health in de huisartsenpraktijk tijdens de coronapandemie. Organisatie van zorg op afstand in coronatijd]. Netherlands Institute for Health Services Research. 2020. URL: <https://www.nivel.nl/nl/publicatie/toename-gebruik-e-health-de-huisartsenpraktijk-tijdens-de-coronapandemie-organisatie-van> [accessed 2020-08-26]
10. Krijgsman J, Swinkels I, van Jong LB, de Out J, Out K, Friele R, et al. More than technology: eHealth monitor 2016. NIVEL. 2016 Oct 6. URL: <https://nivel.nl/nl/publicatie/meer-dan-techniek-ehealth-monitor-2016> [accessed 2019-08-15]
11. Wouters M, Swinkels I, Sinnige J, de Jong JJ, Brabers A, van Lettow L, et al. Kies bewust voor eHealth. NIVEL. 2017 Nov. URL: https://www.nivel.nl/sites/default/files/bestanden/Nictiz_eHealth_monitor_2017.pdf [accessed 2019-08-15]
12. Titzler I, Saruhanjan K, Berking M, Riper H, Ebert DD. Barriers and facilitators for the implementation of blended psychotherapy for depression: a qualitative pilot study of therapists' perspective. *Internet Interv* 2018 Jun;12:150-164 [FREE Full text] [doi: [10.1016/j.invent.2018.01.002](https://doi.org/10.1016/j.invent.2018.01.002)] [Medline: [30135779](https://pubmed.ncbi.nlm.nih.gov/30135779/)]
13. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. *JMIR Mhealth Uhealth* 2020 Feb 20;8(2):e15935. [doi: [10.2196/15935](https://doi.org/10.2196/15935)]
14. Remote care closer? Digital care after the corona crisis. Raad Volksgezondheid & Samenleving. 2020 Aug 27. URL: <https://www.raadvv.nl/documenten/publicaties/2020/08/27/zorg-op-afstand-dichterbij> [accessed 2020-09-05]

15. van der Vaart R, Atema V, Evers AW. Guided online self-management interventions in primary care: a survey on use, facilitators, and barriers. *BMC Fam Pract* 2016 Mar 09;17:27 [FREE Full text] [doi: [10.1186/s12875-016-0424-0](https://doi.org/10.1186/s12875-016-0424-0)] [Medline: [26961547](https://pubmed.ncbi.nlm.nih.gov/26961547/)]
16. Verhaak P, van Beljouw I, Have JT. De opmars van de POH-GGZ. *Tijdschrift voor Praktijkondersteuning* 2010 Oct;5(5):130-134. [doi: [10.1007/bf03089425](https://doi.org/10.1007/bf03089425)]
17. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreur K, E-COMPARED consortium. Attitudes towards digital treatment for depression: a European stakeholder survey. *Internet Interv* 2017 Jun;8:1-9 [FREE Full text] [doi: [10.1016/j.invent.2017.01.001](https://doi.org/10.1016/j.invent.2017.01.001)] [Medline: [30135823](https://pubmed.ncbi.nlm.nih.gov/30135823/)]
18. Conijn B, van der Poel K, Abello K, Boon B. Infographic e-mental health by GP and POH-GGZ. Trimbos. 2015. URL: <https://www.trimbos.nl/aanbod/webwinkel/product/af1364-infographic-e-mental-health-door-huisarts-en-poh-ggz> [accessed 2019-12-01]
19. Lokman S, van der Poel A, Smeets O, Martin K. E-mental health in de huisartsenpraktijk. Trimbos. 2016. URL: <https://www.trimbos.nl/aanbod/webwinkel/product/inf010-e-mental-health-in-de-huisartsenpraktijk> [accessed 2019-08-15]
20. Kip H, Sieverink F, van Gemert-Pijnen LJ, Bouman YH, Kelders SM. Integrating people, context, and technology in the implementation of a web-based intervention in forensic mental health care: mixed-methods study. *J Med Internet Res* 2020 May 26;22(5):e16906 [FREE Full text] [doi: [10.2196/16906](https://doi.org/10.2196/16906)] [Medline: [32348285](https://pubmed.ncbi.nlm.nih.gov/32348285/)]
21. Wind L, Milo M. E-mental-health bij de huisarts. In: *De dokter en de patiënt met psychische problemen*. Houten: Bohn Stafleu van Loghum; Oct 2018.
22. Feijt MA, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived drivers and barriers to the adoption of eMental health by psychologists: the construction of the levels of adoption of eMental health model. *J Med Internet Res* 2018 Apr 24;20(4):e153 [FREE Full text] [doi: [10.2196/jmir.9485](https://doi.org/10.2196/jmir.9485)] [Medline: [29691215](https://pubmed.ncbi.nlm.nih.gov/29691215/)]
23. Aref-Adib G, McCloud T, Ross J, O'Hanlon P, Appleton V, Rowe S, et al. Factors affecting implementation of digital health interventions for people with psychosis or bipolar disorder, and their family and friends: a systematic review. *Lancet Psychiatry* 2019 Mar;6(3):257-266. [doi: [10.1016/s2215-0366\(18\)30302-x](https://doi.org/10.1016/s2215-0366(18)30302-x)]
24. Broeders, DW, Das HD, Jennissen RP, Tiemeijer WL, De Visser M. Van verschil naar potentieel. Een realistisch perspectief op de sociaaleconomische gezondheidsverschillen. The Netherlands Scientific Council for Government Policy (WRR). 2018. URL: <https://www.wrr.nl/publicaties/policy-briefs/2018/08/27/van-verschil-naar-potentieel-een-realistisch-perspectief-op-de-sociaaleconomische-gezondheidsverschillen> [accessed 2019-07-21]
25. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. *J Med Internet Res* 2017 Apr 27;19(4):e136 [FREE Full text] [doi: [10.2196/jmir.6731](https://doi.org/10.2196/jmir.6731)] [Medline: [28450271](https://pubmed.ncbi.nlm.nih.gov/28450271/)]
26. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. *Int J Environ Res Public Health* 2019 Feb 21;16(4):645 [FREE Full text] [doi: [10.3390/ijerph16040645](https://doi.org/10.3390/ijerph16040645)] [Medline: [30795623](https://pubmed.ncbi.nlm.nih.gov/30795623/)]
27. van Houwelingen CT, Ettema RG, Antonietti MG, Kort HS. Understanding older people's readiness for receiving telehealth: mixed-method study. *J Med Internet Res* 2018 Apr 06;20(4):e123 [FREE Full text] [doi: [10.2196/jmir.8407](https://doi.org/10.2196/jmir.8407)] [Medline: [29625950](https://pubmed.ncbi.nlm.nih.gov/29625950/)]
28. Peek ST, Wouters EJ, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJ. Factors influencing acceptance of technology for aging in place: a systematic review. *Int J Med Inform* 2014 Apr;83(4):235-248 [FREE Full text] [doi: [10.1016/j.ijmedinf.2014.01.004](https://doi.org/10.1016/j.ijmedinf.2014.01.004)] [Medline: [24529817](https://pubmed.ncbi.nlm.nih.gov/24529817/)]
29. van Deursen A. Digitale ongelijkheid in Nederland anno 2018. 2018. URL: <https://research.utwente.nl/en/publications/digitale-ongelijkheid-in-nederland-anno-2018> [accessed 2019-07-21]
30. de Veer AJ, Peeters JM, Brabers AE, Schellevis FG, Rademakers JJ, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. *BMC Health Serv Res* 2015 Mar 15;15:103 [FREE Full text] [doi: [10.1186/s12913-015-0765-8](https://doi.org/10.1186/s12913-015-0765-8)] [Medline: [25889884](https://pubmed.ncbi.nlm.nih.gov/25889884/)]
31. Bernard R, Sabariego C, Cieza A. Barriers and facilitation measures related to people with mental disorders when using the web: a systematic review. *J Med Internet Res* 2016 Jun 09;18(6):e157 [FREE Full text] [doi: [10.2196/jmir.5442](https://doi.org/10.2196/jmir.5442)] [Medline: [27282115](https://pubmed.ncbi.nlm.nih.gov/27282115/)]
32. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. *J Med Internet Res* 2016 Nov 22;18(11):e309 [FREE Full text] [doi: [10.2196/jmir.6511](https://doi.org/10.2196/jmir.6511)] [Medline: [27876684](https://pubmed.ncbi.nlm.nih.gov/27876684/)]
33. Kim H, Xie B. Health literacy in the eHealth era: a systematic review of the literature. *Patient Educ Couns* 2017 Jun;100(6):1073-1082. [doi: [10.1016/j.pec.2017.01.015](https://doi.org/10.1016/j.pec.2017.01.015)] [Medline: [28174067](https://pubmed.ncbi.nlm.nih.gov/28174067/)]
34. Bolman C. eHealth bij kwetsbare groepen een uitdaging. Open Universiteit. 2019 Jan 25. URL: <https://www.ou.nl/documents/40554/724769/Oratieboekje+Catherine+Bolman+DEF+15012019.pdf/48046c78-622a-400f-213d-46995a221b46> [accessed 2019-02-07]
35. Swinkels I. E-health verhoogt risico op ongelijkheid. *Huisarts en Wetenschap* 2017 Sep 12;60(9):477. [doi: [10.1007/s12445-017-0290-6](https://doi.org/10.1007/s12445-017-0290-6)]
36. van der Els-Schoenmaker D, Hopmans E, Jansen S, Korte A. E-health kan toegankelijkheid zorg belemmeren. *Medisch Contact*. 2020 Jan 15. URL: <https://tinyurl.com/yvnnma84> [accessed 2020-10-07]

37. Plantinga S, Kaal M. Hoe mediawijs is Nederland? Mediawijzer. Kantar Public. 2018 Sep 17. URL: <https://www.mediawijzer.net/wp-content/uploads/sites/6/2018/09/Rapport-Mediawijsheid-volwassenen-2018.pdf> [accessed 2019-07-21]
38. Vrieling J. Pro Deo: Ongevraagd juridisch advies voor vrijheidsstrijders en andere fanatici. Netherlands: Universitaire Pers Leuven; 2018.
39. Beuning J. EHealth: possibilities, use and views [mogelijkheden, gebruik en opvattingen]. CBS. 2019 Apr 19. URL: <https://www.cbs.nl/nl-nl/achtergrond/2019/16/ehealth-mogelijkheden-gebruik-en-opvattingen> [accessed 2019-07-16]
40. Wentzel J, van der Vaart R, Bohlmeijer ET, 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](https://doi.org/10.2196/mental.4534)] [Medline: [26860537](https://pubmed.ncbi.nlm.nih.gov/26860537/)]
41. Osma J, Sprenger M, Mettler T. Introduction of e-mental health in national health systems – A health professionals' perspective. *Health Policy Technol* 2017 Dec;6(4):436-445. [doi: [10.1016/j.hlpt.2017.07.001](https://doi.org/10.1016/j.hlpt.2017.07.001)]
42. Fishbein M, Ajzen I. Predicting and Changing Behavior: The Reasoned Action Approach. New York: Psychology Press; 2010.
43. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q* 2003;27(3):425-478. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
44. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q* 2012;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
45. Fleuren MA, Paulussen TG, Van Dommelen P, Van Buuren S. Towards a measurement instrument for determinants of innovations. *Int J Qual Health Care* 2014 Oct 20;26(5):501-510 [FREE Full text] [doi: [10.1093/intqhc/mzu060](https://doi.org/10.1093/intqhc/mzu060)] [Medline: [24951511](https://pubmed.ncbi.nlm.nih.gov/24951511/)]
46. Rogers E. Diffusion of Innovations, 5th Edition. New York, NY: The Free Press; 2003.
47. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. *Int J Med Inform* 2017 May;101:75-84. [doi: [10.1016/j.ijmedinf.2017.02.002](https://doi.org/10.1016/j.ijmedinf.2017.02.002)] [Medline: [28347450](https://pubmed.ncbi.nlm.nih.gov/28347450/)]
48. IJzerman RV, van der Vaart R, Evers AW. Internet-based cognitive behavioral therapy among psychologists in a medical setting: a survey on implementation. *J Med Internet Res* 2019 Aug 09;21(8):e13432 [FREE Full text] [doi: [10.2196/13432](https://doi.org/10.2196/13432)] [Medline: [31400101](https://pubmed.ncbi.nlm.nih.gov/31400101/)]
49. Chao C. Factors determining the behavioral intention to use mobile learning: an application and extension of the UTAUT model. *Front Psychol* 2019 Jul 16;10:1652 [FREE Full text] [doi: [10.3389/fpsyg.2019.01652](https://doi.org/10.3389/fpsyg.2019.01652)] [Medline: [31379679](https://pubmed.ncbi.nlm.nih.gov/31379679/)]
50. Peels D, Mudde A, Bolman C, Golsteijn R, de Vries H, Lechner L. Correlates of the intention to implement a tailored physical activity intervention: perceptions of intermediaries. *Int J Environ Res Public Health* 2014 Feb 10;11(2):1885-1903 [FREE Full text] [doi: [10.3390/ijerph110201885](https://doi.org/10.3390/ijerph110201885)] [Medline: [24518647](https://pubmed.ncbi.nlm.nih.gov/24518647/)]
51. Vis C, Mol M, Kleiboer A, Bührmann L, Finch T, Smit J, et al. Improving implementation of eMental health for mood disorders in routine practice: systematic review of barriers and facilitating factors. *JMIR Ment Health* 2018 Mar 16;5(1):e20 [FREE Full text] [doi: [10.2196/mental.9769](https://doi.org/10.2196/mental.9769)] [Medline: [29549072](https://pubmed.ncbi.nlm.nih.gov/29549072/)]
52. Thompson RL, Higgins CA, Howell JM. Personal computing: toward a conceptual model of utilization. *MIS Q* 1991 Mar;15(1):125. [doi: [10.2307/249443](https://doi.org/10.2307/249443)]
53. Kivi M, Eriksson MC, Hange D, Petersson E, Björkelund C, Johansson B. Experiences and attitudes of primary care therapists in the implementation and use of internet-based treatment in Swedish primary care settings. *Internet Interv* 2015 Sep;2(3):248-256. [doi: [10.1016/j.invent.2015.06.001](https://doi.org/10.1016/j.invent.2015.06.001)]
54. Becker EM, Jensen-Doss A. Computer-assisted therapies: examination of therapist-level barriers to their use. *Behav Ther* 2013 Dec;44(4):614-624. [doi: [10.1016/j.beth.2013.05.002](https://doi.org/10.1016/j.beth.2013.05.002)] [Medline: [24094787](https://pubmed.ncbi.nlm.nih.gov/24094787/)]
55. Chismar W, Wiley-Patton S. Predicting internet use: applying the extended technology acceptance model to the healthcare environment. In: *E-Health Systems Diffusion and Use: The Innovation, the User and the Use IT Model*. Pennsylvania: Idea Group Publishing; 2006.
56. Feijt MA, de Kort YA, Westerink JH, Bierbooms JJ, Bongers IM, IJsselstein WA. Assessing professionals' adoption readiness for eMental health: development and validation of the eMental health adoption readiness scale. *J Med Internet Res* 2021 Sep 17;23(9):e28518 [FREE Full text] [doi: [10.2196/28518](https://doi.org/10.2196/28518)] [Medline: [34533469](https://pubmed.ncbi.nlm.nih.gov/34533469/)]
57. Berndt N, Bolman C, de Vries H, Segaar D, van Boven I, Lechner L. Smoking cessation treatment practices: recommendations for improved adoption on cardiology wards. *J Cardiovasc Nurs* 2013;28(1):35-47. [doi: [10.1097/JCN.0b013e318231f424](https://doi.org/10.1097/JCN.0b013e318231f424)] [Medline: [22048618](https://pubmed.ncbi.nlm.nih.gov/22048618/)]
58. Zijlstra DN, Bolman CA, Muris JW, de Vries H. How to persuade more primary care professionals to adopt a valued smoking cessation referral aid: a cross-sectional study of facilitators and barriers. *BMC Prim Care* 2022 Sep 07;23(1):227 [FREE Full text] [doi: [10.1186/s12875-022-01843-3](https://doi.org/10.1186/s12875-022-01843-3)] [Medline: [36071372](https://pubmed.ncbi.nlm.nih.gov/36071372/)]
59. De Veirman A. Determinants of the use of eMental Health by the mental health counsellors working in general practitioner practices. Open Universiteit. 2020 Dec 06. URL: <https://research.ou.nl/> [accessed 2022-11-30]
60. Tabachnick B, Fidell L. Using Multivariate Statistics. Boston, United States: Allyn and Bacon; 2001.
61. Field A. Discovering Statistics Using SPSS 4th edition. Thousand Oaks, California: Sage Publications Limited; 2020.
62. Braun P, Drüge M, Hennemann S, Nitsch FJ, Staeck R, Apolinário-Hagen J. Acceptance of e-mental health services for different application purposes among psychotherapists in clinical training in Germany and Switzerland: secondary analysis

- of a cross-sectional survey. *Front Digit Health* 2022 Feb 28;4:840869 [FREE Full text] [doi: [10.3389/fdgh.2022.840869](https://doi.org/10.3389/fdgh.2022.840869)] [Medline: [35295621](https://pubmed.ncbi.nlm.nih.gov/35295621/)]
63. Sander J, Bolinski F, Diekmann S, Gaebel W, Günther K, Hauth I, et al. Online therapy: an added value for inpatient routine care? Perspectives from mental health care professionals. *Eur Arch Psychiatry Clin Neurosci* 2022 Feb 16;272(1):107-118 [FREE Full text] [doi: [10.1007/s00406-021-01251-1](https://doi.org/10.1007/s00406-021-01251-1)] [Medline: [33725165](https://pubmed.ncbi.nlm.nih.gov/33725165/)]
64. Keil M, Beranek PM, Konsynski BR. Usefulness and ease of use: field study evidence regarding task considerations. *Decision Support Syst* 1995 Jan;13(1):75-91. [doi: [10.1016/0167-9236\(94\)e0032-m](https://doi.org/10.1016/0167-9236(94)e0032-m)]
65. 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 Oct 26;11(1):146 [FREE Full text] [doi: [10.1186/s13012-016-0510-7](https://doi.org/10.1186/s13012-016-0510-7)] [Medline: [27782832](https://pubmed.ncbi.nlm.nih.gov/27782832/)]
66. Greenhalgh T, Wherton J, Papoutsis C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017 Nov 01;19(11):e367 [FREE Full text] [doi: [10.2196/jmir.8775](https://doi.org/10.2196/jmir.8775)] [Medline: [29092808](https://pubmed.ncbi.nlm.nih.gov/29092808/)]
67. Nuijen J, Kenter A, Ringoir L. Organisatie en uitvoering van de functie POH-GGZ in de praktijk. Trimbos Instituut. 2021. URL: <https://www.trimbos.nl/wp-content/uploads/2021/12/AF1936-Organisatie-en-uitvoering-van-de-functie-POH-GGZ-in-de-praktijk.pdf> [accessed 2022-11-16]
68. Gaebel W, Lukies R, Kerst A, Stricker J, Zielasek J, Diekmann S, et al. Upscaling e-mental health in Europe: a six-country qualitative analysis and policy recommendations from the eMEN project. *Eur Arch Psychiatry Clin Neurosci* 2021 Sep 11;271(6):1005-1016. [doi: [10.1007/s00406-020-01133-y](https://doi.org/10.1007/s00406-020-01133-y)] [Medline: [32393997](https://pubmed.ncbi.nlm.nih.gov/32393997/)]

Abbreviations

- EMH:** e-mental health
GP: general practitioner
MHC: mental health counselor
MIDI: Measurement Instrument for Determinants of Innovation
RAA: reasoned action approach
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Mavragani; submitted 06.11.21; peer-reviewed by M Feijt, W Zhang, C El Morr; comments to author 16.02.22; revised version received 16.06.22; accepted 05.10.22; published 20.12.22.

Please cite as:

De Veirman AEM, Thewissen V, Spruijt MG, Bolman CAW

Factors Associated With Intention and Use of e-Mental Health by Mental Health Counselors in General Practices: Web-Based Survey
JMIR Form Res 2022;6(12):e34754

URL: <https://formative.jmir.org/2022/12/e34754>

doi: [10.2196/34754](https://doi.org/10.2196/34754)

PMID: [36538357](https://pubmed.ncbi.nlm.nih.gov/36538357/)

©Ann E M De Veirman, Viviane Thewissen, Matthijs G Spruijt, Catherine A W Bolman. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Analysis of Patient Cues in Asynchronous Health Interactions: Pilot Study Combining Empathy Appraisal and Systemic Functional Linguistics

Elena Rey Velasco^{1,2}, BSc, MSc; Hanne Sæderup Pedersen³, BA, MA, PhD; Timothy Skinner¹, BSc, MD, PhD; Impact Diabetes B2B Collaboration Group⁴

¹Department of Psychology, University of Copenhagen, Copenhagen, Denmark

²Liva Healthcare, Copenhagen, Denmark

³Department of Nordic Studies and Linguistics, University of Copenhagen, Denmark

⁴See Acknowledgements, Dublin, Ireland

Corresponding Author:

Elena Rey Velasco, BSc, MSc

Department of Psychology

University of Copenhagen

Øster Farimagsgade 2A

Copenhagen, 1353

Denmark

Phone: 45 35324800

Email: erv@psy.ku.dk

Abstract

Background: Lifestyle-related diseases are among the leading causes of death and disability. Their rapid increase worldwide has called for low-cost, scalable solutions to promote health behavior changes. Digital health coaching has proved to be effective in delivering affordable, scalable programs to support lifestyle change. This approach increasingly relies on asynchronous text-based interventions to motivate and support behavior change. Although we know that empathy is a core element for a successful coach-user relationship and positive patient outcomes, we lack research on how this is realized in text-based interactions. Systemic functional linguistics (SFL) is a linguistic theory that may support the identification of empathy opportunities (EOs) in text-based interactions, as well as the reasoning behind patients' linguistic choices in their formulation.

Objective: This study aims to determine whether empathy and SFL approaches correspond and complement each other satisfactorily to study text-based communication in a health coaching context. We sought to explore whether combining empathic assessment with SFL categories can provide a means to understand client-coach interactions in asynchronous text-based coaching interactions.

Methods: We retrieved 148 text messages sent by 29 women who participated in a randomized trial of telecoaching for the prevention of gestational diabetes mellitus (GDM) and postnatal weight loss. We conducted a pilot study to identify users' explicit and implicit EOs and further investigated these statements using the SFL approach, focusing on the analysis of transitivity and thematic analysis.

Results: We identified 164 EOs present in 42.37% (3478/8209) of the word count in the corpus. These were mainly negative (n=90, 54.88%) and implicit (n=55, 60.00%). We distinguished opening, content and closing messages structures. Most of the wording was found in the content (n=7077, 86.21%) with a declarative structure (n=7084, 86.30%). Processes represented 22.4% (n=1839) of the corpus, with half being material (n=876, 10.67%) and mostly related to food and diet (n=196, 54.92%), physical activity (n=96, 26.89%), and lifestyle goals (n=40, 11.20%).

Conclusions: Our findings show that empathy and SFL approaches are compatible. The results from our transitivity analysis reveal novel insights into the meanings of the users' EOs, such as their seek for help or praise, often missed by health care professionals (HCPs), and on the coach-user relationship. The absence of explicit EOs and direct questions could be attributed to low trust on or information about the coach's abilities. In the future, we will conduct further research to explore additional linguistic features and code coach messages.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR) ACTRN12620001240932; <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=380020>

(*JMIR Form Res* 2022;6(12):e40058) doi:[10.2196/40058](https://doi.org/10.2196/40058)

KEYWORDS

telehealth; telecoaching; asynchronous communication; empathy; systemic functional linguistics; communication; health promotion; coding; linguistic analysis; user experience; coach-user interaction; tool development; lifestyle-related disease

Introduction

Noncommunicable diseases (NCDs) account for 73.6% of deaths worldwide. These lifestyle-related diseases, such as cardiovascular diseases (CVDs), some cancers, respiratory diseases, and diabetes, are among the most common causes of death and disability [1]. The rapid increase in NCD rates is a global disease burden in both developed and developing nations. However, we now know that these NCDs can be prevented, or substantially delayed, by changes in lifestyle (eg, factors such as diet, physical activity, stress, and sleep), as shown in numerous trials [2-4]. Research has demonstrated that with appropriate individual and group support, individuals can achieve significant weight loss and sustainable changes in lifestyle [5,6]. Nonetheless, the high incidence of NCDs and the limited resources to deliver best-practice behavior change programs make large-scale prevention programs challenging. The need for cost-effective alternatives has led some countries, such as the United States [7] and the United Kingdom [8], to seek new strategies to promote health behavior changes.

To address the challenges of scalability and cost-effectiveness, NCD prevention programs are increasingly using digital technologies. One technology that facilitates access to prevention programs is telehealth, which is the use of video or audio technologies to deliver a health intervention. Telehealth has the potential to reduce health care costs and increase the scope of these programs, as it can substitute or supplement in-person visits when personal attendance is not possible (eg, patients living in rural areas [9,10]). Using telehealth to deliver face-to-face behavior change programs has been shown to be as effective as in-person programs for NCDs [11,12]. For these reasons, the field of telehealth has experienced substantial growth over recent years, and the COVID-19 pandemic has accelerated the process, with programs for mental health, rehabilitation, and medical consultations showing rapid increases in usage [13-15].

In the context of disease prevention programs, this approach is increasingly referred to as telehealth coaching in order to distinguish it from the delivery of more traditional telehealth services. Telehealth coaching uses an integrative health coaching (IHC) approach. IHC connects the coaching intervention with the individual's personal values and sense of purpose [16]. Instead of being instructed on how to reach their goals, the coach provides the user, or person being coached, with the knowledge, skills, and confidence to perform autonomously [17]. These telehealth coaching programs combine multimodalities of digital technology to support people in achieving their lifestyle goals in a synchronous or asynchronous form. Traditional synchronous interactions use real-time, face-to-face meetings, telephone

calls, or video calls [18], and asynchronous interactions consist of the exchange of texts, audio, or video messages that the user can access and review later [19]. At the same time, health coaching allows for only human [14,20], only automated [21,22], or hybrid [23-25] modalities. Although all of these have shown positive results, it is still not clear which one is more effective [26]. A recent meta-analysis showed that automated digital interventions (ADIs) are a good addition to weight loss coaching interventions and results are more effective when the coaching program duration is shorter [27].

Generally, digital health coaching interventions follow a prespecified framework, such as manuals or guidelines, based on the current evidence on behavior change [28] and psychosocial theories [29,30]. Evaluating whether coaches are delivering a program as intended is key to ensuring a homogenous and effective intervention, and telehealth coaching poses unique challenges in this regard due to its multiple modalities [31]. Although there is increasing research exploring the fidelity of such programs, current research work has focused on synchronous, face-to-face interventions delivered by coaches [32-34]. This research typically quantifies the behavior change techniques (BCTs) delivered by the coach and to some extent the way in which these interventions are realized. State-of-the-art findings show a predominant focus on the coaches' performance and users' outcomes without accounting for the users' cues and responses [35], in addition to inconsistencies in fidelity reporting [36]. With increasing drivers for efficiency and the use of responsive artificial intelligence (AI) systems, the use of asynchronous interactions to support health coaching is growing. However, there is little research on these asynchronous interactions and a clear need to understand their nature and how to optimize them. Although asynchronous interventions are delivered through audio or video messages on a digital platform, the most common form of interaction is through the exchange of text messages. A coaching platform can be automatized to send scheduled messages (eg, reminders). There is a body of research on the use of automated messages to remind, prompt, or nudge healthier behavior, which demonstrates their potential [37], effectiveness [38], and language used [39,40] in text-based behavior change interventions. However, these messages represent a 1-way communication from the coaching platform to the individual. The users participating in these telehealth coaching programs also communicate directly to their coach or AI coaching platform. Their text messages can be responded to by an AI-based system (eg, chatbots) or by their coach (ie, individually crafted correspondence).

Nonetheless, there is a wealth of literature on the effectiveness of traditional synchronous, face-to-face patient-provider interactions where researchers share an overall concern for the

quality of asynchronous consultations [41], as well as for the quality of the relationship developed [42]. In this context, the concepts of empathy, sympathy, and compassion in health care are crucial in the patient-provider relationship but sometimes confused with one another [43]. Compassion is a deep awareness of another person's suffering, along with the wish to relieve it [44], whereas empathy is the cognitive skill to understand and "feel with" the patient. Some authors have identified the dimensions of cognitive, affective (relegated to sympathy), and emotional within the definition of empathy [45]. Other authors, such as Piasecki, see clinical empathy as "the ability to understand and participate in another person's feelings and emotional state, while sympathy describes the listener's feelings without understanding or sharing the patient's emotions" [46]. The positive effect of clinical empathy on patient outcomes has been documented across psychological, sociological, therapeutic, and behavioral disciplines [47,48] and should be preserved in text-based, asynchronous interventions. An empathetic response is important for building a therapeutic alliance in psychotherapy, and effective relational skills are essential in behavior change programs for promoting health outcomes. Thereby, an empathetic frame is a good start point when coding asynchronous messages. There are a number of tools in patient-provider communication for identifying opportunities for empathetic responses [49], showing how providers often miss these opportunities [50], and providing advice to prevent it [51]. According to a review by Epstein et al [52], patient-centered communication (PCC) comprises "(1) eliciting and understanding the patient's perspective—concerns, ideas, expectations, needs, feelings and functioning, (2) Understanding the patient within his or her unique psychosocial context, (3) Reaching a shared understanding of the problem and its treatment with the patient that is concordant with the patient's values, and (4) Helping patients to share power and responsibility by involving them in choices to the degree that they wish." Epstein's arguments are present when expressing empathy in a health care context. However, this approach has not been informed by our understandings of language and, in particular, the functions of language.

Pounds [53] presented an empathy appraisal approach, supported by previous discourse analysis studies based on systemic functional linguistics (SFL), to explore the expressions of empathy in PCC. Nonetheless, it is surprising that her approach does not incorporate SFL into this patient-provider communication analysis. According to SFL theory, developed by Halliday and Matthiessen [54], language in itself has a communicative and a meaning potential that is realized through language production, and that language in itself is social semiotics, an approach to communication that aims to comprehend how individuals in particular social contexts interact through a variety of means. The goal of studying communication from this angle is to classify the semiotic decisions that communicators are able to make [55]. The empathy opportunities (EOs) that Pounds considered in her empathy appraisal may be further informed by these choices. Suchman [56] defined implicit EOs as "patient statements from which a clinician might infer an underlying emotion that has not been explicitly expressed" and explicit EOs as "statements about situations or concerns that might plausibly be associated with

an emotion." To fully grasp the meaning potential of the patients' linguistic choices whenever they express an EO, however, we must first understand the 3 SFL metafunctions that construe that meaning: the ideational metafunction, which describes the speaker's inner and outer experience; the interpersonal metafunction, which concerns the relationship between the speaker and the recipient as well as between the speaker and their message; and the textual metafunction, which is used to interpret the text as a text and not just as a cluster of words or clauses [57]. The transitivity system is a component of the ideational metafunction and goes further than the distinction between transitive and intransitive verbs. A transitivity analysis explores how the speaker construes their experience of the world. The processes, participants implicated, and circumstances of this experience are all part of the transitivity system. Processes are realized by the verbal group of the clause and can be classified as material, mental, relational, verbal, existential, or behavioral [58]. We provide a further description of the process categories in the Methods section. Several researchers have chosen this approach for a quantitative analysis of written discourse in literature, news, and social media texts [59-61]. Matthiessen's [62] work adds valuable insights into the use of SFL in health care contexts and PCC. Pounds and De Pablos-Ortega [63] foresee the combination of the empathy appraisal approach and SFL categories to better understand patients' (or users') perspectives and to improve doctors' (or experts') communicative strategies in online counseling. Additionally, the experiential metafunction, which is embedded in the ideational metafunction, describes how the speaker uses language to communicate their perception of themselves and the world. For example, Fosgerau et al [64] examined the choices of patients with depression in the transitivity system. Furthermore, this system is the most basic SFL grouping used to quantify the experiential meaning expressed in a text message-based interaction systematically. Thus, a combination of the empathetic qualities' identification in a message and its functional grammar analysis may provide a start point for coding asynchronous messages. Results from this approach would subsequently lead to the identification of an appropriate coaching response.

Thereby, in this paper, we seek to explore whether Pounds' empathy appraisal and SFL approaches have utility in coding asynchronous text messages in a health context. To that end, we conduct a pilot study to analyze a data set of messages posted by users of a telehealth coaching program. We then discuss how the findings may be used to inform optimal coaching responses to those messages.

Methods

Study Design

We coded a sample of 148 messages sent by 29 women from March 7 to June 21, 2021, on a telehealth coaching platform. The study population was an Irish cluster that belongs to an ongoing randomized trial on a telehealth coaching intervention for the prevention of gestational diabetes mellitus (GDM) and postnatal weight loss in 800 women in Australia, Ireland, the United Kingdom, and Spain (Bump2Baby and Me, protocol

registration no. ACTRN12620001240932) [65]. The analyzed messages were the first 148 messages sent by the first 29 participants enrolled in the study, who had thus been in the intervention for a period between 0 and 15 weeks. Participants (users) were onboarded after a synchronous initial consultation with their health coach, and then, they received an average of 15 automated messages between enrolment and birth, which included educational material on lifestyle, well-being, and nutrition. Users also received nonautomated messages from their coach, which accounted for an average of 4 weekly, 4 biweekly, and 3 monthly tailored messages before birth. Coach messages included comments on the users' progress and lifestyle goals, as well as providing educational content and counseling. These communications were based on a predefined structure and a framework grounded on the BCT taxonomy [28] and the motivational interviewing approach [66]. We imported these 148 user-sent messages to NVivo 12 Plus (QSR International), a qualitative analysis software program [67], and then coded them according to the empathy appraisal [53] and SFL [54] categories explained later. Author ERV performed 2 rounds of the coding process for all categories and discussed the issues with a second coder (author HSP).

Ethical Considerations

The Bump2Baby and Me trial, where the authors are authorized researchers, is the source of the data set that was examined. Ethical approval was obtained, and all study participants provided written informed consent for the use of their data for research purposes, provided the findings were presented anonymously. Ethical approval was granted for all study sites (Dublin: National Maternity Hospital Ethics Committee ref EC18.2020; Bristol: Wales Research Ethics Committee ref 21/WA/0022; Granada: CEIM/CEI Provincial de Granada; Melbourne: Monash Health Human Research Ethics Committee ref RES-20-0000-892A). The data used in the study belongs

only to the Irish arm of the study (Dublin). More information concerning these ethical considerations can be found in the published study protocol [65].

Empathy Appraisal Categories

We assessed empathy according to the "appraisal" dimensions of empathy in doctor-patient interactions described by Pounds [53], where patients' expression of feelings and views are categorized as the following EOs:

- Explicit expressions of negative feelings, such as an emotive behavior or a mental state ("I cried when I found out").
- Implicit expression of negative feelings through reference to a negative experience, such as fear, confusion, anxiety, or sadness ("It's been 3 days and I haven't heard back from my GP").
- Explicit expression of negative judgment (others or self; "She is such an irresponsible person").
- Implicit expression of negative judgment (others or self; "I could have done better").
- Explicit or implicit expression of positive self-judgment ("I am eating healthier than ever!").
- Explicit expression of negative appreciation (things, events, actions; "The dinner was so boring").
- Implicit expression of negative appreciation (things, events, actions; "I am not sure this is something for me").

Message Structure

We used a message structure to explore how each message was organized and whether it affected participants' expressions of empathy. The main categories were *opening*, *content*, and *closing*, according to previous research on written messaging dynamics [68,69] to illustrate the text structure or "reading path" [70,71]. During the analysis, we created 2 more categories: *full structure* (Textbox 1), for messages using all 3 categories, and *single structure* (Textbox 2), for messages using only 1 of them.

Textbox 1. Full-structure examples.

Example 1

Opening: "Hi (coach name), hope you're well. Quick question for you.

Content: I weigh myself every Monday morning for the study and I've actually lost weight over the last few weeks. Just 0.15kg. Should I be worried as I read from 15 weeks I should be putting on a pound a week!

Closing: Thanks a mill!"

Example 2

Opening: "Hi (coach name) hope your week is going well (emoji)

Content: so far mine is. Nausea has eased big time in the past 10 which is great and I've been having my evening meal. Still need to work on time out for a book etc (emoji) a work in progress. It would be great if you could send me some stretching to over the next few weeks to try keep the body somewhat limber. Find my hips can be a bit creekly or sore in the morning so maybe something to assist?

Closing: Thank you (participant name)"

Textbox 2. Single-structure examples.

Examples

- "I think I would like to re configure my goals regarding exercise. If I could measure my steps that would probably be a good start to keep tabs on myself? What do you think?"
- "Thank you very much for all the information (coach name)"
- "Pilates starts this eve with elbowroom (emoji)" (attached image)

Sentence Structure

We based the sentence structure categories on Halliday's [54] systemic functional grammar according to speech function. In a declarative sentence, the subject comes before the finite (verb). In an interrogative sentence, the finite comes before the subject. Lastly, the subject is implicit in an imperative sentence [54].

Throughout the analysis, we found that some sentences had the speech function of a question realized by a declarative structure. Halliday [54] previously described this phenomenon regarding the relationship between the sentence structure and the 4 speech functions *offer*, *command*, *statement*, and *question*. As a result, we created a fourth category to account for it (Table 1).

Table 1. Sentence structure categories and examples.

Sentence structure category	Example
Declarative	"I signed up for a 4 week yoga class"
Interrogative	"What do you think?"
Imperative	"Please send them to me"
Declarative structure, question function	"I would like to check with you whether you have got any video of pelvic floor exercises"

Processes

We used Halliday's [54] classification to define the process categories. A process is realized by the verb and contributes to the speaker's construal of experience. We coded each clause according to the *material*, *mental*, *relational*, *behavioral*, *verbal*, and *existential* process categories (Table 2): *Material* processes

construe the actions of doing and happening. *Mental* processes account for sensing. *Relational* processes are used to characterize and identify. *Behavioral* processes represent outer manifestations of human inner workings, such as consciousness and physiological processes. *Verbal* processes refer to the language form and use, such as saying and meaning. Lastly, *existential* processes represent existence or happening [54].

Table 2. Process categories and examples.

Process category	Example
Material	<ul style="list-style-type: none"> "We made a pumpkin cake"
Mental	<ul style="list-style-type: none"> "I have just read your book" "She is considering your offer"
Relational	<ul style="list-style-type: none"> "The weather was very nice" "I have a blue coat"
Behavioral	<ul style="list-style-type: none"> "I will have a look"
Verbal	<ul style="list-style-type: none"> "We talked about the meeting"
Existential	<ul style="list-style-type: none"> "There is a shop around the corner"

Transitivity Analysis

Halliday's [72] concept of transitivity supplements the differentiation between transitive and nontransitive verbs. This differentiation depends on the presence or absence of an object that completes the process meaning [72]. Through the choices in the transitivity system, the speaker construes their experiences of the external world and the internal world of their consciousness. This system considers the participants involved, as well as the surrounding circumstances [73]. Thereby, transitivity allowed us to explore the construals of experience in the corpus by identifying the processes and participants [74].

the telehealth coaching program details, users had been on the program for a mean of 80.76 days (SD 30.47) and had sent a mean of 2.27 messages (SD 1.19) at the time of our analysis. In contrast, coaches had sent a mean of 7.62 messages (SD 1.82). The most common goals set by the users were related to physical activity (n=27, 93.1%), diet (n=24, 82.8%), and the number of steps (n=21, 72.4%). Because this was a coaching program for pregnant women, weight was not a frequent goal (n=7, 24.1%) and coaches were encouraged not to promote it. Users could manually add any lifestyle-related goal into the life goals category (n=15, 51.7%) on the platform, such as "meditate in the morning," "go to bed before midnight," or "read a book for 20mins."

Results

Participants' Demographics and Program Details

Table 3 shows a description of the users' demographics and program details. The mean age was 37.59 years (SD 3.69), and the BMI was overall normal (mean 25.82, SD 5.68). Regarding

We present an overview of the coding results in Table 4 as the number of coded references (occurrences), word count, and word count percentages. We identified 164 EOs, accounting for 42.37% (3478/8209) of the corpus. Negative empathic statements were the most present (n=2026, 24.68%), mostly through an implicit approach (n=1442, 17.57%) as an *implicit*

negative appreciation of things, events, or actions (n=987, 12.02%). This implicit approach rate was similar to that of the *explicit or implicit expression of positive self-judgment* category (n=1481, 18.04%). We did not identify any explicit expression of negative judgment about others or self.

Content was the predominant structural component (n=7077, 86.21%). Nearly half of the messages (n=4011, 48.86%) included all 3 structural components (*opening, content, and closing*), while 5.26% (n=432) were identified as a *single message*, including 1 of the components (*opening, content, or closing*). In some cases, a user sent more than 1 message at the same time, resulting in the structural components being divided. We conducted a separate analysis comparing full- and single-structured messages that showed no differences for EOs, sentence structure, and processes.

Sentence structure coding revealed that the preferred sentence structure was *declarative* in this corpus (n=7084, 86.30%). This indicated that participants used these message exchanges to narrate, describe, or state rather than to request for information, guidance, or support. However, 9.43% (774/8209) of the corpus was not coded, because it did not meet the definition of a clause as previously explained (greetings, thanking, laughter, emojis, links in between independent sentences, vocatives), and was labeled as “other.”

Further, *processes* accounted for 22.40% (n=1839) of the corpus and were *material* in almost half of the cases (n=876, 10.67%), followed by *relational* (n=495, 6.03%). In Table 5, we show the *process* occurrences in percentages (%) for each process category identified in the EOs expressed. Overall, all the processes were evenly spread in both positive and negative EO categories. *Material* (n=224, 43.2%) and *relational* (n=192,

37.0%) processes were the most recurrent for expressing EOs, often combined in the same EO category. Participants used *material* and *relational* processes similarly to express an *explicit negative EO* (n=101, 45.1%, and n=81, 42.2%, respectively), mostly for *explicit expression of negative appreciation* (eg “I was *working* [material] in the office last week and my diet was [relational] terrible”; n=69, 30.8%, and n=54, 28.1%, respectively) and *explicit or implicit expression of positive self-judgment* (eg, “On Friday I *did* [material] my Pilates classes and it was [relational] great after, as a miracle my back pain *disappeared* [material]”; n=101, 45.1%, and n=79, 41.2%, respectively). In both cases, participants introduced the situation with the material process and communicated their emotions about it with the relational process. In addition, *mental* and *behavioral* processes were used more often for *negative* (eg, “I *forgot* [mental] to take my multivitamin for 3 days last week” and “Things aren’t the same since before childbirth sometimes when I *sneeze* [behavioral]”; n=36, 62%, vs n=22, 28%, and n=16, 64%, vs n=9, 36%, respectively) than for *positive* EOs (eg, “I *feel* [mental] my sleep is getting better but I *think* [mental] that might be due to increasing my walking distance” and “...also *listening* [behavioral] to my body when I need rest and a cup of tea”). *Existential* (eg, “...however *there has been* [existential] a day or two were I didn’t snack and that reflected in my energy levels and mood”) and verbal processes (eg “I have to *admit* [verbal] that our portion sizes would be much larger than these”) were marginally identified in 6 (1%) and 14 (3%) of all EOs in a similar proportion for negative and positive expressions (n=3, 50%, each and n=7, 50%, each, respectively). However, when the expressions were negative, participants only used these processes for the *explicit expression of negative appreciation* category.

Table 3. Participants’ demographics and program details.

Characteristic/detail	Participants
Age (years), mean (SD)	37.59 (3.69)
BMI, mean (SD)	25.82 (5.68)
Program details, mean (SD)	
Days on program	80.76 (30.47)
Coach sent messages	7.62 (1.82)
User sent messages	2.27 (1.19)
Goals of participants, n (%)	
Weight	7 (24.1)
Physical activity	27 (93.1)
Number of steps	21 (72.4)
Diet	24 (82.8)
Life	15 (51.7)

Table 4. Coding results expressed in number of occurrences, word count, and percentage of total word count.

Category	Occurrences, n (%)	Word count (N=8209), n (%)
EOs^a	164 (100)	3478 (42.37)
Implicit expression of negative feelings	12 (7.32)	351 (4.28)
Implicit expression of negative appreciation (things, events, actions)	39 (23.78)	987 (12.02)
Implicit expression of negative judgement (others or self)	4 (2.44)	104 (1.27)
Pooled implicit negative EOs	55 (33.54)	1442 (17.57)
Explicit expression of negative feelings	15 (9.15)	267 (3.25)
Explicit expression of negative appreciation (things, events, actions)	20 (12.20)	317 (3.86)
Explicit expression of negative judgement (others or self)	0	0
Pooled explicit negative EOs	35 (21.34)	584 (7.11)
Pooled negative EOs	90 (54.88)	2026 (24.68)
Explicit or implicit expression of positive self-judgement	74 (45.12)	1481 (18.04)
Message structure (n=148 messages)		
Opening	75 (34.40)	430 (5.24)
Content	96 (44.04)	7077 (86.21)
Closing	47 (21.56)	270 (3.29)
Full structure (pooled opening, content, and closing)	38 (25.68)	4011 (48.86)
Single (opening, content, or closing)	28 (18.92)	432 (5.26)
Sentence structure (n=734 sentences)		
Declarative	697 (94.96)	7084 (86.30)
Declarative, question function	4 (0.54)	76 (0.93)
Imperative	14 (1.91)	88 (1.07)
Interrogative	19 (2.59)	187 (2.28)
Other	0	774 (9.43)
Process	1025 (100)	1839 (22.40)
Behavioral	34 (3.32)	87 (1.06)
Existential	10 (0.98)	24 (0.29)
Material	430 (41.95)	876 (10.67)
Mental	180 (17.56)	287 (3.50)
Relational	325 (31.71)	495 (6.03)
Verbal	46 (4.49)	85 (1.04)

^aEO: empathy opportunity.

Table 5. Percentage (%) of occurrences per process category identified for each EO^a category.

Process (n=519 occurrences, 50.63%) found in the identified EOs (n=164)	Behavioral (n=25, 4.8%), n (%)	Existential (n=6, 1.2%), n (%)	Material (n=224, 43.2%), n (%)	Mental (n=58, 11.2%), n (%)	Relational (n=192, 37.0%), n (%)	Verbal (n=14, 2.7%), n (%)
EOs (n=164 occurrences)						
Explicit expression of negative appreciation (things, events, actions)	3 (12.0)	3 (50.0)	69 (30.8)	13 (22.4)	54 (28.1)	6 (43.0)
Explicit expression of negative judgement (others or self)	2 (8)	0	7 (3.1)	2 (3.4)	8 (4.2)	1 (7.0)
Explicit expressions of negative feelings	4 (16.0)	0	25 (11.2)	7 (12.1)	19 (9.9)	0
<i>Pooled explicit negative EOs</i>	9 (36.0)	3 (50.0)	101 (45.1)	21 (36.2)	81 (42.2)	7 (50.0)
Implicit expression of negative appreciation (things, events, actions)	5 (20.0)	0	16 (7.1)	3 (5.2)	17 (8.9)	0
Implicit expression of negative judgement (others or self)	0	0	0	0	0	0
Implicit expressions of negative feelings	2 (8.0)	0	7 (3.1)	11 (19.0)	15 (7.8)	0
<i>Pooled implicit negative EOs</i>	7 (28.0)	0	22 (9.8)	15 (25.9)	33 (17.2)	0
<i>Pooled negative EOs</i>	16 (64.0)	3 (50.0)	123 (54.9)	36 (62.1)	113 (58.9)	7 (50.0)
Explicit or implicit expression of positive self-judgement	9 (36.0)	3 (50.0)	101 (45.1)	22 (37.9)	79 (41.1)	7 (50.0)

^aEO: empathy opportunity.

Transitivity Analysis

When we performed a transitivity analysis, the participant roles varied according to the process type. As shown in Table 4, *material* processes dominated the text corpus (n=876, 10.67%), followed by *relational* (n=495, 6.03%). *Relational* processes are used for either characterizing, including a carrier and an attribute as components of the system, or identifying, involving a value and a token. *Material* processes, on the other hand, include an actor (participant), and some demand a goal, while others do not [54]. In addition to these grammatical roles, we categorized findings from this analysis thematically to supplement the meanings expressed. We present the results from transitivity and thematic analyses in Tables 6-11. Most (n=300, 92.2%) *relational* processes were *attributive* (eg, “Your links were very helpful”). The remaining 7.8% (n=25) were *identifying* (“My starting weight was 51.5kg”). The most frequent themes were food and diet (n=63, 19.3%), well-being (n=60, 18.2%), and physical activity (n=44, 13.5%). Similarly, *material* processes frequently (n=217, 70.4%) had the user as

the *actor*, and although their goals were widely spread, the most common categories were food and diet (n=196, 54.9%), physical activity (n=96, 26.9%), and goals (n=40, 11.2%). For example, “I *open* the dates *put* a bit of peanut butter in them, then *put* them in the freezer to harden” and “I’ve *added* a pelvic floor exercise goal.” In contrast, *mental* processes (n=287, 3.5%) involve a *senser* and a *phenomenon* in the transitivity system. This corpus showed a predominance (n=84, 93.3%) of the user as the *senser* and food and diet (n=24, 26.1%), well-being (n=16, 18.9%) and physical activity (n=13, 14.8%) as the *phenomenon* (eg, “I have *included* new snacks like olives” and “I *decided* to have a go with cross trainer”). In *verbal* clauses, a *sayer* directs a message to a *receiver*. In this corpus, despite its low occurrence (n=85, 1.0%), the most frequent *sayer* was the user (n=19, 55.9%) and the *receiver* was usually a health care professional (HCP; n=5, 38.5%; eg, “I *talked* with my GP about the pains”). The most common thematic, as in the other processes, was food and diet (n=9, 20.5%), with an identical occurrence to well-being (n=9, 20.5%).

Table 6. Transitivity analysis results for material processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Actor (n=308)	
User	217 (70.4)
Not human	63 (20.7)
We	11 (3.6)
Another person	6 (1.8)
Coach	5 (1.6)
User's HCP ^a	4 (1.3)
User's partner	2 (0.6)
Goal (n=357)	
Food and diet	196 (54.9)
Physical activity	96 (26.9)
Goals	40 (11.2)
Other (something, nothing, anything, things)	17 (4.6)
Other (place, object, pain, work, mood, medicine, body part)	5 (1.4)
Message (user or coach sent)	2 (0.5)
Person (user, coach, baby, HCP)	1 (0.2)
App	1 (0.2)

^aHCP: health care professional.

Table 7. Transitivity analysis results for relational processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Attributive	
With a carrier	266 (81.7)
Without a carrier	34 (10.4)
Identifying	25 (7.8)
Themes (n=325)	
Food and diet	63 (19.3)
Well-being	60 (18.2)
Physical activity	44 (13.5)
Goals	30 (9.1)
Pregnancy and baby	27 (8.1)
Pain	25 (7.8)
Stress	21 (6.4)
Work	16 (4.7)
App	14 (4.4)
Coach messages	9 (2.7)
Mood and emotions	8 (2.4)
User messages	5 (1.4)
Weather	3 (1.0)
App	1 (0.2)

Table 8. Transitivity analysis results for mental processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Senser (n=90)	
User	84 (93.3)
Coach	3 (3.2)
Doctor	2 (2.1)
User's partner	1 (1.1)
Phenomenon (n=90)	
Food and diet	24 (26.1)
Well-being	16 (18.2)
Physical activity	13 (14.8)
Goals	9 (10.2)
App	9 (10.2)
Pain	9 (10.2)
Planning	4 (4.6)
Baby	3 (3.4)
Coach messages	2 (2.3)

Table 9. Transitivity analysis results for behavioral processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Behaver (n=15)	
User	13 (94.3)
You	1 (2.9)
We	1 (2.9)
Themes (n=35)	
Food and diet	13 (38.2)
Physical activity	5 (14.7)
Pain	3 (8.8)
App	3 (8.8)
Goals	3 (8.8)
Pregnancy and baby	3 (8.8)
Well-being	3 (8.8)
Coach messages	1 (2.9)

Table 10. Transitivity analysis results for verbal processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Sayer (n=34)	
User	19 (55.9)
Coach	12 (35.3)
HCP ^a	2 (5.9)
We	1 (2.9)
Receiver (n=13)	
HCP	5 (38.5)
User	4 (30.8)
Coach	4 (30.8)
Themes (n=46)	
Food and diet	9 (20.5)
Well-being	9 (20.5)
Coach messages	7 (15.4)
Pain	6 (12.8)
Physical activity	6 (12.8)
Goals	4 (7.7)
User messages	2 (5.1)
App	1 (2.6)
Pregnancy and baby	1 (2.6)

^aHCP: health care professional.

Table 11. Transitivity analysis results for existential processes.

Processes, grammatical roles, and themes	Occurrences, n (%)
Themes (n=10)	
Physical activity	3 (30.0)
Food and diet	3 (30.0)
Stress	2 (20.0)
App	1 (10.0)
Pregnancy and baby	1 (10.0)

Discussion

Principal Findings

Empathy and SFL approaches, as we hypothesized in the introduction, can be successfully combined. Our findings show that the SFL categories we explored in the transitivity analysis correspond to and supplement Pounds' EOs. Our findings reveal interesting meanings that originate from the user's linguistic choices whenever they express an EO, particularly when these are "hidden" in an implicit form. Since HCPs frequently overlook these in patient-provider communication, identifying and responding to them optimally is critical for successful health interventions. To the best of our knowledge, no other researchers have previously presented this novel perspective.

Overall, our results show that the users expressed negative EOs more often than positive ones (74 vs 90), of which 60% (55 of 90) were expressed implicitly. Given the context of our data set—a coaching program in which pregnant women communicate with a coach who supports them throughout their journey—the existence of negative EOs is not surprising. We frequently use negative statements to draw attention to a problem that we expect the receiver to empathize with or assist us with. Positive EOs, on the other hand, are less common because they do not serve the purpose of seeking support. However, they do provide a chance for the coach to praise the user's behavior [53]. Moreover, the user's preference for implicit EOs could be due to a polite relationship with their coaches, which would prevent them from making too negative statements. The absence of *explicit expressions of negative judgment about others or self* could support this interpretation. Such insights could be useful for coaches to detect empathic expressions and support

users further. Moreover, our results from the message and sentence structure analyses indicated that most of the wording used was found in the content section of the message with a predominant use of a declarative structure. We explored a relationship between the empathy categories and the message structure, and between the empathy categories and the sentence structure. Such analyses showed no variability in the data across categories; hence, we chose not to include them in this paper. Nevertheless, the predominance of this sentence structure finding is expected, as the use of statements prevails in lengthier and more monological, narrational stretches of communication, which most of these messages were. The scarce presence of interrogative sentences (2.28%) shows that these users were not posing questions and asking for help. Nonetheless, a more qualitative analysis of interrogative occurrences could provide a deeper understanding of these linguistic choices. We added a fourth category during our analysis to account for those sentences where a question (interrogative function) was atypically realized through a statement (declarative structure), representing 0.93% of the sentences in the corpus. A functional interpretation of this phenomenon could be that users were moderating their queries to be less imposing and less direct. The short interaction time (3 months) can also explain this declarative choice, indicating an insufficient time allowed to develop a coach-user relationship. At the start of the program, users met their coach during a synchronous call followed by a small number of purely asynchronous interactions (the mean for coach-sent messages was 7.62 and user-sent messages was 2.27). Participants could be afraid or not feel the confidence to actively ask for information or help due to an insufficient or too polite relationship with their coach. Another reason could be low trust on or missing information about the coach's ability to support them.

Our transitivity analysis showed that the users were the main participants in the clause (eg, material processes described their lifestyle actions, such as food and diet, physical activity, and goals), such as cooking, eating, or exercising, as this was a lifestyle, goal setting-based coaching program. Additionally, these results were in line with the characteristics of the EOs detected: material processes were predominately used for *expressions of positive self-judgment* (eg, "I am *eating* healthy and *doing* long walks every day") or *implicit expressions of negative appreciation* (eg, "I only *had* one Panadol for pain management, but it *does not work* that well"); mental processes disclosed negative feelings (eg, "I *feel* extremely tired and struggling to get 10 thousand step per day"), and negative appreciations were realized through attributive, relational processes (eg, "my snacking *has been* desperate"). These are highly interesting findings for health behavior change programs that have the potential to contribute to promoting user outcomes [75]. The information shared by the user, being explicit or implicit, helps the coach understand the user's perspectives and coaching needs. Although explicit expressions are easy to detect, a more efficient detection and understanding of implicit expressions will contribute to better coaching support. Our insights will provide guidance on the most empathic coach responses and serve to determine the optimal text message-based coach responses in telehealth interventions. Because of its relationship to the EO categories, transitivity

analysis opens a range of opportunities for improving patient care overall. Our perspectives for these findings include further exploration of the empathy and the linguistic elements (SFL) found in the coaches' optimal responses and their connection with user outcomes. Moreover, an association with the BCTs used by the coaches in response to these messages could provide additional insights to boost the impact of digital lifestyle text-based interventions.

Comparison With Prior Work

As we previously described in the Introduction section of this paper, prior work has mainly studied empathy and linguistics in text-based communication separately. Empathy is an important element in the patient-provider relationship that improves patient outcomes [47]. Some authors have measured empathy in a digital setting with surveys [76], with different indexes or scales [77], or as a predefined element in broader coding systems [78]. Other research work has focused on a computational approach to automatize empathy detection (eg, in digital mental health services [79,80]). With regard to linguistics, there has been an increasing interest in digital communication [81], and researchers have applied different linguistic perspectives, such as digital conversation analysis (CA) [82] and SFL. Both CA and SFL perspectives have been applied to digital contexts, such as social media and digital consultations [83]. However, Pounds [53] was the first one to recently define and demonstrate the use of empathy appraisal categories in text-based, patient-provider interactions. The appraisal framework generally studies the meaning negotiation among the speakers, using every utterance to align or misalign with others. In SFL, this framework describes the linguistic resources that the speakers use to construe their social experience and build an intersubjectivity with the recipient, contributing to the interpersonal metafunction [84]. Furthermore, according to Martin and White [85], the appraisal system is organized in 3 domains used to negotiate and modulate emotions, judgements, and valuations: *engagement*, *attitude*, and *graduation*. The *attitude* is the system of meanings represented by the feelings expressed. *Graduation* intensifies or diminishes this representation of meanings. *Engagement* reflects the commitment of the speaker to the appraisal expressed. These appraisal system domains are further explained by Martin [86] as an expansion of the theoretical and descriptive focus of SFL described by Halliday and serve to analyze the speakers' feelings. Pounds' empathy appraisal categories are grounded on this research work and were later suggested for their combination with SFL categories by Pounds herself and De-Pablos Ortega [42]. We contacted Pounds and De-Pablos Ortega for research collaboration. However, they confirmed that they had discontinued their work on the topic. Thereby, we are the first to code a text-based health interaction using both an empathy and an SFL approach. This pilot study served to assess whether these 2 methodologies were compatible, with promising results. We will include further features previously used in text-message coding, such as sentiment analysis, in our future research. Some software can perform automatic sentiment analysis (eg, through word rating [87]), while other authors have resorted to more elaborate machine learning and algorithms for more accurate results [88,89]. Furthermore, coding for

additional elements (eg, emojis or modality) could inform the negotiation of interpersonal roles, such as the user-coach relationship, and be paired with the EO displayed [73].

Limitations

Given the small sample size (n=148), our results should be carefully observed. We aimed to test the combination of empathy and linguistic approaches for text messages analysis. Our findings are preliminary and part of a broader project that will continue exploring methodological possibilities in asynchronous communication analysis.

Conclusion

Our transitivity analysis supports the combination of an empathy and a linguistic (SFL) approach. The processes and their related elements correlate with the empathy categories identified in the corpus. These are promising results for future coding in

asynchronous, online interactions. Our study findings shed light on the empathy and linguistic characteristics present in text message-based coaching. We draw attention to the meanings of patient EOs, such as implicitly seeking help or praise, because research shows that HCPs frequently miss these opportunities. Their identification and management have significant implications for the coach-user relationship and to improve coach training in the future. Our next steps will be to study the coaches' messages and to explore the coach-user relationship-building process. We will code the coaches' messages for linguistic choices (SFL) and how they respond to the EOs presented by the users. Additionally, we will link our results with user outcomes in this lifestyle coaching program during pregnancy when at risk of GDM. This future research will allow for the formulation of optimal coach empathic responses.

Acknowledgments

ERV performed data collection and analysis, synthesized the results, and prepared and edited the manuscript until submission. HSP contributed to the data analysis and interpretation processes, provided reliability coding, and participated in the manuscript reviewing process. TS conceived the research, supported the interpretation and discussion of results, and reviewed the final manuscript. This research is part of an industrial PhD project sponsored by Innovation Fund Denmark, the University of Copenhagen, and Liva Healthcare. The Impact Diabetes B2B (Bump2Baby and Me) project has received funding from the European Union's Horizon 2020 research and innovation program (grant agreement no. 847984). The project also acknowledges collaborative funding from the National Health and Medical Research Council (grant no. 1194234). We would like to acknowledge the contribution of the members of the Impact Diabetes B2B Collaboration Group: Associate Professor Sharleen L O'Reilly (University College Dublin), Dr Aisling A Geraghty (University College Dublin), Dr Faisal Zahoor (University College Dublin), Prof Fionnuala M McAuliffe (University College Dublin), Associate Professor Mary Codd (University College Dublin), Associate Professor Ricardo Segurado (University College Dublin), Prof Helle T Maindal (Aarhus University), Nanna Husted Jensen (Aarhus University), Dr Anna Davies (University of Bristol), Associate Professor Christy Burden (University of Bristol), Prof Jane E Norman (University of Bristol), Prof Karsten Vrangbæk (University of Copenhagen), Dr Laura Elina Pirhonen (University of Copenhagen), Prof Enrique Herrera-Viedman (University of Granada), Dr Mercedes Bermúdez (University of Granada), Prof Cristina Campoy (University of Granada), Prof Alberto Puertas (University of Granada), Dr Francisca S Molina (University of Granada), Dr Ditte Hjorth Laursen (Liva Healthcare), Katie Angotti (Liva Healthcare), Stig Jørgensen (Liva Healthcare), Prof Karen J Campbell (Deakin University), Associate Professor Rachel Laws (Deakin University), Associate Professor Vincent L Versace (Deakin University), Prof Helena Teede (Monash University), Dr Cheryce L Harrison (Monash University), Associate Professor Jacqueline Boyle (Monash University), Dr Georgia Soldartis (Monash University).

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the presence of sensitive information, such as names and places, in the user messages analyzed. Researchers who wish to access the full data set may be granted access upon reasonable request to the corresponding author.

Conflicts of Interest

ERV is employed at the company that provides coaching services for the study trial where the research was conducted (Liva Healthcare).

References

1. World Health Organization. Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. 2020. URL: <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/gho-leading-causes-of-death> [accessed 2022-11-29]
2. Sakakibara BM, Kim AJ, Eng JJ. A systematic review and meta-analysis on self-management for improving risk factor control in stroke patients. *Int J Behav Med* 2017 Feb;24(1):42-53 [FREE Full text] [doi: [10.1007/s12529-016-9582-7](https://doi.org/10.1007/s12529-016-9582-7)] [Medline: [27469998](https://pubmed.ncbi.nlm.nih.gov/27469998/)]

3. Alageel S, Gulliford MC, McDermott L, Wright AJ. Multiple health behaviour change interventions for primary prevention of cardiovascular disease in primary care: systematic review and meta-analysis. *BMJ Open* 2017 Jun 15;7(6):e015375 [FREE Full text] [doi: [10.1136/bmjopen-2016-015375](https://doi.org/10.1136/bmjopen-2016-015375)] [Medline: [28619779](https://pubmed.ncbi.nlm.nih.gov/28619779/)]
4. Almutairi N, Hosseinzadeh H, Gopaldasani V. The effectiveness of patient activation intervention on type 2 diabetes mellitus glycemic control and self-management behaviors: a systematic review of RCTs. *Prim Care Diabetes* 2020 Mar;14(1):12-20. [doi: [10.1016/j.pcd.2019.08.009](https://doi.org/10.1016/j.pcd.2019.08.009)] [Medline: [31543458](https://pubmed.ncbi.nlm.nih.gov/31543458/)]
5. Liu Y, Croft JB, Wheaton AG, Kanny D, Cunningham TJ, Lu H, et al. Clustering of five health-related behaviors for chronic disease prevention among adults, United States, 2013. *Prev Chronic Dis* 2016 May 26;13:E70 [FREE Full text] [doi: [10.5888/pcd13.160054](https://doi.org/10.5888/pcd13.160054)] [Medline: [27236381](https://pubmed.ncbi.nlm.nih.gov/27236381/)]
6. Craddock KA, ÓLaighin G, Finucane FM, Gainforth HL, Quinlan LR, Ginis KAM. Behaviour change techniques targeting both diet and physical activity in type 2 diabetes: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2017 Mar 08;14(1):18 [FREE Full text] [doi: [10.1186/s12966-016-0436-0](https://doi.org/10.1186/s12966-016-0436-0)] [Medline: [28178985](https://pubmed.ncbi.nlm.nih.gov/28178985/)]
7. Ely EK, Gruss SM, Luman ET, Albright AL. Response to comment on Ely et al. A national effort to prevent type 2 diabetes: participant-level evaluation of CDC's National Diabetes Prevention Program. *Diabetes Care* 2017;40:1331-1341. *Diabetes Care* 2017 Nov 23;40(11):e163-e163. [doi: [10.2337/dci17-0036](https://doi.org/10.2337/dci17-0036)] [Medline: [29061592](https://pubmed.ncbi.nlm.nih.gov/29061592/)]
8. Valabhji J, Barron E, Bradley D, Bakhai C, Fagg J, O'Neill S, et al. Early outcomes from the English National Health Service Diabetes Prevention Programme. *Diabetes Care* 2020 Jan;43(1):152-160 [FREE Full text] [doi: [10.2337/dc19-1425](https://doi.org/10.2337/dc19-1425)] [Medline: [31719054](https://pubmed.ncbi.nlm.nih.gov/31719054/)]
9. Nelson R. Telemedicine and telehealth: the potential to improve rural access to care. *Am J Nurs* 2017 Jun;117(6):17-18. [doi: [10.1097/01.NAJ.0000520244.60138.1c](https://doi.org/10.1097/01.NAJ.0000520244.60138.1c)] [Medline: [28541980](https://pubmed.ncbi.nlm.nih.gov/28541980/)]
10. Speyer R, Denman D, Wilkes-Gillan S, Chen Y, Bogaardt H, Kim J, et al. Effects of telehealth by allied health professionals and nurses in rural and remote areas: a systematic review and meta-analysis. *J Rehabil Med* 2018 Mar 28;50(3):225-235 [FREE Full text] [doi: [10.2340/16501977-2297](https://doi.org/10.2340/16501977-2297)] [Medline: [29257195](https://pubmed.ncbi.nlm.nih.gov/29257195/)]
11. Wu C, Wu Z, Yang L, Zhu W, Zhang M, Zhu Q, et al. Evaluation of the clinical outcomes of telehealth for managing diabetes: a PRISMA-compliant meta-analysis. *Medicine (Baltimore)* 2018 Oct;97(43):e12962 [FREE Full text] [doi: [10.1097/MD.00000000000012962](https://doi.org/10.1097/MD.00000000000012962)] [Medline: [30412116](https://pubmed.ncbi.nlm.nih.gov/30412116/)]
12. Berry RM, Ashby JS, Gnilk PB, Matheny KB. A comparison of face-to-face and distance coaching practices: coaches' perceptions of the role of the working alliance in problem resolution. *Consult Psychol J: Pract Res* 2011 Dec;63(4):243-253. [doi: [10.1037/a0026735](https://doi.org/10.1037/a0026735)]
13. Shigekawa E, Fix M, Corbett G, Roby DH, Coffman J. The current state of telehealth evidence: a rapid review. *Health Aff (Millwood)* 2018 Dec;37(12):1975-1982. [doi: [10.1377/hlthaff.2018.05132](https://doi.org/10.1377/hlthaff.2018.05132)] [Medline: [30633674](https://pubmed.ncbi.nlm.nih.gov/30633674/)]
14. An S, Song R. Effects of health coaching on behavioral modification among adults with cardiovascular risk factors: systematic review and meta-analysis. *Patient Educ Couns* 2020 Oct;103(10):2029-2038 [FREE Full text] [doi: [10.1016/j.pec.2020.04.029](https://doi.org/10.1016/j.pec.2020.04.029)] [Medline: [32448627](https://pubmed.ncbi.nlm.nih.gov/32448627/)]
15. Nelson LA, Mayberry LS, Wallston K, Kripalani S, Bergner EM, Osborn CY. Development and usability of REACH: a tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. *JMIR Hum Factors* 2016 Sep 08;3(2):e23 [FREE Full text] [doi: [10.2196/humanfactors.6029](https://doi.org/10.2196/humanfactors.6029)] [Medline: [27609738](https://pubmed.ncbi.nlm.nih.gov/27609738/)]
16. Malecki HL, Gollie JM, Scholten J. Physical activity, exercise, whole health, and integrative health coaching. *Phys Med Rehabil Clin N Am* 2020 Nov;31(4):649-663. [doi: [10.1016/j.pmr.2020.06.001](https://doi.org/10.1016/j.pmr.2020.06.001)] [Medline: [32981584](https://pubmed.ncbi.nlm.nih.gov/32981584/)]
17. Bodenheimer T. Coaching patients to be active, informed partners in their health. *Fam Syst Health* 2020 Jun;38(2):190-192. [doi: [10.1037/fsh0000494](https://doi.org/10.1037/fsh0000494)] [Medline: [32525353](https://pubmed.ncbi.nlm.nih.gov/32525353/)]
18. Thiyagarajan A, Grant C, Griffiths F, Atherton H. Exploring patients' and clinicians' experiences of video consultations in primary care: a systematic scoping review. *BJGP Open* 2020;4(1):bjgpopen20X101020 [FREE Full text] [doi: [10.3399/bjgpopen20X101020](https://doi.org/10.3399/bjgpopen20X101020)] [Medline: [32184212](https://pubmed.ncbi.nlm.nih.gov/32184212/)]
19. Seppen BF, den Boer P, Wiegel J, Ter Wee MM, van der Leeden M, de Vries R, et al. Asynchronous mHealth interventions in rheumatoid arthritis: systematic scoping review. *JMIR Mhealth Uhealth* 2020 Nov 05;8(11):e19260 [FREE Full text] [doi: [10.2196/19260](https://doi.org/10.2196/19260)] [Medline: [33151161](https://pubmed.ncbi.nlm.nih.gov/33151161/)]
20. Linden A, Butterworth S, Prochaska J. Motivational interviewing-based health coaching as a chronic care intervention. *J Eval Clin Pract* 2010 Mar;16(1):166-174. [doi: [10.1111/j.1365-2753.2009.01300.x](https://doi.org/10.1111/j.1365-2753.2009.01300.x)] [Medline: [20367828](https://pubmed.ncbi.nlm.nih.gov/20367828/)]
21. Klein M, Mogles N, van Wissen A. Intelligent mobile support for therapy adherence and behavior change. *J Biomed Inform* 2014 Oct;51:137-151 [FREE Full text] [doi: [10.1016/j.jbi.2014.05.005](https://doi.org/10.1016/j.jbi.2014.05.005)] [Medline: [24858491](https://pubmed.ncbi.nlm.nih.gov/24858491/)]
22. Hansel B, Giral P, Gambotti L, Lafourcade A, Peres G, Filipecki C, et al. A fully automated web-based program improves lifestyle habits and HbA1c in patients with type 2 diabetes and abdominal obesity: randomized trial of patient e-coaching nutritional support (the ANODE Study). *J Med Internet Res* 2017 Nov 08;19(11):e360 [FREE Full text] [doi: [10.2196/jmir.7947](https://doi.org/10.2196/jmir.7947)] [Medline: [29117929](https://pubmed.ncbi.nlm.nih.gov/29117929/)]
23. Wang L, Fagan C, Yu C. Popular mental health apps (MH apps) as a complement to telepsychotherapy: guidelines for consideration. *J Psychother Integ* 2020 Jun;30(2):265-273. [doi: [10.1037/int0000204](https://doi.org/10.1037/int0000204)] [Medline: [2020](https://pubmed.ncbi.nlm.nih.gov/2020/)]

24. White JS, Toussaert S, Thrul J, Bontemps-Jones J, Abroms L, Westmaas JL. Peer mentoring and automated text messages for smoking cessation: a randomized pilot trial. *Nicotine Tob Res* 2020 Mar 16;22(3):371-380. [doi: [10.1093/ntr/ntz047](https://doi.org/10.1093/ntr/ntz047)] [Medline: [30892616](https://pubmed.ncbi.nlm.nih.gov/30892616/)]
25. Damschroder LJ, Buis LR, McCant FA, Kim HM, Evans R, Oddone EZ, et al. Effect of adding telephone-based brief coaching to an mHealth app (Stay Strong) for promoting physical activity among veterans: randomized controlled trial. *J Med Internet Res* 2020 Aug 04;22(8):e19216 [FREE Full text] [doi: [10.2196/19216](https://doi.org/10.2196/19216)] [Medline: [32687474](https://pubmed.ncbi.nlm.nih.gov/32687474/)]
26. Singh HK, Kennedy GA, Stupans I. Does the modality used in health coaching matter? A systematic review of health coaching outcomes. *Patient Prefer Adherence* 2020 Aug;14:1477-1492 [FREE Full text] [doi: [10.2147/PPA.S265958](https://doi.org/10.2147/PPA.S265958)] [Medline: [32904668](https://pubmed.ncbi.nlm.nih.gov/32904668/)]
27. Berry MP, Sala M, Abber SR, Forman EM. Incorporating automated digital interventions into coach-delivered weight loss treatment: a meta-analysis. *Health Psychol* 2021 Aug;40(8):534-545. [doi: [10.1037/hea0001106](https://doi.org/10.1037/hea0001106)] [Medline: [34618500](https://pubmed.ncbi.nlm.nih.gov/34618500/)]
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](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
29. Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. *Health Psychol Rev* 2019 Mar;13(1):110-130. [doi: [10.1080/17437199.2018.1534071](https://doi.org/10.1080/17437199.2018.1534071)] [Medline: [30295176](https://pubmed.ncbi.nlm.nih.gov/30295176/)]
30. Pourrazavi S, Kouzekanani K, Bazargan-Hejazi S, Shaghagh A, Hashemiparast M, Fathifar Z, et al. Theory-based e-health literacy interventions in older adults: a systematic review. *Arch Public Health* 2020 Aug 10;78(1):72 [FREE Full text] [doi: [10.1186/s13690-020-00455-6](https://doi.org/10.1186/s13690-020-00455-6)] [Medline: [32793345](https://pubmed.ncbi.nlm.nih.gov/32793345/)]
31. Sineath A, Lambert L, Verga C, Wagstaff M, Wingo BC. Monitoring intervention fidelity of a lifestyle behavioral intervention delivered through telehealth. *Mhealth* 2017 Aug 25;3:35 [FREE Full text] [doi: [10.21037/mhealth.2017.07.04](https://doi.org/10.21037/mhealth.2017.07.04)] [Medline: [28894745](https://pubmed.ncbi.nlm.nih.gov/28894745/)]
32. Zengul A, Evans E, Hall A, Qu H, Willig A, Cherrington A, et al. Telehealth behavioral intervention for diabetes management in adults with physical disabilities: intervention fidelity protocol for a randomized controlled trial. *JMIR Res Protoc* 2021 Sep 10;10(9):e31695 [FREE Full text] [doi: [10.2196/31695](https://doi.org/10.2196/31695)] [Medline: [34505835](https://pubmed.ncbi.nlm.nih.gov/34505835/)]
33. Dogar O, Boehnke JR, Lorencatto F, Sheldon TA, Siddiqi K. Measuring fidelity to behavioural support delivery for smoking cessation and its association with outcomes. *Addiction* 2020 Feb;115(2):368-380 [FREE Full text] [doi: [10.1111/add.14804](https://doi.org/10.1111/add.14804)] [Medline: [31496033](https://pubmed.ncbi.nlm.nih.gov/31496033/)]
34. Hurlocker MC, Madson MB, Schumacher JA. Motivational interviewing quality assurance: a systematic review of assessment tools across research contexts. *Clin Psychol Rev* 2020 Dec;82:101909 [FREE Full text] [doi: [10.1016/j.cpr.2020.101909](https://doi.org/10.1016/j.cpr.2020.101909)] [Medline: [32920371](https://pubmed.ncbi.nlm.nih.gov/32920371/)]
35. Chai LK, Collins CE, May C, Brown LJ, Ashman A, Burrows TL. Fidelity and acceptability of a family-focused technology-based telehealth nutrition intervention for child weight management. *J Telemed Telecare* 2021 Mar;27(2):98-109. [doi: [10.1177/1357633X19864819](https://doi.org/10.1177/1357633X19864819)] [Medline: [31390947](https://pubmed.ncbi.nlm.nih.gov/31390947/)]
36. Salloum RG, Rojewski AM, Piper ME, Blalock JA, Borrelli B, Boyce LM, et al. Reporting treatment fidelity in behavioral tobacco treatment clinical trials: scoping review and measurement recommendations. *Nicotine Tob Res* 2022 Mar 01;24(2):150-159 [FREE Full text] [doi: [10.1093/ntr/ntab140](https://doi.org/10.1093/ntr/ntab140)] [Medline: [34197617](https://pubmed.ncbi.nlm.nih.gov/34197617/)]
37. Suffoletto B. Text message behavioral interventions: from here to where? *Curr Opin Psychol* 2016 Mar 01;9:16-21 [FREE Full text] [doi: [10.1016/j.copsyc.2015.09.012](https://doi.org/10.1016/j.copsyc.2015.09.012)] [Medline: [26665157](https://pubmed.ncbi.nlm.nih.gov/26665157/)]
38. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. *JAMA* 2015;314(12):1255-1263. [doi: [10.1001/jama.2015.10945](https://doi.org/10.1001/jama.2015.10945)] [Medline: [26393848](https://pubmed.ncbi.nlm.nih.gov/26393848/)]
39. Paulus T, Warren A, Lester J. Applying conversation analysis methods to online talk: a literature review. *Discourse, Context Media* 2016 Jun;12:1-10 [FREE Full text] [doi: [10.1016/j.dcm.2016.04.001](https://doi.org/10.1016/j.dcm.2016.04.001)]
40. Villarreal A. *Constructing Meaning: A Systemic Functional Review of Translated Health Texts and Community Narratives*. Ann Arbor, MI: ProQuest; 2017.
41. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, et al. Are there interactional differences between telephone and face-to-face psychological therapy? A systematic review of comparative studies. *J Affect Disord* 2020 Mar 15;265:120-131 [FREE Full text] [doi: [10.1016/j.jad.2020.01.057](https://doi.org/10.1016/j.jad.2020.01.057)] [Medline: [32090733](https://pubmed.ncbi.nlm.nih.gov/32090733/)]
42. Pounds G, De Pablos -Ortega C. Patient-centred communication in British, Italian and Spanish 'Ask-the-Expert' healthcare websites. *Commun Med* 2015 Nov 07;12(2-3):225-241. [doi: [10.1558/cam.27231](https://doi.org/10.1558/cam.27231)] [Medline: [29048864](https://pubmed.ncbi.nlm.nih.gov/29048864/)]
43. Jeffrey D. Empathy, sympathy and compassion in healthcare: is there a problem? Is there a difference? Does it matter? *J R Soc Med* 2016 Dec 06;109(12):446-452 [FREE Full text] [doi: [10.1177/0141076816680120](https://doi.org/10.1177/0141076816680120)] [Medline: [27923897](https://pubmed.ncbi.nlm.nih.gov/27923897/)]
44. Chochinov HM. Dignity and the essence of medicine: the A, B, C, and D of dignity conserving care. *BMJ* 2007 Jul 28;335(7612):184-187 [FREE Full text] [doi: [10.1136/bmj.39244.650926.47](https://doi.org/10.1136/bmj.39244.650926.47)] [Medline: [17656543](https://pubmed.ncbi.nlm.nih.gov/17656543/)]
45. Guidi C, Traversa C. Empathy in patient care: from 'clinical empathy' to 'empathic concern'. *Med Health Care Philos* 2021 Dec 01;24(4):573-585 [FREE Full text] [doi: [10.1007/s11019-021-10033-4](https://doi.org/10.1007/s11019-021-10033-4)] [Medline: [34196934](https://pubmed.ncbi.nlm.nih.gov/34196934/)]

46. Piasecki M. *Clinical Communication Handbook*. Hoboken, NJ: Blackwell Science; 2003.
47. Howick J, Moscrop A, Mebius A, Fanshawe TR, Lewith G, Bishop FL, et al. Effects of empathic and positive communication in healthcare consultations: a systematic review and meta-analysis. *J R Soc Med* 2018 Jul;111(7):240-252 [[FREE Full text](#)] [doi: [10.1177/0141076818769477](https://doi.org/10.1177/0141076818769477)] [Medline: [29672201](https://pubmed.ncbi.nlm.nih.gov/29672201/)]
48. Bonvicini KA, Perlin MJ, Bylund CL, Carroll G, Rouse RA, Goldstein MG. Impact of communication training on physician expression of empathy in patient encounters. *Patient Educ Couns* 2009 Apr;75(1):3-10 [[FREE Full text](#)] [doi: [10.1016/j.pec.2008.09.007](https://doi.org/10.1016/j.pec.2008.09.007)] [Medline: [19081704](https://pubmed.ncbi.nlm.nih.gov/19081704/)]
49. Bylund CL, Makoul G. Examining empathy in medical encounters: an observational study using the Empathic Communication Coding System. *Health Commun* 2005 Oct;18(2):123-140. [doi: [10.1207/s15327027hc1802_2](https://doi.org/10.1207/s15327027hc1802_2)] [Medline: [16083407](https://pubmed.ncbi.nlm.nih.gov/16083407/)]
50. Hsu I, Saha S, Korthuis PT, Sharp V, Cohn J, Moore RD, et al. Providing support to patients in emotional encounters: a new perspective on missed empathic opportunities. *Patient Educ Couns* 2012 Sep;88(3):436-442 [[FREE Full text](#)] [doi: [10.1016/j.pec.2012.06.015](https://doi.org/10.1016/j.pec.2012.06.015)] [Medline: [22818767](https://pubmed.ncbi.nlm.nih.gov/22818767/)]
51. Muñoz-Miquel A. Empathy, emotions and patient-centredness: a case study on communication strategies. *HJLCB* 2019 Nov 01;59(1):71-89. [doi: [10.7146/hjlc.v59i1.116990](https://doi.org/10.7146/hjlc.v59i1.116990)]
52. Epstein RM, Franks P, Fiscella K, Shields CG, Meldrum SC, Kravitz RL, et al. Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. *Soc Sci Med* 2005 Oct;61(7):1516-1528. [doi: [10.1016/j.socscimed.2005.02.001](https://doi.org/10.1016/j.socscimed.2005.02.001)] [Medline: [16005784](https://pubmed.ncbi.nlm.nih.gov/16005784/)]
53. Pounds G. Empathy as 'appraisal': a new language-based approach to the exploration of clinical empathy. *J Appl Linguistics Prof Pract* 2010 Dec 22;7(2):145-168. [doi: [10.1558/japl.v7i2.145](https://doi.org/10.1558/japl.v7i2.145)]
54. Halliday M, Matthiessen CMIM. *Halliday's Introduction to Functional Grammar*. 4th ed. London, UK: Routledge; 2014.
55. Halliday M. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London, UK: Edward Arnold; 1978.
56. Suchman AL, Markakis K, Beckman HB, Frankel R. A model of empathic communication in the medical interview. *JAMA* 1997 Feb 26;277(8):678. [doi: [10.1001/jama.1997.03540320082047](https://doi.org/10.1001/jama.1997.03540320082047)]
57. Banks D. Systemic functional linguistics as a model for text analysis. *ASp* 2002 Dec 01(35-36):23-34. [doi: [10.4000/asp.1584](https://doi.org/10.4000/asp.1584)]
58. Martin J, Matthiessen CMIM, Painter C. *Working with Functional Grammar*. London, UK: Edward Arnold; 1997.
59. Zhang Y. Transitivity analysis of Hillary Clinton's and Donald Trump's first television debate. *Int J Appl Linguist Engl Lit* 2017 Oct 10;6(7):65. [doi: [10.7575/aiac.ijalel.v.6n.7p.65](https://doi.org/10.7575/aiac.ijalel.v.6n.7p.65)]
60. Wahyudin AY. An analysis of process type used in news item. *TEKNOSASTIK* 2016;14(1):1-8. [doi: [10.33365/ts.v14i1.79](https://doi.org/10.33365/ts.v14i1.79)]
61. Klaisingto P. A transitivity analysis of the experience of depression as expressed on Thai Twitter. *J Liberal Arts* 2021;13(1):188-215.
62. Matthiessen CMIM. Applying systemic functional linguistics in healthcare contexts. *Text Talk* 2013;33(4-5):437-466. [doi: [10.1515/text-2013-0021](https://doi.org/10.1515/text-2013-0021)]
63. Pounds G. Patient-centred communication in ask-the-expert healthcare websites. *Appl Linguistics* 2018;39(2):117-134. [doi: [10.1093/applin/amv073](https://doi.org/10.1093/applin/amv073)]
64. Fosgerau CF, Davidsen AS. Patients' representations of depressive symptoms and physicians' responses in clinical encounters. *CAM* 2019 Jul 03;15(1):28-39. [doi: [10.1558/cam.32486](https://doi.org/10.1558/cam.32486)]
65. O'Reilly SL, Burden C, Campoy C, McAuliffe FM, Teede H, Andresen J, IMPACT DIABETES B2B Collaboration Group. Bump2Baby and Me: protocol for a randomised trial of mHealth coaching for healthy gestational weight gain and improved postnatal outcomes in high-risk women and their children. *Trials* 2021 Dec 28;22(1):963 [[FREE Full text](#)] [doi: [10.1186/s13063-021-05892-4](https://doi.org/10.1186/s13063-021-05892-4)] [Medline: [34963483](https://pubmed.ncbi.nlm.nih.gov/34963483/)]
66. Hardcastle SJ, Fortier M, Blake N, Hagger MS. Identifying content-based and relational techniques to change behaviour in motivational interviewing. *Health Psychol Rev* 2017 Mar;11(1):1-16. [doi: [10.1080/17437199.2016.1190659](https://doi.org/10.1080/17437199.2016.1190659)] [Medline: [27189713](https://pubmed.ncbi.nlm.nih.gov/27189713/)]
67. QSR International. NVivo. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home> [accessed 2022-11-29]
68. Bernicot J, Volckaert-Legrier O, Goumi A, Bert-Erboul A. Forms and functions of SMS messages: a study of variations in a corpus written by adolescents. *J Pragmat* 2012 Sep;44(12):1701-1715. [doi: [10.1016/j.pragma.2012.07.009](https://doi.org/10.1016/j.pragma.2012.07.009)]
69. Herring SC, Androutsopoulos J. Computer-mediated discourse 2.0. In: *The Handbook of Discourse Analysis*. New York, NY: John Wiley & Sons; 2015.
70. McHoul AW. *Telling how Texts Talk: Essays on Reading and Ethnomethodology*. London, UK: Routledge & K. Paul; 1982.
71. Have PT. Structuring writing for reading: hypertext and the reading body. *Hum Stud* 1999 Oct;22(2-4):273-298. [doi: [10.1023/a:1005496619385](https://doi.org/10.1023/a:1005496619385)]
72. Wang H. The problems of transitivity studies and its solution. *JSS* 2015;03(10):170-176. [doi: [10.4236/jss.2015.310023](https://doi.org/10.4236/jss.2015.310023)]
73. Halliday MAK, Matthiessen CMIM. *An Introduction to Functional Grammar*. 3 ed. London, UK: Edward Arnold; 2004.
74. Morley G. *Syntax in Functional Grammar: An Introduction to Lexicogrammar in Systemic Linguistics*. London, UK: Continuum International; 2000.

75. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
76. Nambisan P. Information seeking and social support in online health communities: impact on patients' perceived empathy. *J Am Med Inform Assoc* 2011 May 01;18(3):298-304 [FREE Full text] [doi: [10.1136/amiajnl-2010-000058](https://doi.org/10.1136/amiajnl-2010-000058)] [Medline: [21486888](https://pubmed.ncbi.nlm.nih.gov/21486888/)]
77. Shen L. On a scale of state empathy during message processing. *West J Commun* 2010 Oct 29;74(5):504-524. [doi: [10.1080/10570314.2010.512278](https://doi.org/10.1080/10570314.2010.512278)]
78. Malik SH, Coulson NS. Coping with infertility online: an examination of self-help mechanisms in an online infertility support group. *Patient Educ Couns* 2010 Nov;81(2):315-318. [doi: [10.1016/j.pec.2010.01.007](https://doi.org/10.1016/j.pec.2010.01.007)] [Medline: [20144521](https://pubmed.ncbi.nlm.nih.gov/20144521/)]
79. Sharma A, Miner A, Atkins D, Althoff T. A computational approach to understanding empathy expressed in text-based mental health support. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP); November 2020; Online. [doi: [10.18653/v1/2020.emnlp-main.425](https://doi.org/10.18653/v1/2020.emnlp-main.425)]
80. Yao H, De SA, Foster A, Galynker I, Lok B. Toward automated evaluation of empathetic responses in virtual human interaction systems for mental health scenarios. 2020 Presented at: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (IVA 2020); October 20-22, 2020; Virtual. [doi: [10.1145/3383652.3423916](https://doi.org/10.1145/3383652.3423916)]
81. Taiwo R, editor. *Handbook of Research on Discourse Behavior*. Digital Communication: Language Structures/Social Interaction (2 Volumes). Hershey, PA: IGI Global; 2010.
82. Virtanen MT, Kääntä L. At the intersection of text and conversation analysis: analysing asynchronous online written interaction. *AFinLA Theme* 2018 Dec 14(11):137-155. [doi: [10.30660/afinla.69081](https://doi.org/10.30660/afinla.69081)]
83. Gardner S, Alsop S. *Systemic Functional Linguistics in the Digital Age*. Sheffield, UK: Equinox; 2016.
84. Oteiza T. The appraisal framework and discourse analysis. In: *The Routledge Handbook of Systemic Functional Linguistics*. London, UK: Routledge; 2017.
85. Martin J, White P. *The Language of Evaluation: Appraisal in English*. New York, NY: Springer; 2005.
86. Martin JR. Evolving systemic functional linguistics: beyond the clause. *Funct Linguist* 2014 Apr 10;1(3):1-24. [doi: [10.1186/2196-419x-1-3](https://doi.org/10.1186/2196-419x-1-3)]
87. NVIVO. Automatically Detect and Code Sentiment. URL: <https://help-nv.qsrinternational.com/20/win/Content/coding/auto-detect-code-sentiment.htm?Highlight=sentiment> [accessed 2022-11-29]
88. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J* 2014 Dec;5(4):1093-1113 [FREE Full text] [doi: [10.1016/j.asej.2014.04.011](https://doi.org/10.1016/j.asej.2014.04.011)]
89. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *J Med Internet Res* 2013 Nov 01;15(11):e239 [FREE Full text] [doi: [10.2196/jmir.2721](https://doi.org/10.2196/jmir.2721)] [Medline: [24184993](https://pubmed.ncbi.nlm.nih.gov/24184993/)]

Abbreviations

- AI:** artificial Intelligence
BCT: behavior change technique
CA: conversation analysis
EO: empathy opportunity
GDM: gestational diabetes mellitus
HCP: health care professional
IHC: integrative health coaching
NCD: noncommunicable disease
PCC: patient-centered communication
SFL: systemic functional linguistics

Edited by A Mavragani; submitted 07.06.22; peer-reviewed by Z Demjen, A Holsting; comments to author 02.11.22; revised version received 21.11.22; accepted 21.11.22; published 20.12.22.

Please cite as:

Rey Velasco E, Pedersen HS, Skinner T, Impact Diabetes B2B Collaboration Group
 Analysis of Patient Cues in Asynchronous Health Interactions: Pilot Study Combining Empathy Appraisal and Systemic Functional Linguistics

JMIR Form Res 2022;6(12):e40058

URL: <https://formative.jmir.org/2022/12/e40058>

doi: [10.2196/40058](https://doi.org/10.2196/40058)

PMID: [36538352](https://pubmed.ncbi.nlm.nih.gov/36538352/)

©Elena Rey Velasco, Hanne Sæderup Pedersen, Timothy Skinner, Impact Diabetes B2B Collaboration Group. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Navigating the Systemic Conditions of a Digital Health Ecosystem in Alberta, Canada: Embedded Case Study

Chad Saunders^{1,2}, MBA, PhD; Devon Currie², MSc; Shane Virani², MSc; Jill De Groot², MA

¹Entrepreneurship & Innovation, Haskayne School of Business, University of Calgary, Calgary, AB, Canada

²Ward of the 21st Century, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:

Chad Saunders, MBA, PhD
Entrepreneurship & Innovation
Haskayne School of Business
University of Calgary
2500 University Drive NW
Calgary, AB, T2N 1N4
Canada
Phone: 1 403 220 6075
Email: wsaunder@ucalgary.ca

Abstract

Background: Digital health promises numerous value-creating outcomes. These include improved health, reduced costs, and the creation of lucrative markets, which, in turn, provide high-quality employment, productivity growth, and a climate that attracts investment. For this value creation and capture, the activities of a diverse set of stakeholders within a digital health ecosystem require coordination. However, the antecedents of the coordination needed for an effective digital health ecosystem are not well understood.

Objective: The purpose of this study was to investigate the systemic conditions of the digital health ecosystem in Alberta, Canada, as critical antecedents to ecosystem coordination from the perspective of the authors as applicants to an innovative digital health funding program embedded within the larger digital health ecosystem of innovators or entrepreneurs, health system leaders, support partners, and funders.

Methods: We employed a qualitative embedded case study of the systemic conditions within the digital health ecosystem in Alberta, Canada (main case) using semistructured interviews with 36 stakeholders representing innovators or entrepreneurs, health system leaders, support partners, and funders (subcases). The interviews were conducted over a 2-month period between May 26 and July 22, 2021. Data were coded for key themes and synthesized around 5 propositions developed from academic publications and policy reports.

Results: The findings indicated varying levels of support for each proposition, with moderate support for accessing real problems, data, training, and space for evaluations. However, the most fundamental gap appears to be in ecosystem navigation, in particular, the absence of intermediaries (eg, individuals, organizations, and technology) to provide guidance on the available support services and dependencies among the various ecosystem actors and programs.

Conclusions: Navigating the systemic conditions of the digital health ecosystem is extremely challenging for entrepreneurs, especially those without prior health care experience, and this remains an issue even for those with such experience. Policy interventions aimed at increasing collaboration among ecosystem support providers, along with tools and incentives to ensure coordination, are essential as the ecosystem and those dependent on it grow.

(*JMIR Form Res* 2022;6(12):e36265) doi:[10.2196/36265](https://doi.org/10.2196/36265)

KEYWORDS

digital health; entrepreneurial ecosystem; systemic conditions; policy

Introduction

Background

Alberta Innovates is Alberta's largest research and innovation agency. In the fall of 2020, they launched a new initiative called the Health Innovation Platform Partnerships (HIPP). The objective of HIPP was "...to build a health innovation ecosystem that is robust, coordinated, and a competitive advantage for Alberta innovators in the health industry" [1]. Recognizing the need to foster more coordination among ecosystem stakeholders, the approach taken with this initiative was quite different from prior funding programs. In this regard, the grant was structured with incentives for applicants to not only provide value-generating activities to clients but also incentivize coordination with other ecosystem actors. There were 2 stages of funding, with the initial stage being an open call for proof-of-concept proposals. These were narrowed down to 11 applicants, who were then invited to submit a full application 6 months later at stage 2, informed by their stage 1 findings. What follows is a case analysis of the experience of one of the applicants, as they validated key ecosystem assumptions underlying their proposed initiative. Specifically, the authors constituted a team at the University of Calgary that submitted an application for a Digital Health Collaboratorium as part of the Ward of the 21st Century (W21C) Research and Innovation Centre, an established health systems research initiative with an overarching mandate to improve patient safety and quality of care [2].

Entrepreneurial Ecosystems

Entrepreneur ecosystems are defined as "a set of interdependent actors and factors coordinated in such a way that they enable productive entrepreneurship" [3]. The assessment of digital health entrepreneurial ecosystems is based on a reconceptualization of the theoretical model developed by Stam and Spigel [4], as shown in Figure 1. Building on a comprehensive review of existing research, including related concepts such as industrial districts, clusters, and innovation systems, Stam [3] created a theoretical model containing 10 key ecosystem elements. The framework conditions included social (ie, informal and formal institutions) and physical conditions enabling or constraining entrepreneurial activity. Systemic conditions are at the core of the ecosystem, representing networks of entrepreneurs, leadership, finance, talent, knowledge, and support services. The presence of these elements and the interactions between them predominantly determine the

success of the ecosystem [3] and thus serve as the focus of this study.

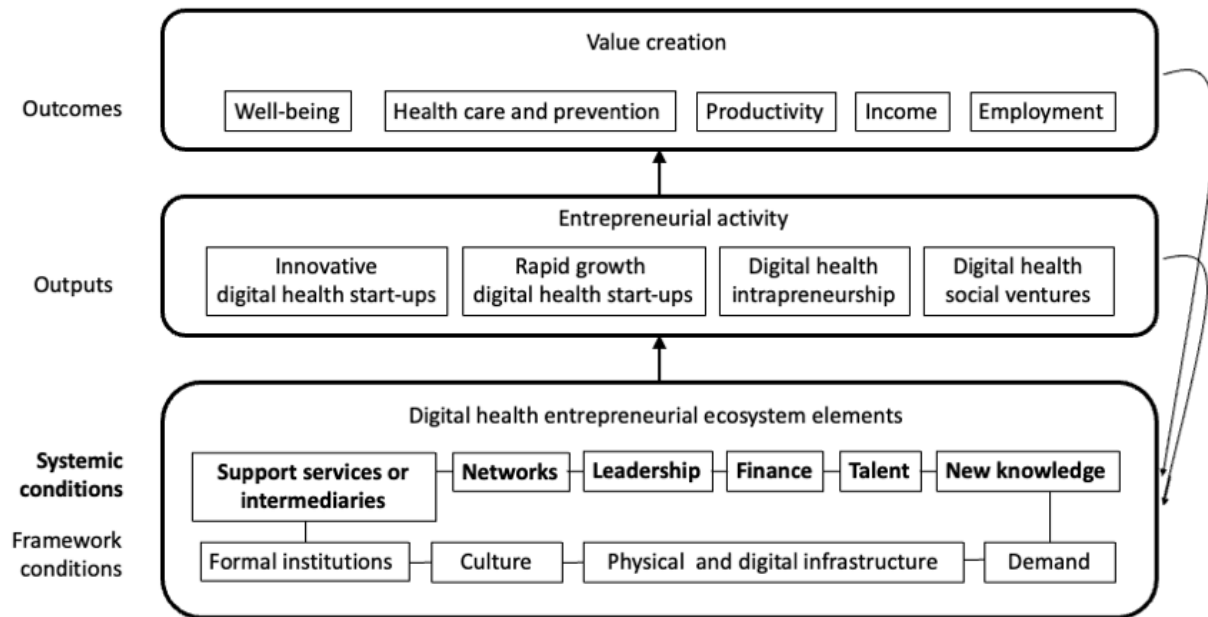
As Stam [3] points out, networks of entrepreneurs facilitate information flow, which, in turn, enables the effective distribution of people and funding. Leadership provides direction and role models for the entrepreneurial ecosystem and is critical for establishing and maintaining ecosystem health. Access to finance is crucial to support the ongoing entrepreneurial activities with their inherent risks and fuel a diverse and skilled group of talented workers. Finally, a supply of support services by a variety of intermediaries can substantially lower the entry barriers for new entrepreneurial projects and facilitate product and service introduction.

Stam and Spigel [4] explained the following:

The new model includes insights from the previous literature (i.e., the aspects that have been deemed important elements of entrepreneurial ecosystems), but most importantly it provides more causal depth...including the upward and downward causation, and intra-layer causal relations. Upward causation reveals how the fundamental causes of new value creation are mediated by intermediate causes, while downward causation shows how outcomes and outputs of the system over time also feed back into the system conditions.

We foreground the systemic conditions as critical antecedents of entrepreneurial activity within the digital health entrepreneurial ecosystem. The purpose of this study was to explore a set of propositions based on the systemic conditions within a digital health ecosystem for key stakeholders, including digital health innovators or entrepreneurs, investors, and health system leaders.

Digital health promises value creation outcomes, including improved health, reduced costs, and the creation of lucrative markets, which, in turn, provide high-quality employment, productivity growth, and attract investment [5]. However, our understanding of how this occurs within entrepreneurial ecosystems generally [3] and within a digital health ecosystem specifically remains limited [6]. The theoretical model (Figure 1) indicates that for value creation and capture, the activities of a diverse set of stakeholders within a digital health ecosystem would need coordination. However, the antecedents of the coordination needed for an effective digital health ecosystem are not well understood [6].

Figure 1. Theoretical model of digital health entrepreneurial ecosystems.

Proposition Development

Overview

This work was conducted through the O'Brien Institute for Public Health's W21C Research and Innovation Centre at the University of Calgary. W21C serves as a research and beta test site for novel approaches to health care delivery, human factors research, and innovative medical technologies. Experience with health care innovators and entrepreneurs through prior work at W21C provided insight into the struggle of identifying real problems within the health care system. Even if problems are deemed legitimate, there is a further obstacle of prioritizing them against one another. It is very challenging for innovators and entrepreneurs to gain visibility into where problem-solving opportunities exist, and this is exacerbated if the innovator or entrepreneur is from outside the health domain. This issue has been broadly recognized within Alberta; in 2012, the Strategic Clinical Networks (SCNs) were created to bring together a diverse set of stakeholders to both identify and rank problems to facilitate the implementation of solutions [7]. Similar challenges have been identified in other jurisdictions of Canada. In Ontario, the approach of "build it and they will come" used by many innovators and entrepreneurs is not working for any of the ecosystem stakeholders [6]. Ultimately, it leads to solutions for nonproblems or solutions that are not feasible to implement in practice. This is an area where digital health is seen as a facilitator of the coordination needed and where the emergence of digital platforms in health care that effectively mediate 2-sided markets by providing technology that connects those that need a service with those that deliver the service lags behind other industries where these digital platforms are well established [8]. This leads to our first proposition.

Proposition 1: Digital Health Innovators and Entrepreneurs Struggle to Access Real Problems

Digital health innovators and entrepreneurs are often tasked with developing and implementing new digital health

interventions to achieve better health outcomes, among other benefits. However, a critical challenge within health care—specifically for digital health technologies—is how to demonstrate *better* and the implications of these improvements for those who adopt the new technology. Evaluation approaches such as randomized control trials, which are required in other health technology assessments, are often impractical or not applicable to digital health innovations [9]. Although institutions such as the World Health Organization provide useful guidelines for monitoring and evaluating digital health interventions [10], most interventions require access to health care data that the innovator or entrepreneur lacks. This is often exacerbated by disconnects between health information legislation and consumer expectations around privacy and security surrounding the use of their data for digital health interventions, as these are increasingly more likely to be operated by companies such as Apple or Oracle rather than traditional health care providers [11,12]. The data generated using digital health technologies are mediated by a plethora of businesses, such as software or device companies, apps, and cloud hosting services, whose information governance processes are complicated and can further obscure visibility into the data needed to guide innovation [12]. This leads to our second proposition.

Proposition 2: Digital Health Innovators and Entrepreneurs Struggle to Access Relevant Data

As noted earlier, although organizations such as the World Health Organization provide guidance on the evaluation of digital health technologies [10] and various regulatory bodies have streamlined their review of specific digital health technologies (eg, Federal Drug Administration [13] and Health Canada [14]), getting help often remains elusive for many innovators and entrepreneurs. Ironically, the proliferation of Academic Medical Centers (AMCs), incubators, accelerators, and regional economic development initiatives focused on helping innovators and entrepreneurs has led to the unintended consequence of making it *more* difficult to navigate these increasingly complex and specialized ecosystem offerings [15].

The adoption of digital technologies in health care is additionally slowed by complex bureaucracy, laborious administrative approval processes, excessive risk assessment, understaffed health information technology departments, and a general reluctance to implement new apps in the clinical workflow. This makes the health care ecosystem incredibly complex and difficult to navigate [16]. Furthermore, entrepreneurs and innovators must identify and collaborate with clinical end users at the site (eg, physicians, nurses, etc) while also designing their solutions to concurrently satisfy pain points for a variety of other decision makers (eg, procurement specialists, practice managers, patients, health IT departments, and security or privacy officers) [16]. Thus, innovators and entrepreneurs are often at a loss to understand how to navigate the business side of their enterprise while also facing the dual challenge of navigating the health care system. This leads to our third proposition.

Proposition 3: Digital Health Innovators and Entrepreneurs Struggle to Understand Where to Start and Where They Need to Go on Their Innovation Journey

A critical systemic condition of an entrepreneurial ecosystem is the provision of talent, which in the context of digital health necessitates personnel trained in digital health to realize the expected improvement in outcomes [17]. There is a further need to increase interorganizational knowledge sharing to facilitate the creation of digital health learning ecosystems. This highlights that the knowledge and experience surrounding technology adoption and implementation is particularly valuable for members of other organizations contemplating similar digitally enabled transformations [18]. Although digital health learning ecosystems rely on formal mechanisms and processes to provide their foundation, they are most effective when supported by informal networks [18]. This leads to our fourth proposition.

Proposition 4: The Breadth of Training Provided Does Not Adequately Meet the Needs of Digital Health Innovators and Entrepreneurs

Emerging digital technologies offer enormous potential to improve quality, reduce costs, and increase patient centeredness in health care. Although AMCs play a key role in advancing medical care through cutting-edge medical research, traditional models for invention, validation, and commercialization have been designed around biomedical initiatives at AMCs. This makes them unsuitable for new digital health technologies [19]. However, AMCs are uniquely positioned. They house cross-disciplinary expertise in health and technology and train the next generation of health professionals, thus providing the opportunity to connect academics and clinicians from across a variety of disciplines with innovators and entrepreneurs [15]. This leads to our fifth proposition.

Proposition 5: Digital Health Innovators and Entrepreneurs Lack a “Space” to Evaluate Their Offerings

In the next section, we have presented an embedded case study to evaluate the set of propositions developed earlier based on the systemic conditions within a digital health ecosystem for key stakeholders, including digital health innovators or entrepreneurs, investors, and health system leaders.

Methods

Recruitment

To explore these propositions, a qualitative embedded case study of the systemic conditions within the digital health ecosystem in Alberta was conducted using semistructured interviews with 36 stakeholders, representing 31% (11/36) innovators or entrepreneurs, 19% (7/36) health system leaders, 39% (14/36) support partners, and 11% (4/36) funders (Table 1). These roles were selected using a theoretical sampling approach to match the systemic conditions of the entrepreneurial ecosystem framework guiding this study (Figure 1).

Participants were recruited from within the University of Calgary network, partnered organizations, and the community at large, leveraging the diverse connections of the W21C. The participants were selected using theoretical sampling based on their roles, interests, and goals in the development of digital health solutions, as guided by the theoretical framework (Figure 1). To ensure adequate coverage, the participants were also selected to cover a range of organizational sizes, activity scopes, and venture phases of entrepreneurs or support providers (Table 1). Interviews were conducted over a 2-month period between May 26 and July 22, 2021, by trained researchers following a pilot-tested semistructured interview protocol. The interviews were recorded (with consent from the participants) and then transcribed verbatim for data analysis. Thematic analysis methods were used on the interview data to address each proposition, and the resulting themes were synthesized using a Gioia data structure diagram [20].

Throughout the analysis, we were attuned to the presence or absence of particular aspects of the underlying assumptions that the participants raised and the depth of insight that they provided as indicators of whether a particular theme was relevant. Our qualitative approach focused on theoretical generalizability and not statistical generalizability. As such, themes were relevant based on a number of criteria, including their prevalence (ie, counts), the depth of coverage (including the time spent on the topic as a proxy for the importance) that the participants afforded the topic, and our assessment of the theoretical relevance of the topic in providing novel or nuanced understanding of the underlying assumption we were attempting to evaluate.

The findings from this study have been presented from the perspective of the authors as applicants to the innovative digital health funding program (ie, HIPPP) embedded within the larger digital health ecosystem of innovators or entrepreneurs, health system leaders, support partners, and funders.

Table 1. Summary of the study participants by ecosystem role.

ID	Ecosystem role	Size	Scope	Phase
P02	Innovator or entrepreneur	Medium	International	Established
P03	Innovator or entrepreneur	Micro	AB ^a	Startup
P04	Innovator or entrepreneur	Micro	AB	Startup
P08	Innovator or entrepreneur	Small	National	Startup
P09	Innovator or entrepreneur	Small	National	Established
P11	Innovator or entrepreneur	Medium	International	Established
P14	Innovator or entrepreneur	Micro	AB	Startup
P15	Innovator or entrepreneur	Micro	AB	Startup
P18	Innovator or entrepreneur	Small	National	Startup
P21	Innovator or entrepreneur	Small	International	Startup
P31	Innovator or entrepreneur	Small	International	Startup
P24	Health system leader	Large	AB	Hospital or community
P25	Health system leader	Large	AB	Hospital
P26	Health system leader	Large	AB	Hospital
P28	Health system leader	Large	NS ^b	Hospital or community
P33	Health system leader	Large	BC ^c	Hospital or community
P35	Health system leader	Large	ON ^d	Hospital or community
P36	Health system leader	Large	NL ^e	Hospital or community
P01	Support partners	Small	National	All
P05	Support partners	Small	Alberta	All
P06	Support partners	Large	International	All
P07	Support partners	Small	International	Startup
P10	Support partners	Small	AB	Startup
P12	Support partners	Micro	AB	Startup
P13	Support partners	Large	AB	Established
P19	Support partners	Small	AB	Startup or rapid growth
P22	Support partners	Small	AB	All
P23	Support partners	Small	AB	All
P29	Support partners	Small	AB	Startup or rapid growth
P30	Support partners	Small	AB	Startup or rapid growth
P32	Support partners	Small	AB	Startup or rapid growth
P34	Support partners	Small	National	Startup or rapid growth
P16	Funders	Small	National	Startup or rapid growth
P17	Funders	Small	AB	Startup or rapid growth
P20	Funders	Small	AB	Rapid growth
P27	Funders	Large	National	Startup or rapid growth

^aAB: Alberta.^bNS: Nova Scotia.^cBC: British Columbia.^dON: Ontario.^eNL: Newfoundland and Labrador.

Ethical Considerations

This study involved interaction with human participants and was approved by the University of Calgary's Conjoint Faculties Research Ethics Board under the ethics ID REB21-0242. All the extracted data were anonymized before being analyzed. The participants were not compensated, and their participation was entirely voluntary.

Results

Overview

The findings have been organized around 5 propositions developed from the broader literature on the navigation of ecosystems by entrepreneurs, within the context of digital health. These propositions have been presented as aggregate dimensions supported by first-order concepts and second-order themes (Figure 2) and summarized using a Gioia data structure diagram [20].

Table 2 provides a summary of the level of support for the findings described in detail in the subsequent section.

Figure 2. Data structure.

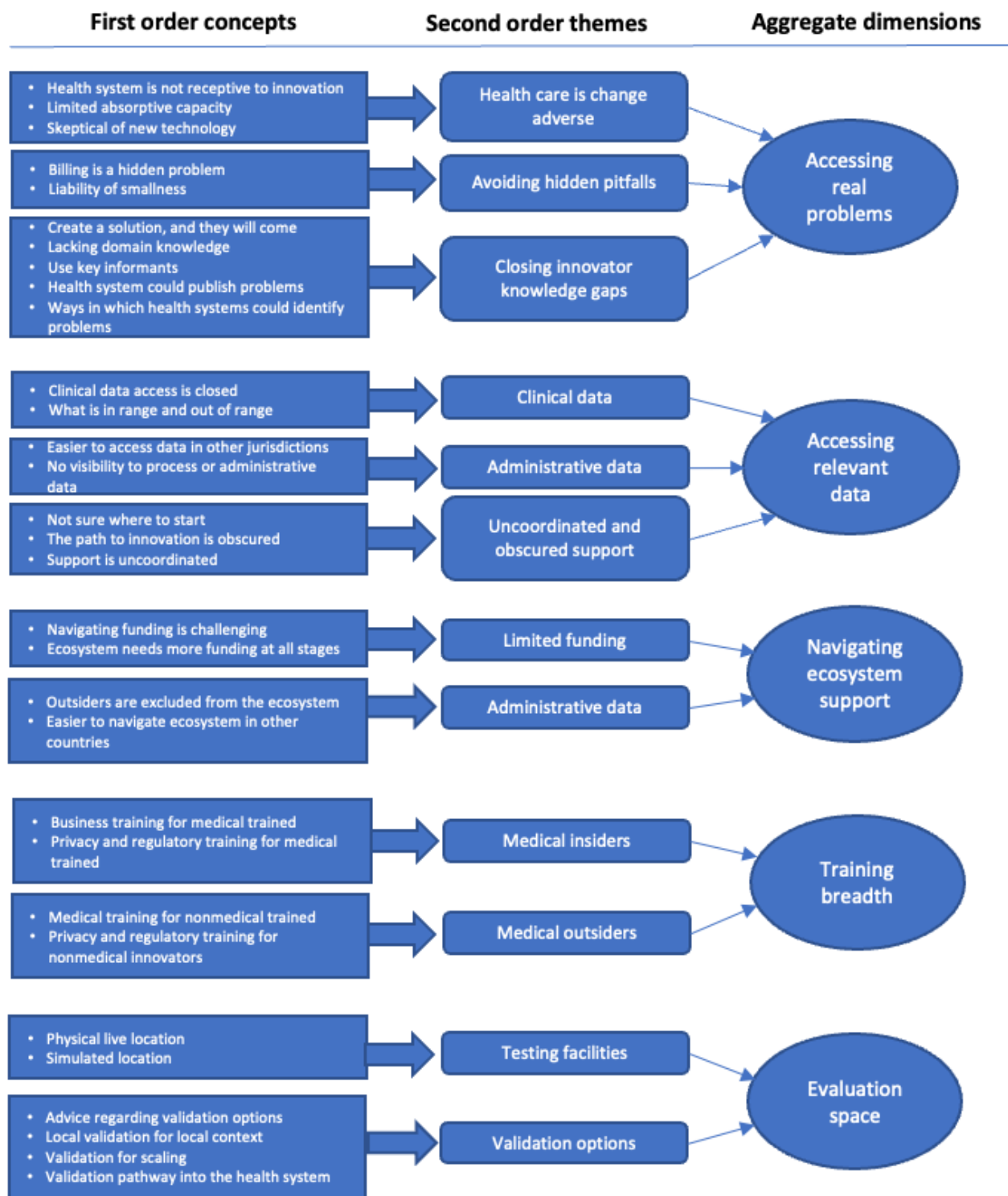


Table 2. Summary of the level of support for the propositions.

Proposition	Level of support
Proposition 1: digital health innovators and entrepreneurs struggle to access real problems	Moderate
Proposition 2: digital health innovators and entrepreneurs struggle to access relevant data	Moderate
Proposition 3.: digital health innovators and entrepreneurs struggle to understand where to start and where they need to go on their innovation journey	Strong
Proposition 4: the breadth of training provided does not adequately meet the needs of digital health innovators and entrepreneurs	Moderate
Proposition 5: digital health innovators and entrepreneurs lack a “space” to evaluate their offerings	Moderate

Proposition 1: Digital Health Innovators and Entrepreneurs Struggle to Access Real Problems

We anticipated that there would be a discrepancy between what digital health innovators or entrepreneurs and health system leaders perceive as problems. This disconnect would, in turn, lead to lost time and wasted resources for health innovators or entrepreneurs as they try to identify problems to solve.

Health System Not Receptive to Innovation

A total of 55% (6/11) of innovators or entrepreneurs expressed the theme that the health system is not receptive to innovation. They cited aspects such as not hearing back from representatives within the health system and trying and failing to work with the health system for years and that it is easier to get products into the private areas of the health care system. Not only is the prevailing sentiment that the health system is not receptive to innovation, but comments also showcased innovators or entrepreneurs avoiding the public health system because of this lack of receptivity. One of the innovators or entrepreneurs noted the following:

...we have by design decided to focus our efforts on the private sector, and not the public sector. [P31, innovator or entrepreneur]

Limited Absorptive Capacity

A total of 50% (7/14) support partners echoed innovators or entrepreneurs' belief that the health system is not receptive to innovation, commenting that it is not designed to understand its own challenges and adopts technology slowly. They noted that the health system's role is to provide health care services, not to test new technologies, and suggested some explanations for why the health system is resistant to innovation. Examples of barriers to innovation include the protection of patient privacy and the fact that health systems in Canada are generally focused on multinational companies, encouraging the promotion of a procurement model of technology adoption:

...when you're starting out and looking to see if something can work, you don't need to apply a procurement model to it. And unfortunately, that's what we do in Alberta. [P22, support partner]

One of the health system leaders agreed, emphasizing that health systems currently do not have the ability to trial new innovations and do not have the incentives to move away from the current vetted technologies:

We have these products that we feel are our ideal state all the time, and that to detract or maybe push

away from innovation or new concepts or new design. Because if we have something that's vetted and supported then why would we change it even though something might be better? [P24, health system leader]

Skeptical of New Technology

A total of 29% (2/7) of health system leaders raised this theme, saying that health care workers can be skeptical of new innovations. This is broadly because they can end up being responsible for teaching patients how to use the software, leading to the belief that digital health technology will ultimately make their lives more difficult. Therefore, it is paramount to maintain engagement with clinicians throughout the innovation process to gain their support:

I think the problem that every digital program is up against is it has a bad reputation, right? To the medical staff. It's a no brainer that it's better for accuracy and decreasing med errors and that kind of thing. But the reputation it has is this is going to make my life harder, basically. It's going to...I don't get paid for it. I'm going to spend more hours getting things done. [P33, health system leader]

Billing Is a Hidden Problem

Innovators or entrepreneurs expressed the theme that it is difficult to adopt innovation within the health system because of technological barriers, including the use of outdated technology (ie, older operating systems and internet browsers) and the difficulty in integrating novel products into the billing infrastructure:

How do we get it in the doctor's hands and how do they get compensated for using that tool as part of it? Because we're operating outside of the norm. Normally, we schedule an appointment, you go in and see your doctor. She spends 30 minutes with us and then codes it to this bill code. So, if I now text her my longitudinal data and she spends 20 minutes assessing that in an application, can she still bill for that? How does that happen? [P18, innovator or entrepreneur]

Liability of Smallness

In addition, 35% (5/14) of support partners identified the challenges with product integration. These include gaps in knowledge related to the existing solutions between frontline and operations staff in the health system, a lack of bandwidth or capacity to bring innovations into the health system, and the

difficulty in getting innovation into health systems without a strong company or clinical reputation. One of the support partners noted the following:

...if you're a new software startup in Calgary, as an example, [large healthcare organization] probably isn't going to give you the time of day until you have some sort of reputation. [P06, support partner]

Health system leaders also explained that modernizing the health system is a top priority and that there is a mismatch between system operations and what is currently unfolding within the research literature. One of the health system leaders expressed the following:

...sounds so trivial when I say it. The problem that we have been solving and are still solving is just becoming part of the 21st century! [P35, health system leader]

Create a Solution and They Will Come

In support of the first proposition, 22% (4/18) of funders and support partners, indicated that innovators or entrepreneurs from outside the health care space can try to solve pain points that do not exist or to develop a solution before the problem is clearly defined. One of the support partners described this in the following manner:

...the tendency is to come up with a mouse trap and then look around to see if anyone wants it. The tendency is solution first. [P19, support partner]

One of the funders shared their perspective on this:

So I would say for, so with a lot of the companies that we see coming to us from the health sector, a lot of them have novel technology, but aren't necessarily sure how to commercialize or where the best, business opportunity is for them. [P16, funder]

Lacking Domain Knowledge

A total of 10% (2/21) of health system leaders and support partners, commented that innovators or entrepreneurs without a health care background can struggle to navigate the health system, especially with regard to procurement:

If it is a truly legitimate problem within the health system, really understanding it, if you're not within the health system can be challenging. Navigating the procurement. [P32, support partner]

There's no point in a small startup working in a lab, and we do see this all the time and I'm sure you've experienced it, they're in their labs with their techie people and then they bring you this thing and you go, "This is a hospital. No, that wouldn't work here." They tend to come from...They're passionate around technology and not from that hospital or healthcare operational understanding. I think if more of those companies had that sort of engagement and advice early on, they'd be creating better products too. [P35, health system leader]

Use Key Informants

However, 45% (5/11) of innovators or entrepreneurs stated that they seek the help of a clinical adviser or have personal access to a clinician to help identify problems or evaluate the usefulness of innovations.

Health System Could Publish Problems

A total of 18% (2/11) of innovators or entrepreneurs and 14% (1/7) of health system leaders suggested that health systems could aid innovation by publishing their needs using a portal or website. One of the innovators or entrepreneurs mentioned that this method had already worked successfully for them in another country:

In a hospital in Berlin, because we did some testing with them, and they effectively published...I don't know if it's a hospital, or maybe it's some medical association, but they published their most critical use cases. Or maybe areas where they need technology to help. I know there is in Berlin. Anybody has access to it, and they get an idea...And they even have a contact, that if you want to understand a little bit more you can talk to somebody. [P02, innovator or entrepreneur]

Ways Health System Could Identify Problems

Moreover, 9% (1/11) of innovators or entrepreneurs and 7% (1/14) of support partners suggested 2 ways in which the health system could better identify problems: surveying clinicians to identify their needs and documenting clinical workflows so that innovators or entrepreneurs can see the current process and imagine an improvement:

If the technical as well as the clinical leaders had a list of things that they're wasting time on or things that they would love to embrace. If some of that was documented, I feel that would be golden for some of us in the industry in order to really feed off of them. [P11, innovator or entrepreneur]

Proposition 2: Digital Health Innovators and Entrepreneurs Struggle to Access Relevant Data

We anticipated that digital health innovators or entrepreneurs would generally be unaware of where to get access to the data they need to validate and test their innovations and would be uncertain how to access and manage the data when they did locate them.

Clinical Data Access Is Closed

Supporting the second proposition, 45% (5/11) of innovators or entrepreneurs mentioned that they faced challenges accessing the data needed to support their innovation journey. One of the innovators or entrepreneurs stated the following:

In healthcare, you find lots of challenges. The first challenge is you don't get any access anywhere. No data access, you have nothing. [P03, innovator or entrepreneur]

Easier to Access Data in Other Jurisdictions

Support partners and health system leaders echoed the data needs of innovators or entrepreneurs. They commented that, in general, health systems could improve access to their data for innovators or entrepreneurs, as has been done in other regions.

Other jurisdictions have anonymized data in a way that is sufficient enough to provide access to patient data. Why can't we do it? Why can't we just follow one of the models that have already been established as successful in other jurisdictions? [P32, support partner]

No Visibility to Process or Administrative Data

Most commonly, innovators or entrepreneurs said that they would benefit from more data to assess the market for their products. They mentioned that these data may be used to assess details such as the prevalence of clinic visits and the necessary information that would be valuable to clinicians when making decisions. A total of 11% (4/36) of participants noted that these additional data would allow them to better understand the potential growth of their innovations. One of the innovators or entrepreneurs said the following:

I think having access to administrative hospital data, so you can link your very specific data collection with general hospitalization data, because I think it would speak to the ability to expand. [P15, innovator or entrepreneur]

What Is in Range and Out of Range

In addition, innovators or entrepreneurs commented that they would benefit from knowing what metrics the health system requires. For example, an innovator or entrepreneur highlighted the difficulty in knowing what the normal versus critical biometrics are as well as how accurate measurements should be:

...we have to find out to understand if the condition is critical or not? How accurate it should be, specifically during those critical events? [P02, innovator or entrepreneur]

Although innovators or entrepreneurs typically had more access to data than anticipated, they still expressed a need for additional data access to better market and scale their products. Furthermore, the consensus among support partners and health system leaders was that the innovation ecosystem would be enhanced by allowing innovators or entrepreneurs more access to health system data, not simply clinical data but also administrative and process data.

Proposition 3: Digital Health Innovators and Entrepreneurs Struggle to Understand Where to Start and Where They Need to Go on Their Innovation Journey

We anticipated that digital health innovators or entrepreneurs would face challenges in knowing where to start and how to get help with their innovations and that this challenge would be more pronounced for innovators or entrepreneurs without experience in the health system.

Not Sure Where to Start

Overall, we found strong support for our proposition that innovators or entrepreneurs struggle to navigate the digital health ecosystem. A total of 47% (17/36) of participants (innovators or entrepreneurs, support partners, health system leaders, and funders) referenced the theme that innovators or entrepreneurs require more guidance through the digital health ecosystem. From the innovator or entrepreneur perspective, 55% (6/11) of participants expressed that they were unsure where they could get assistance and that they would have benefited from help earlier on in their innovation journey:

We could have saved ourselves a lot of time, money, and effort if someone had said, 'Okay, yeah, you guys are here, but maybe you should be doing this,' And not someone who is advising us for their own gains. [P08, innovator or entrepreneur]

"Outsiders" (Nonhealth) Are Excluded From the Ecosystem

Echoing the comments from innovators or entrepreneurs, 50% (7/14) of support partners presented the idea that the ecosystem is difficult for those starting from outside the health care system to understand, lacks a clear path forward for innovators or entrepreneurs, and has overlaps between support providers that force innovators or entrepreneurs to decide between multiple options:

They receive calls for expressions of interest from 10 different organizations and then kinda go okay well I can't work with all. I don't have the capacity cause I'm a start up. This one sounds interesting but I don't know how it connects to the rest of them and I don't know that I have an overall navigator saying we should go through this first and this and this. Then they get the result from whatever platform they work with, and they don't know what the next step might be. [P01, support partner]

The Path to Innovation Is Obscured

These issues are apparent to health system leaders, with 43% (3/7) of them saying that innovators or entrepreneurs need better guidance and that the process of getting innovation into the health system could be more transparent:

I think having a clear process, which we're trying to move to, and like [my colleague] said, it's taken us 20 years and it's not perfect. We're trying to get there. It helps the innovators too, because I think it should help the innovators and those companies and things because they can see the process. They can also see if a decision maybe hasn't gone the way that they wanted, who did it? How to appeal or discuss. [P25, health system leader]

Support Is Uncoordinated

In addition to the navigation challenges that innovators or entrepreneurs face, the digital health support ecosystem currently relies on ad hoc interactions between support partners and could be better connected. A total of 44% (11/25) of participants—a mix of funders, health system leaders, and support

partners—indicated that the cooperation between ecosystem support partners was informal. Their comments centered on the perception that the support ecosystem requires more collaboration. This is because of the reliance on unstructured referrals between support partners, who make recommendations based on personal knowledge that can be lost when these individuals leave the support organization:

I hold a lot of relationships already and let alone two years from now. And it's kind of the story of these types of roles is that, those relationships basically have to start again from scratch because someone leaves their role. [P10, support partner]

Although the support partner ecosystem remains disconnected, the ecosystem members are aware of this problem and want to improve collaboration. Overall, 33% (6/18) of participants, including both funders and support partners, made the point that the ecosystem needs more collaboration, with 3 support partners suggesting not only that there needs to be better handoffs between support partners but also that they are, in fact, eager to collaborate more:

I think that my personal view is that we have major holes in this game here. I think that we have many, many different organizations that need to work a lot better together and have you know better handoffs and better synergy between the service offerings. [P01, support partner]

Navigating Funding Is Challenging

One of the funders pointed out that it is not clear which organization provides funding at which stage. Another funder said that venture capitalists do not have the technical ability to evaluate start-ups that are subject to regulations. However, investors are attracted to the shorter timelines required to commercialize digital health products. One of the support partners noted that innovators need to be aware of the due diligence requirements that funders will require and that innovators are often unaware of these and do not adequately think about them in advance.

Ecosystem Needs More Funding at All Stages

A total of 27% (3/11) of innovators or entrepreneurs, 14% (2/14) of support partners, and 25% (1/4) of funders raised the theme that early-stage funding is lacking in the digital health ecosystem. Innovators or entrepreneurs mentioned that additional early-stage funding could be used to help them develop business plans and improve growth. Support partners mentioned that early-stage funding is lacking overall and that innovators or entrepreneurs are often unsure of what the exact requirements are for funding during the early stages. One of the funders emphasized the notable gap that Alberta does not have any incubators or accelerators with attached funding:

...we don't have any incubators or accelerators that have funding attached to them, which is typical in Ontario, Quebec and British Columbia. [P20, funder]

Moreover, 8% (2/25) of innovators or entrepreneurs and support partners described the theme that the ecosystem is lacking funding at all stages. From the innovators or entrepreneurs'

perspective, raising funds is difficult, often requiring help from friends or family, and there need to be better connections with angel investors or venture capital firms. This support partner mentioned that the most common request they receive is to assist in finding funders or grants:

...raising money is very, very difficult. You have to have an investor community that's willing to take a chance. And for us, it was friends, families, some local angels. [P21, innovator or entrepreneur]

Easier to Navigate Ecosystem in Other Countries

In addition, 50% (2/4) of support partners and 9% (1/11) of innovators or entrepreneurs expressed that it is difficult to get products sold in Canada, relative to other countries, owing to regulations and a lack of capital:

...it's like a cultural risk aversion for companies to try out new solutions or solutions from younger, smaller companies. So they tend to get their first customers outside of Alberta and outside of Canada. [P19, support partner]

Overall, navigating the digital health ecosystem in Alberta is challenging for innovators or entrepreneurs. However, this need is recognized by ecosystem players who have the desire for better collaboration between support partners, including more transparency around the available funding.

Proposition 4: The Breadth of Training Provided Does Not Adequately Meet the Needs of Digital Health Innovators and Entrepreneurs

We anticipated that digital health innovators or entrepreneurs would need a variety of training specifically tailored to the unique needs of the digital health industry.

Business Training

A total of 53% (19/36) of participants referenced areas where innovators or entrepreneurs could use additional training in the digital health ecosystem. The most prevalent area of need was business training, with 10 references from innovators or entrepreneurs, support partners, and funders to areas of need such as general business training, marketing, and pitching ideas to investors. Business training was also an area of need for clinical innovators or entrepreneurs, with 36% (4/11) of participants referencing their need for training to better understand app development or get their product to market:

We both come from the researcher side, so we have that squared away, but just trying to integrate that, and our research ideas into the actual business and money side of it, getting the product out there into the market. [P15, innovator or entrepreneur]

The other piece that we don't have much training in and bluntly, it is something that I don't really blame the clinical group for not knowing, but they don't understand what the operation side is. For all they have to submit, paperwork or somebody signs off on their budget, they don't really understand the black box know of the inner workings. And as somebody who's been learning that on the go over the past little

while, it is not complicated, but if you've never had it opened up, you don't understand the rules and things that supply chain management works under. [P25, health system leader]

Privacy Training

Although business training was a prominent focus of innovator or entrepreneur recommendations, regulatory or privacy training was identified as a need by 14% (2/14) of support partners and 25% (1/4) of funders, and training in health economics was identified as a need by 14% (1/7) of health system leaders and 7% (1/14) of support partners. Innovators or entrepreneurs' training needs are ubiquitous and diverse, highlighting the need for a customized referral process and improved collaboration between the existing training providers. A total of 29% (2/7) of health system leaders and 14% (2/14) of support partners mentioned that privacy regulations are a barrier to innovation because they restrict access to data and that privacy regulations can be surprising to innovators or entrepreneurs who are unfamiliar with the requirements:

I think that people are becoming more savvy to the issue of privacy and healthcare information sensitivity. But that is still something that surprises some people in terms of what kinds of safeguards they need to have in place, and why they can't just be the direct link between a doctor's office, who may have results, and the individual. [P06, support partner]

Regulatory Training

However, 29% (2/7) of health system leaders mentioned that while privacy regulations are a challenge to innovation, the public wants more access to their data and is less concerned about privacy. Innovation could be improved by allowing more access to patient data, which is something that the general public is becoming more comfortable with:

Every time we do these surveys across Canada, the public expects their data is going to be available across their healthcare providers and within their circle of care, yet we haven't facilitated that in any way. We've actually made it almost impossible. I think we need more opportunities for interoperability data exchange, whether it's through data sharing agreements or some reworking of health information custodians. [P35, health system leader]

Proposition 5: Digital Health Innovators and Entrepreneurs Lack a "Space" to Evaluate Their Offerings

We hypothesized that digital health innovators or entrepreneurs lacked locations to evaluate and validate their digital health products. A total of 82% (9/11) of innovators or entrepreneurs and 14% (2/14) of support partners, identified themes related to the validation needs of innovators or entrepreneurs.

Physical Living Laboratory Location

Overall, 36% (4/11) of innovators or entrepreneurs mentioned the need for a designated place to validate that a product works

and is safe to use. This validation involves testing the accuracy, reliability, and safety of a device:

You need that place where you take something, you build a functional unit, but then you need to really vet it. Does it really work? Does it really do what you intend it to do? And did you understand the requirements correctly. [P15, innovator or entrepreneur]

Simulated Location

A distinction was also made between a physical space where innovations and products could be evaluated in a real-world "living laboratory" setting and a simulated environment. Several support partners noted the following:

I know with [large health organization], they're also working on this big synthetic data initiative. I don't know if you've heard of that, but that's certainly an opportunity. It's a new thing, but it seems quite promising and it seems like it actually generates the kind of data that a machine learning algorithm would actually be able to be trained on. That being said, if your machine learning algorithm requires actual images, diagnostic images, it's a little bit different than just lab results. But nevertheless, it's hard to create synthetic data around those kinds of complicated bits of data. But certainly you can start somewhere, get the ball rolling, and then as people become more comfortable with it, start prying open those doors. [P32, support partner]

I think simulations have a big role to play here and is something that you could easily build in a very confidential manner of, here's a simulated, it's not a real. And whether you do that through low tech or high tech, you could do it with actors in a room. 'This is how I would talk to a patient. This is why I couldn't ask them. This is that.' But you could also do it as virtual simulations, where you're basically recreating a simulated environment. 'This is what it's like at the average maternity ward in Alberta.' Or maternity ward's over 20 beds say. 'When we talk about this being a problem, here's a picture of the thing that is a problem.' And you can immediately say why. [P34, support partner]

Advice Regarding Options

A total of 27% (3/11) of innovators or entrepreneurs and 7% (1/14) of support partners identified the need for advice regarding validation possibilities and potential next steps. This need involves identifying the type of validation an innovator or entrepreneur might need. For example, the level of rigor of a clinical trial is compared with that of the validation of a minimum viable product or initial prototype used to economically validate key business and clinical assumptions:

Knowing the different options, because obviously a more rigorous clinical trial would be more expensive. Sometimes mobile-held applicants don't require that compared to a medical device. So being able to know what the minimum threshold might be, would be useful

in terms of being efficient with funding. [P14, innovator or entrepreneur]

Local Validation for Local Context

In addition, 27% (3/11) of innovators or entrepreneurs and 7% (1/14) of support partners discussed the theme of needing testing, either in general or locally, representing innovators or entrepreneurs' need for better access to local validation to help them demonstrate value of their interventions to the health system:

International validation isn't enough for them, right? They want to be able to see local proof of concept. They want to be able to see people who they know, who are familiar to them put a seal of faith on these technologies. [P31, innovator or entrepreneur]

Validation for Scaling

A total of 36% (4/11) of innovators or entrepreneurs expressed the need for more validation with clinicians or in clinics. Innovators or entrepreneurs spoke about the benefits of observing clinical workflows, validating the use case with clinicians, and having access to clinicians who can give advice to improve the product:

Definitely, can we call them subject matter experts? So, it's people who really understand field. Because you can read a lot of papers, but if you still don't know what it all means, and it's hard to learn everything, you will still end up building a product that doesn't have good applicability. [P02, innovator or entrepreneur]

Furthermore, 18% (2/11) of innovators or entrepreneurs communicated the need to validate their product with the end users to verify usability or adherence. Innovators or entrepreneurs need to ensure that their products are something that a clinician or a member of the general public would feel comfortable using:

...the challenge I'm going to have on the consumer side is consumer perception about what we're doing. Are they going to use the product? [P18, innovator or entrepreneur]

Validation Pathway Into the Health System

Overall, 14% (2/14) of support partners mentioned that it is difficult for innovators or entrepreneurs to validate their products within the health system because of its size and the lack of a formalized mechanism to do so:

...a more formalized mechanism to participate in those validation opportunities with healthcare providers and healthcare organizations [is needed]. [P05, support partner]

Discussion

Principal Findings

Our exploration of the systemic conditions within a digital health ecosystem for key stakeholders, including digital health innovators or entrepreneurs, investors, and health system leaders,

illustrates the role of these conditions in predominantly determining the success of the ecosystem. Significant gaps were identified in each component of the systemic conditions within the digital health ecosystem of Alberta. However, the most foundational gap appears to be in the context of navigating the ecosystem in the absence of intermediaries tasked with providing guidance to digital health entrepreneurs around the available support services and dependencies among the various ecosystem actors and programs offered. These findings provide support for the underlying premise of the HIPP program, which motivated this study, to incentivize greater coordination among ecosystem support providers.

Comparison With Prior Work

Our findings add to the emerging literature on ecosystems [3], specifically within a digital health context [6]. The set of propositions guiding this study was developed from the literature; however, our findings provide additional, often nuanced, insights into each of these themes. For accessing real problems, we support a prior work that identifies this issue, specifically within the Alberta context [7]. The Alberta SCNs created in 2012 have a mandate for consolidating the real problems they are facing and ranking them for use by the ecosystem. However, as the findings from this study show, a decade into that mandate, there remain concerns over the transparency of those problems, as they do not appear to be well known or understood within the digital health entrepreneurial community. The SCNs are by their very nature not consumer focused, even if they continue to deliver on increasing the accessibility and transparency of the real problems they face, leaving a significant gap on the consumer and prevention side of the identification of real problems. In many respects, our findings parallel, at least at the leadership and operational levels, the challenges identified in Ontario [6].

Within the context of information, we provide further evidence of the challenges in accessing clinical and administrative data [12]. In Alberta, great strides have been made with respect to administrative data access, and initiatives at both regional and provincial levels have streamlined this access. However, transparency of these processes remains a challenge for many innovators or entrepreneurs. From the perception of our participants, it is often more dependent on the informal process of finding a contact that takes an interest in your project to help you navigate the administrative burdens of discovering what data are available and how to request access to them.

This notion of challenges around navigation was a recurring theme across every proposition. This issue was particularly acute for funding opportunities within the digital health ecosystem of Alberta. An issue that was highlighted in this study is the unique challenges that innovators or entrepreneurs from outside health care face when transitioning into the health care industry. This is particularly disconcerting in Alberta, where health care is presented as a destination industry for individuals in sectors that are in decline to repurpose their talents and experience. This connects to the broader discussion of certain groups being excluded from a digital health ecosystem because of a lack of necessary infrastructure, social disadvantage,

economic factors, health status, lack of skills or interest, or inadequate recognition of their needs [11].

Although policy makers have recognized the lack of coordination within ecosystems and have attempted to provide concierge services, one-stop points of contact, and centralized triage for innovators or entrepreneurs, these have not generally been successful. However, the findings from this study and our experience with the HIPP program provide additional insights into the design criteria for such navigation services in the future. Specifically, it is nearly impossible to centralize such services because no organization or government holds a sufficient mandate or resources to do so. For example, a program spearheaded at the federal level could easily be undermined by a lack of provincial or municipal referrals. The HIPP approach provided a key insight in that the incentives were structured such that the participants were rewarded for not only delivering the services they promised but also contributing to benefits that accrued to the ecosystem overall. Thus, balancing group and individual incentives is essential for successful coordination, as the group or ecosystem-level incentives are generally lacking. These incentives need to be flexible to allow for the broad inclusion of stakeholders, and they also need to incentivize a reduction in duplication. Currently, all incentives are for support providers to deliver as many services as possible without regard for how well they can deliver those services, whether those services integrate with the services of other providers, or whether such services are already better deployed by other ecosystem stakeholders. For organic ecosystem coordination to flourish, there needs to be equal incentives to give up or better integrate services as there is to create new services. Finally, a common approach to ecosystem orchestration is the use of web-based platforms that attempt to document all the services offered by the ecosystem combined with navigation tools to help innovators or entrepreneurs self-serve at least a portion of their search and selection processes. There are prominent examples of these at the provincial (eg, Alberta Innovates' Support Finder) and federal (eg, Government of Canada's Business Benefits Finder) levels, where such data aggregation is attempted, but they usually suffer the same fate after the initial funding runs out—the data quickly become stale, and the usefulness rapidly declines, or they only cover parts of the ecosystem that are relevant to the sponsor's mandate. A successful approach would require a governance structure in which such platforms are not owned by a specific ecosystem stakeholder; instead, the platform should be owned by a consortium of ecosystem players. This greatly aligns the incentives and ability to coordinate activities in a structured manner and keep data current. Such 2-sided markets are still challenging to maintain, but incentivizing the supply side (ie, support providers) to cooperate quickly leads to reasons for the demand side (ie, innovators or entrepreneurs) to show up, which, in turn, reinforces the supply side.

On the training side, one approach to address this gap is to provide differentiated training options for medical insiders versus outsiders. Medical insiders generally needed business and privacy or regulatory training, while medical outsiders needed additional training on the medical system, billing, and procurement. Interestingly, being an insider did not always accompany an understanding of some of these topics, so some level of baselining is needed to assess digital health innovators or entrepreneurs' readiness.

Finally, there is a need for evidence supporting the claims that digital health offerings provide real improvements over the status quo [16]. Of particular interest to W21C, as a support provider located within an AMC, are the future roles of AMCs in supporting digital health ecosystems. AMCs provide cross-disciplinary expertise in health and technology and train the next generation of health professionals, thus providing the opportunity to connect academics and clinicians across a variety of disciplines with innovators and entrepreneurs [15]. Demand for both living laboratories and simulated environments is strong among digital health innovators or entrepreneurs. The ability to evaluate their processes, algorithms, devices, and software in an environment that provides systematic feedback is essential [14]. This issue was particularly acute for digital health innovators or entrepreneurs, as one of the advantages of digital health innovations is the rapid rate at which they can be scaled. This is in sharp contrast to many health care innovations such as drugs, biomedical interventions, and medical devices, which are subject to substantially more regulatory requirements [12].

Limitations

The findings of this study are based on the experience of the digital health ecosystem in Alberta, Canada, which may not be generalizable to other contexts. However, as discussed in the comparison with prior work, there are common themes that appear to transcend jurisdictions. Although the research team selected ecosystem actors who broadly represent several factors of theoretical importance, there is a possibility that key stakeholder perspectives were omitted.

Conclusions

Navigating the systemic conditions of the digital health ecosystem is extremely challenging for innovators or entrepreneurs without prior health care experience, and this remains an issue even for those with such experience. Policy interventions aimed at increasing collaboration among ecosystem support providers, along with tools and incentives to ensure coordination, are essential as ecosystems grow. By improving the systemic conditions highlighted in this study, the Alberta digital health ecosystem can increase its competitiveness and foster greater innovation, talent, opportunities, choices, and access to digital health care across the country [21].

Acknowledgments

The authors gratefully acknowledge the input of the study participants and the Ward of the 21st Century support team. This study was funded by an Alberta Innovates Health Innovation Platform Partnerships grant.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Health Innovation Platform Partnerships. Alberta Innovates. URL: <https://albertainnovates.ca/programs/health-innovation-platform-partnerships/> [accessed 2022-01-07]
2. Ward of the 21st Century Research and Innovation Centre. W21C. URL: <https://www.w21c.org/> [accessed 2022-01-07]
3. Stam E. Entrepreneurial ecosystems and regional policy: a sympathetic critique. *Eur Plan Stud* 2015 Jun 08;23(9):1759-1769. [doi: [10.1080/09654313.2015.1061484](https://doi.org/10.1080/09654313.2015.1061484)]
4. Stam E, Spiegel B. Entrepreneurial ecosystems. Utrecht School of Economics. 2016. URL: <https://econpapers.repec.org/paper/usetkiwps/1613.htm> [accessed 2022-01-07]
5. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851 [FREE Full text] [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
6. Desveaux L, Soobiah C, Bhatia RS, Shaw J. Identifying and overcoming policy-level barriers to the implementation of digital health innovation: qualitative study. *J Med Internet Res* 2019 Dec 20;21(12):e14994 [FREE Full text] [doi: [10.2196/14994](https://doi.org/10.2196/14994)] [Medline: [31859679](https://pubmed.ncbi.nlm.nih.gov/31859679/)]
7. Ambler KA, Leduc MA, Wickson P. Innovating to achieve service excellence in Alberta Health Services. *CMAJ* 2019 Dec 04;191(Suppl):S52-S53 [FREE Full text] [doi: [10.1503/cmaj.190598](https://doi.org/10.1503/cmaj.190598)] [Medline: [31801769](https://pubmed.ncbi.nlm.nih.gov/31801769/)]
8. Hermes S, Riasanow T, Clemons EK, Böhm M, Krcmar H. The digital transformation of the healthcare industry: exploring the rise of emerging platform ecosystems and their influence on the role of patients. *Bus Res* 2020 Sep 11;13(3):1033-1069. [doi: [10.1007/s40685-020-00125-x](https://doi.org/10.1007/s40685-020-00125-x)]
9. Greaves F, Joshi I, Campbell M, Roberts S, Patel N, Powell J. What is an appropriate level of evidence for a digital health intervention? *Lancet* 2019 Dec 22;392(10165):2665-2667. [doi: [10.1016/S0140-6736\(18\)33129-5](https://doi.org/10.1016/S0140-6736(18)33129-5)] [Medline: [30545779](https://pubmed.ncbi.nlm.nih.gov/30545779/)]
10. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. Geneva, Switzerland: World Health Organization; 2016. URL: <https://apps.who.int/iris/handle/10665/252183> [accessed 2022-01-07]
11. Lupton D. Digital health now and in the future: findings from a participatory design stakeholder workshop. *Digit Health* 2017 Nov 9;3:2055207617740018 [FREE Full text] [doi: [10.1177/2055207617740018](https://doi.org/10.1177/2055207617740018)] [Medline: [29942616](https://pubmed.ncbi.nlm.nih.gov/29942616/)]
12. Pagliari C. Digital health and primary care: past, pandemic and prospects. *J Glob Health* 2021 Jul 02;11:01005 [FREE Full text] [doi: [10.7189/jogh.11.01005](https://doi.org/10.7189/jogh.11.01005)] [Medline: [34221352](https://pubmed.ncbi.nlm.nih.gov/34221352/)]
13. Guidances with Digital Health Content. U.S. Food and Drug Administration. URL: <https://www.fda.gov/medical-devices/digital-health-center-excellence/guidances-digital-health-content> [accessed 2022-01-07]
14. Guidance Document: Software as a Medical Device (SaMD): Definition and Classification. Health Canada. 2019 Oct 3. URL: <https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/software-medical-device-guidance-document.html> [accessed 2022-01-07]
15. Desveaux L, Kelley LT, Bhatia RS, Jamieson T. Catalyzing digital health innovation in Ontario: the role of an academic medical centre. *Health Policy* 2020 Nov;16(2):55-68 [FREE Full text] [doi: [10.12927/hcpol.2020.26353](https://doi.org/10.12927/hcpol.2020.26353)] [Medline: [33337314](https://pubmed.ncbi.nlm.nih.gov/33337314/)]
16. Van Winkle B, Solad Y, Vaswani N, Rosner BI. Navigating the digital health ecosystem to bridge the gap from innovation to transformation: a NODE. *Health perspective on digital evidence. Digit Biomark* 2019 Jun 25;3(2):83-91 [FREE Full text] [doi: [10.1159/000500194](https://doi.org/10.1159/000500194)] [Medline: [32095768](https://pubmed.ncbi.nlm.nih.gov/32095768/)]
17. Curioso WH. Building capacity and training for digital health: challenges and opportunities in Latin America. *J Med Internet Res* 2019 Dec 18;21(12):e16513 [FREE Full text] [doi: [10.2196/16513](https://doi.org/10.2196/16513)] [Medline: [31850849](https://pubmed.ncbi.nlm.nih.gov/31850849/)]
18. Cresswell K, Sheikh A, Franklin BD, Krasuska M, The Nguyen H, Hinder S, et al. Interorganizational knowledge sharing to establish digital health learning ecosystems: qualitative evaluation of a national digital health transformation program in England. *J Med Internet Res* 2021 Aug 19;23(8):e23372 [FREE Full text] [doi: [10.2196/23372](https://doi.org/10.2196/23372)] [Medline: [34420927](https://pubmed.ncbi.nlm.nih.gov/34420927/)]
19. DePasse JW, Chen CE, Sawyer A, Jethwani K, Sim I. Academic Medical Centers as digital health catalysts. *Healthc (Amst)* 2014 Sep;2(3):173-176. [doi: [10.1016/j.hjdsi.2014.05.006](https://doi.org/10.1016/j.hjdsi.2014.05.006)] [Medline: [26250503](https://pubmed.ncbi.nlm.nih.gov/26250503/)]
20. Gioia DA, Corley KG, Hamilton AL. Seeking qualitative rigor in inductive research: notes on the Gioia methodology. *Organ Res Methods* 2012 Jul 24;16(1):15-31. [doi: [10.1177/1094428112452151](https://doi.org/10.1177/1094428112452151)]
21. Digital Health Care Market Study. Competition Bureau of Canada. 2020. URL: <https://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/04555.html> [accessed 2022-01-07]

Abbreviations

AMC: Academic Medical Center

HIPP: Health Innovation Platform Partnerships

SCN: Strategic Clinical Network

W21C: Ward of the 21st Century

Edited by A Mavragani; submitted 14.01.22; peer-reviewed by H Durrani, K Chigudu, DE Brown; comments to author 04.11.22; revised version received 21.11.22; accepted 28.11.22; published 21.12.22.

Please cite as:

Saunders C, Currie D, Virani S, De Groot J

Navigating the Systemic Conditions of a Digital Health Ecosystem in Alberta, Canada: Embedded Case Study

JMIR Form Res 2022;6(12):e36265

URL: <https://formative.jmir.org/2022/12/e36265>

doi: [10.2196/36265](https://doi.org/10.2196/36265)

PMID: [36542428](https://pubmed.ncbi.nlm.nih.gov/36542428/)

©Chad Saunders, Devon Currie, Shane Virani, Jill De Groot. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Comparing a Fitbit Wearable to an Electrocardiogram Gold Standard as a Measure of Heart Rate Under Psychological Stress: A Validation Study

Joel Gagnon¹, PhD; Michelle Khau², MSc; Léandre Lavoie-Hudon², BA; François Vachon¹, PhD; Vicky Drapeau³, PhD; Sébastien Tremblay¹, PhD

¹School of Psychology, Faculty of Social Sciences, Laval University, Québec, QC, Canada

²Faculty of Social Sciences, Laval University, Québec, QC, Canada

³Quebec Heart and Lung Institute Research Center, Department of Physical Education, Faculty of Educational Sciences, Centre Nutrition, santé et société (NUTRISS), Institute of Nutrition and Functional Foods (INAF), Laval University, Québec, QC, Canada

Corresponding Author:

Joel Gagnon, PhD

School of Psychology, Faculty of Social Sciences

Laval University

Pavillon Félix-Antoine-Savard, 1144

2325, rue des Bibliothèques

Québec, QC, G1V 0A6

Canada

Phone: 1 8193830645

Email: joel.gagnon.2@ulaval.ca

Abstract

Background: Wearable devices collect physiological and behavioral data that have the potential to identify individuals at risk of declining mental health and well-being. Past research has mainly focused on assessing the accuracy and the agreement of heart rate (HR) measurement of wearables under different physical exercise conditions. However, the capacity of wearables to sense physiological changes, assessed by increasing HR, caused by a stressful event has not been thoroughly studied.

Objective: This study followed 3 objectives: (1) to test the ability of a wearable device (Fitbit Versa 2) to sense an increase in HR upon induction of psychological stress in the laboratory; (2) to assess the accuracy of the wearable device to capture short-term HR variations caused by psychological stress compared to a gold-standard electrocardiogram (ECG) measure (Biopac); and (3) to quantify the degree of agreement between the wearable device and the gold-standard ECG measure across different experimental conditions.

Methods: Participants underwent the Trier Social Stress Test protocol, which consists of an oral phase, an arithmetic stress phase, an anticipation phase, and 2 relaxation phases (at the beginning and the end). During the stress protocol, the participants wore a Fitbit Versa 2 and were also connected to a Biopac. A mixed-effect modeling approach was used (1) to assess the effect of experimental conditions on HR, (2) to estimate several metrics of accuracy, and (3) to assess the agreement: the Bland-Altman limits of agreement (LoA), the concordance correlation coefficient, the coverage probability, the total deviation index, and the coefficient of an individual agreement. Mean absolute error and mean absolute percent error were calculated as accuracy indices.

Results: A total of 34 university students were recruited for this study (64% of participants were female with a mean age of 26.8 years, SD 8.3). Overall, the results showed significant HR variations across experimental phases. Post hoc tests revealed significant pairwise differences for all phases. Accuracy analyses revealed acceptable accuracy according to the analyzed metrics of accuracy for the Fitbit Versa 2 to capture the short-term variations in psychological stress levels. However, poor indices of agreement between the Fitbit Versa 2 and the Biopac were found.

Conclusions: Overall, the results support the use of the Fitbit Versa 2 to capture short-term stress variations. The Fitbit device showed acceptable levels of accuracy but poor agreement with an ECG gold standard. Greater inaccuracy and smaller agreement were found for stressful experimental conditions that induced a higher HR. Fitbit devices can be used in research to measure HR variations caused by stress, although they cannot replace an ECG instrument when precision is of utmost importance.

(JMIR Form Res 2022;6(12):e37885) doi:[10.2196/37885](https://doi.org/10.2196/37885)

KEYWORDS

Fitbit device; wearable; heart rate; measurement accuracy; criterion validity; interdevice agreement; psychological stress; stress; physiological; behavioral; mental health; well-being

Introduction

The health and well-being of the population are a growing concern for clinicians and researchers. The World Health Organization (WHO) reported that noncommunicable diseases (eg, heart disease, stroke, cancer, diabetes, and chronic lung disease) were the cause of 71% of deaths worldwide in 2017 [1]. In addition, the exceptional regulatory measures taken to fight the COVID-19 pandemic have had a negative impact on the psychological health of citizens as made evident by a surge in incidences of psychological crises [2]. The extraordinary context generated by the COVID-19 pandemic has posed unprecedented challenges for governments [3,4] and has led to a redefinition of human activities within our societies, namely an acceleration in the digital transformation of health care [5]. The limitations of the current human health management model, which operates primarily downstream (ie, once problems emerge), are increasingly apparent. Moreover, the prevalence of health problems in the world far exceeds the current capacity of professionals and health services [6]. With the ever-increasing presence of technology in human life, researchers have taken an interest in the use of consumer wearable devices, such as smartwatches, armbands, rings, and other accessories, designed to be worn all day, to collect real-time physiological (eg, cardiac activity, skin temperature) and behavioral measures (eg, frequency of physical activity, step count, and sleep patterns). Moreover, these measures can be used to identify individuals at risk of declining mental health and well-being. Importantly, this information may provide relevant insight into the early detection and prevention of disease and well-being deterioration [7,8].

Regarding the validity and reliability of wearable activity monitors, systematic reviews have shown that these devices are somewhat accurate and stable to estimate heart rate (HR) and step count in adults. However, they provide an unreliable estimate for energy expenditure under different activities [9-12]. First, as several studies have pointed out, the sensors used to detect HR in most wearables (including Fitbit) are more sensitive to motion-induced artifacts (signal interference) than electrocardiogram (ECG) technology [13-15]. Accordingly, insufficient pressure and sensor-skin contact, as well as too much pressure such that blood flow was constricted, can affect HR measures. Despite these shortcomings, wearable activity monitors can provide important insight into physiological patterns. Importantly, small longitudinal studies have found support for the use of wearables to measure stress levels among adults [16,17]. Sano et al [18] conducted a monthlong longitudinal study of 201 university students to evaluate the possibility to predict mental health and stress using data collected with Q-sensor and Motion Logger wearable devices. The results revealed electrodermal activity (ie, skin conductance and temperature) as a predictor of mental health and stress.

Additionally, several studies have revealed the variability in HR to be a valid indicator of stress and have applied this measure in the study of major depressive disorder, stress resilience, stress regulation, and recovery from mental and physical stress [19-22], although to our knowledge, only one study has examined the feasibility of using wearable activity monitors to measure HR as a direct indicator of stress [23]. However, a limitation of this study was the lack of comparison measures. As such, the validity of the relationship between wearable activity monitors measured HR and stress, as well as the potential applications, remain unclear. Moreover, to our knowledge, no laboratory study has assessed the relationship between HR and psychological stress using a wearable activity monitor. Thus, the aim of this study was threefold: (1) to test the ability of a wearable activity monitor, specifically a Fitbit device, to sense an increase in HR upon induction of psychological stress; (2) to assess the accuracy (ie, the closeness of the agreement between the result of a measurement and a true value of the thing being measured [24]) of a Fitbit device to capture a physiological change (increased HR) caused by psychological stress compared to a gold-standard ECG measure; and (3) to quantify the degree of agreement (ie, the degree of concordance or extent to which one measure can replace another) between the Fitbit device and a gold-standard ECG measure across different experimental conditions.

Methods**Recruitment**

Participants were recruited through the mailing list from the authors' university. Eligibility criteria were being between 18 and 65 years of age; being registered as a full-time student; having access to a smartphone; absence of current or past, non-BMI-related, pathology (somatic, psychiatric, or both); not taking painkillers, medications that affect the heart rhythm, or medication for major depression or other mood disorders; not being pregnant or breastfeeding; and understanding French (spoken and written).

Ethical Considerations

This study was approved by the Human Research Ethics Committee of Université Laval (2020-053/10-11-2020). French-speaking participants provided written informed consent in French prior to participation in the study. All study data were deidentified to protect the privacy and confidentiality of participants. Upon completion of the study, the participants received a CAD \$40 (US \$29.31) monetary compensation.

Participants

A total of 34 healthy university students were recruited for this study. Participant demographic data are presented in [Table 1](#). The study occurred during the winter 2021 and summer 2021 semesters.

Table 1. Participant demographic data (N=34).

Demographics	Study sample
Age (years), mean (SD)	26.8 (8.5)
Gender, n (%)	
Male	13 (35)
Female	21 (58)
N/A ^a	2 (5)
Ethnicity, n (%)	
Caucasian	23 (67)
African	4 (11)
Hispanic	3 (8)
N/A	1 (2)
Middle Eastern	2 (5)
Education level, n (%)	
Bachelor	14 (41)
Master	10 (26)
Doctoral	8 (23)
N/A	3 (8)
Program of study, n (%)	
Health Science	8 (23)
Science and Engineering	5 (14)
Languages	3 (8)
Arts and Humanities	3 (8)
Psychology	3 (8)
Social Science	2 (5)
Administration	2 (5)
Education	1 (2)
N/A	7 (20)

^aN/A: not applicable.

Fitbit Device as an HR Monitor

Fitbit (Fitbit Inc) is one of the most popular wearable activity monitors and the most frequently studied [9]. While Fitbit's market shares have diminished from its peak in 2018, as of 2019, the company remains in the top 5 wearable companies by shipment volume and market share [25]. Recent studies have investigated the capacity of Fitbit devices to measure HR under different exercise conditions. Benedetto et al [26] in their controlled assessment of the Fitbit Charge 2 accuracy in measuring HR found wide variability in precision during different intensities on a stationary bike. In another study, Thomson et al [27] compared the HR measurement of the Apple Watch and the Fitbit Charge HR 2 with an electrocardiogram (ECG) among healthy young adults across different treadmill exercise intensities. The results showed diminished accuracy with increased exercise intensities for all devices, while the Fitbit had comparably greater relative error rates (ranging from 4.91% for very light exercise to 13.04% for very vigorous

exercise) compared to the Apple Watch. Indeed, a general lack of accuracy at higher exercise intensities has been repeatedly reported in the literature for Fitbit devices [28-30]. Nevertheless, Fitbit data may still be useful for other purposes, such as providing a proxy measure of psychological stress, as investigated within the context of the present study.

Experimental Procedure

Students interested in the study were invited to the laboratory. Upon arrival, participants were asked to read and sign an informed consent document. Participants who consented to the study were then asked to complete a web-based self-reported questionnaire to gather sociodemographic information. Next, they were asked to install the wearable activity monitor's (Fitbit) mobile app on their cell phone. In order to assess Fitbit detection of stress-induced change in HR, the participants were then given a Fitbit Versa 2 to wear on their nondominant hand during the experiment. Participants were asked to sit in front of 3 cameras. Once seated, the research assistant installed 3 electrodes, located

on the left and right of the chest, as well as below the ribs on the left of the abdomen of the participant. Once the electrodes were placed, the research assistant started the physiological recordings and made a visual inspection of the signal to ensure that the electrodes made good contact with the skin. Afterward, the research assistant started the stress protocol (described below). At the end of the stress protocol, the research assistant debriefed the participant and explained the purpose of the protocol. This step was important to ensure that the participant would not experience anger toward the research assistant when leaving. The experimental procedure took on average 1 hour and 30 minutes to complete.

Stress Protocol

The Trier Social Stress Test (TSST) was used to induce psychological stress. The TSST is a standardized psychosocial stress test that has been extensively used by researchers worldwide [31]. The TSST has been recognized to be an especially successful way of triggering stress [32]. The TSST consists of a waiting period, stress period, and rest period which can be divided into 5 experimental phases: relaxation, anticipation, oral, arithmetic, and relaxation. During the waiting period (the relaxation phase), the participant was left alone in the room and was told to relax for 5 minutes. The stress period was divided into 3 parts. First, a 3-minute anticipatory stress period (anticipation phase) during which the participant was asked to prepare a speech about why they would be a good candidate for their dream job. Second, a 5-minute speech task (oral phase) during which they delivered their speech in front of the research assistant. The research assistant was instructed to prompt the participant to continuously talk for the entire 5 minutes. If the participant were to stop talking before the end of the condition, the research assistant would use verbal prompts to pressure the participant to continue talking. Third, immediately after the speech task, the participant was asked to verbally perform a 5-minute mental arithmetic task (arithmetic phase). For this task, the participant was required to continuously subtract 13 from the number 1687. If the participant made a mistake or hesitated for more than 3 seconds, the research assistant triggered a loud buzzer and instructed the participant to start again from the initial number. Following the arithmetic task, the participant was asked to relax for 5 minutes (the relaxation phase). During the TSST, the participant was filmed from 3 angles (front, 45° left, and 45° right) and was informed that these recordings would be analyzed by 2 language analysis experts.

Material and Measures

Fitbit Versa 2

The experimental device was the wrist-worn Fitbit Versa 2, Version 35.72.1.9 (Fitbit Inc). Fitbit HR data were retrieved from the Fitabase platform (Small Steps Labs) and then stored in a secure S3 bucket maintained by the authors' university for analysis.

Electrocardiography

Surface electrodes were used for ECG recordings using a Biopac MP150 acquisition system for physiological data acquisition. The electrodes recorded electrical impulses from the heart. Data

were sampled at 1000 Hz. The recording was performed by a NeuroScan system (NeuroScan Inc, SynAmps). Heartbeats per minute (bpm) were calculated as an indicator of physiological stress arousal.

Statistical Analysis

Descriptive statistics for HR measurements (mean and standard deviation) were calculated for the Fitbit Versa 2 and the Biopac across experimental phases (relaxation, anticipation, oral, and arithmetic). The relaxation phase at the beginning and the one at the end of the experimental protocol were merged to create a single relaxation phase for the analyses.

First, to assess the ability of the Fitbit Versa 2 to sense an increase in HR, differences across experimental phases were examined using a mixed-effect model that was constructed using the “lme4 [33]” package available in R (R Foundation for Statistical Computing). The model included the experimental phases as a fixed effect, whereas the participants and the interaction between participants and experimental phases were treated as random effects. For the overall model, a Satterthwaite adjustment was used to compute the degrees of freedom. Partial η^2 was computed using the “effectsize [34]” package available in R. Post hoc tests were conducted using the “emmeans [35]” package available in R with the Kenward-Roger method to compute the degrees of freedom, and the *P*-values were adjusted using the Tukey method.

Second, to assess the accuracy of the Fitbit Versa 2 measured HR compared to a gold-standard ECG measure, mean absolute error (MAE), and mean absolute percentage error (MAPE) between the Biopac and the Fitbit Versa 2 were calculated as overall measurement error. The clinically acceptable difference (CAD) was set as 10 bpm, such that differences in MAE of less than 10 bpm were regarded as clinically insignificant, thus showing good accuracy for the Fitbit Versa 2. This was based on the American National Standard of “Cardiac monitors, heart rate meters, and alarms” that permit “readout error of no greater than $\pm 10\%$ of the input rate or ± 5 bpm, whichever is greater [36].” A MAPE threshold of 10% was used to assess the accuracy of the Fitbit Versa 2 [37,38].

Third, to quantify the degree of agreement between the Fitbit Versa 2 and the Biopac across different experimental phases, 5 metrics of agreement were calculated. First, limits of agreement between the Biopac and the Fitbit Versa 2 were evaluated using a mixed-effect model to account for the effect of the participant, the experimental phases, and time (ie, repeated measures) as recommended by Parker et al [39]. Bias-corrected and accelerated bootstrapping with 5000 replications were used to estimate the 95% CI. The analysis was conducted using the “SimplyAgree [40]” package in R. Second, the concordance correlation coefficient (CCC) with 95% CI was estimated from a linear mixed model using the appropriate intraclass correlation coefficient [41]. The CCC indicates the proportion of the total variability accounted for (1) by the participant, (2) the experimental phase, and (3) their interaction. The CCC is a standardized coefficient taking values from 1 (perfect disagreement) to 1 (perfect agreement). In other words, a CCC of 1 indicates the absence of variability in the device across

participant and experimental phases [42]. In this study, the following guidelines [43] were used to interpret the CCC: <0.90 (poor), 0.90 to 0.95 (moderate), 0.95 to 0.99 (substantial), and >0.99 (perfect). The CCC was calculated using the “ccrm” package in R [44]. Third, the coverage probability index proposed by Lin et al [45] was estimated by calculating the probability that the between-device differences lie within the boundary of the predefined CAD. As such, a larger probability indicates closer agreement. A mixed-effect modeling approach was used to calculate the coverage probability index, which required the range of CADs and the mean square deviation. The mean squared deviation is the expected squared difference between readings by 2 different devices on the same individual performing the same activity at the same time. Similar to Parker et al [42], the mean squared deviation was obtained based on the mixed-effect model. Fourth, the total deviation index was estimated based on the mean squared deviation from the mixed-effect model. This index provides the boundary within which the differences between devices will be contained $p \times 100\%$ of the time. The predefined CAD of ± 10 bpm was used to interpret whether the interval signified agreement. An interval contained between the CAD would indicate that the 2 devices can be used interchangeably. Finally, the coefficient of individual agreement (CIA [46,47]) was calculated. The CIA is a scaled coefficient that quantifies the magnitude of variability between devices compared to the replication variability within devices. A CIA value of 1 indicates that using different devices

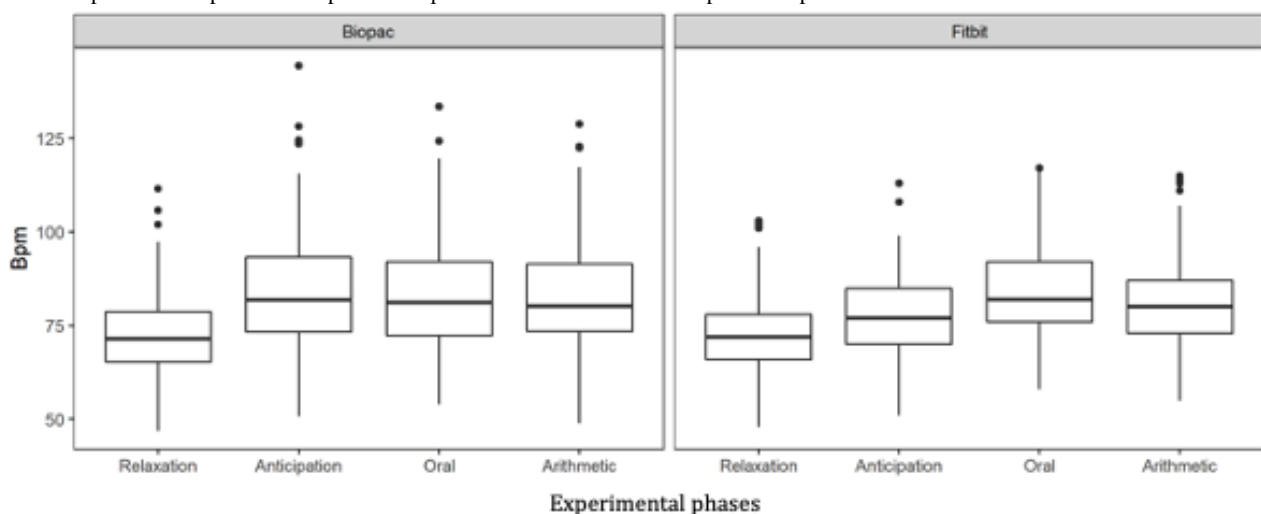
makes no difference to the variability of repeated measurements taken under the same conditions within the same subject. Following past studies [42,48,49], the CIA was calculated based on the mean squared deviation as the disagreement index. A CIA >0.80 is considered acceptable [46,48,49]. A 95% CI was calculated using a bootstrapping procedure with 5000 replications. All statistical analyses were conducted using R version 4.0.3 [50].

Results

HR Mean Differences Across Experimental Phases

Figure 1 displays boxplots of the bpm across experimental phases and between devices. The results from the mixed-effect model revealed significant variations from the Fitbit Versa 2 HR measurements among experimental phases ($F_{3,103}=44.03$; $P<.001$; $\eta^2_{\text{partial}}=0.56$, 90% CI 0.45-0.64). A post hoc Tukey test revealed that the mean HR from the relaxation phase was significantly lower than all other experimental phases (anticipation, oral, and arithmetic) at $P<.001$. Furthermore, the mean HR from the anticipation phase was significantly lower than the oral ($P<.001$) and arithmetic ($P=.02$) phases. Finally, the mean HR from the oral phase was significantly higher than the arithmetic phase ($P=.02$). Overall, the results revealed the capacity of the Fitbit Versa 2 to detect short-term variations in levels of psychological stress.

Figure 1. Boxplots of the bpm across experimental phases and between devices. bpm: beats per minute.



Accuracy

Analysis of the measurement error between the Biopac and the Fitbit Versa 2 showed an overall MAE of 5.87 (SD 6.57, 95% CI 3.57-8.16) bpm, which is below the predefined CAD of ± 10

bpm showing good accuracy of the Fitbit Versa 2. Moreover, the results revealed an overall MAPE of 7.24% below the predefined threshold of 10% for acceptable accuracy. Table 2 shows the MAE (and SD) and MAPE for each experimental phase.

Table 2. Accuracy of the Fitbit Versa 2 across experimental phases.

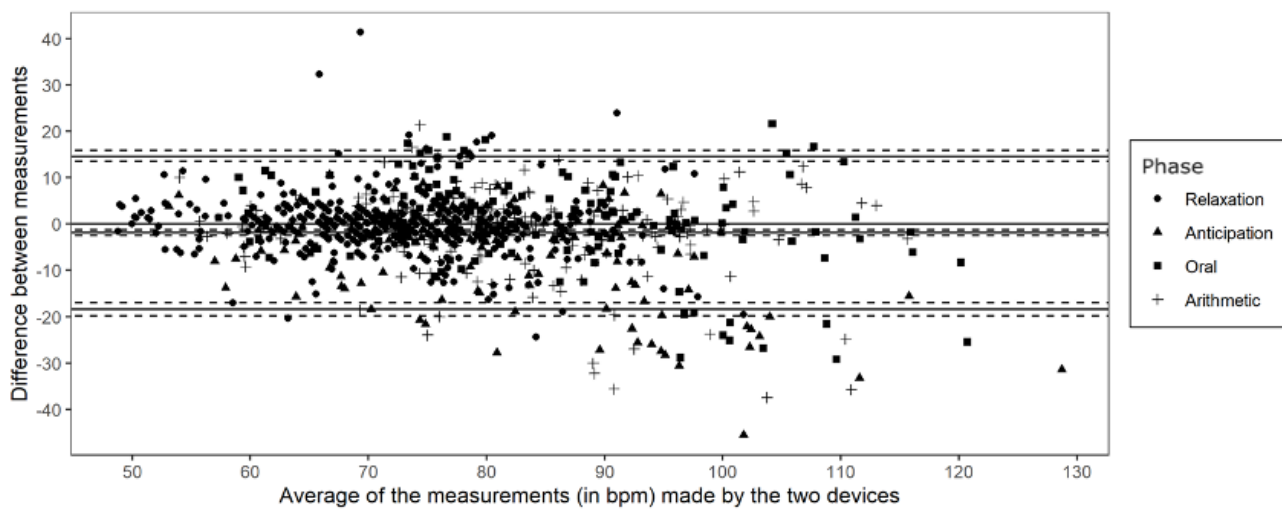
Phase	Mean absolute error (SD)	95% CI	Mean absolute percentage error (%)
Relaxation	4.32 (4.92)	2.03-6.61	6.08
Anticipation	8.94 (8.92)	6.65-11.24	9.83
Oral	6.60 (6.33)	4.30-8.89	7.88
Arithmetic	6.13 (6.88)	3.84-8.43	7.16

Agreement

Results from the mixed-effect limits of agreement method revealed a mean bias of -1.91 (95% LoA -18.37 to 14.58). Figure 2 shows the corresponding Bland-Altman plot with its 95% LoA. The CCC was estimated to be 0.76 (95% CI 0.66 - 0.83). The coverage probability index with a CAD of ± 10 bpm was estimated to be 0.72 (95% CI 0.66 - 0.81). The

mixed-effect model estimated a mean squared deviation of 85.69 (95% CI 57.84 - 111.51). The total deviation index was calculated to be 18.14 (95% CI 14.91 - 20.70). Prior to analyzing the CIA, the residual error variance was calculated. The results showed a Bland-Altman repeatability coefficient of 15.11 , which signifies an approximate 95% probability that the repeated bpm values are within 15 bpm of each other. The CIA was estimated to be 0.69 (95% CI 0.57 - 0.84).

Figure 2. Bland-Altman plot. Mean bias and limits of agreement are shown by the full lines, while confidence intervals are shown by the dashed lines. bpm: beats per minute.



The variance component estimates of the mixed-effect model were evaluated to find the principal sources of disagreement (Table 3). Results showed substantial within-subject variability ($\sigma^2=116.57$). Moreover, the variability of the experimental phases ($\sigma^2=22.58$), within-subject residual ($\sigma^2=29.72$), and the subject-phase interaction ($\sigma^2=14.47$) was high. To better understand the effect of the specific experimental phase on the

agreement between devices, mixed-effect models were analyzed for each phase separately.

In sum, results revealed that, when compared with a gold-standard device, the Fitbit Versa 2 shows overall poor agreement on all metrics analyzed. Further analyses conducted for each experimental phase revealed adequate agreement during the relaxation phase, whereas the preparation phase showed the worst agreement between the 2 devices (Table 3).

Table 3. Metrics of agreement between the Fitbit Versa 2 and the gold-standard ECG across experimental phases.

Phase	LoA ^a (95% LoA)	Concordance correlation coefficient (95% CI)	Coverage probability index (95% CI)	Total deviation index (95% CI)	Coefficient of individual agreement (95% CI)
Relaxation	-0.27 (-8.70 to 8.16)	0.78 (0.67 to 0.84)	0.85 (0.79 to 0.90)	13.76 (11.85 to 15.69)	0.98 (0.88 to 0.99)
Preparation	-6.70 (-20.36 to 6.97)	0.56 (0.44 to 0.67)	0.55 (0.49 to 0.67)	25.94 (20.18 to 29.93)	0.52 (0.35 to 0.69)
Oral	0.65 (-11.08 to 12.38)	0.81 (0.71 to 0.87)	0.74 (0.68 to 0.83)	17.37 (14.17 to 19.74)	0.61 (0.41 to 0.79)
Arithmetic	-1.32 (-12.95 to 10.31)	0.74 (0.58 to 0.87)	0.71 (0.62 to 0.88)	18.42 (12.70 to 22.23)	0.59 (0.38 to 0.86)

^aLoA: limits of agreement.

Discussion

Main Study Findings

Regarding the first objective of testing the ability of a Fitbit device to sense an increase in HR upon induction of psychological stress, results from a mixed-effect model revealed that the HR measurements from the Fitbit Versa 2 showed

significant mean differences across all experimental phases. Regarding the second objective of assessing the accuracy of a Fitbit device to capture change in HR compared to a gold-standard ECG, the MAE index and MAPE results showed acceptable accuracy across all phases. Regarding the third objective of quantifying the degree of agreement between the Fitbit device and the gold-standard ECG measure, the 95% CI was fairly large and lies outside of the predefined CAD,

indicating that the 2 devices did not reach the desired LoA. Moreover, the 95% CI indicated that HR measurements could be underestimated by almost 19 bpm. Visual inspection of the Bland-Altman plot showed small differences between the 2 devices for bpm <90 and greater differences for higher bpm.

Unsurprisingly, the differences in HR detected across experimental phases were markedly more important between the relaxation condition and the oral and arithmetic conditions, a pattern that is echoed in previous studies that have investigated HRV using the TSST [51-53]. These results are also in line with longitudinal research that found support for the use of wearable devices to measure stress levels [17-19,26] and indicate that short-term variations in levels of psychological stress can be detected using a Fitbit device. Additionally, although this is the first study to quantify the accuracy of a Fitbit device under experimentally induced psychological stress, accuracy estimates from Fitbit devices under different exercise intensities have been published [27-29]. Regarding the overall MAE, results from this study revealed similar accuracy to what has been found in a past study comparing a Fitbit device to an ECG gold standard [25]. In this study, MAPE estimates show a loss in accuracy under the stressful phases compared with the relaxation phase. Similar patterns were found in past studies where lower MAPEs were associated with light exercise and higher MAPEs with more vigorous exercise [25,28]. Overall, the evidence from several studies including the present one showed that under a normal or relaxed state, Fitbit devices provide accurate HR measurements. However, a loss in the accuracy of these devices can be observed, especially under high HR-inducing physical or psychological stress. Finally, though Bland-Altman LoA revealed a small mean bias compared to previous studies on exercise intensities [25], results regarding the degree of agreement between both devices echo previous findings revealing that wearable devices tend to not perform well compared to gold-standard devices at higher bpm conditions [25,26,54]. Interestingly, the highest mean bias was found in the anticipation phase. The high variability in HR measurements across participants from both devices (especially from the Biopac) may partly explain this result. This variability may emerge from individual differences, in coping with anticipation of a stressful event, especially since the majority of participants were female. Indeed, previous research indicates that men tend to show higher levels of stress than women during the anticipation of a psychosocial stress task [55].

The CCC indices found in this study consistently showed poor agreement between the devices. When compared to past research on physical activity, the CCC found in this study revealed better agreement than what has been found by Thomson et al and worse agreement than the results from Wang et al [56]. Evaluation of the variance components of the mixed-effect model showed important between-subjects variability which is not surprising given the nature of the stressful phases used in this study. For example, some participants may experience more stress during a verbal task than others. A review found that 30% to 50% of people have a fear of public speaking with 40% reporting anxiety about being negatively evaluated by others [57]. Moreover, interparticipant variability in HR changes is echoed in a previous study that used Fitbit-measured HR as an

indicator of stress [24]. As such, to account for this expected variability, the CIA was computed as it is less dependent on the between-subjects variability compared to the CCC [47,58]. However, the repeatability coefficient of Bland-Altman was found to be unacceptably high and warrants caution when interpreting the CIA. Based on past guidelines suggesting a value of at least 0.80 to conclude good agreement [46,49], the overall estimated CIA in this study suggested poor agreement between the devices. The only CIA that reached a good agreement was for the relaxation phase indicating the similarities between the two devices for low bpm. This result provides further evidence that Fitbit devices tend to show greater precision for low bpm conditions for physical activities [25,26].

The estimated overall coverage probability index (CPI) was well below the predefined 0.95 threshold to represent reasonable agreement, suggesting unsatisfactory agreement between devices. Unsurprisingly, the lowest CPI estimate was found in the anticipation phase, which showed the largest between-subjects and between-devices variation in bpm. Results from the total deviation index (TDI) indicated that differences between the Biopac and the Fitbit Versa 2 are expected to lie within ± 18.14 bpm 95% of the time. Compared with the predefined CAD of ± 10 bpm, all TDI values showed poor agreement and were too large to conclude that the 2 devices could be used interchangeably. Overall, the indices of agreement computed in this study showed that the HR measurements from the Fitbit Versa 2 vary significantly from an ECG gold standard, especially for higher bpm. In light of these results, it appears that although the Fitbit Versa 2 can capture short-term variations in bpm under different stress and relaxation conditions, the precision of these variations is questionable.

Limitations

This study contains limitations that need to be acknowledged to fully appreciate its results. First, the sample size of 34 participants may be considered small and did not specifically exclude participants using substances of abuse that may affect their heart rate (eg, nicotine and alcohol). However, our recruitment criteria and size are comparable to previous similar studies that also did not specifically exclude individuals using substances of abuse and have sample sizes that range from 15 to 50 participants [26-29]. Second, the levels of psychological stress were experimentally induced in a controlled laboratory setting, and further research is needed to test whether these results also apply in natural living conditions. However, the psychological stress and relaxation conditions were induced using a well-validated protocol (TSST). Moreover, efforts were made to ensure rigor in analyses, namely through the use of 5 different metrics estimated with their corresponding 95% CI to determine agreement between the Fitbit Versa 2 and the gold-standard device. Third, we used the Biopac as the gold-standard device for measuring HR. While this ECG-based instrument provides medical-grade HR data, it involves the use of electrodes which, when placed incorrectly, can generate noise in the signal and even lead to less accurate data [59]. Despite this potential limitation, the authors believe it was important to have a gold-standard device with which to compare the Fitbit device for the purpose of concurrent validation. Fourth, we did not consider the skin color or the skin photosensitivity, factors

that have previously been suggested to affect the signal resolution of sensors that use photoplethysmography technology such as Fitbit. In this study, over 30% of the participants were non-Caucasian, this could have also affected the accuracy of Fitbit readings [60,61]. Similarly, individual variables of body mass index, prior level of physical activity, and presence of symptoms of psychological disorders (with or without a diagnosis) were beyond the scope of this research and therefore not considered during analyses. However, these variables should be considered in future research interested in quantifying the impacts of individual variables on HR measurements. Finally, while the Fitbit Versa 2 was found to be able to capture short-term stress variation, longitudinal studies are needed before concluding on the potential of this device to capture mid-to-long-term stress levels to predict psychological distress and diminished well-being. Nevertheless, a strength remains that this is the first study to quantify the accuracy of a Fitbit device under experimentally induced psychological stress and can serve as an important foundation for future research regarding wearable activity monitors and psychological stress.

Conclusions

With the ubiquity of wearable devices and the growing interest to use the data they provide in the health sector, research is needed to test the reliability and validity of these instruments. To our knowledge, this is the first study to test the accuracy and agreement of a wearable device (Fitbit Versa 2) under different psychological stress-inducing experimental conditions. Results showed that the short-term variations in psychological stress levels were successfully captured by the Fitbit Versa 2. Moreover, MAE and MAPE estimates were all below the predefined threshold of ± 10 bpm, indicating acceptable accuracy of the Fitbit Versa 2. However, across the 5 metrics of agreement analyzed, results revealed poor agreement between the HR measurement from the Fitbit device and the Biopac. Importantly, the results of this study have implications in advancing research involving the use of wearable devices as it provides preliminary evidence that the HR measurement from the Fitbit Versa 2 can be used to detect psychological stress among a nonclinical adult population.

Acknowledgments

This work was supported by Mitacs through the Mitacs Accelerate program (grant IT13707) with the contribution of Lucine and Berger-Levrault.

Data Availability

The data sets and the R code generated and analyzed during this study are available at [62].

Authors' Contributions

JG, FV, VD, and ST contributed to the study concept and design. Analysis and interpretation of data were done by JG, FV, VD, and ST. JG, MK, and LLH drafted the manuscript. FV, VD, and ST contributed during the critical revision of the manuscript. JG performed the statistical analysis. All authors approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

References

1. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388(10053):1659-1724.
2. Yang H, Xian X, Hu J, Millis JM, Zhao H, Lu X, et al. Public psychosocial and behavioral responses in the first wave of COVID-19 pandemic: a large survey in China. *Front Psychiatry* 2021 Jul 28;12:676914. [doi: [10.3389/fpsy.2021.676914](https://doi.org/10.3389/fpsy.2021.676914)]
3. Anderson RM, Hollingsworth TD, Baggaley RF, Maddren R, Vegvari C. COVID-19 spread in the UK: the end of the beginning? *Lancet* 2020 Aug;396(10251):587-590. [doi: [10.1016/s0140-6736\(20\)31689-5](https://doi.org/10.1016/s0140-6736(20)31689-5)]
4. Anderson M, Mckee M, Mossialos E. Covid-19 exposes weaknesses in European response to outbreaks. *BMJ* 2020 Mar 18;368:m1075. [doi: [10.1136/bmj.m1075](https://doi.org/10.1136/bmj.m1075)] [Medline: [32188590](https://pubmed.ncbi.nlm.nih.gov/32188590/)]
5. Viner RM, Bonell C, Drake L, Jourdan D, Davies N, Baltag V, et al. Reopening schools during the COVID-19 pandemic: governments must balance the uncertainty and risks of reopening schools against the clear harms associated with prolonged closure. *Arch Dis Child* 2021 Feb 03;106(2):111-113 [FREE Full text] [doi: [10.1136/archdischild-2020-319963](https://doi.org/10.1136/archdischild-2020-319963)] [Medline: [32747375](https://pubmed.ncbi.nlm.nih.gov/32747375/)]
6. Patel V, Prince M. Global mental health: a new global health field comes of age. *JAMA* 2010 May 19;303(19):1976-1977 [FREE Full text] [doi: [10.1001/jama.2010.616](https://doi.org/10.1001/jama.2010.616)] [Medline: [20483977](https://pubmed.ncbi.nlm.nih.gov/20483977/)]
7. Joshi M, Ashrafian H, Aufegger L, Khan S, Arora S, Cooke G, et al. Wearable sensors to improve detection of patient deterioration. *Expert Rev Med Dev* 2019 Feb 06;16(2):145-154 [FREE Full text] [doi: [10.1080/17434440.2019.1563480](https://doi.org/10.1080/17434440.2019.1563480)] [Medline: [30580650](https://pubmed.ncbi.nlm.nih.gov/30580650/)]

8. Müller R, Hamacher D, Hansen S, Oschmann P, Keune PM. Wearable inertial sensors are highly sensitive in the detection of gait disturbances and fatigue at early stages of multiple sclerosis. *BMC Neurol* 2021 Sep 04;21(1):337 [FREE Full text] [doi: [10.1186/s12883-021-02361-y](https://doi.org/10.1186/s12883-021-02361-y)] [Medline: [34481481](https://pubmed.ncbi.nlm.nih.gov/34481481/)]
9. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. *JMIR Mhealth Uhealth* 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: [10.2196/18694](https://doi.org/10.2196/18694)] [Medline: [32897239](https://pubmed.ncbi.nlm.nih.gov/32897239/)]
10. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act* 2015 Dec 18;12:159 [FREE Full text] [doi: [10.1186/s12966-015-0314-1](https://doi.org/10.1186/s12966-015-0314-1)] [Medline: [26684758](https://pubmed.ncbi.nlm.nih.gov/26684758/)]
11. 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](https://doi.org/10.2196/10527)] [Medline: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)]
12. 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](https://doi.org/10.2196/mhealth.9754)] [Medline: [29650506](https://pubmed.ncbi.nlm.nih.gov/29650506/)]
13. Schäfer A, Vagedes J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. *Int J Cardiol* 2013 Jun 05;166(1):15-29. [doi: [10.1016/j.ijcard.2012.03.119](https://doi.org/10.1016/j.ijcard.2012.03.119)] [Medline: [22809539](https://pubmed.ncbi.nlm.nih.gov/22809539/)]
14. Sweeney KT, Ward TE, McLoone SF. Artifact removal in physiological signals—practices and possibilities. *IEEE Trans Inf Technol Biomed* 2012 May;16(3):488-500. [doi: [10.1109/TITB.2012.2188536](https://doi.org/10.1109/TITB.2012.2188536)] [Medline: [22361665](https://pubmed.ncbi.nlm.nih.gov/22361665/)]
15. Maeda Y, Sekine M, Tamura T. Relationship between measurement site and motion artifacts in wearable reflected photoplethysmography. *J Med Syst* 2011 Oct 7;35(5):969-976. [doi: [10.1007/s10916-010-9505-0](https://doi.org/10.1007/s10916-010-9505-0)] [Medline: [20703691](https://pubmed.ncbi.nlm.nih.gov/20703691/)]
16. Reingle Gonzalez JM, Jetelina KK, Bishopp SA, Livingston MD, Perez RA, Gabriel KP. The feasibility of using real-time, objective measurements of physiological stress among law enforcement officers in Dallas, Texas. *PIJPSM* 2019 Aug 12;42(4):701-710. [doi: [10.1108/pijpsm-12-2018-0184](https://doi.org/10.1108/pijpsm-12-2018-0184)]
17. Runge N, Haarman S, Fisher M. Using Fitbit fitness trackers to measure teacher stress and coping. *Int J Soc Policy Educ* 2020;2:56-70.
18. Sano A, Taylor S, McHill AW, Phillips AJ, Barger LK, Klerman E, et al. Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *J Med Internet Res* 2018 Jun 08;20(6):e210 [FREE Full text] [doi: [10.2196/jmir.9410](https://doi.org/10.2196/jmir.9410)] [Medline: [29884610](https://pubmed.ncbi.nlm.nih.gov/29884610/)]
19. Thayer JF, Ahs F, Fredrikson M, Sollers JJ, Wager TD. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neurosci Biobehav Rev* 2012 Feb;36(2):747-756. [doi: [10.1016/j.neubiorev.2011.11.009](https://doi.org/10.1016/j.neubiorev.2011.11.009)] [Medline: [22178086](https://pubmed.ncbi.nlm.nih.gov/22178086/)]
20. Schiweck C, Piette D, Berckmans D, Claes S, Vrieze E. Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. A systematic review. *Psychol Med* 2018 Aug 23;49(2):200-211. [doi: [10.1017/s0033291718001988](https://doi.org/10.1017/s0033291718001988)]
21. Pulpulos MM, Baeken C, De Raedt R. Cortisol response to stress: the role of expectancy and anticipatory stress regulation. *Horm Behav* 2020 Jan;117:104587. [doi: [10.1016/j.yhbeh.2019.104587](https://doi.org/10.1016/j.yhbeh.2019.104587)] [Medline: [31639385](https://pubmed.ncbi.nlm.nih.gov/31639385/)]
22. Dong S, Lee M, Park H, Youn I. Stress resilience measurement with heart-rate variability during mental and physical stress. 2018 Presented at: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 18-21, 2018; Honolulu, HI p. 5290-5293. [doi: [10.1109/embc.2018.8513531](https://doi.org/10.1109/embc.2018.8513531)]
23. Pakhomov SVS, Thuras PD, Finzel R, Eppel J, Kotlyar M. Using consumer-wearable technology for remote assessment of physiological response to stress in the naturalistic environment. *PLoS One* 2020 Mar 25;15(3):e0229942 [FREE Full text] [doi: [10.1371/journal.pone.0229942](https://doi.org/10.1371/journal.pone.0229942)] [Medline: [32210441](https://pubmed.ncbi.nlm.nih.gov/32210441/)]
24. Is there a difference between measurement. National Physical Laboratory. URL: <https://www.npl.co.uk/resources/q-a/difference-accuracy-uncertainty> [accessed 2021-11-08]
25. Wilde D. Report: Google's recently acquired Fitbit accounted for 6% of wearable market in 2019. 9to5Google. URL: <https://9to5google.com/2020/03/11/report-googles-recently-acquired-fitbit-accounted-for-6-of-wearable-market-in-2019/> [accessed 2021-10-13]
26. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the Fitbit Charge 2 for monitoring heart rate. *PLoS One* 2018;13(2):e0192691 [FREE Full text] [doi: [10.1371/journal.pone.0192691](https://doi.org/10.1371/journal.pone.0192691)] [Medline: [29489850](https://pubmed.ncbi.nlm.nih.gov/29489850/)]
27. Thomson EA, Nuss K, Comstock A, Reinwald S, Blake S, Pimentel RE, et al. Heart rate measures from the Apple Watch, Fitbit Charge HR 2, and electrocardiogram across different exercise intensities. *J Sports Sci* 2019 Jan 18;37(12):1411-1419. [doi: [10.1080/02640414.2018.1560644](https://doi.org/10.1080/02640414.2018.1560644)]
28. Cadmus-Bertram L, Gangnon R, Wirkus EJ, Thraen-Borowski KM, Gorzelitz-Liebhauser J. The accuracy of heart rate monitoring by some wrist-worn activity trackers. *Ann Intern Med* 2017 Apr 18;166(8):610-612 [FREE Full text] [doi: [10.7326/L16-0353](https://doi.org/10.7326/L16-0353)] [Medline: [28395305](https://pubmed.ncbi.nlm.nih.gov/28395305/)]
29. Jo E, Lewis K, Directo D, Kim MJ, Dolezal BA. Validation of biofeedback wearables for photoplethysmographic heart rate tracking. *J Sports Sci Med* 2016 Sep;15(3):540-547 [FREE Full text] [Medline: [27803634](https://pubmed.ncbi.nlm.nih.gov/27803634/)]

30. Boudreaux B, Hebert E, Hollander D. Validity of wearable activity monitors during cycling and resistance exercise. *Med Sci Sports Exerc* 2018;50(3):624-633. [doi: [10.1249/mss.0000000000001471](https://doi.org/10.1249/mss.0000000000001471)]
31. Narvaez Linares N, Charron V, Ouimet A, Labelle P, Plamondon H. A systematic review of the Trier Social Stress Test methodology: issues in promoting study comparison and replicable research. *Neurobiol Stress* 2020 Nov;13:100235 [FREE Full text] [doi: [10.1016/j.ynstr.2020.100235](https://doi.org/10.1016/j.ynstr.2020.100235)] [Medline: [33344691](https://pubmed.ncbi.nlm.nih.gov/33344691/)]
32. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull* 2004 May;130(3):355-391. [doi: [10.1037/0033-2909.130.3.355](https://doi.org/10.1037/0033-2909.130.3.355)] [Medline: [15122924](https://pubmed.ncbi.nlm.nih.gov/15122924/)]
33. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):1-48.
34. Ben-Shachar M, Lüdtke D, Makowski D. effectsize: estimation of effect size indices and standardized parameters. *J Open Source Softw* 2020 Dec;5(56):2815. [doi: [10.21105/joss.02815](https://doi.org/10.21105/joss.02815)]
35. Lenth R, Buerkner P, Herve M, Love J, Riebl H, Singmann H. Emmeans: estimated marginal means, aka least-squares means. The Comprehensive R Archive Network. URL: <https://CRAN.R-project.org/package=emmeans> [accessed 2022-12-01]
36. ANSI/AAMI EC13-2002—cardiac monitors, heart rate meters, and alarms. ANSI. URL: <https://webstore.ansi.org/standards/aami/ansiaamiec132002> [accessed 2021-10-19]
37. Nelson MB, Kaminsky LA, Dickin DC, Montoye AHK. Validity of consumer-based physical activity monitors for specific activity types. *Med Sci Sports Exerc* 2016 Aug;48(8):1619-1628. [doi: [10.1249/MSS.0000000000000933](https://doi.org/10.1249/MSS.0000000000000933)] [Medline: [27015387](https://pubmed.ncbi.nlm.nih.gov/27015387/)]
38. Müller AM, Wang NX, Yao J, Tan CS, Low ICC, Lim N, et al. Heart rate measures from wrist-worn activity trackers in a laboratory and free-living setting: validation study. *JMIR Mhealth Uhealth* 2019 Oct 02;7(10):e14120 [FREE Full text] [doi: [10.2196/14120](https://doi.org/10.2196/14120)] [Medline: [31579026](https://pubmed.ncbi.nlm.nih.gov/31579026/)]
39. Parker RA, Weir CJ, Rubio N, Rabinovich R, Pinnock H, Hanley J, et al. Application of mixed effects limits of agreement in the presence of multiple sources of variability: exemplar from the comparison of several devices to measure respiratory rate in COPD patients. *PLoS One* 2016 Dec 14;11(12):e0168321 [FREE Full text] [doi: [10.1371/journal.pone.0168321](https://doi.org/10.1371/journal.pone.0168321)] [Medline: [27973556](https://pubmed.ncbi.nlm.nih.gov/27973556/)]
40. Caldwell A. SimplyAgree: flexible and robust agreement and reliability analyses. The Comprehensive R Archive Network. URL: <https://CRAN.R-project.org/package=SimplyAgree> [accessed 2022-12-01]
41. Carrasco JL, Phillips BR, Puig-Martinez J, King TS, Chinchilli VM. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Comput Methods Programs Biomed* 2013 Mar;109(3):293-304. [doi: [10.1016/j.cmpb.2012.09.002](https://doi.org/10.1016/j.cmpb.2012.09.002)] [Medline: [23031487](https://pubmed.ncbi.nlm.nih.gov/23031487/)]
42. Parker RA, Scott C, Inácio V, Stevens NT. Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners. *BMC Med Res Methodol* 2020 Jun 12;20(1):154 [FREE Full text] [doi: [10.1186/s12874-020-01022-x](https://doi.org/10.1186/s12874-020-01022-x)] [Medline: [32532218](https://pubmed.ncbi.nlm.nih.gov/32532218/)]
43. McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA Client Rep HAM2005-062. 2005. URL: <https://doczz.net/doc/8245728/a-proposal-for-strength-of-agreement-criteria-for-lines> [accessed 2022-06-10]
44. Carrasco J, Martinez J. cccrm: concordance correlation coefficient for repeated (and non-repeated) measures. The Comprehensive R Archive Network. URL: <https://CRAN.R-project.org/package=cccrm> [accessed 2022-12-01]
45. Lin L, Hedayat AS, Sinha B, Yang M. Statistical methods in assessing agreement. *J Am Stat Assoc* 2002 Mar;97(457):257-270. [doi: [10.1198/016214502753479392](https://doi.org/10.1198/016214502753479392)]
46. Haber M, Barnhart HX. A general approach to evaluating agreement between two observers or methods of measurement from quantitative data with replicated measurements. *Stat Methods Med Res* 2008 Apr 14;17(2):151-169. [doi: [10.1177/0962280206075527](https://doi.org/10.1177/0962280206075527)] [Medline: [17698934](https://pubmed.ncbi.nlm.nih.gov/17698934/)]
47. Barnhart HX, Likhnygina Y, Kosinski AS, Haber M. Comparison of concordance correlation coefficient and coefficient of individual agreement in assessing agreement. *J Biopharm Stat* 2007 Jul 05;17(4):721-738. [doi: [10.1080/10543400701329497](https://doi.org/10.1080/10543400701329497)] [Medline: [17613650](https://pubmed.ncbi.nlm.nih.gov/17613650/)]
48. Pan Y, Gao J, Haber M, Barnhart HX. Estimation of coefficients of individual agreement (CIAs) for quantitative and binary data using SAS and R. *Comput Meth Progr Biomed* 2010 May;98(2):214-219. [doi: [10.1016/j.cmpb.2009.12.002](https://doi.org/10.1016/j.cmpb.2009.12.002)]
49. Haber M, Gao J, Barnhart H. Evaluation of agreement between measurement methods from data with matched repeated measurements via the coefficient of individual agreement. *J Data Sci JDS* 2010;8(3):457-469. [doi: [10.6339/jds.2010.08\(3\).619](https://doi.org/10.6339/jds.2010.08(3).619)]
50. A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
51. Hellhammer J, Schubert M. The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test. *Psychoneuroendocrinology* 2012 Jan;37(1):119-124. [doi: [10.1016/j.psyneuen.2011.05.012](https://doi.org/10.1016/j.psyneuen.2011.05.012)] [Medline: [21689890](https://pubmed.ncbi.nlm.nih.gov/21689890/)]
52. Mozos OM, Sandulescu V, Andrews S, Ellis D, Bellotto N, Dobrescu R, et al. Stress detection using wearable physiological and sociometric sensors. *Int J Neur Syst* 2016 Dec 28;27(02):1650041. [doi: [10.1142/s0129065716500416](https://doi.org/10.1142/s0129065716500416)]
53. Shiban Y, Diemer J, Brandl S, Zack R, Mühlberger A, Wüst S. Trier social stress test in vivo and in virtual reality: dissociation of response domains. *Int J Psychophysiol* 2016 Dec;110:47-55. [doi: [10.1016/j.ijpsycho.2016.10.008](https://doi.org/10.1016/j.ijpsycho.2016.10.008)] [Medline: [27742258](https://pubmed.ncbi.nlm.nih.gov/27742258/)]

54. Al-Kaisey AM, Koshy AN, Ha FJ, Spencer R, Toner L, Sajeev JK, et al. Accuracy of wrist-worn heart rate monitors for rate control assessment in atrial fibrillation. *Int J Cardiol* 2020 Feb 01;300:161-164. [doi: [10.1016/j.ijcard.2019.11.120](https://doi.org/10.1016/j.ijcard.2019.11.120)] [Medline: [31787389](https://pubmed.ncbi.nlm.nih.gov/31787389/)]
55. Kirschbaum C, Wüst S, Hellhammer D. Consistent sex differences in cortisol responses to psychological stress. *Psychosom Med* 1992;54(6):648-657. [doi: [10.1097/00006842-199211000-00004](https://doi.org/10.1097/00006842-199211000-00004)]
56. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiol* 2017 Jan 01;2(1):104-106. [doi: [10.1001/jamacardio.2016.3340](https://doi.org/10.1001/jamacardio.2016.3340)] [Medline: [27732703](https://pubmed.ncbi.nlm.nih.gov/27732703/)]
57. Garcia-Leal C, Graeff FG, Del-Ben CM. Experimental public speaking: contributions to the understanding of the serotonergic modulation of fear. *Neurosci Biobehav Rev* 2014 Oct;46 Pt 3:407-417. [doi: [10.1016/j.neubiorev.2014.09.011](https://doi.org/10.1016/j.neubiorev.2014.09.011)] [Medline: [25277282](https://pubmed.ncbi.nlm.nih.gov/25277282/)]
58. Barnhart H. A review on assessing agreement. In: Wiley StatsRef: Statistics Reference Online. Hoboken, New Jersey: American Cancer Society; 2018:1-30.
59. Pettersen TR, Fållun N, Norekvål TM. Improvement of in-hospital telemetry monitoring in coronary care units: an intervention study for achieving optimal electrode placement and attachment, hygiene and delivery of critical information to patients. *Eur J Cardiovasc Nurs* 2014 Dec 04;13(6):515-523. [doi: [10.1177/1474515113515585](https://doi.org/10.1177/1474515113515585)] [Medline: [24304659](https://pubmed.ncbi.nlm.nih.gov/24304659/)]
60. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas* 2007 Mar;28(3):R1-39. [doi: [10.1088/0967-3334/28/3/R01](https://doi.org/10.1088/0967-3334/28/3/R01)] [Medline: [17322588](https://pubmed.ncbi.nlm.nih.gov/17322588/)]
61. Butler MJ, Crowe JA, Hayes-Gill BR, Rodmell PI. Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin. *Physiol Meas* 2016 May;37(5):N27-N37. [doi: [10.1088/0967-3334/37/5/N27](https://doi.org/10.1088/0967-3334/37/5/N27)] [Medline: [27100666](https://pubmed.ncbi.nlm.nih.gov/27100666/)]
62. Supplementary materials for: comparing a fitbit wearable to an electrocardiogram gold standard as a measure of heart rate under psychological stress: a validation study. Open Science Framework. URL: <https://osf.io/qpv9d/> [accessed 2022-12-12]

Abbreviations

- bpm:** beats per minute
- CAD:** clinically acceptable difference
- CCC:** concordance correlation coefficient
- CIA:** coefficient of individual agreement
- CP:** coverage probability
- ECG:** electrocardiogram
- HR:** heart rate
- LoA:** limits of agreement
- MAE:** mean absolute error
- MAPE:** mean absolute percentage error
- TDI:** total deviation index
- TSST:** Trier Social Stress Test
- WHO:** World Health Organization

Edited by A Mavragani; submitted 10.03.22; peer-reviewed by M Attridge, RP Rajkumar, N Bowles; comments to author 12.10.22; revised version received 28.10.22; accepted 29.10.22; published 21.12.22.

Please cite as:

Gagnon J, Khau M, Lavoie-Hudon L, Vachon F, Drapeau V, Tremblay S

Comparing a Fitbit Wearable to an Electrocardiogram Gold Standard as a Measure of Heart Rate Under Psychological Stress: A Validation Study

JMIR Form Res 2022;6(12):e37885

URL: <https://formative.jmir.org/2022/12/e37885>

doi: [10.2196/37885](https://doi.org/10.2196/37885)

PMID: [36542432](https://pubmed.ncbi.nlm.nih.gov/36542432/)

©Joel Gagnon, Michelle Khau, Léandre Lavoie-Hudon, François Vachon, Vicky Drapeau, Sébastien Tremblay. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 21.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Voice App Design for Heart Failure Self-management: Proof-of-Concept Implementation Study

Antonia Barbaric^{1,2,3}, MASc; Cosmin Munteanu^{4,5}, PhD; Heather Ross^{6,7,8}, MHSc, MD; Joseph A Cafazzo^{1,2,3,9,10}, PEng, PhD

¹Centre for Digital Therapeutics, Techna Institute, University Health Network, Toronto, ON, Canada

²Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

³Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

⁴Institute for Communication, Culture, Information, and Technology, University of Toronto, Mississauga, ON, Canada

⁵Technologies for Aging Gracefully Lab, University of Toronto, Toronto, ON, Canada

⁶Ted Rogers Centre for Heart Research, University Health Network, Toronto, ON, Canada

⁷Department of Medicine, University of Toronto, Toronto, ON, Canada

⁸Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada

⁹Department of Computer Science, University of Toronto, Toronto, ON, Canada

¹⁰Healthcare Human Factors, Techna Institute, University of Toronto, Toronto, ON, Canada

Corresponding Author:

Antonia Barbaric, MASc

Centre for Digital Therapeutics

Techna Institute

University Health Network

Toronto General Hospital - RFE Building, 4th floor

190 Elizabeth St

Toronto, ON, M5G2C4

Canada

Phone: 1 416 340 4800 ext 4765

Email: antonia.barbaric@mail.utoronto.ca

Abstract

Background: Voice user interfaces are becoming more prevalent in health care and are commonly being used for patient engagement. There is a growing interest in identifying the potential this form of interface has on patient engagement with digital therapeutics (DTx) in chronic disease management. Making DTx accessible through an alternative interaction model also has the potential to better meet the needs of some patients, such as older adults and those with physical and cognitive impairments, based on existing research.

Objective: This study aimed to evaluate how participants with heart failure interacted with a voice app version of a DTx, *Medly*, through a proof-of-concept implementation study design. The objective was to understand whether the voice app would enable the participants to successfully interact with the DTx, with a focus on acceptability and feasibility.

Methods: A mixed methods concurrent triangulation design was used to better understand the acceptability and feasibility of the use of the *Medly* voice app with the study participants (N=20) over a 4-week period. Quantitative data included engagement levels, accuracy rates, and questionnaires, which were analyzed using descriptive statistics. Qualitative data included semistructured interviews and were analyzed using a qualitative descriptive approach.

Results: The overall average engagement level was 73% (SD 9.5%), with a 14% decline between results of weeks 1 and 4. The biggest difference was between the average engagement levels of the oldest and youngest demographics, 84% and 43%, respectively, but these results were not significant—Kruskal-Wallis test, $H(2)=3.8$ ($P=.14$). The *Medly* voice app had an overall accuracy rate of 97.8% and was successful in sending data to the clinic. From an acceptability perspective, the voice app was ranked in the 80th percentile, and overall, the users felt that the voice app was not a lot of work (average of 2.1 on a 7-point Likert scale). However, the overall average score for whether users would use it in the future declined by 13%. Thematic analysis revealed the following: the theme feasibility of clinical integration had 2 subthemes, namely users adapted to the voice app's conversational style and device unreliability, and the theme voice app acceptability had 3 subthemes, namely the device integrated well within

household and users' lives, users blamed themselves when problems arose with the voice app, and voice app was missing specific, desirable user features.

Conclusions: In conclusion, participants were largely successful in using the *Medly* voice app despite some of the barriers faced, proving that an app such as this could be feasible to be deployed in the clinic. Our data begin to piece together the patient profile this technology may be most suitable for, namely those who are older, have flexible schedules, are confident in using technology, and are experiencing other medical conditions.

(*JMIR Form Res* 2022;6(12):e40021) doi:[10.2196/40021](https://doi.org/10.2196/40021)

KEYWORDS

heart failure; self-management; digital therapeutics; voice-activated technology; smart speaker; formative evaluation; mobile phone

Introduction

Background

Chronic diseases are the leading cause of death and disability worldwide, with >41 million people dying every year owing to these diseases [1]. Cardiovascular diseases, such as myocardial infarctions and high blood pressure, are responsible for most chronic disease-related deaths (17.9 million people) [1]. Patient self-care is considered essential in the prevention and management of chronic diseases [2], as studies have shown the benefits of this approach, which include improved health outcomes, decreased clinic visits, and decreased health costs [3]. Mobile health is a type of digital health technology that involves the use of mobile devices for medical and public health practices [4] and enables the integration of self-care support into a patient's routine [5]. Mobile health apps are one of the most popular tools for helping patients with chronic conditions manage their health at home [6]. However, the use of conversational agents for health-related purposes is an emerging field of research [7], and early evidence suggests that they may also be effective for the self-management of chronic diseases [8].

Conversational agents are a type of dialogue in the field of human-computer interaction and can either be voice based or typing based [9]. With voice user interfaces (VUIs), users can interact with a computing system using only speech. An example of VUIs is voice apps. The primary advantage of implementing VUIs in any environment is simplicity because it does not require the user to interact with a hand-held technology, as we are typically accustomed to. Some examples of how VUIs are being used in a clinical setting include improving physician note transcription, supporting patient registration processes, improving patient engagement with chronic disease management programs, and aging in place [10]. In a home setting, voice apps are designed to help patients manage their chronic conditions independently [11-15] and most often include informational and assistive services such as general educational content, reminders, and tracking tools. The research disseminated so far has limited efficacy in supporting final conclusions because the studies are still in development and piloting phases. As a result, there is a growing interest in investigating the feasibility of using voice apps to encourage patient engagement, specifically for chronic disease management.

Heart Failure

Previous research has begun to investigate the feasibility of voice-activated technology for monitoring patients with heart failure (HF) [16]. HF is a cardiovascular disease that develops when the heart muscle becomes damaged or weak [17], making it difficult to pump enough blood to meet the body's needs [18]. When this happens, fluid builds up in various parts of the body (such as the legs and ankles), creates congestion in the lungs, and leads to a lack of oxygen being delivered to the rest of the body [19]. The 2 most common causes of HF are high blood pressure and coronary artery disease; other risk factors include obesity, smoking, high cholesterol, and previous health conditions (past myocardial infarctions and heart defects at birth) [20]. It is estimated that 64.3 million people are living with HF worldwide [21].

To date, there have been limited studies investigating the potential of using a voice app for HF self-management. Some voice apps include basic functionality to help patients manage their conditions, such as asking preappointment clinical screening questions, scheduling appointments, and setting medication reminders [11,12]. Other, more recent studies have investigated using voice apps to monitor patients' conditions through a series of symptom questions related to HF [13,16]. Feasibility was an outcome that all the studies investigated, and the results concluded that it is worthwhile to investigate how this technology can be used as an alternative platform to manage HF.

Medly

Medly is an evidence-based, HF self-management program that was developed by the University Health Network (UHN) and is implemented as part of the standard of care at UHN's Ted Rogers Center of Excellence for Heart Failure clinic [22]. The program is deployed as a mobile app, and patients access it daily using their mobile phones to log clinically relevant physiological measurements (weight, blood pressure, and heart rate) and HF-related symptoms. All patients input the same measurements and are asked the same symptom-related questions despite the stage of their HF.

The *Medly* algorithm generates an automated self-care message for the patient based on the data inputted and the patient's medical history (determined when the patient is onboarded to the *Medly* program). The *Medly* program was deployed as a voice app as part of a previous work, and a usability study was

performed with the voice app at the UHN's Heart Failure Clinic [23].

Objectives

The purpose of the previous usability study was mainly focused on whether the *Medly* voice app functioned as intended; feedback on the voice app design and data regarding user experience were collected. Given that the usability study took place in a controlled laboratory setting and focused on the voice app design, we sought to perform a proof-of-concept implementation study in the intended environment. The *Medly* voice app was used as a case study to investigate the broader application of voice apps for chronic disease management. The goal of this study was to determine whether voice apps can be a practical alternative for enabling patients to receive a digital therapeutic. A total of 2 constructs from the implementation framework (acceptability and feasibility) by Proctor et al [24] guided our research question: What is the acceptability and feasibility of a voice application for patients, through the use of a smart speaker, for a home chronic disease management platform? If the study findings concluded that the voice app is acceptable to patients and feasible to be deployed in a real-world setting, the inclusion of this technology to deploy digital therapeutics could add benefit to the current models of care by offering patients multiple ways to interact with these types of programs.

Methods

Participant Recruitment

This study asked patients with HF to interact with the *Medly* voice app in their homes for a 4-week period. The *Medly* voice app was accessed through an Amazon Alexa (Amazon.com, Inc) device; each participant was provided a device to use for the study duration. The participants were considered eligible if they had been diagnosed with HF by a physician at the UHN's HF clinic and were prescribed the *Medly* program. The participants were also required to speak and read English adequately to understand the voice prompts in the *Medly* app. The *Medly* nurse coordinator first provided a brief overview of the research study to interested patients before introducing them to the study coordinator. If they agreed to participate, written informed consent was obtained by the study coordinator before onboarding.

Given that this study was designed as a proof of concept, a small sample size was used to gather preliminary evidence that provided insights into the success of this intervention. A total of 20 participants were recruited for the study based on similar guidance provided for pilot studies [25]. Of the 20 participants, 7 (35%) were recently onboarded (within the last 2 months) to the *Medly* program at the time when the study was being conducted. The *Medly* nurse coordinator recommended a cutoff of 2 months, given their experience with how long it typically takes for patients to settle in comfortably with the app.

All the participants were required to perform a double entry of their *Medly* measurements for the 4-week duration; more specifically, they were asked to first input their *Medly* measurements on the smartphone app before interacting with

the voice app. Each participant received a gift card to compensate for their time participating in the study.

Ethics Approval

Ethics approval was obtained from the UHN Research Ethics Board (20-6095).

Study Outcome Measures

The evaluation of the *Medly* voice app was influenced by the implementation outcomes framework of Proctor et al [24] by focusing on 2 outcomes, specifically, acceptability and feasibility. Acceptability is defined as the perception among patients that the *Medly* voice app is agreeable or satisfactory, and feasibility is described as the extent to which the *Medly* voice app can be successfully used by patients.

Data Collection

Data were gathered through 3 questionnaires, namely System Usability Scale (SUS) [26], National Aeronautics and Space Administration (NASA)-Task Load Index (TLX) [27], and Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) [28], and semistructured interviews. Information regarding how often the voice app misheard and incorrectly recorded data was retrieved from the voice app server. During the interviews, the participants were asked about their overall experience and satisfaction with using the voice app. Other quantitative data were also collected: engagement levels (defined as the number of days the user inputted their data using the voice app divided by the total study duration—28 days) and accuracy rates (calculated by comparing the measurements inputted on the smartphone app with those recorded on the voice app). The following data were used to deduce whether the voice app was deemed acceptable by users: engagement levels, SUS, and semistructured interviews; similarly, feasibility was identified through the following: engagement levels, accuracy rates, NASA-TLX, UTAUT2 (through an effort expectancy lens), and semistructured interviews.

The study coordinator performed an onboarding session over the phone with each participant to help them set up and access the *Medly* voice app and provided them with an instruction manual (Multimedia Appendix 1). The participants were then asked a few questions regarding how comfortable they were using technology to help the study coordinator understand their comfort levels with technology (Multimedia Appendix 2). As the *Medly* smartphone app is part of the standard of care at UHN, the participants were made aware that they needed to perform a double entry of their *Medly* measurements for the 4-week duration and were told to prioritize the *Medly* smartphone app, namely to input measurements on the phone first and to follow guidance only from the smartphone app. Semistructured interviews were conducted at the end of weeks 1 and 4 and took place with the study coordinator over the phone. Questionnaires were sent out electronically at the end of weeks 2 and 4 so that the participants had privacy and felt comfortable sharing their honest thoughts and opinions.

Study Analysis and Statistical Tests

A mixed methods, triangulation convergence model was used to draw conclusions [29]. Descriptive statistics for the

standardized questionnaire responses were calculated and recorded using Microsoft Excel. Graphical representations of engagement levels were also created using Microsoft Excel. The responses from the SUS questionnaire were analyzed as per standard protocol [26], and averages were calculated for the NASA-TLX and UTAUT2 questionnaires, both overall and question specific. Data were categorized in different ways using various attributes (age, whether they were recently onboarded to the *Medly* program, whether they had any prior experience using a smart speaker, and comfort level with technology). Given the data characteristics, nonparametric statistical tests were conducted with each attribute (treated as an independent variable) for engagement levels and scores from SUS, NASA-TLX, and UTAUT2; a *P* value of <.05 was used to indicate statistical significance.

For the qualitative data, interview transcripts were analyzed and coded by the study coordinator (AB). Themes from the interviews were identified using an inductive, qualitative descriptive approach [30]. Once these themes were generated, a deductive approach was used to categorize them under the guidance of implementation outcomes framework (with a focus on the acceptability and feasibility constructs) by Proctor et al [24]. The transcripts and coding were organized using Microsoft Word. Owing to the small sample size and lack of power and statistical significance in the results, more emphasis was placed on the qualitative analysis, whereas the quantitative data and interpretations were used to support the qualitative findings.

Results

Characteristics of Study Participants

A total of 20 patients were recruited for the study, with a fairly even split between sexes (female: 9/20, 45%; male: 11/20, 55%) and an average age of 57.8 (SD 13.1) years. None of the participants were aged <20 years, 10% (2/20) of users were aged between 21 and 40 years, 35% (7/20) were aged between 41 and 60 years, and 55% (11/20) were aged between 61 and 80 years.

All the recruited patients were required to be enrolled in the *Medly* program, with a mix of recent onboarders (7/20, 35%) and those who had been enrolled in the program for longer (13/20, 65%). The participants were also asked about their comfort levels with technology and whether they had used a smart speaker before, and 90% (18/20) of users provided responses. Regarding comfort levels, of the 18 patients, 1 (6%) patient was very uncomfortable, 0 (0%) were somewhat uncomfortable, 6 (33%) were neutral, 2 (11%) were somewhat comfortable, and 9 (50%) were very comfortable. Regarding prior use of a smart speaker, of the 18 patients, 7 (39%) indicated that they had interacted with a smart speaker before, whereas the remainder (n=11, 61%) had not.

Quantitative Data

Engagement Levels and Accuracy Rates

The overall engagement level for the entire study population during the 4-week period was 73%, with noticeable drops in engagement as the weeks progressed (Table 1) and an overall decline of 14% when comparing the average engagement levels of weeks 1 and 4.

Table 1. Average engagement levels over the 4-week study duration.

Week	Engagement level ^a (%), average (SD)	Days missed, average (SD)
1	80.7 (11.3)	1.4 (0.11)
2	75.0 (5.8)	1.8 (0.06)
3	70.7 (7.9)	2.0 (0.08)
4	67.1 (8.1)	2.3 (0.08)

^aOverall average engagement level is 73.4% (SD.9.5%).

Over the 4-week duration (28 days), 9 entries (out of 411) were incorrect measurements submitted using the *Medly* voice app, indicating an overall accuracy rate of 97.8%. The errors varied between weight and blood pressure measurements. A subset (4/20, 20%) of participants was not able to successfully submit their correct readings, which led to the 9 errors that were recorded.

In addition to calculating the overall engagement levels, descriptive statistics were calculated, and the attributes mentioned previously were used to compare the results among the subgroups in the study population. The results are shown in [Multimedia Appendix 3](#). Although some trends were identified, the statistical tests indicated no significant differences between the groups.

There was no difference in the average engagement levels between the recently onboarded (n2) and existing *Medly* patients (n1; Mann-Whitney $U=45$, $n_1=13$, $n_2=7$; $P=.99$). Similar to the findings related to the entire study population, engagement levels were lower in the fourth week than in the first week for both groups. Average engagement levels increased as the age groups increased, with the oldest demographic (aged 61-80 years) having the best engagement level of 84.1%, approximately double the overall engagement level of the youngest age group in the study—Kruskal-Wallis test, $H(2)=3.8$ ($P=.14$). Those aged 61 to 80 years were the most consistent throughout the 4-week duration and had the smallest difference among the weekly average engagement levels.

A similar trend was observed when comparing participants based on their described comfort levels with technology

(statistical test results were not significant). Those who were very confident consistently used the technology more through the 4 weeks than those who reported less confidence, with a 13.6% overall difference (Mann-Whitney $U=23.5$, $n_1=6$, $n_2=12$; $P=.72$). There were also consistently higher engagement levels in the group that had never interacted with smart speakers before than in the group that had, with a 7.6% difference (Mann-Whitney $U=38$, $n_1=6$, $n_2=12$; $P=.86$). Both groups steadily declined in engagement as the weeks progressed, with similar overall differences between averages of weeks 1 and 4.

Acceptability of the Medly Voice App

Findings from the SUS questionnaire paired with those from the semistructured interviews were used to better understand the acceptability of using the voice app version of the *Medly* program.

The responses from the SUS questionnaire from the second week resulted in an overall average score of 69 (out of 100), ranking the voice app in the 53rd percentile based on previous studies. By contrast, the average score from the fourth week was 77 (out of 100), ranking it in the 80th percentile based on previous studies. These data indicated an overall increase in the level of satisfaction with using the *Medly* voice app (by 27%) in the study population. The difference in the averages for each individual question between weeks 2 and 4 was also calculated, with the last question in the survey having the biggest difference of 13%. The participants felt that as time went on, they needed to learn more things about the voice app to successfully interact with it (consistent with the NASA-TLX cognitive load results). Response distributions in the results of weeks 2 and 4 were fairly similar for all the questions (Figure S1 in [Multimedia Appendix 4](#)).

Average SUS scores were also calculated based on the different patient characteristics (age, *Medly* status, comfort levels, and familiarity with interacting with a smart speaker). Overall, the scores were similar in range for all the characteristics. However, the largest range in the data was identified in the age groups, with the oldest (61–80 age group) demographic providing the lowest score (72 out of 100), ranking it in the 62nd percentile, whereas the middle-aged demographic provided an average score of 87.5 (out of 100), ranking it in the 96th percentile. The average score from the youngest demographic was 77.5, ranking it in the 80th percentile. The Kruskal-Wallis test showed that these findings had no significant difference— $H(2)=0.89$ ($P=.64$).

Feasibility of the Medly Voice App

The NASA-TLX questionnaire was used in this study to better assess the workload perceived by the study participants when using the *Medly* voice app. A 4% increase was seen in the average scores between the results of weeks 2 and 4, indicating a slightly higher workload. Although the averages for each of the questions were fairly low, questions relating to (1) success rates; (2) how hard they needed to work to accomplish the task; and (3) feelings of discouragement, irritation, and stress scored worse than the rest of the questions. The results are shown in Figure S2 in [Multimedia Appendix 4](#). The participants also felt less successful with using the *Medly* voice app at the end of the

study than they did at the end of week 2 (22% difference in the results).

When analyzing the scores based on the different age groups, it was found that the youngest demographic felt that they needed to work the most (highest average of 2.67) when compared with the middle-aged (average of 1.61) and oldest demographics (average of 2.12); the results of the Kruskal-Wallis test was not insignificant— $H(2)=0.039$ ($P=.98$). It was also found that those who were newly onboarded to the *Medly* program felt more rushed when using the voice app and less successful when inputting their measurements as compared with those who had been on the *Medly* program for a longer time (approximately 15% difference in scores for each question); Mann-Whitney test was not significant ($U=25.5$, $n_1=12$, $n_2=6$; $P=.73$). The difference in the average scores for those who described themselves as less confident when using technology consistently gave poorer scores for each of the questions, indicating that they had a more difficult time than those who described themselves as confident; the Mann-Whitney test was also not significant ($U=11$, $n_1=12$, $n_2=6$; $P=.55$; Table S1 in [Multimedia Appendix 4](#)).

In summary, the descriptive statistics showed that the youngest age group felt that they needed to work the most, the study population collectively felt that they needed to put in slightly more effort as time went on, and those who were less familiar with technology had more difficulty using the voice app than those who were more confident.

The UTAUT2 questionnaire was used to better understand participants' thoughts regarding facilitating conditions, effort expectancy, habit, and behavioral intention when it came to using the voice app. The biggest difference between the results of weeks 2 and 4 was regarding whether they would use the *Medly* voice app in the future, with a 13% decline in the average score. The oldest demographic was the least keen on using it in the future, whereas the middle-aged demographic was the most interested in future use; the Kruskal-Wallis test indicated these results to be not statistically significant— $H(2)=1.88$ ($P=.39$). When asked whether the voice app became a habit, those who had used the technology before agreed more than those who had not (19% difference in the responses), although this test was also not statistically significant (Mann-Whitney $U=38$, $n_1=7$, $n_2=13$; $P=.86$).

Overall, all the participants felt that the voice app required low effort to use and that it was easy for them to operate. They were less certain about whether using the voice app had become a habit for them (this can be supported by engagement levels) and were least certain about whether they would use the voice app in the future, as shown in Table S2 in [Multimedia Appendix 4](#).

Qualitative Data

The interview themes were classified using implementation outcomes by Proctor et al [24], specifically focusing on the *feasibility* and *acceptability* constructs to answer the research question. The themes (1) feasibility of clinical integration and (2) voice app acceptability are presented in the subsequent sections, each with their own set of accompanying subthemes.

Feasibility of Clinical Integration

The feasibility of clinical integration was influenced by several factors; in our findings, the 2 subthemes of (1) users adapting to the voice app's conversational style and (2) device unreliability helped determine the potential that this technology has to be integrated into existing workflows and practice. Whether the users are able to adapt to the voice app and the extent to which the device is considered unreliable will identify the feasibility of the voice app being realistically used in the clinical environment. Further details regarding these 2 subthemes are provided in the subsequent sections.

Users Adapting to the Voice App's Conversational Style

Most participants found the device setup and instructions fairly straightforward but at times struggled to successfully log their measurements on the *Medly* voice app. When the participants struggled, they adjusted the way they spoke instead of continuing in their natural manner in hope that the voice app would understand them better:

I learned how to get into her rhythm as opposed to her getting into my rhythm. [Participant 04]

Specific strategies were used to change their speaking style, which most often involved modifying the volume, tone, pace, and style of their speech. Different strategies seemed to work better for different participants, specifically with the pace at which they spoke:

Now I just say 116.4 pounds (faster) and there's absolutely no issues with her now. [Participant 12]

Of course I would either make sure to be speaking directly at it or elevate my voice or something like that. [Participant 15]

I want to record one hundred, but it's very typical to say "a hundred" and not "one hundred," but I notice it doesn't pick up on that. [Participant 17]

Once the participants changed their conversational tone when speaking to the voice app, they began to notice difficulties in the interaction because it no longer felt like a natural conversation:

It's like when you talk to someone foreign or you know from another country or another language and you try to say a few words for them to understand it. [Participant 12]

I try to, like, separate each word, almost like I had to speak robotic. [Participant 18]

I have to be serious, slow and sure of how I say the numbers. [Participant 17]

Another interaction strategy adopted by most participants involved using the touchscreen capabilities of the device. In most cases, this alternative input was the favorable approach over using voice because it was simpler to use and, most importantly, faster:

I got into a routine which allowed me to go through it as quickly as possible, and that routine would be that I would speak the results for weight, blood pressure and heart rate, and then I would interact

directly on the touch screen for symptoms so we didn't have to wait for her. So yes, every time I use the touch screen it works fine and the fact that I could use a touch screen and it would work even though she hadn't finished speaking is a big plus for me. [Participant 15]

Interactions were found to be most successful when the participants did not multitask on other items:

You can multitask if you really want to, but that's what I think mistakes can be made easier. [Participant 02]

I knew the questions that were going to be asked after a while, but I still listened. Only because you know I'd rather do it right than wrong if I can. [Participant 08]

Despite the learning curve experienced by most participants, the mitigation strategies described earlier support the feasibility of deploying a voice app, such as *Medly*, in the clinic because of the perseverance displayed by these participants to make the interaction easier for themselves over time.

Device Unreliability

Almost all the participants experienced some level of difficulty when they interacted with the voice app. Sometimes, the voice app froze, and the session ended abruptly; at other times, it would not provide the user with an opportunity to correct any of the wrong measurements:

You can go back and correct it, right, but sometimes it gives you a little bit of a hassle so I have to start over. [Participant 02]

Then she just shut down...When she couldn't get the measurements or something, she would just turn off. [Participant 04]

The participants also described instances where the voice app was unable to correctly pick up the information they were saying, making them feel frustrated, annoyed, panicked, and discouraged to the point where they no longer wanted to use the device that day:

Yeah, I'd wake up in a great mood and oftentimes it was so frustrating that it made me cranky afterwards. Yeah, it really switched my mood. One time she repeated it to me and I thought she got it alright and then she repeated it and said that I fainted and I had not fainted, so I panicked. [Participant 18]

When the voice app was unable to pick up the correct measurements, the participants often felt the need to speak louder. This was considered to be problematic specifically in situations where a participant may not be feeling well and does not have the ability to project their voice. As explained by one of the participants, with the smartphone, they were able to share information without needing to exert a lot of energy:

I would never want it to not be on my phone when I go into the hospital and I have a hard time talking. If my blood pressure is through the roof or it's way too low from retaining water, it's so hard to speak and I love that I could just throw my phone at the

doctor and be like “look, this was [my data] two days ago”...I really like that. [Participant 18]

Although the voice app seemed feasible to deploy from a patient interaction perspective, the users also experienced difficulties when interacting with the device for various tech-related reasons. Understanding the causes and frequencies of these malfunctions will help identify when and where it is appropriate to use voice apps such as *Medly*.

Voice App Acceptability

This theme described the extent to which the study participants found the *Medly* voice app satisfactory. This level of acceptability included not only the participants' thoughts but also other factors that may have influenced their experience, as described by the following subthemes: (1) the device integrated well within household and users' lives, (2) the users blamed themselves when problems arose with the voice app, and (3) the voice app was missing specific features desired by the users.

Device Integration in the Household

In addition to using the device to access the *Medly* voice app, many participants also found that they used it for other purposes during their time in the study. Over the 4 weeks, some participants described the device as a companion, with one of the participants noting the following:

She became like a buddy. I know it's little quirks, specifically when it makes mistakes...I would say for people that live on their own or whatever it can become like a friend, right? [Participant 08]

Some participants also described their experience interacting with the device as “pleasant,” and others specifically felt the need to use manners and be polite while conversing with it:

And I've gotten along with Alexa just fine. It was so cute. I was inputting on Medly and I did it with Alexa at the same time and at the end I said “Alexa, thank you” and she said “you bet”...One night I said, “oh Alexa goodnight” and she said “night night, sleep well.” [Participant 08]

The device became a companion not only for the users but also for their family members and friends:

She did give my granddaughter a knock knock joke the other night. [The grandkids] have fun with her by asking what the weather is or something like that. [Participant 10]

This interaction is an example of how easily the device can fit in and become integrated within a space in the household. While in common areas, the users have noted using the device for other activities, such as the following:

I let it play music for me or I ask what's the weather like today and I do the CTV News first thing in the morning, so yeah, I think it's a great thing. [Participant 02]

Having the device in common spaces also served as a reminder for some participants who had difficulty remembering to perform their *Medly* measurements. Others also mentioned that

because the device was placed in a common space, they would be more inclined to use *Medly* on it:

Seeing the monitor right there on the counter I feel like it definitely encourages and motivates me and is a visual reminder as opposed to the app on the phone to actually do it. [Participant 11]

At first I thought it would be my phone. But probably you know, now it's Alexa. She sits right there, so probably Alexa. [Participant 02]

Some participants also placed the device in other places in their house, such as the bedroom. In these cases as well, they found the setup useful:

I use it at night time when I'm going to bed like you know, relaxing music. [Participant 06]

Furthermore, in some cases, the voice app was more preferred when compared with the smartphone:

I'm in my bedroom and I have a bathroom in the room, so when I go in the bathroom to weigh myself, I do my blood pressure at the same time. So ideally that is where I talk to [Alexa]...over the last week it's been working and I really like that because then I'm done and then I can go right back to bed after I take my pills so it doesn't make my mind wake up. [Participant 03]

I'm sort of having concussion symptoms and the phone makes me nauseous. So at the moment, I prefer only having to do it with Alexa. [Participant 14]

Despite the benefits of the device integrating well within different spaces in the household, there are drawbacks that can exist when keeping the device in a public space. Most participants noted the importance of having a quiet space to focus and successfully submit their readings:

Honestly like I did it more often when I didn't have my son because everything here he likes to speak over me...He would repeat 'Alexa' behind me. [Participant 18]

Like if my husband would walk into the kitchen as I was doing it, I would shoo him away, literally. [Participant 08]

Users Blamed Themselves When Problems Arose With the Voice App

Although some participants experienced frustration when the device abruptly stopped working or incorrectly heard them, often times (especially in the first week), the users felt that it was their fault when a mistake happened:

I wasn't annoyed by it. I just thought, oh, I'm not speaking clearly or loudly enough, or you know. [Participant 08]

Well again, I go back to the learning curve in the first week. There was some frustration, but you can't blame that on Alexa, that was all me. [Participant 05]

These reflections indicated that the users were generally understanding of the voice app and had some patience when interacting with it.

Missing but Desired Voice App Features

The participants shared some of the features they valued in devices that programs such as *Medly* can be offered on. In particular, the users preferred to interact with a device that is fast and can quickly record their data for the day. In some instances, the users compared the capability of the voice app with that of Bluetooth, indicating that Bluetooth is a much faster and simpler process:

It's just really cumbersome, like the whole process. And I guess part of that is because the [smartphone] app is so easy. And I think it could get even easier if I got the Bluetooth blood pressure and scale. [Participant 04]

To me, honestly, because they want it in the morning, the smartphone is much faster. [Participant 09]

Most users also expressed concern about how they would use the voice app should they go on an overnight trip. A device that is small enough to be portable when traveling was desired and often mentioned:

The only thing I don't like about it is it is big and bulky so it is not something I would be too inclined to want to travel with. So yeah, so for me the mobility issue would be a bit of a concern if I had to rely on it. [Participant 13]

Discussion

Principal Findings

This manuscript presents the findings from a proof-of-concept implementation study for a voice app designed for patients with HF using a mixed methods approach. To our knowledge, this is the first evaluation of a voice app used for helping patients manage an advanced chronic condition at home. To date, studies have only reported on accuracy and acceptability levels in a controlled laboratory environment; however, these findings are still consistent with the results presented in this paper [11,14]. Although the SUS scores were higher in week 4 than in week 2, engagement levels declined by 14% between the start and end of the study. The participants felt that they needed to use a higher cognitive load in week 4 than in week 2 (4% increase), and the average rating regarding whether they would use it in the future decreased by 13%. An accuracy rate of 97.8% indicates that the participants were able to successfully log their measurements most of the time, which may have led to the higher SUS score. Some qualitative findings can be potential reasons why engagement levels declined. In particular, from a feasibility perspective, the device was at times unreliable, and the users had to work (to varying efforts) to adapt to the flow of the conversation. Although this may have been tolerable in the first few weeks, over time, it may have become tiresome, depending on how quickly the users adapted. Similarly, because the users often blamed themselves when mistakes arose, this could have created a negative association with the voice app, and over time, the users may have begun to feel discouraged from using it.

To better understand the voice app's acceptability and feasibility of implementation, we sought to identify any noticeable

differences between the participants in terms of engagement levels. Although our quantitative data are not statistically significant, our observed findings are similar to those presented by Ware et al [31], namely the finding that engagement levels were highest in the older age group and progressively lower in the younger age groups. This finding is also consistent with other research that specifically focused on the use of voice-based conversational agents among the older adult population [4,32-37]. Although the oldest group had the highest engagement levels, the middle-aged demographic (aged 41-60 years) had the highest average SUS score, indicating that they were the most accepting of the voice app. Although we cannot conclude any findings definitely based on these observations, it provides a starting point for future work.

One of the most common responses provided by the participants during interviews was the notion that the voice app takes a long time to complete and, in particular, takes longer than the *Medly* smartphone app. The users often described being rushed out the door in the mornings, in which case they appreciated being able to use the smartphone app to quickly input their measurements. This type of lifestyle and response was observed less with the older demographics, who generally seemed to have more patience and understanding when interacting with the voice app. There were also specific cases in which the voice app actually proved to be more useful than the smartphone. One of the participants was experiencing concussion-type symptoms and, as a result, had limited screen time, so the voice app worked well for them. Another participant often felt fatigue as one of the side effects of their medications and experienced difficulties navigating the *Medly* smartphone app in the mornings. In this case, they also appreciated how much easier it was to perform the required tasks using the *Medly* voice app. Similar sentiments were echoed by other participants who realized that they can successfully record their readings when speaking in a relaxed, nonstrenuous manner. Although this worked well for some participants, one of the participants in a similar situation had a different experience, specifically because the voice app was unable to decipher their speech when they were feeling unwell owing to their weak and fragile voice. As a result, further advancements are required to better recognize sound, specifically when users are unable to exert large amounts of energy while speaking. Similar technical limitations have also been outlined in other studies on voice apps [15].

The findings from this study also show how well integrated the device became in many households and the potential benefits this may have for participants. Owing to the versatility of the device, it quickly became a part of many users' daily routines, from listening to music to asking for dinner recipes, and even started turning into a companion. Not only did the device provide social support, but it also served as a visual reminder to perform their *Medly* measurements. A participant noted that they would be more inclined to use the *Medly* voice app simply because it was in a common space they frequent in their house. Therefore, the natural integration of the device into users' lives over the 4 weeks shows the possibility that it may make it more convenient for some to perform their *Medly* measurements and may encourage and motivate others who often forget.

These findings help begin to uncover the “profile” of the patient demographic this technology would be most suitable for. We suspect that those who are older adults (aged >60 years), feel more confident in using technology, and have less busy schedules have an easier time, are more successful, and are consistent when interacting with the voice app. In addition, those with multimorbidity can benefit from using this platform, especially because of the common side effects they may experience from their conditions.

Comparison With Prior Work

To our knowledge, this study is part of only a few studies that have investigated the use of a voice app for a chronic disease in the intended environment for a prolonged period (4-week duration). Similarly, this work is one of the firsts to study a voice app that is designed to be personalized to individual patients (output responses depend on the parameters set when the patient is onboarded to the program). A systematic review performed by Bérubé et al [38] specifically focused on voice-based conversational agents for chronic health conditions and found only 2 voice apps designed as conversational agents for HF [39,40]. Both studies were primarily focused on the system architecture and accuracy of speech recognition, and one of the studies relied on the smartphone to implement the voice-based assistant. Other studies have focused on the acceptance and feasibility of voice apps for HF through preliminary assessments, such as survey responses based on usability studies performed in controlled environments [11,12]; the results of all these studies showed the promise that this technology has in the field of chronic disease management, especially for HF. Finally, 2 more recent studies investigated the engagement [13] and feasibility [16] of an HF-related voice app for a longer duration (90 days). The study performed by Apergi et al [13] showed higher engagement with the older patient cohort (similar to this study’s results), and Shara et al [16] reported favorable perception and high comfort levels in their study population. This study begins to uncover the potential that a voice app platform has for a program, such as *Medly*, and provides a basis for future work to explore who may benefit the most from this platform and why.

Limitations

Multiple limitations were identified over the course of the study and, as a result, should be acknowledged to better understand the impact of the findings.

First, because there were numerous questionnaires and interviews, the study team was mindful of the potential for social

desirability bias [41]. As a result, the participants were encouraged to speak honestly and were given the opportunity to disclose their thoughts through questionnaires privately instead of over the phone. Second, because this study was a proof of concept for a voice app in its intended environment, the sample size was not statistically powered, and most of the findings were interpreted in a qualitative manner. Future work should design studies with statistical significance (including using a validated questionnaire to capture user comfort levels with technology) to better understand who this may be most beneficial for. Third, specific study factors could have impacted the participant’s thoughts, experiences, and feedback. The users were aware that the study duration was only a 4-week period and, as a result, may have had higher engagement levels than if they were asked to use the voice app for a longer period. The participants were also required to perform a double entry of their measurements; the study results may have differed if users were only required to use the voice app. Fourth, because the inclusion criteria were general enough to include any patient enrolled in the program, selection bias likely occurred during recruitment. In this case, there may have been missed opportunities to include a greater variety of demographics in the study, especially those who primarily spoke languages other than English. Finally, because most participants in this study had never interacted with a smart speaker before, their thoughts and feedback may have been influenced by the fact that they were interacting with a novel technology. As a result, their thoughts on the device itself could be reflected in their responses, even though any VUI device could have been used in the study.

Conclusions

This study used a mixed methods approach to investigate the acceptability and feasibility of deploying a voice app for digital therapeutics used in chronic disease management. Overall, our findings conclude that the participants were largely successful in using the *Medly* voice app despite some of the barriers faced, proving that an app such as this could be feasible to be deployed in the clinic for future use. Our data begin to piece together the patient profile that this technology may be most suitable for. Future work should involve a statistically powered study that investigates the following demographics: those who are older (>60 years), have less busy schedules, exhibit high confidence levels when using technology, or experience symptoms (such as fatigue or headaches) from chronic conditions.

Acknowledgments

The authors wish to thank the patients who participated in this study. They thank the *Medly* nurse coordinators Mary O’Sullivan, Sarvati Bhatt, Eva Pavic, Tina Carriere, and Annabelle Fontanilla for helping with the recruitment process. They also wish to express their gratitude to Quynh Pham and Patrick Ware for helping guide this research project’s methodology as well as to Cait Nuun, Madison Taylor and Denise Ng for their research ethics board expertise and guidance.

Conflicts of Interest

JAC and HR are part of the team that founded the *Medly* system under the intellectual property policies of the University Health Network and may benefit from future commercialization of this technology.

Multimedia Appendix 1

Instruction manual for the *Medly* voice app.

[[DOCX File , 492 KB - formative_v6i12e40021_app1.docx](#)]

Multimedia Appendix 2

Baseline questionnaire for the participants.

[[DOCX File , 22 KB - formative_v6i12e40021_app2.docx](#)]

Multimedia Appendix 3

Overall and weekly average engagement levels based on various patient characteristics.

[[DOCX File , 23 KB - formative_v6i12e40021_app3.docx](#)]

Multimedia Appendix 4

Data showcasing the positive (a) and negative (b) attribute question and results from the System Usability Scale questionnaire, with week 2 data on top and week 4 data on the bottom (Figure S1). National Aeronautics and Space Administration (NASA)-Task Load Index score distributions in the results of weeks 2 and 4 (top and bottom, respectively; Figure S2). Average scores for each NASA-Task Load Index question (Table S1). Average scores for each of the constructs from the Unified Theory of Acceptance and Use of Technology 2 questionnaire (Table S2).

[[DOCX File , 134 KB - formative_v6i12e40021_app4.docx](#)]

References

1. Noncommunicable diseases. World Health Organization. 2022 Sep 16. URL: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases> [accessed 2022-05-30]
2. Dickson VV, Clark RA, Rabelo-Silva ER, Buck HG. Self-care and chronic disease. *Nurs Res Pract* 2013;2013:827409-827402 [[FREE Full text](#)] [doi: [10.1155/2013/827409](https://doi.org/10.1155/2013/827409)] [Medline: [24349772](https://pubmed.ncbi.nlm.nih.gov/24349772/)]
3. Seto E, Leonard K, Cafazzo J, Masino C, Barnsley J, Ross H. Self-care and quality of life of heart failure patients at a multidisciplinary heart function clinic. *J Cardiovasc Nurs* 2011;26(5):377-385. [doi: [10.1097/JCN.0b013e31820612b8](https://doi.org/10.1097/JCN.0b013e31820612b8)] [Medline: [21263339](https://pubmed.ncbi.nlm.nih.gov/21263339/)]
4. Ware P, Ross HJ, Cafazzo JA, Laporte A, Seto E. Implementation and evaluation of a smartphone-based telemonitoring program for patients with heart failure: mixed-methods study protocol. *JMIR Res Protoc* 2018 May 03;7(5):e121 [[FREE Full text](#)] [doi: [10.2196/resprot.9911](https://doi.org/10.2196/resprot.9911)] [Medline: [29724704](https://pubmed.ncbi.nlm.nih.gov/29724704/)]
5. Irfan Khan A, Gill A, Cott C, Hans PK, Steele Gray C. mHealth tools for the self-management of patients with multimorbidity in primary care settings: pilot study to explore user experience. *JMIR Mhealth Uhealth* 2018 Aug 28;6(8):e171 [[FREE Full text](#)] [doi: [10.2196/mhealth.8593](https://doi.org/10.2196/mhealth.8593)] [Medline: [30154073](https://pubmed.ncbi.nlm.nih.gov/30154073/)]
6. Mahmood A, Kedia S, Wyant DK, Ahn S, Bhuyan SS. Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digit Health* 2019 Oct 10;5:2055207619882181. [doi: [10.1177/2055207619882181](https://doi.org/10.1177/2055207619882181)] [Medline: [31656632](https://pubmed.ncbi.nlm.nih.gov/31656632/)]
7. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc* 2018 Sep 01;25(9):1248-1258 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
8. Griffin AC, Xing Z, Khairat S, Wang Y, Bailey S, Arguello J, et al. Conversational agents for chronic disease self-management: a systematic review. *AMIA Annu Symp Proc* 2020;2020:504-513 [[FREE Full text](#)] [Medline: [33936424](https://pubmed.ncbi.nlm.nih.gov/33936424/)]
9. Alnefaie A, Singh S, Kocaballi B, Prasad M. An overview of conversational agent: applications, challenges and future directions. In: *Proceedings of the 17th International Conference on Web Information Systems and Technologies - WEBIST. 2021 Presented at: The 17th International Conference on Web Information Systems and Technologies - WEBIST,; Oct 26-28, 2021; Online.* [doi: [10.5220/0010708600003058](https://doi.org/10.5220/0010708600003058)]
10. Brownstein J, Lannon J, Lindenauer S. 37 Startups building voice applications for healthcare. *MobiHealthNews*. 2018 Aug 3. URL: <https://www.mobihealthnews.com/content/37-startups-building-voice-applications-healthcare> [accessed 2022-05-30]
11. Jadczyk T, Kiwic O, Khandwalla RM, Grabowski K, Rudawski S, Magaczewski P, et al. Feasibility of a voice-enabled automated platform for medical data collection: CardioCube. *Int J Med Inform* 2019 Sep;129:388-393. [doi: [10.1016/j.ijmedinf.2019.07.001](https://doi.org/10.1016/j.ijmedinf.2019.07.001)] [Medline: [31445282](https://pubmed.ncbi.nlm.nih.gov/31445282/)]
12. Dojchinovski D, Ilijevski A, Gusev M. Interactive home healthcare system with integrated voice assistant. In: *Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). 2019 Presented at: 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO); May 20-24, 2019; Opatija, Croatia.* [doi: [10.23919/mipro.2019.8756983](https://doi.org/10.23919/mipro.2019.8756983)]
13. Apergi LA, Bjarnadottir MV, Baras JS, Golden BL, Anderson KM, Chou J, et al. Voice interface technology adoption by patients with heart failure: pilot comparison study. *JMIR Mhealth Uhealth* 2021 Apr 01;9(4):e24646 [[FREE Full text](#)] [doi: [10.2196/24646](https://doi.org/10.2196/24646)] [Medline: [33792556](https://pubmed.ncbi.nlm.nih.gov/33792556/)]

14. Cheng A, Raghavaraju V, Kanugo J, Handrianto Y, Shang Y. Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In: Proceedings of the 15th IEEE Annual Consumer Communications & Networking Conference (CCNC). 2018 Presented at: 15th IEEE Annual Consumer Communications & Networking Conference (CCNC); Jan 12-15, 2018; Las Vegas, NV, USA. [doi: [10.1109/ccnc.2018.8319283](https://doi.org/10.1109/ccnc.2018.8319283)]
15. Sezgin E, Noritz G, Elek A, Conkol K, Rust S, Bailey M, et al. Capturing at-home health and care information for children with medical complexity using voice interactive technologies: multi-stakeholder viewpoint. *J Med Internet Res* 2020 Feb 13;22(2):e14202 [FREE Full text] [doi: [10.2196/14202](https://doi.org/10.2196/14202)] [Medline: [32053114](https://pubmed.ncbi.nlm.nih.gov/32053114/)]
16. Shara N, Bjarnadottir MV, Falah N, Chou J, Alqutri HS, Asch FM, et al. Voice activated remote monitoring technology for heart failure patients: study design, feasibility and observations from a pilot randomized control trial. *PLoS One* 2022 May 6;17(5):e0267794 [FREE Full text] [doi: [10.1371/journal.pone.0267794](https://doi.org/10.1371/journal.pone.0267794)] [Medline: [35522660](https://pubmed.ncbi.nlm.nih.gov/35522660/)]
17. What is cardiovascular disease? American Stroke Association. URL: <https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease> [accessed 2022-10-04]
18. Kannel WB, Belanger AJ. Epidemiology of heart failure. *Am Heart J* 1991 Mar;121(3):951-957. [doi: [10.1016/0002-8703\(91\)90225-7](https://doi.org/10.1016/0002-8703(91)90225-7)]
19. What is heart failure? American Heart Association. URL: <https://www.heart.org/en/health-topics/heart-failure/what-is-heart-failure> [accessed 2022-02-17]
20. Heart F. Heart failure. American Heart Association. URL: <https://www.heart.org/en/health-topics/heart-failure> [accessed 2022-10-04]
21. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. *Eur J Heart Fail* 2020 Aug;22(8):1342-1356 [FREE Full text] [doi: [10.1002/ejhf.1858](https://doi.org/10.1002/ejhf.1858)] [Medline: [32483830](https://pubmed.ncbi.nlm.nih.gov/32483830/)]
22. Ware P, Ross HJ, Cafazzo JA, Boodoo C, Munnery M, Seto E. Outcomes of a heart failure telemonitoring program implemented as the standard of care in an outpatient heart function clinic: pretest-posttest pragmatic study. *J Med Internet Res* 2020 Feb 08;22(2):e16538 [FREE Full text] [doi: [10.2196/16538](https://doi.org/10.2196/16538)] [Medline: [32027309](https://pubmed.ncbi.nlm.nih.gov/32027309/)]
23. Barbaric A, Munteau C, Ross H, Tong K, Cafazzo J. Design, development, and usability evaluation of a voice app experience for heart failure management. *JMIR Formative Res* 2022 Oct. [doi: [10.2196/41628](https://doi.org/10.2196/41628)]
24. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar 19;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
25. Johanson GA, Brooks GP. Initial scale development: sample size for pilot studies. *Educ Psychol Measurement* 2009 Dec 18;70(3):394-400. [doi: [10.1177/0013164409355692](https://doi.org/10.1177/0013164409355692)]
26. System Usability Scale (SUS). usability.gov. URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> [accessed 2022-05-30]
27. Hart S, Staveland L. Development of NASA-TLX (task load index): results of empirical and theoretical research. *Human Mental Workload* 1988;52:139-183.
28. Nordhoff S, Louw T, Innamaa S, Lehtonen E, Beuster A, Torrao G, et al. Using the UTAUT2 model to explain public acceptance of conditionally automated (L3) cars: a questionnaire study among 9,118 car drivers from eight European countries. *Transportation Res Part F Traffic Psychol Behav* 2020 Oct;74:280-297. [doi: [10.1016/j.trf.2020.07.015](https://doi.org/10.1016/j.trf.2020.07.015)]
29. Yu CH. *Organizational Res Method* 2008 Aug 15;12(4):801-804. [doi: [10.1177/1094428108318066](https://doi.org/10.1177/1094428108318066)]
30. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. *Glob Qual Nurs Res* 2017 Nov 24;4:2333393617742282 [FREE Full text] [doi: [10.1177/2333393617742282](https://doi.org/10.1177/2333393617742282)] [Medline: [29204457](https://pubmed.ncbi.nlm.nih.gov/29204457/)]
31. Ware P, Dorai M, Ross HJ, Cafazzo JA, Laporte A, Boodoo C, et al. Patient adherence to a mobile phone-based heart failure telemonitoring program: a longitudinal mixed-methods study. *JMIR Mhealth Uhealth* 2019 Feb 26;7(2):e13259 [FREE Full text] [doi: [10.2196/13259](https://doi.org/10.2196/13259)] [Medline: [30806625](https://pubmed.ncbi.nlm.nih.gov/30806625/)]
32. Demiris G, Rantz M, Skubic M, Aud M, Tyrer H. Home-based assistive technologies for elderly: attitudes and perceptions. *AMIA Annu Symp Proc* 2005:935 [FREE Full text] [Medline: [16779222](https://pubmed.ncbi.nlm.nih.gov/16779222/)]
33. Demiris G, Rantz MJ, Aud MA, Marek KD, Tyrer HW, Skubic M, et al. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. *Med Inform Internet Med* 2004 Jun 12;29(2):87-94. [doi: [10.1080/14639230410001684387](https://doi.org/10.1080/14639230410001684387)] [Medline: [15370989](https://pubmed.ncbi.nlm.nih.gov/15370989/)]
34. Choi Y, Demiris G, Thompson H. Feasibility of smart speaker use to support aging in place. *Innov Aging* 2018:560.
35. Rodríguez MD, Beltrán J, Valenzuela-Beltrán M, Cruz-Sandoval D, Favela J. Assisting older adults with medication reminders through an audio-based activity recognition system. *Pers Ubiquit Comput* 2020 Jun 09;25(2):337-351. [doi: [10.1007/s00779-020-01420-4](https://doi.org/10.1007/s00779-020-01420-4)]
36. Pollack M. Intelligent technology for an aging population: the use of AI to assist elders with cognitive impairment. *AI Magazine* 2005;26(2):9-24.
37. Shalini S, Levins T, Robinson E, Lane K, Park G, Skubic M. Development and Comparison of Customized Voice-Assistant Systems for Independent Living Older Adults. Cham: Springer; 2019.

38. Bérubé C, Schachner T, Keller R, Fleisch E, V Wangenheim F, Barata F, et al. Voice-based conversational agents for the prevention and management of chronic and mental health conditions: systematic literature review. *J Med Internet Res* 2021 Mar 29;23(3):e25933 [FREE Full text] [doi: [10.2196/25933](https://doi.org/10.2196/25933)] [Medline: [33658174](https://pubmed.ncbi.nlm.nih.gov/33658174/)]
39. Galescu L, Allen J, Ferguson G, Quinn J, Swift M. Speech recognition in a dialog system for patient health monitoring. In: *Proceedings of the 2009 IEEE International Conference on Bioinformatics and Biomedicine Workshop*. 2009 Presented at: 2009 IEEE International Conference on Bioinformatics and Biomedicine Workshop; Nov 01-04, 2009; Washington, DC, USA. [doi: [10.1109/bibmw.2009.5332111](https://doi.org/10.1109/bibmw.2009.5332111)]
40. Lobo J, Ferreira L, Ferreira A. CARMIE: a conversational medication assistant for heart failure. In: *International Journal of E-Health and Medical Communications (IJEHMC)*. Pennsylvania, United States: IGI Global; 2017.
41. Grimm P. Social desirability bias. In: *Wiley International Encyclopedia of Marketing*. Hoboken, New Jersey, United States: Wiley; 2010.

Abbreviations

HF: heart failure

NASA: National Aeronautics and Space Administration

SUS: System Usability Scale

TLX: Task Load Index

UHN: University Health Network

UTAUT2: Unified Theory of Acceptance and Use of Technology 2

VUI: voice user interface

Edited by A Mavragani; submitted 13.06.22; peer-reviewed by Z Xing, I Madujibeya; comments to author 22.09.22; revised version received 13.10.22; accepted 26.10.22; published 21.12.22.

Please cite as:

Barbaric A, Munteanu C, Ross H, Cafazzo JA

A Voice App Design for Heart Failure Self-management: Proof-of-Concept Implementation Study

JMIR Form Res 2022;6(12):e40021

URL: <https://formative.jmir.org/2022/12/e40021>

doi: [10.2196/40021](https://doi.org/10.2196/40021)

PMID: [36542435](https://pubmed.ncbi.nlm.nih.gov/36542435/)

©Antonia Barbaric, Cosmin Munteanu, Heather Ross, Joseph A Cafazzo. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 21.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Mobile Phone Text Messaging Intervention to Manage Fatigue for People With Multiple Sclerosis, Spinal Cord Injury, and Stroke: Development and Usability Testing

Kerri A Morgan^{1*}, OTR/L, ATP, PhD; Alex W K Wong^{2,3,4*}, PhD, DPhil; Kim Walker^{1*}, OTR/L, ATP, OTD; Rachel Heeb Desai^{1*}, OTR/L, OTD; Tina M Knepper^{1*}, HA/HIM, BS; Pamela K Newland^{5*}, CMSRN, RN, PhD

¹Program in Occupational Therapy, St. Louis School of Medicine, Washington University, St. Louis, MO, United States

²Center for Rehabilitation Outcomes Research, Shirley Ryan AbilityLab, Chicago, IL, United States

³Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

⁴Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

⁵Goldfarb School of Nursing, Barnes Jewish College, St. Louis, MO, United States

* all authors contributed equally

Corresponding Author:

Kerri A Morgan, OTR/L, ATP, PhD

Program in Occupational Therapy

St. Louis School of Medicine

Washington University

5200 Berthold Avenue Suite B

St. Louis, MO, 63110

United States

Phone: 1 314 286 1659

Fax: 1 314 531 3985

Email: morgank@wustl.edu

Abstract

Background: Fatigue significantly affects daily functioning in persons with disabilities. Fatigue management can be challenging, and the information provided during routine physician visits to manage fatigue can be overwhelming. One way to address fatigue is to increase knowledge, skills, and confidence for self-management (ie, patient activation). Self-management programs have shown promising effects in targeting fatigue in persons with disabilities. However, satisfaction with self-management programs is low for persons with disabilities, and tailoring interventions to personalized needs has been recommended. SMS text messaging is increasingly being used to implement health behavior change interventions in a person's natural environment. Little has been done to link mobile health approaches with patient activation and self-management to address fatigue in persons with disabilities.

Objective: This study aimed to develop and test a mobile phone-based fatigue self-management SMS text messaging intervention targeting patient activation in 3 groups of persons with disabilities: persons with multiple sclerosis, persons who had a stroke, and persons with a spinal cord injury.

Methods: We used evidence-based resources and input from a consumer advisory board (CAB; composed of 2 participants from each of the 3 disability groups) and a neurologist to develop the intervention. The study was conducted using a 4-step process: development of the initial SMS text messaging library and categorization of the content into 9 content areas, review and modification of the SMS text messages by the neurologist and CAB, integration of the content library into a digital platform, and utility testing by CAB members.

Results: A total of 6 CAB participants rated SMS text messages covering 9 domain areas of fatigue self-management with good clarity (mean ratings=3.5-5.0 out of 5) and relevance (mean ratings=3.2-5.0 out of 5). Overall, SMS text messaging content was reported by CAB participants as helpful, clear, and well suited for a mobile health intervention. The CAB reached consensus on the time of day that SMS text messages should be sent (morning) and their frequency (once per day). This feedback led the research team to narrow down the program to deliver 48 SMS text messages, 1 per day, Monday through Thursday only, a total of 4 SMS text messages per week, over a 12-week period. The final set of SMS text messages was programmed into a digital platform with a predefined delivery schedule. The usability of the intervention was high, with 55 (83%) out of 66 responses endorsing the highest rating.

Conclusions: This study demonstrates a step-by-step process for developing a fatigue self-management SMS text messaging intervention for persons with disabilities. For this population, whose access to health services is often limited, this intervention provides an alternative delivery model to increase access to fatigue information and deliver content that aligns with the person's needs.

(*JMIR Form Res* 2022;6(12):e40166) doi:[10.2196/40166](https://doi.org/10.2196/40166)

KEYWORDS

fatigue; disability; mobile health; mHealth; patient activation

Introduction

Background

Fatigue significantly impacts the daily functioning of persons with disabilities [1] and decreases participation in major life activities [2-4]. Persons with disabilities reported fatigue as a common and burdensome symptom. Approximately 50% to 80% of people with multiple sclerosis (MS), a spinal cord injury (SCI), or a stroke experience fatigue [2,5,6], a rate 2 to 3 times more prevalent than in the general population and significantly higher than in older adults with other medical conditions [7]. Particularly for persons with disabilities, fatigue is related to other symptoms such as sleep or pain [8] and negatively impacts psychological well-being and quality of life [9-11]. Fatigue also leads to absenteeism and affects the work and productivity of individuals with physical disabilities and chronic disease [12-14]. Fatigue is estimated to cost employers US \$136 billion annually in health-related lost work and productivity in the general population [15].

The management of chronic symptoms can be challenging. Self-management programs have shown promising effects in improving self-efficacy and the ability to manage symptoms of persons with disabilities [1,16,17]. Although fatigue is common and distressing across a wide range of chronic disability conditions, no fatigue self-management intervention can be used across different disability groups [18]. A previous report indicated that fatigue management could benefit from a general transdiagnostic approach, with the goal of focusing on individual needs rather than a specific disease [19]. Thus, developing a general fatigue self-management intervention that can be applied across various groups may benefit a larger population living with various disabilities. Satisfaction with current self-management interventions for persons with disabilities is low because the experience of fatigue is unique to the patient experience, and tailoring toward personalized needs has thus been recommended [20]. One possible solution to address these challenges is to increase patient activation—knowledge, skills, and confidence for self-management [21]. Evidence suggests that increasing levels of patient activation improve health outcomes and care experiences and reduce health care costs [22-24]. Mobile health (mHealth) tools, especially SMS text messages using mobile phones, appear to be effective in improving activation and self-management behaviors in chronic disabling conditions [25-27]. SMS text messaging interventions are also increasingly being used to implement self-management programs for persons with disabilities because of their high reach, high accessibility, and relatively low-cost communication strategies [16,28-32]. However, there is limited research linking

mHealth approaches with patient activation and self-management to address fatigue in persons with disabilities.

Objective

This study aimed to describe the process of developing and testing the utility of a fatigue self-management SMS text messaging intervention based on patient activation for persons with disabilities. Engaging end users in designing a fatigue self-management SMS text messaging intervention in the early phase of technology development may improve user experiences, accessibility, and usability of the technology [33]. Thus, the design of this study incorporated individuals with MS, SCI, and stroke to provide valuable feedback for developing the fatigue self-management SMS text messaging intervention.

Methods

Participants

This study included an interdisciplinary investigator team with unique clinical expertise in MS, SCI, and stroke. All investigators had experience in disability research to develop the fatigue management SMS text messaging intervention for persons with disabilities. The investigator team organized a consumer advisory board (CAB) that included 2 persons from each disability group (MS, SCI, and stroke). Inclusion criteria for CAB participants were as follows: (1) aged >18 years, (2) had a disability for at least 1 year, (3) reported fatigue that was well managed, (4) had the ability to read and speak English at an eighth-grade level, and (5) were willing to use their own mobile phone and SMS to trial the SMS text messaging intervention. Exclusion criteria included the following: (1) evidence of an acute condition (eg, relapse), (2) sleep apnea, (3) inability to provide consent, (4) terminal cancer, and (5) pregnancy. We used purposive sampling to maximize different ages, genders, races, and ethnicities in an attempt to represent a diverse spectrum of persons with disabilities. The final sample was composed of 6 persons with disabilities, which was deemed by the investigator team as large enough for utility testing but small enough to conduct advisory board sessions in a focus group format [34,35]. A larger sample size is required for future pilot testing. This study also included a physiatrist who had provided medical rehabilitation care for >20 years to provide feedback as an expert health care provider on developing the content and format of the fatigue self-management SMS text messaging intervention.

Procedures

Overview

This study followed a user-centric co-design approach, meaning that participants helped to shape the intervention and their study experience [36]. We used step-by-step procedures to develop and test the utility of the fatigue self-management SMS text messaging intervention. The steps included the following: (1) development of the initial SMS text messaging library for the self-management intervention, (2) review and modification of SMS text messages, (3) integration of the content library into the SMS text messaging system, and (4) utility testing of the SMS text messaging prototype.

Step 1: Development of Initial Text Messaging Library for Fatigue Self-management

The investigator team formed an initial library of SMS text messages based on evidence-based resources from the National MS Society [37,38], the American Stroke Association [39], and the National Spinal Cord Injury Association [40]. The content collected was assessed by the research team and categorized into 9 key content areas for the library of targeted fatigue self-management SMS text messages (eg, sleep, energy conservation, and simplifying activities). The investigator team, composed of 3 licensed occupational therapists, also reviewed the content for accuracy and appropriateness. The investigator team further customized SMS text messages into patient activation measure (PAM) levels 1 to 2 and 3 to 4, where level 1 to 2 SMS text messages focused on building knowledge and increasing awareness and level 3 to 4 SMS text messages focused on increasing skills or maintaining behaviors for self-management [41]. The 4 PAM levels were collapsed into 2 levels to balance tailoring with feasibility (ie, the development of 2 series of SMS text messages was more achievable than that of 4, given our study time frame and funding). Furthermore, collapsing of the PAM levels has been suggested in previous studies based on psychometric assessments.

Step 2: Review and Modification of SMS Text Messages

Health Care Expert Opinion

A psychiatrist with extensive care experience with persons with disabilities reviewed the SMS text messaging library for accuracy and appropriateness and participated in an in-depth interview conducted by the investigator team to determine the SMS text messaging library content customized to the fatigue self-management needs of individuals with MS, SCI, and stroke. The interview lasted approximately 1.5 hours and was conducted according to a discussion guide iteratively developed by the investigator team. Detailed notes were taken and summarized during the interviews. The investigator team refined the SMS text messaging library content based on the summary notes.

CAB Feedback

The investigator team identified, screened, and consented 6 participants to join the CAB. A total of 2 advisory board sessions were held via Zoom videoconferencing (Zoom Video Communications) for approximately 90 minutes each to solicit their feedback on the refined SMS text messaging library program. A session moderator, a note taker, and 2 additional

research team members helped to facilitate these sessions. Before the first meeting, the investigator team sent a document displaying all the potential SMS text messages included in the program to the participants. Participants were asked to rate the clarity and relevance of all SMS text messages on a 5-point Likert scale (with 5 being the highest rating). The document also included a section in which participants could note their overall thoughts or explain their ratings of certain SMS text messages. During the meeting, the investigator team provided all the participants with background information about the project and patient activation. Subsequently, the investigator team elicited feedback regarding general thoughts and concerns about the refined SMS text messaging library program. The first half of the potential SMS text messages was discussed in the first session, and the remaining half of the SMS text messages was discussed in the second meeting. During the second meeting, the research team also asked the participants about the delivery of SMS text messages, such as the number of SMS text messages per day that would be appropriate and the best time to send them (morning or afternoon). Owing to scheduling conflicts, the first meeting was held twice (first time with 2 participants and repeated for the second time with 4 participants). Both sessions were audio-recorded and transcribed verbatim using Trint [42], a professional transcription service. The transcripts were checked for accuracy by members of the research team.

The original text library began with 57 SMS text messages. Following consultation with the psychiatrist, we added the ninth content category of “fatigue education and awareness” and added 16 SMS text messages (for a total of 73 SMS text messages). Further refinement was made to the text library based on the CAB’s clarity and relevance ratings of each text on a 5-point Likert scale. If the mean relevance or clarity rating for a text was between 3.51 and 3.99, the investigator team modified the SMS text message to improve the phrasing or level of detail. If the mean relevance or clarity rating for a text was ≤ 3.5 , the text was removed, leading to 61 SMS text messages. The investigator team finalized the text library with the highest-rated texts across each category. The final 48 texts also aligned with the CAB’s recommendation of 1 text per day during the 12 weeks (ie, Monday through Thursday only, 1 text per day for a total of 4 texts per week, with a question asking about the application of text content delivered on Friday). The number of texts was kept consistent across PAM levels 1 to 2 and 3 to 4.

Step 3: Integration of the Content Library Into the Text Messaging System

We used a digital platform designed by Epharmix (Epharmix, Inc) [43] to deliver the SMS text messaging intervention. We used the predeveloped intervention builder from Epharmix to set up the SMS text messaging logic. A research staff member programmed the SMS text messages and intervention logistics into the Epharmix system. Epharmix complementarily provided the intervention builder and some basic technical assistance. Information was sent and collected via SMS text messages, transmitted between the cellular carrier and the intervention builder using the Secure Sockets Layer, and stored in a secure and encrypted university MySQL [44] database server

environment. A business associate’s agreement between Epharmix and the university was set up to ensure that electronic patient health information was kept secure. The investigators accessed the database using the Epharmix website.

In the Epharmix system, SMS text messages were grouped into message types, including a welcome message, PAM level 1 to 2 weekly SMS text messages, PAM level 3 to 4 weekly SMS

text messages, and check-in SMS text messages asking participants to rate the impact of their fatigue on their daily life and report their use of the tips that week (Figure 1). The system allows research staff to examine participants’ information and monitor their responses to the check-in SMS text messages via the dashboard (Figures 2 and 3). Figures 4 and 5 show screenshots of how the SMS text messages were displayed on a mobile device.

Figure 1. Available message prescriptions for participants.

4. Choose Interventions

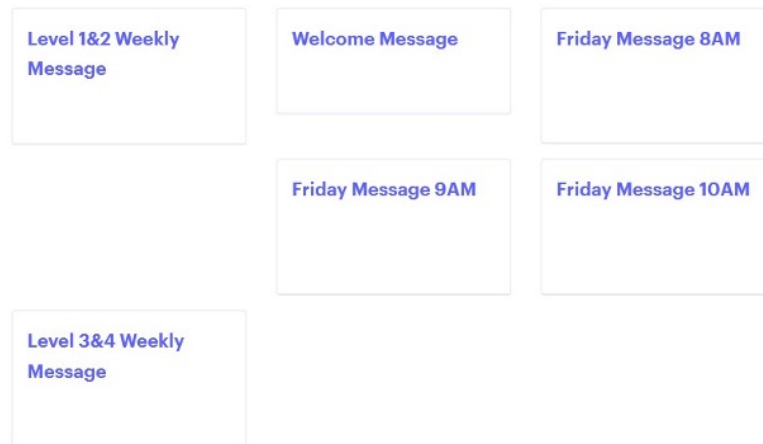


Figure 2. Participant dashboard. MRN: medical record number.

Test #1 Test #1

Phone	Date of Birth
MRN none given	Notes no notes provided

Prescriptions

Intervention	Status	Contact	Start Date	Modify	Data
Welcome Message 🌟	Pending		01/31/2022	✎	📄
Friday Message 8AM 🌟	Pending		01/31/2022	✎	📄
Friday Message 9AM 🌟	Pending		01/31/2022	✎	📄
Friday Message 10AM 🌟	Pending		01/31/2022	✎	📄
Level 3&4 Weekly Message 🌟	Pending		01/31/2022	✎	📄

Figure 3. Participant responses to posed weekly question.



#	date	Response
1	12/17/2021	3
2	12/10/2021	3
3	12/03/2021	2
4	11/26/2021	3
5	11/19/2021	3
6	11/12/2021	3
7	11/05/2021	2
8	10/29/2021	2
9	10/22/2021	3

Figure 4. Example of daily fatigue management tip.

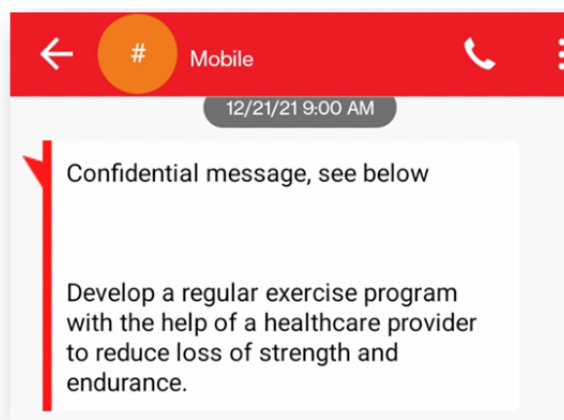
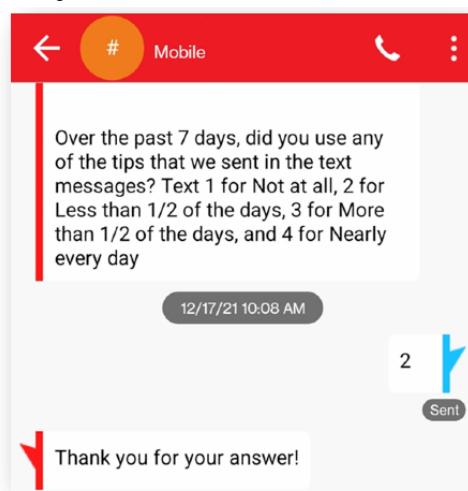


Figure 5. Example of posed weekly question and response.



Step 4: Utility Testing of the Text Messaging Prototype

Once the entire fatigue management SMS text messaging intervention was programmed into the Epharmix intervention builder platform, the 6 CAB members tested the SMS text messaging prototype by receiving SMS text messages on their mobile devices for 1 week. The investigator team sent a usability survey to all CAB members immediately after they completed the 1-week trial. The survey consisted of 11 questions with a 1-7 Likert scale response (1=strongly disagree to 7=strongly agree) and 5 additional qualitative questions. Qualitative questions included whether CAB members thought the timing of the SMS text messages was appropriate, whether they felt that the frequency of the SMS text messages sent was appropriate, whether they had any difficulties responding to the check-in questions, whether they had any further comments regarding the SMS text messaging intervention, and whether they approved of the SMS text messaging intervention. Quantitative questions on the usability survey were adapted from the extended version of the Technology Acceptance Model, which has been shown to accurately account for a large percentage of variance in user usefulness perceptions [45].

Data Analysis

In step 2, the research team reviewed detailed notes from the expert physician, identified recommendations that aligned with the goals of the intervention, and made the version for the CAB to review. The CAB member ratings for each SMS text message were also collected and combined into 1 master spreadsheet for analysis. The investigator team computed the mean clarity and relevance ratings and identified the highest-rated and lowest-rated SMS text messages. Transcriptions of the advisory board sessions produced by Trint were checked for accuracy by 2 trained graduate students. Using a content analysis approach [46], the investigator team deductively developed a formal codebook based on responses from the semistructured questions. A total of 2 graduate students used the codebook to independently code all transcribed texts and met weekly to discuss and reconcile coding discrepancies. In addition, the advisory board qualitative data were triangulated with the data from the usability survey and the physician's input, and

establishing interrater reliability was therefore deemed as unnecessary [47]. We summarized the feedback into the final coding results, which were incorporated into the modifications made to the final version of the text messaging prototype. We used NVivo (version 12; QSR International), a qualitative data analysis software, to analyze the qualitative data [48]. In step 4, all CAB utility ratings were tracked in a spreadsheet and reviewed for the frequency of ratings. Comments shared in the open-ended questions were reviewed for themes.

Ethics Approval

Institutional review board approval from Washington University in St. Louis was obtained before the study (20210319).

Consent, Data Security, and Compensation

All participants were provided with a letter of information, considered a waiver of written consent, before participating, which described the details and purpose of the study. Study data were deidentified and collected through secure and Health Insurance Portability and Accountability Act-compliant platforms (eg, university email and videoconferencing). Participants were financially compensated US \$50 per CAB meeting for up to 4 meetings in the form of a written check that was mailed to them.

Results

Step 1: Development of Initial Text Messaging Library for Fatigue Self-management

A total of 9 focus areas (8 identified by the research team and 1 added after consultation with the psychiatrist) were developed to be delivered over 12 weeks. A total of 57 SMS text messages were initially created for PAM levels 1 to 2 and 3 to 4 (Tables 1 and 2) from the current evidence and resources on fatigue management. The 57 SMS text messages were increased to 73 SMS text messages following consultation with the psychiatrist, refined to 61 after the CAB review, and decreased to 48 by the investigator team to fit within the 12-week time frame (ie, 1 text per weekday, Monday through Thursday only, for a total of 4 texts per week).

Table 1. Summary of SMS text category development and number of SMS text messages.

Weekly topic	Operational definition	Number of text messages			
		Original content development	After a physiatrist consult	After consumer advisory board reviews	Final
Fatigue education and awareness (week 1) ^a	Learn about fatigue and how to be individually aware of its impact, as it is different for each person	0	6	5	4
Behavioral techniques (weeks 2 and 3)	Change certain behaviors to reduce fatigue	14	16	11	8
Energy conservation (weeks 4 and 5)	Prevent wasteful use of energy to minimize fatigue in daily activities	7	8	8	8
Environment and assistive technology (week 6)	Alter your surroundings and use adaptive equipment to preserve energy and decrease fatigue	8	8	6	4
Simplifying activities (week 7)	Break down chores and tasks into smaller pieces so that they are easier to manage and do not cause greater fatigue	6	6	4	4
Diet and hydration (week 8)	Understand the impact that hydration and healthy food can have on fatigue	3	6	6	4
Effective communication (week 9)	Ensure that the people around you listen and understand your ideas and concepts regarding how fatigue impacts you	4	5	4	4
Exercise and physical activity (weeks 10 and 11)	Determine the right amount and type of exercise or physical activity to combat fatigue	10	11	10	4
Sleep (week 12)	Develop healthy sleep patterns to decrease fatigue	5	7	7	8
Total SMS text messages for all patient activation measure levels 1-4	N/A ^b	57	73	61	48 ^c

^aThis topic was added following consultation with the physiatrist; before that, we only had 8 topics. The number of SMS text messages was kept consistent across the patient activation measure levels 1 to 2 and 3 to 4.

^bN/A: not applicable.

^cOne text per day, Monday through Thursday only, a total of 4 SMS text messages per week for 12 weeks.

Table 2. Examples of SMS text messages by patient activation measure levels.

Weekly topic	Example SMS text messages for PAM ^a levels 1-2	Example SMS text messages for PAM levels 3-4
Fatigue education and awareness (week 1)	Recognize that there can be different causes of fatigue, such as pain, daily activities, and stress. Figuring out what causes your fatigue can be a helpful step in learning how to manage it.	Maintain communication with your health professional to learn how to recognize what is causing your fatigue.
Behavioral techniques (weeks 2 and 3)	Recognize that there are different types of fatigue (such as physical, emotional, and mental or cognitive). Understanding different types of fatigue can help you manage it.	Talk to a health professional about your fatigue. They can help you find strategies to manage symptoms associated with your type of fatigue.
Energy conservation (weeks 4 and 5)	Ask for help with tiring activities. Tell family or friends if you need help with tasks that are difficult for you.	Keep track of which activities fatigue you and plan those activities during the day when you typically have more energy. This will allow you to complete more tasks throughout the day.
Environment and assistive technology (week 6)	Learn about different assistive technology options and ergonomic techniques that can help conserve energy.	Consider using one or more assistive technology or ergonomic technique that works best for you. This can help to reduce strain and stress on your body.
Simplifying activities (week 7)	Changing the way you do certain activities or changing your expectations of yourself can give you more energy.	Think about ways your fatigue can be managed. Try simplifying daily tasks in keeping with your capabilities to manage fatigue.
Diet and hydration (week 8)	Recognize that nutrition plays a role in combating fatigue. Learn about healthy diet choices that can impact energy levels.	Eating a diet with whole grains, nuts, seeds, and lean proteins will keep your body fueled regularly and help you beat fatigue.
Effective communication (week 9)	Talk with family and friends about your fatigue and be honest about how it impacts your daily life. This reality may also make you more aware of your own fatigue.	Make an effort to honestly communicate with your family and friends about how fatigue impacts your daily life consistently (maybe schedule a weekly check-in).
Exercise and physical activity (weeks 10 and 11)	Realize that a regular exercise routine can reduce loss of strength and endurance and improve fatigue symptoms.	Consistently set an exercise goal and plan it into your weekly schedule. This should make following it easier.
Sleep (week 12)	Treat symptoms that may interfere with sleep, such as spasticity or urinary problems. This can help to reduce fatigue.	Speak with your health care provider about any symptoms, such as spasticity or urinary problems, that may be interfering with your sleep and increasing your fatigue.

^aPAM: patient activation measure.

Step 2: Review and Modification of SMS Text Messages

Health Care Expert Opinion

Modifications based on the expert physician's opinion were made to address two main points: (1) the SMS text messaging intervention should introduce basic fatigue education in its first week and (2) the wording of SMS text messages and the sources from which the content is developed should be more "mainstream." The first point led to the addition of "fatigue education and awareness" as a new category placed at the beginning of the program to educate about fatigue, different types of fatigue, and their causes, resulting in 9 focus areas. The second point led to the review of more mainstream sources for examples of a layman's presentation and wording of evidence-based information on fatigue (eg, *Time* magazine and *Men's Health* magazine) and an improvement in the readability of the SMS text messages.

CAB Feedback

A total of 6 persons with disabilities participated as advisory board members to provide feedback on the SMS text messaging intervention. The advisory board was evenly divided by sex and reported an average age of 47 (SD 13) years (Table 3). Regarding the quantitative feedback, the range in relevance ratings for individual SMS text messages was from 3.2 to 5.0,

and the range in clarity ratings was from 3.5 to 5.0 across 9 focus areas. Table 4 outlines a range of quantitative ratings for various example SMS text messages, accompanied by the action taken (ie, remove, modify, or keep).

We summarized the qualitative advisory board feedback based on the 9 SMS text message focus areas (ie, fatigue education and awareness, behavioral techniques, energy conservation, environment and assistive technology, simplifying activities, diet and hydration, effective communication, exercise and physical activity, and sleep; Table 5). Participants provided a wide range of qualitative feedback related to the clarity and relevance of the SMS text messages. The participants also described their personal experiences with some tips or suggestions in the SMS text messages they used for fatigue self-management. A common theme expressed by participants across all areas was the need for more examples (eg, for the diet and hydration area, participants explained that examples of healthy snacks would be helpful). Overall, participants reported that the SMS text messaging content was helpful, clear, and well suited for a mHealth intervention for persons with MS, SCI, and stroke. The advisory board noted that the content seemed particularly relevant for newly diagnosed individuals, as some of the information was not new knowledge for them and served as more of a reminder.

Table 3. Overview of advisory board demographics (N=6).

Participant characteristics	Value
Age (years), mean (SD; range)	46.67 (12.6; 25-59)
Gender, n (%)	
Male	3 (50)
Female	3 (50)
Race, n (%)	
Black or African American	1 (17)
White	5 (83)
Ethnicity, n (%)	
Hispanic or Latino origin	0 (0)
Non-Hispanic or Latino origin	6 (100)
Diagnosis, n (%)	
Multiple sclerosis	2 (33)
Spinal cord injury	2 (33)
Stroke	2 (33)

Table 4. Examples of the range of quantitative advisory board feedback and changes made.

Content categories	Initial SMS text message example	Clarity rating (1-5), mean (SD)	Relevance rating (1-5), mean (SD)	Action taken
Fatigue education and awareness	“Take your fatigue medication on time because sometimes fatigue can occur without any warning.” (PAM ^a level 1-2)	5.0 (0.7)	3.2 (0.1)	Removed
Behavioral techniques	“Participate in cognitive behavioral therapy (CBT) to cope with difficult situations with the help of a health care provider.” (PAM level 3-4)	4.2 (0.5)	3.8 (0.3)	Modified by replacing “cognitive behavioral therapy” with “problem-solving skills”
Energy conservation	“Use grocery delivery services to conserve energy that would be spent going to the store.” (PAM level 1-2)	4.7 (0.3)	3.8 (0.4)	Modified so that grocery delivery is an example rather than a focus of the text
Environment and assistive technology	“Plan ahead for your return back to work or school and think about modifying activities and tasks that are more difficult so that you can continue to do them.” (PAM level 3-4)	3.5 (0.6)	3.8 (0.6)	Removed
Simplifying activities	“Taking rest breaks between demanding tasks can help you combat fatigue.” (PAM level 1-2)	4.8 (0.4)	4.7 (0.5)	Kept
Diet and hydration	“Eat healthy foods that are high in iron. Plan to add these food choices to your meal to promote more long-term energy: spinach, legumes, pumpkin seeds, turkey, broccoli, tofu, & fish.” (PAM level 3-4)	5.0 (0.1)	4.7 (0.4)	Kept
Effective communication	“Recognize that joining a support group can help to decrease your frustration with having long-term fatigue.” (PAM level 1-2)	4.3 (0.5)	4.2 (0.4)	Kept
Exercise and physical activity	“Active wheeling or walking to places you would have normally traveled to in a vehicle or by parking a bit farther away will provide some exercise or physical activity.” (PAM level 1-2)	4.2 (0.3)	3.8 (0.4)	Modified by replacing reference to parking with a general statement about using active wheeling or walking instead of driving
Sleep	“Monitor the temperature of your sleeping space to make sure it is not too hot. Heat can make it more difficult to fall and stay asleep and can increase fatigue.” (PAM level 3-4)	4.7 (0.1)	5 (0.5)	Kept

^aPAM: patient activation measure.

Table 5. Examples of qualitative advisory board feedback.

Content category	Initial SMS text example	Summarized feedback	Example quotes
Fatigue education and awareness	“Figure out which type of fatigue you may be experiencing by talking to a health professional. Your fatigue may look different from other people’s fatigue, and they can help you figure out what works best for you.” [PAM ^a level 3-4]	This statement should be delivered earlier in the week.	“This statement should come earlier.” [58-year-old White woman]
Behavioral techniques	“Schedule meaningful activities into your daily routine to help fight fatigue.” [PAM level 3-4]	The phrase “meaningful activities” is confusing and vague.	“What are some examples...that would be meaningful to somebody that would help them reduce fatigue? Because I had a hard time. If I sat down on the sofa and watched TV all afternoon, I might reduce my fatigue, but I don’t think that’s good for anybody.” [44-year-old White woman]
Energy conservation	“Planning visits and knowing what challenges you will face when arriving will help to manage energy use.” [PAM level 1-2]	The term “visits” is too vague.	“Can you be a bit more specific than ‘visits’?” [58-year-old White woman]
Environment and assistive technology	“Be aware that different body positions can cause strain, which may increase feelings of fatigue.” [PAM level 1-2]	Something we just figure out on our own.	“It just didn’t seem like it really was relevant because it’s like, you kind of live it and you’ve kind of just got to test out the waters to see what works good for you.” [37-year-old Black or African American man]
Simplifying activities	“Reducing the number of transfers you make daily can help you to reduce your fatigue.” [PAM level 1-2]	The text needs to specify to whom transferring is relevant.	“Can you specify what kind of person needs to be concerned about transferring?” [58-year-old White woman]
Diet and hydration	“Eat a balanced diet regularly to reduce fatigue.” [PAM level 1-2]	Examples would be helpful.	“Perhaps include a list of healthy snack ideas. Sometimes people with disabilities are too fatigued to eat full meals.” [57-year-old White woman]
Effective communication	“Talk with family and friends about your fatigue and be honest about how it impacts your daily life. This reality may also make you more aware of your own fatigue.” [PAM level 1-2]	This text message tip is easier said than done.	“This would be a delicate subject, talking to family and friends about fatigue. Other people usually think sitting down for 10 minutes is all I need. They have no frame of reference.” [57-year-old White woman]
Exercise and physical activity	“Find self-care activities that are important to you. This will help to reduce your fatigue levels.” [PAM level 1-2]	Examples would be helpful.	“Pretty vague. Do you mean putting on makeup and styling your hair?” [57-year-old White woman]
Sleep	“Electronic devices’ artificial blue light can suppress the sleep hormone melatonin, which may make it more difficult for you to fall asleep. Track your screen time throughout the day to determine if it makes a difference in your sleep.” [PAM level 1-2]	The 2 statements need a clearer connection between each other.	“I don’t see the correlation between the first and second sentence. It’s the second one that impacts sleep.” [58-year-old White woman]

^aPAM: patient activation measure.

Table 6 lists the additional delivery and logistics codes with example quotes. The participants agreed that SMS text messages should be grouped by topic for each week. They also agreed that SMS text messages should be delivered at the same time every day, preferably in the morning. Regarding personalization of the SMS text message content, participants felt that SMS text messages with introductory content, reminders, and general statements would be best suited for individuals with a newly

diagnosed disability. In contrast, the content should be more personalized, tailored, and strategy based for those who have lived with their disability for a longer period. The CAB has also reached a consensus on the preference for only 1 SMS text message per day. This feedback led the investigator team to narrow down the program to deliver 1 SMS text message per day for 12 weeks. Therefore, the research team eliminated the SMS text messages with the lowest relevance and clarity ratings.

Table 6. Examples of delivery and logistics feedback.

Delivery and logistics codes	Topic description	Consensus	Example quote
Format	In what order the content should be delivered	Structured and grouped together	"I like it more structured, too, where, you know, you can be like, every Wednesday, you get something about working out or stuff like that or, 'did you work out today?'" [57-year-old White woman]
Personalization	How personalized the content should be	Broad and simple is better for newly diagnosed individuals; more personalized and strategy based is better for individuals who have lived with their disability for a longer period	"If you just had a stroke or you were just diagnosed with MS or you just got a spinal cord injury, then you know, this is all great. But if you're further along, like, I don't need a reminder to take medicine, and I don't want it either." [58-year-old White woman]
Timing	At what time the content should be delivered	Once per day in the morning	"I think if they were at the same time every day, it would be more coordinated. Sometime in the morning." [59-year-old White man]

Step 3: Integration of the Content Library Into the Text Messaging System

In most cases, research and Epharmix staff worked together and obtained the functionality that the designed SMS text messaging intervention required. However, we encountered a few limitations with programming. The first limitation incurred during programming was that the intervention builder application only allowed 50 scripts (the individual SMS text messages sent to the users), but the designed SMS text messaging intervention required 135 scripts. The Epharmix staff were able to open the permissions and allow up to 200 scripts.

The next limitation was that the Epharmix allowed only 3 different sets of SMS text messages. However, this project had 6 different sets of scripts to deliver the 12-week fatigue self-management SMS text messaging intervention. The first set was a single message delivered once as a welcome message to all the participants before beginning the intervention phase. The second set of scripts consisted of 48 separate daily fatigue management tips delivered at the same preset time once a day on Monday through Thursday for participants who scored a PAM level of 1 to 2. The third set of scripts was a different set of 48 daily fatigue management tips delivered at a preset time once a day on Monday through Thursday for participants who scored a PAM-13 level of 3 to 4. Scripts fourth to sixth were single SMS text messages set up to deliver check-ins on Fridays at 3 different time points and with 3 different SMS text messages. The first Friday message posed a question asking the participants to rate how their fatigue had impacted their daily life over the past week (1=not at all to 5=very much). The second message asked the participants to rate how often they had used the provided fatigue tips over the past week (1=not at all to 4=nearly every day). The third message was a motivational one that thanked them for their participation and commended them on a job well done. The Epharmix staff was able to open permissions, allowing the system to accept 6 sets of SMS text messages.

The next set of challenges came during the research staff and CAB testing of the developed SMS text messaging intervention. Every SMS text message delivered began with a "confidential

message, see below." The wording confused CAB members, and they did not know who was sending the SMS text messages. Unfortunately, the SMS text message was automatically generated by the system, and it could not be deleted or edited. We also discovered that the SMS text messages sent to participants did not come from the same phone number, which made the program's SMS text messages harder to identify as being related to the study. Therefore, we decided to educate all research participants before the intervention phase that SMS text messages would come from different phone numbers, and they would always begin with the "confidential message" wording.

Several more limitations of the intervention builder application did not necessarily impact the participants but influenced how we managed the study. We could not repeat the SMS text messages if the participant did not respond, and there was no way for us to know whether the SMS text messages were opened or read. To overcome these limitations, the research coordinator planned to closely follow the required weekly responses and conduct follow-up phone calls during the next testing phase with future research participants in the full 12-week program.

Step 4: Utility Testing of the Text Messaging Prototype

After completing the 1-week trial of the SMS text messaging intervention, the CAB members provided feedback on their experiences using the usability survey. Members rated the first 11 questions of the survey using a 1-7 Likert rating scale. [Table 7](#) presents the questions and responses. The CAB members supported the developed formatting and usability of SMS text messaging. There was a single rating of 1, which was related to the question, "Whenever I made a mistake using the text messaging program, I could recover easily and quickly." The CAB member who reported this rating commented, "I'm sure it's easy to recover from a mistake—hit the back key on my phone and was able to answer the question—but I'm not sure if this was the correct thing to do. I most likely missed the instruction—what to do if you make a mistake." The low rating was not necessarily because the participant disagreed with the ease of recovering from a mistake but because they were unsure whether the method they used was correct.

Table 7. Advisory board member perspectives on usability (scale of 1=strongly disagree to 7=strongly agree, or not applicable).

Usability statement	Value, mean (SD)
The SMS (text) program format was easy to use	7 (0)
It was easy for me to learn to use the SMS (text) program	7 (0)
The questions posed in the SMS (text) program were easy to read, understand, and respond to	6.5 (0.8)
Whenever I made a mistake using the SMS (text) program, I could recover easily and quickly (missing=3)	5 (3.5)
I like the SMS (text) program	7 (0)
The information in the SMS (text) program was well organized	6.7 (0.8)
I feel comfortable using the SMS (text) program in social settings	7 (0)
The amount of time involved in using the SMS (text) program has been fitting for me	7 (0)
I would use the SMS (text) program again	7 (0)
Overall, I am satisfied with the SMS (text) program	7 (0)
The SMS (text) program has all the functions and capabilities I expected it to have	6.8 (0.4)

The 5 additional open-ended questions in the usability survey primarily received positive feedback, supporting the usability of the SMS text messaging intervention. The first question, “Did you feel like the timing (in the morning) of the daily text messages was appropriate?” received responses including the following: “fit my schedule well,” “I work nights, but I found the messages when I woke up,” and “it was the middle of my workday or seemed like it—but I work 6 am to 3 pm, more than likely it’s OK for everyone.” The second question, “Did you feel like the frequency of how often the text messages were sent was appropriate” received a single response: “A couple more would have been appropriate also, like maybe in the evening.” The next question, “Did you have any difficulties responding to the check-in questions posed on Fridays,” also received a single response: “The number 1 was already entered as the response, and the first time, I accidentally sent that as my response. The second time, I was able to send my own response.” The fourth question, “Do you approve of the text program format that has been developed” received 2 responses: “Did a nice job, liked all the content,” and “I liked how y’all developed it—informative but not overwhelming.”

Discussion

Principal Findings

Fatigue is a common chronic condition for persons with disabilities [49]. Education on fatigue self-management is crucial for supporting patient activation to reduce fatigue [50]. This study described the 4-step process used to develop the fatigue self-management SMS text messaging intervention. The expert physician and CAB provided insight into the clarity, relevance, and content of the fatigue self-management SMS text messaging intervention for persons with MS, SCI, and stroke. Participants in the CAB found the content of the SMS text messages relevant, providing support, motivation, and accountability via a simple and convenient mode of communication. Participants also indicated that 1 weekly check-in message and 4 fatigue self-management SMS text messages per week were appropriate for intervention delivery. Our findings concur with those of an earlier qualitative study; participants in both studies showed positive experiences with SMS text messaging interventions

and perceived the incorporation of SMS text messages as an adjunctive tool to support health management [51]. Previous mHealth SMS text message-based interventions have used a variety of methods to develop message content. Some recent mHealth interventions have used message libraries adapted from previous successful trials or developed based on feedback from expert working groups without consultation from end users [52,53]. However, recent studies have incorporated co-design or participatory methods, and this approach appears to be becoming the standard. Examples of end user or consumer involvement in the development stages include real-time ratings, daily qualitative telephone interviews, crowdsourcing, usability testing, and mixed co-design workshops comprising both health professionals and consumers [27,54-56].

This study incorporated the concept of patient activation into the development of the fatigue self-management SMS text messaging intervention for individuals with various disabilities. This approach is similar to another ongoing trial in Southeast Asia that proposed the development and testing of a mobile app-based self-management program to empower people with knowledge and skills to manage metabolic syndromes (eg, diabetes or hypertension) [57]. A unique finding of our study is that our process involved both medical staff and persons with disabilities as part of the team to guide the development of the digital intervention. Participants also evaluated all SMS text messages to increase their relevance and clarity before the SMS text messages were incorporated into the digital platform. Participants made further suggestions to support intervention delivery and improve user engagement. One suggestion was to provide a structured format wherein we would send SMS text messages at the same time throughout the intervention. We incorporated this suggestion by programming all SMS text messages to be presented in a structured sequence. Another suggestion was to personalize SMS text messages based on the chronicity of an individual’s disability. Interventions for persons with a recent onset of disease or following a new injury should emphasize psychoeducation, which enables people to learn and adjust to a new disability. In contrast, interventions for people in the chronic stage should emphasize building strategies to cope with barriers encountered while living in the community

[58,59]. Our SMS text messaging intervention is focused on increasing self-management knowledge and skills by providing patient activation-based health tips to manage fatigue. This intervention would likely be appropriate for persons with various chronic stages of disability. Further research is required to support this argument.

Challenges were encountered when integrating SMS text messages into the digital platform. A key lesson learned in this study is to foster close collaboration between the technology industry and medicine to develop or adapt technology [60]. We also confirmed the importance of involving the end user to help guide the development and optimization of the technology to be used in medicine. This practice should become standard both within health care and across many other sectors. With support from our technology partner, we could adapt several features of intervention technology to meet the needs of our target population without sacrificing functionality. We also conducted usability testing in which our CAB members trialed the SMS text messaging intervention for a week. They endorsed all items with positive responses, suggesting the adequate usability of the SMS text messaging intervention. Future research is needed to examine the usability of this program in a larger disability group.

Additional Study Limitations

A potential limitation is that our SMS text messaging intervention included SMS text messages from 9 domain areas. It is unclear which domain areas are most effective. A future trial may ask participants to rate which domains of SMS text messages they find most helpful in managing their fatigue. We designed the intervention content by structuring various domain areas across 12 weeks. We anticipate that not all content will benefit all participants at one point in time. A future study may explore whether adding a preparation session to identify the areas needed and prescribing the relevant SMS text messages for each participant would lead to a better outcome.

Conclusions

This study demonstrates a robust method for developing a fatigue self-management SMS text messaging intervention for people with MS, SCI, and stroke. With input and support from a medical expert, persons with disabilities, and the technology vendor, we developed an intervention delivered through technology to people with various disabilities at different patient activation levels. The next steps include pilot-testing the fatigue self-management SMS text messaging intervention with people with MS, SCI, and stroke to examine its flexibility and explore its initial effects.

Acknowledgments

The authors would like to thank Dr T Huskey and the consumer advisory board members for their time and effort in reviewing and providing feedback on the development of the fatigue self-management SMS text messaging intervention. The authors would also like to thank Megen Devine for her assistance in editing this manuscript. The contents of this manuscript were developed under a grant from The Foundation for Barnes-Jewish Hospital Institute of Clinical and Translational Sciences (award 5562). The funders did not have a role in planning (eg, identifying the study design) or executing (eg, collection, analysis, or interpretation of the data) the study described in this manuscript. The funders also did not have a role in writing the manuscript or deciding to submit the manuscript.

Data Availability

The underlying data sets generated and analyzed during this study cannot be sufficiently deidentified and, therefore, cannot be made publicly available because of ethical considerations. Deidentified data can be made available from the corresponding author upon reasonable request for the purpose of further research.

Authors' Contributions

KAM conceived and designed the study and assisted with data analysis and interpretation. She drafted the paper and provided final approval of the draft submitted for publication. AWKW was involved in conceiving and designing the study and interpreting the data. He provided intellectual content related to all sections of the paper and provided final approval of the draft submitted for publication. KW was involved in the data collection. She contributed intellectual content to the Methods section and provided final approval of the draft submitted for publication. RHD was involved in providing technical support for the acquisition of data. She contributed intellectual content to the Methods and Results sections and provided final approval of the draft submitted for publication. TMK was involved in interpreting the data. She provided final approval of the draft submitted for publication. PKN was involved in conceiving the design of the study and interpretation of the data. She provided intellectual content related to all sections of the paper and provided final approval of the draft submitted for publication.

Conflicts of Interest

None declared.

References

1. Hirsche RC, Williams B, Jones A, Manns P. Chronic disease self-management for individuals with stroke, multiple sclerosis and spinal cord injury. *Disabil Rehabil* 2011;33(13-14):1136-1146. [doi: [10.3109/09638288.2010.523103](https://doi.org/10.3109/09638288.2010.523103)] [Medline: [20919875](https://pubmed.ncbi.nlm.nih.gov/20919875/)]
2. Anton HA, Miller WC, Townson AF, Imam B, Silverberg N, Forwell S. The course of fatigue after acute spinal cord injury. *Spinal Cord* 2017 Jan;55(1):94-97. [doi: [10.1038/sc.2016.102](https://doi.org/10.1038/sc.2016.102)] [Medline: [27349608](https://pubmed.ncbi.nlm.nih.gov/27349608/)]
3. Cao Y, DiPiro N, Krause JS. Association of secondary health conditions with future chronic health conditions among persons with traumatic spinal cord injury. *Top Spinal Cord Inj Rehabil* 2020;26(4):283-289 [FREE Full text] [doi: [10.46292/sci20-00020](https://doi.org/10.46292/sci20-00020)] [Medline: [33536734](https://pubmed.ncbi.nlm.nih.gov/33536734/)]
4. Chen K, Marsh EB. Chronic post-stroke fatigue: it may no longer be about the stroke itself. *Clin Neurol Neurosurg* 2018 Nov;174:192-197. [doi: [10.1016/j.clineuro.2018.09.027](https://doi.org/10.1016/j.clineuro.2018.09.027)] [Medline: [30266010](https://pubmed.ncbi.nlm.nih.gov/30266010/)]
5. Apolinário-Hagen J, Menzel M, Hennemann S, Salewski C. Acceptance of mobile health apps for disease management among people with multiple sclerosis: web-based survey study. *JMIR Form Res* 2018 Dec 12;2(2):e11977 [FREE Full text] [doi: [10.2196/11977](https://doi.org/10.2196/11977)] [Medline: [30684408](https://pubmed.ncbi.nlm.nih.gov/30684408/)]
6. Craig A, Tran Y, Siddall P, Wijesuriya N, Lovas J, Bartrop R, et al. Developing a model of associations between chronic pain, depressive mood, chronic fatigue, and self-efficacy in people with spinal cord injury. *J Pain* 2013 Sep;14(9):911-920. [doi: [10.1016/j.jpain.2013.03.002](https://doi.org/10.1016/j.jpain.2013.03.002)] [Medline: [23707693](https://pubmed.ncbi.nlm.nih.gov/23707693/)]
7. Barin L, Salmen A, Disanto G, Babičić H, Calabrese P, Chan A, Swiss Multiple Sclerosis Registry (SMSR). The disease burden of Multiple Sclerosis from the individual and population perspective: which symptoms matter most? *Mult Scler Relat Disord* 2018 Oct;25:112-121. [doi: [10.1016/j.msard.2018.07.013](https://doi.org/10.1016/j.msard.2018.07.013)] [Medline: [30059895](https://pubmed.ncbi.nlm.nih.gov/30059895/)]
8. Teshale SM, Molton IR, Jensen MP. Associations among decisional autonomy, fatigue, pain, and well-being in long-term physical disability. *Rehabil Psychol* 2019 Aug;64(3):288-297. [doi: [10.1037/rep0000279](https://doi.org/10.1037/rep0000279)] [Medline: [31033307](https://pubmed.ncbi.nlm.nih.gov/31033307/)]
9. Munce SE, Straus SE, Fehlings MG, Voth J, Nugaeva N, Jang E, et al. Impact of psychological characteristics in self-management in individuals with traumatic spinal cord injury. *Spinal Cord* 2016 Jan;54(1):29-33 [FREE Full text] [doi: [10.1038/sc.2015.91](https://doi.org/10.1038/sc.2015.91)] [Medline: [26055818](https://pubmed.ncbi.nlm.nih.gov/26055818/)]
10. Seves BL, Hoekstra F, Hetinga FJ, Dekker R, van der Woude LH, Hoekstra T. Trajectories of health-related quality of life among people with a physical disability and/or chronic disease during and after rehabilitation: a longitudinal cohort study. *Qual Life Res* 2021 Jan;30(1):67-80 [FREE Full text] [doi: [10.1007/s11136-020-02647-7](https://doi.org/10.1007/s11136-020-02647-7)] [Medline: [32986126](https://pubmed.ncbi.nlm.nih.gov/32986126/)]
11. Wijesuriya N, Tran Y, Middleton J, Craig A. Impact of fatigue on the health-related quality of life in persons with spinal cord injury. *Arch Phys Med Rehabil* 2012 Feb;93(2):319-324. [doi: [10.1016/j.apmr.2011.09.008](https://doi.org/10.1016/j.apmr.2011.09.008)] [Medline: [22289244](https://pubmed.ncbi.nlm.nih.gov/22289244/)]
12. Enns MW, Bernstein CN, Kroeker K, Graff L, Walker JR, Lix LM, CIHR Team in Defining the Burden and Managing the Effects of Psychiatric Comorbidity in Chronic Immunoinflammatory Disease. The association of fatigue, pain, depression and anxiety with work and activity impairment in immune mediated inflammatory diseases. *PLoS One* 2018 Jun 7;13(6):e0198975 [FREE Full text] [doi: [10.1371/journal.pone.0198975](https://doi.org/10.1371/journal.pone.0198975)] [Medline: [29879231](https://pubmed.ncbi.nlm.nih.gov/29879231/)]
13. Rutkowski NA, Sabri E, Yang C. Post-stroke fatigue: a factor associated with inability to return to work in patients less than 60 years-a 1-year follow-up. *PLoS One* 2021 Aug 4;16(8):e0255538 [FREE Full text] [doi: [10.1371/journal.pone.0255538](https://doi.org/10.1371/journal.pone.0255538)] [Medline: [34347804](https://pubmed.ncbi.nlm.nih.gov/34347804/)]
14. van der Hiele K, van Gorp DA, van Egmond EE, Jongen PJ, Reneman MF, van der Klink JJ, et al. Self-reported occupational functioning in persons with relapsing-remitting multiple sclerosis: does personality matter? *J Neurol Sci* 2021 Aug 15;427:117561 [FREE Full text] [doi: [10.1016/j.jns.2021.117561](https://doi.org/10.1016/j.jns.2021.117561)] [Medline: [34216973](https://pubmed.ncbi.nlm.nih.gov/34216973/)]
15. Ricci JA, Chee E, Lorandean AL, Berger J. Fatigue in the U.S. workforce: prevalence and implications for lost productive work time. *J Occup Environ Med* 2007 Jan;49(1):1-10. [doi: [10.1097/01.jom.0000249782.60321.2a](https://doi.org/10.1097/01.jom.0000249782.60321.2a)] [Medline: [17215708](https://pubmed.ncbi.nlm.nih.gov/17215708/)]
16. Dicianno BE, Henderson G, Parmanto B. Design of mobile health tools to promote goal achievement in self-management tasks. *JMIR Mhealth Uhealth* 2017 Jul 24;5(7):e103 [FREE Full text] [doi: [10.2196/mhealth.7335](https://doi.org/10.2196/mhealth.7335)] [Medline: [28739558](https://pubmed.ncbi.nlm.nih.gov/28739558/)]
17. Moss-Morris R, McCrone P, Yardley L, van Kessel K, Wills G, Dennison L. A pilot randomised controlled trial of an Internet-based cognitive behavioural therapy self-management programme (MS Invigor8) for multiple sclerosis fatigue. *Behav Res Ther* 2012 Jun;50(6):415-421. [doi: [10.1016/j.brat.2012.03.001](https://doi.org/10.1016/j.brat.2012.03.001)] [Medline: [22516321](https://pubmed.ncbi.nlm.nih.gov/22516321/)]
18. Whitehead LC, Unahi K, Burrell B, Crowe MT. The experience of fatigue across long-term conditions: a qualitative meta-synthesis. *J Pain Symptom Manage* 2016 Jul;52(1):131-43.e1 [FREE Full text] [doi: [10.1016/j.jpainsymman.2016.02.013](https://doi.org/10.1016/j.jpainsymman.2016.02.013)] [Medline: [27233142](https://pubmed.ncbi.nlm.nih.gov/27233142/)]
19. Menting J, Tack CJ, Bleijenberg G, Donders R, Droogleever Fortuyn HA, Fransen J, et al. Is fatigue a disease-specific or generic symptom in chronic medical conditions? *Health Psychol* 2018 Jun;37(6):530-543. [doi: [10.1037/hea0000598](https://doi.org/10.1037/hea0000598)] [Medline: [29781654](https://pubmed.ncbi.nlm.nih.gov/29781654/)]
20. Nordin Å, Taft C, Lundgren-Nilsson Å, Dencker A. Minimal important differences for fatigue patient reported outcome measures-a systematic review. *BMC Med Res Methodol* 2016 May 26;16:62 [FREE Full text] [doi: [10.1186/s12874-016-0167-6](https://doi.org/10.1186/s12874-016-0167-6)] [Medline: [27387456](https://pubmed.ncbi.nlm.nih.gov/27387456/)]
21. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: [10.1111/j.1475-6773.2005.00438.x](https://doi.org/10.1111/j.1475-6773.2005.00438.x)] [Medline: [16336556](https://pubmed.ncbi.nlm.nih.gov/16336556/)]

22. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. *Health Aff (Millwood)* 2015 Mar;34(3):431-437. [doi: [10.1377/hlthaff.2014.0452](https://doi.org/10.1377/hlthaff.2014.0452)] [Medline: [25732493](https://pubmed.ncbi.nlm.nih.gov/25732493/)]
23. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. *Health Aff (Millwood)* 2013 Feb;32(2):207-214. [doi: [10.1377/hlthaff.2012.1061](https://doi.org/10.1377/hlthaff.2012.1061)] [Medline: [23381511](https://pubmed.ncbi.nlm.nih.gov/23381511/)]
24. Tusa N, Kautiainen H, Elfving P, Sinikallio S, Mäntyselkä P. Relationship between patient activation measurement and self-rated health in patients with chronic diseases. *BMC Fam Pract* 2020 Nov 04;21(1):225 [FREE Full text] [doi: [10.1186/s12875-020-01301-y](https://doi.org/10.1186/s12875-020-01301-y)] [Medline: [33148185](https://pubmed.ncbi.nlm.nih.gov/33148185/)]
25. Gimbel RW, Rennert LM, Crawford P, Little JR, Truong K, Williams JE, et al. Enhancing patient activation and self-management activities in patients with type 2 diabetes using the US Department of Defense mobile health care environment: feasibility study. *J Med Internet Res* 2020 May 26;22(5):e17968 [FREE Full text] [doi: [10.2196/17968](https://doi.org/10.2196/17968)] [Medline: [32329438](https://pubmed.ncbi.nlm.nih.gov/32329438/)]
26. John JR, Tannous WK, Jones A. Outcomes of a 12-month patient-centred medical home model in improving patient activation and self-management behaviours among primary care patients presenting with chronic diseases in Sydney, Australia: a before-and-after study. *BMC Fam Pract* 2020 Aug 08;21(1):158 [FREE Full text] [doi: [10.1186/s12875-020-01230-w](https://doi.org/10.1186/s12875-020-01230-w)] [Medline: [32770944](https://pubmed.ncbi.nlm.nih.gov/32770944/)]
27. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. *Eur J Prev Cardiol* 2014 Apr;21(4):492-499. [doi: [10.1177/2047487312449416](https://doi.org/10.1177/2047487312449416)] [Medline: [22605787](https://pubmed.ncbi.nlm.nih.gov/22605787/)]
28. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. *Annu Rev Public Health* 2015 Mar 18;36:393-415 [FREE Full text] [doi: [10.1146/annurev-publhealth-031914-122855](https://doi.org/10.1146/annurev-publhealth-031914-122855)] [Medline: [25785892](https://pubmed.ncbi.nlm.nih.gov/25785892/)]
29. Lee JA, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. *BMC Med Inform Decis Mak* 2018 Feb 20;18(1):12 [FREE Full text] [doi: [10.1186/s12911-018-0591-0](https://doi.org/10.1186/s12911-018-0591-0)] [Medline: [29458358](https://pubmed.ncbi.nlm.nih.gov/29458358/)]
30. Li SH, Sandler CX, Casson SM, Cassar J, Bogg T, Lloyd AR, et al. Randomised controlled trial of online continuing education for health professionals to improve the management of chronic fatigue syndrome: a study protocol. *BMJ Open* 2017 May 10;7(5):e014133 [FREE Full text] [doi: [10.1136/bmjopen-2016-014133](https://doi.org/10.1136/bmjopen-2016-014133)] [Medline: [28495811](https://pubmed.ncbi.nlm.nih.gov/28495811/)]
31. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. *Arch Phys Med Rehabil* 2019 Jan;100(1):115-127. [doi: [10.1016/j.apmr.2018.07.439](https://doi.org/10.1016/j.apmr.2018.07.439)] [Medline: [30171827](https://pubmed.ncbi.nlm.nih.gov/30171827/)]
32. Wellbeloved-Stone CA, Weppner JL, Valdez RS. A systematic review of telerehabilitation and mHealth interventions for spinal cord injury. *Curr Phys Med Rehabil Rep* 2016 Nov 30;4(4):295-311. [doi: [10.1007/s40141-016-0138-1](https://doi.org/10.1007/s40141-016-0138-1)]
33. Moon NW, Baker PM, Goughnour K. Designing wearable technologies for users with disabilities: accessibility, usability, and connectivity factors. *J Rehabil Assist Technol Eng* 2019 Aug 13;6:2055668319862137 [FREE Full text] [doi: [10.1177/2055668319862137](https://doi.org/10.1177/2055668319862137)] [Medline: [35186318](https://pubmed.ncbi.nlm.nih.gov/35186318/)]
34. Krueger RA, Casey MA. *Focus Groups: A Practical Guide for Applied Research*. 4th edition. Thousand Oaks, CA, USA: Sage Publications; 2009.
35. Turner CW, Lewis JR, Nielsen J. Determining usability test sample size. In: Karwowski W, editor. *International Encyclopedia of Ergonomics and Human Factors*. 2nd edition. Boca Raton, FL, USA: CRC Press; 2006:3084-3088.
36. Vial S, Boudhraâ S, Dumont M. Human-centered design approaches in digital mental health interventions: exploratory mapping review. *JMIR Ment Health* 2022 Jun 07;9(6):e35591 [FREE Full text] [doi: [10.2196/35591](https://doi.org/10.2196/35591)] [Medline: [35671081](https://pubmed.ncbi.nlm.nih.gov/35671081/)]
37. MS Symptoms and Signs of MS. National MS Society. URL: <https://www.nationalmssociety.org/Symptoms-Diagnosis/MS-Symptoms/Fatigue> [accessed 2022-05-02]
38. Newsome SD, Aliotta PJ, Bainbridge J, Bennett SE, Cutter G, Fenton K, et al. A framework of care in multiple sclerosis, part 2: symptomatic care and beyond. *Int J MS Care* 2017;19(1):42-56 [FREE Full text] [doi: [10.7224/1537-2073.2016-062](https://doi.org/10.7224/1537-2073.2016-062)] [Medline: [28243186](https://pubmed.ncbi.nlm.nih.gov/28243186/)]
39. Hinkle JL, Becker KJ, Kim JS, Choi-Kwon S, Saban KL, McNair N, American Heart Association Council on Cardiovascular and Stroke Nursing and Stroke Council. Poststroke fatigue: emerging evidence and approaches to management: a scientific statement for healthcare professionals from the American Heart Association. *Stroke* 2017 Jul;48(7):e159-e170. [doi: [10.1161/STR.000000000000132](https://doi.org/10.1161/STR.000000000000132)] [Medline: [28546322](https://pubmed.ncbi.nlm.nih.gov/28546322/)]
40. Hammell KW, Miller WC, Forwell SJ, Forman BE, Jacobsen BA. Managing fatigue following spinal cord injury: a qualitative exploration. *Disabil Rehabil* 2009;31(17):1437-1445. [doi: [10.1080/09638280802627694](https://doi.org/10.1080/09638280802627694)] [Medline: [19911479](https://pubmed.ncbi.nlm.nih.gov/19911479/)]
41. 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 Pt 1):1005-1026 [FREE Full text] [doi: [10.1111/j.1475-6773.2004.00269.x](https://doi.org/10.1111/j.1475-6773.2004.00269.x)] [Medline: [15230939](https://pubmed.ncbi.nlm.nih.gov/15230939/)]
42. Audio and Video Transcription Service Powered by AI. Trint. URL: <https://trint.com/technology> [accessed 2022-05-13]
43. Epharmix. URL: <https://www.epharmix.com/> [accessed 2022-04-22]
44. Widenius M, Axmark D, DuBois P. *MySQL Reference Manual: Documentation from the Source*. Sebastopol, CA, USA: O'Reilly Media; 2002.
45. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag Sci* 2000 Feb;46(2):186-204 [FREE Full text] [doi: [10.1287/mnsc.46.2.186.11926](https://doi.org/10.1287/mnsc.46.2.186.11926)]

46. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 2nd edition. Thousand Oaks, CA, USA: Sage Publications; 2004.
47. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. *Int J Qual Methods* 2020 Jan 22;19:160940691989922 [FREE Full text] [doi: [10.1177/1609406919899220](https://doi.org/10.1177/1609406919899220)]
48. NVivo (Version 12). QSR International. 2018. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home> [accessed 2022-05-20]
49. Kalapatapu RK, Cristian A. Assessment of fatigue in adults with disabilities. *Phys Med Rehabil Clin N Am* 2009 May;20(2):313-324. [doi: [10.1016/j.pmr.2008.12.001](https://doi.org/10.1016/j.pmr.2008.12.001)] [Medline: [19389613](https://pubmed.ncbi.nlm.nih.gov/19389613/)]
50. Hibbard J. Patient activation and health literacy: what's the difference? How do each contribute to health outcomes. *Stud Health Technol Inform* 2017;240:251-262. [Medline: [28972522](https://pubmed.ncbi.nlm.nih.gov/28972522/)]
51. Lewis E, Hassmén P, Pumpa KL. Participant perspectives of a telehealth trial investigating the use of telephone and text message support in obesity management: a qualitative evaluation. *BMC Health Serv Res* 2021 Jul 09;21(1):675 [FREE Full text] [doi: [10.1186/s12913-021-06689-6](https://doi.org/10.1186/s12913-021-06689-6)] [Medline: [34243772](https://pubmed.ncbi.nlm.nih.gov/34243772/)]
52. Cheung NW, Thiagalingam A, Smith BJ, Redfern J, Barry T, Mercorelli L, et al. Text messages promoting healthy lifestyle and linked with activity monitors stimulate an immediate increase in physical activity among women after gestational diabetes. *Diabetes Res Clin Pract* 2022 Aug;190:109991. [doi: [10.1016/j.diabres.2022.109991](https://doi.org/10.1016/j.diabres.2022.109991)] [Medline: [35835256](https://pubmed.ncbi.nlm.nih.gov/35835256/)]
53. Bae JW, Woo SI, Lee J, Park SD, Kwon SW, Choi SH, et al. mHealth interventions for lifestyle and risk factor modification in coronary heart disease: randomized controlled trial. *JMIR Mhealth Uhealth* 2021 Sep 24;9(9):e29928 [FREE Full text] [doi: [10.2196/29928](https://doi.org/10.2196/29928)] [Medline: [34559058](https://pubmed.ncbi.nlm.nih.gov/34559058/)]
54. Singleton A, Raeside R, Partridge SR, Hayes M, Maka K, Hyun KK, et al. Co-designing a lifestyle-focused text message intervention for women after breast cancer treatment: mixed methods study. *J Med Internet Res* 2021 Jun 14;23(6):e27076 [FREE Full text] [doi: [10.2196/27076](https://doi.org/10.2196/27076)] [Medline: [34125072](https://pubmed.ncbi.nlm.nih.gov/34125072/)]
55. Pathak LE, Aguilera A, Williams JJ, Lyles CR, Hernandez-Ramos R, Miramontes J, et al. Developing messaging content for a physical activity smartphone app tailored to low-income patients: user-centered design and crowdsourcing approach. *JMIR Mhealth Uhealth* 2021 May 19;9(5):e21177 [FREE Full text] [doi: [10.2196/21177](https://doi.org/10.2196/21177)] [Medline: [34009130](https://pubmed.ncbi.nlm.nih.gov/34009130/)]
56. Kruse G, Park ER, Shahid NN, Abrams L, Haberer JE, Rigotti NA. Combining real-time ratings with qualitative interviews to develop a smoking cessation text messaging program for primary care patients. *JMIR Mhealth Uhealth* 2019 Mar 26;7(3):e11498 [FREE Full text] [doi: [10.2196/11498](https://doi.org/10.2196/11498)] [Medline: [30912755](https://pubmed.ncbi.nlm.nih.gov/30912755/)]
57. Daud MH, Ramli AS, Abdul-Razak S, Isa MR, Yusoff FH, Baharudin N, et al. The EMPOWER-SUSTAIN e-Health Intervention to improve patient activation and self-management behaviours among individuals with Metabolic Syndrome in primary care: study protocol for a pilot randomised controlled trial. *Trials* 2020 Apr 05;21(1):311 [FREE Full text] [doi: [10.1186/s13063-020-04237-x](https://doi.org/10.1186/s13063-020-04237-x)] [Medline: [32248825](https://pubmed.ncbi.nlm.nih.gov/32248825/)]
58. Kendall E, Buys N. An integrated model of psychosocial adjustment following acquired disability. *J Rehabil* 1998;64(3):16-20.
59. Tate RL, Broe GA. Psychosocial adjustment after traumatic brain injury: what are the important variables? *Psychol Med* 1999 May;29(3):713-725. [doi: [10.1017/s0033291799008466](https://doi.org/10.1017/s0033291799008466)] [Medline: [10405093](https://pubmed.ncbi.nlm.nih.gov/10405093/)]
60. Al-Tabbaa O, Ankrah S. 'Engineered' university - industry collaboration: a social capital perspective. *Eur Manag Rev* 2019;16(3):543-565. [doi: [10.1111/emre.12174](https://doi.org/10.1111/emre.12174)]

Abbreviations

- CAB:** consumer advisory board
mHealth: mobile health
MS: multiple sclerosis
PAM: patient activation measure
SCI: spinal cord injury

Edited by A Mavragani; submitted 21.06.22; peer-reviewed by S Lopes, A Singleton, R Marshall; comments to author 27.09.22; revised version received 21.10.22; accepted 05.11.22; published 21.12.22.

Please cite as:

Morgan KA, Wong AWK, Walker K, Desai RH, Knepper TM, Newland PK
A Mobile Phone Text Messaging Intervention to Manage Fatigue for People With Multiple Sclerosis, Spinal Cord Injury, and Stroke: Development and Usability Testing
JMIR Form Res 2022;6(12):e40166
URL: <https://formative.jmir.org/2022/12/e40166>
doi: [10.2196/40166](https://doi.org/10.2196/40166)
PMID: [36542466](https://pubmed.ncbi.nlm.nih.gov/36542466/)

©Kerri A Morgan, Alex W K Wong, Kim Walker, Rachel Heeb Desai, Tina M Knepper, Pamela K Newland. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Two-way Automated Text Messaging Support From Community Pharmacies for Medication Taking in Multiple Long-term Conditions: Human-Centered Design With Nominal Group Technique Development Study

Gemma Donovan¹, MPharm, MSc; Nicola Hall², PhD; Felicity Smith³, BPharm, PhD; Jonathan Ling⁴, PhD; Scott Wilkes⁵, PhD

¹Generated Health, London, United Kingdom

²University of Sunderland, Sunderland, United Kingdom

³School of Pharmacy, University College London, London, United Kingdom

⁴Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, United Kingdom

⁵School of Medicine, University of Sunderland, Sunderland, United Kingdom

Corresponding Author:

Gemma Donovan, MPharm, MSc

Generated Health

Mercury House

117 Waterloo Road

London, SE1 8UL

United Kingdom

Phone: 44 345 5050120

Email: gemma.donovan@pshealthgroup.com

Abstract

Background: Reviews of digital communication technologies suggest that they can be effective in supporting medication use; however, their use alongside nondigital components is unclear. We also explored the delivery of a digital communication intervention in a relatively novel setting of community pharmacies and how such an intervention might be delivered to patients with multiple long-term conditions. This meant that despite the large number of intervention examples available in the literature, design questions remained, which we wanted to explore with key stakeholders. Examples of how to involve stakeholders in the design of complex health care interventions are lacking; however, human-centered design (HCD) has been suggested as a potential approach.

Objective: This study aimed to design a new community pharmacy text messaging intervention to support medication use for multiple long-term conditions, with patient and health care professional stakeholders in primary care.

Methods: HCD was used to map the intervention “journey” and identify design questions to explore with patients and health care professionals. Six prototypes were developed to communicate the intervention concept, and a modified version of the Nominal Group Technique was used to gather feedback. Nominal group meetings generated qualitative data using questions about the aspects that participants liked for each prototype and any suggested changes. The discussion was analyzed using a framework approach to transform feedback into statements. These statements were then ranked using a web-based questionnaire to establish a consensus about what elements of the design were valued by stakeholders and what changes to the design were most important.

Results: A total of 30 participants provided feedback on the intervention design concept over 5 nominal group meetings (21 health care professionals and 9 patients) with a 57% (17/30) response rate to the ranking questionnaire. Furthermore, 51 proposed changes in the intervention were generated from the framework analysis. Of these 51 changes, 27 (53%) were incorporated into the next design stage, focusing on changes that were ranked highest. These included suggestions for how text message content might be tailored, patient information materials, and the structure for pharmacist consultation. All aspects that the participants liked were retained in the future design and provided evidence that the proposed intervention concept had good acceptability.

Conclusions: HCD incorporating the Nominal Group Technique is an appropriate and successful approach for obtaining feedback from key stakeholders as part of an iterative design process. This was particularly helpful for our intervention, which combined digital and nondigital components for delivery in the novel setting of a community pharmacy. This approach enabled the collection

and prioritization of useful multiperspective feedback to inform further development and testing of our intervention. This model has the potential to minimize research waste by gathering feedback early in the complex intervention design process.

(*JMIR Form Res* 2022;6(12):e41735) doi:[10.2196/41735](https://doi.org/10.2196/41735)

KEYWORDS

medication adherence; text messaging; human-centered design; complex interventions; community pharmacy

Introduction

Background

Medication nonadherence is a known problem internationally, with 30%-50% of patients not taking medicines as prescribed, particularly in the case of long-term conditions [1]. This means that patients do not attain the health gains expected from medication and represent an avoidable cost to health care systems [2]. Patients are also increasingly managing multiple long-term conditions (MLTCs) [3], resulting in the requirement to manage multiple medicines. This poses additional challenges for patients [4].

Evidence suggests that digital communication technologies can improve medication adherence [5]; however, most interventions are not designed for patients with MLTCs [6]. Methods to develop such interventions are lacking, but updated guidance on complex intervention development from the Medical Research Council (MRC) suggests that approaches such as human-centered design (HCD) could be helpful [7]. Some reviews of digital communication to improve medication adherence have also suggested that their use may be optimized when delivered alongside other components such as face-to-face consultations or telephone appointments [8-10]. However, the contribution of these additional components to overall effectiveness is unclear [6].

Intervention Description

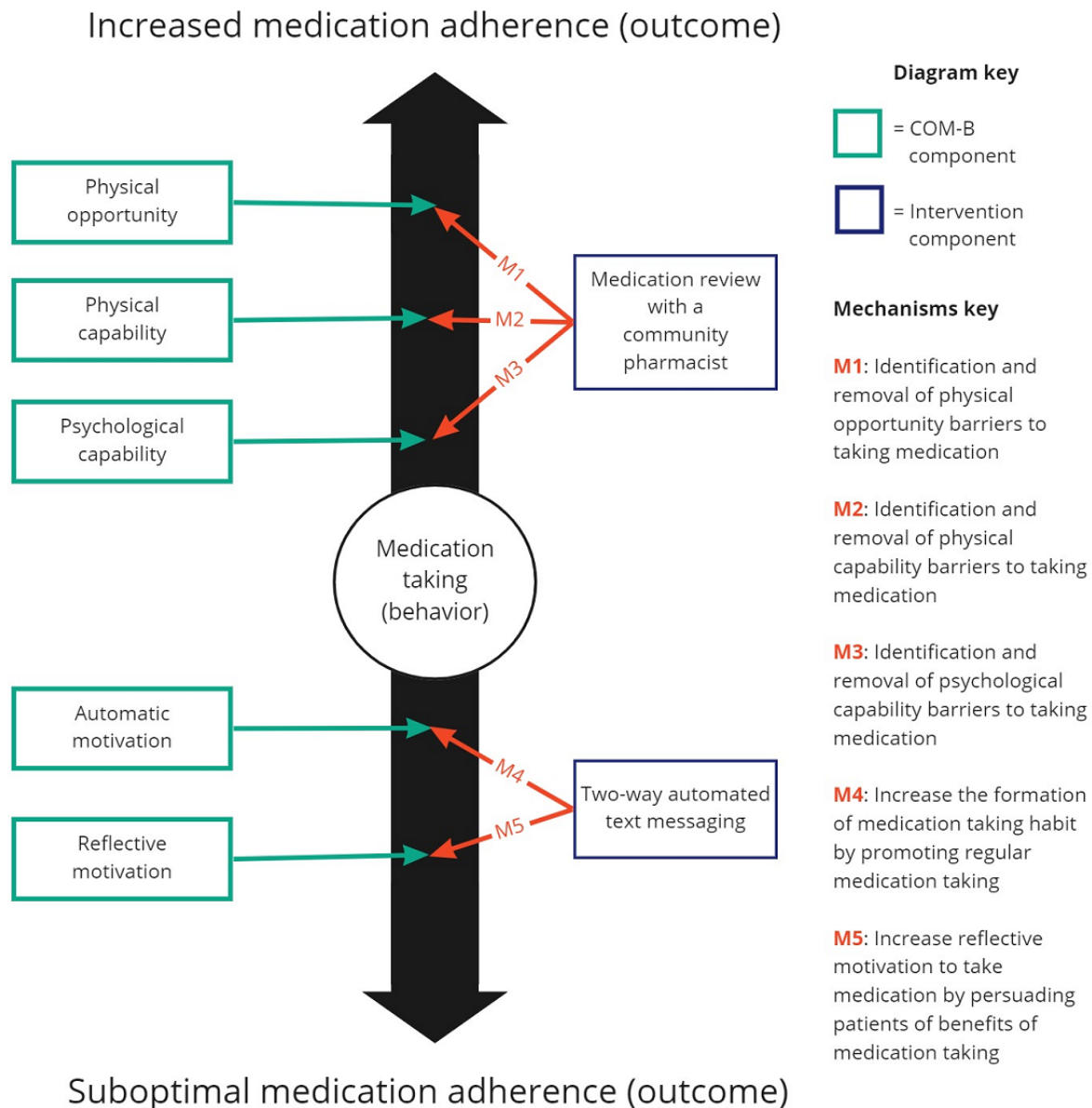
This study describes our first phase of development of a community pharmacy-delivered text messaging intervention to support medication use for patients with MLTCs. Our setting is in the United Kingdom National Health Service (NHS), where approximately 99% of community pharmacies have an NHS

contract to deliver health care services [11,12]. This includes the New Medicines Service [13,14] to support medication taking and, at the time of this project, Medicines Use Reviews [15,16], although these have since been decommissioned.

Our proposed intervention concept aimed to combine an automated two-way text messaging program with community pharmacy support. Our program theory used concepts from the Behavior Change Wheel [17] and was developed using intervention mapping [18]. The Behavior Change Wheel includes the Capability, Opportunity, and Motivation model (COM-B) for behavior performance, with each component able to affect behavior. Our program theory centered on how COM-B is applied to the behavior of “taking medication” to influence the outcome of medication adherence. Similar to others [19], we also considered medication adherence to be an outcome that exists in the spectrum between “suboptimal” and “increased” adherence levels.

Physical opportunity, physical capability, and psychological capability for taking medication by patients were intended to be assessed during a face-to-face pharmacist consultation, with barriers resolved by either the patient or pharmacist, depending on the barriers identified. Two-way automated text messaging aims to support habit formation via the Automatic Motivation component of the COM-B and influence Reflective Motivation for taking medication. An overview of the program theory can be found in [Figure 1](#) and was developed based on our systematic review [6] and experience as health care professionals (HCPs), with input from a project steering group that included patients. However, as examples of similar interventions were lacking and the intervention included both digital and nondigital components, we felt that an HCD design process would be helpful in further developing the intervention concept.

Figure 1. Program theory for how the new intervention combining 2-way automated text messaging delivered from a community pharmacy would work to support medication taking.



miro

Intervention Design Approach

Evidence from a systematic review [20] that informed MRC recommendations on the development of complex interventions suggests that the strength of HCD is its focus on patient experience. However, examples of how prototypes, a key feature of HCD, can support complex health care intervention designs are lacking. Our approach used the HCD toolkit developed by IDEO.org [21], where appropriate tools are recommended for selection depending on the individual intervention. In this study, we describe how we used prototyping to involve patients and HCPs in the intervention design.

We used the guidance for reporting intervention development studies in health research checklist [22] to construct this study, and a copy of the completed checklist is available in [Multimedia Appendix 1](#). This paper aimed to both provide readers with an understanding of our proposed intervention and an example

approach for using HCD and prototypes to develop complex health care interventions.

Methods

Overview

We used the HCD toolkit from IDEO.org [21] to guide the iterative design process and incorporated a modified version of Nominal Group Technique (NGT) [23] to gather and analyze feedback from patient and HCP participants. NGT was chosen to explore consensus on the most important aspects that participants liked and what changes were most important. This consensus was anticipated to be important for informing design changes in a scenario where it would be unfeasible to make all changes or where there may be incompatible changes suggested between participants, which would require resolution.

HCD and Prototype Development

Using the “journey map” tool from the IDEO.org HCD toolkit, a patient journey for the intervention idea was created (Figure 2) along with a series of design questions which were prioritized with input from the project steering committee. The steering committee included researchers, patients, a general practitioner (GP), community pharmacist, and a representative from the software provider (Florence [24]) to advise on how the technology could support intervention design. To explore the identified design questions, prototypes representing ideas related to the design of the intervention were created. These prototypes were used to gather feedback from the patients and HCPs. A list of design questions, prototypes, and participant groups is presented in Table 1. Who was asked to provide feedback on what prototype depended on the design questions to be answered. At this stage in the design process, we only sought feedback on text message content ideas from HCPs to obtain feedback on clinical acceptability, with the intention of receiving feedback on text message content from patients at a future stage in the development process. This was to ensure that we did not spend time creating text message content, which HCPs would not support the delivery of.

Six prototypes were created to support the discussion of the intervention design concept with the patients and HCPs. These included the following: (1) a video of a community pharmacy assistant inviting a patient to receive the new intervention; (2) a questionnaire to tailor the content of the automated 2-way text messaging; (3) an information leaflet for patients; (4) a video of the pharmacist consultation adapted to deliver the new intervention; (5) a document describing how the tailoring questionnaire would determine text messaging content; and (6) a diagram suggesting how community pharmacy and general practice teams might collaborate to support patients during intervention delivery.

A summary of how each prototype was developed can be found in Multimedia Appendix 2 [6,25-33]. The prototypes represented ideas about both intervention and implementation. As good acceptability is known to be important for intervention implementation [34], we wanted to explore this with key stakeholders, in addition to identifying potential changes to the design. Our key stakeholders included patients as end users and community pharmacists as intervention providers. In addition, we included general practice as the care provider responsible for prescribing and monitoring the medication. The involvement of wider primary care in community pharmacy intervention design has also been highlighted as important by others [35,36].

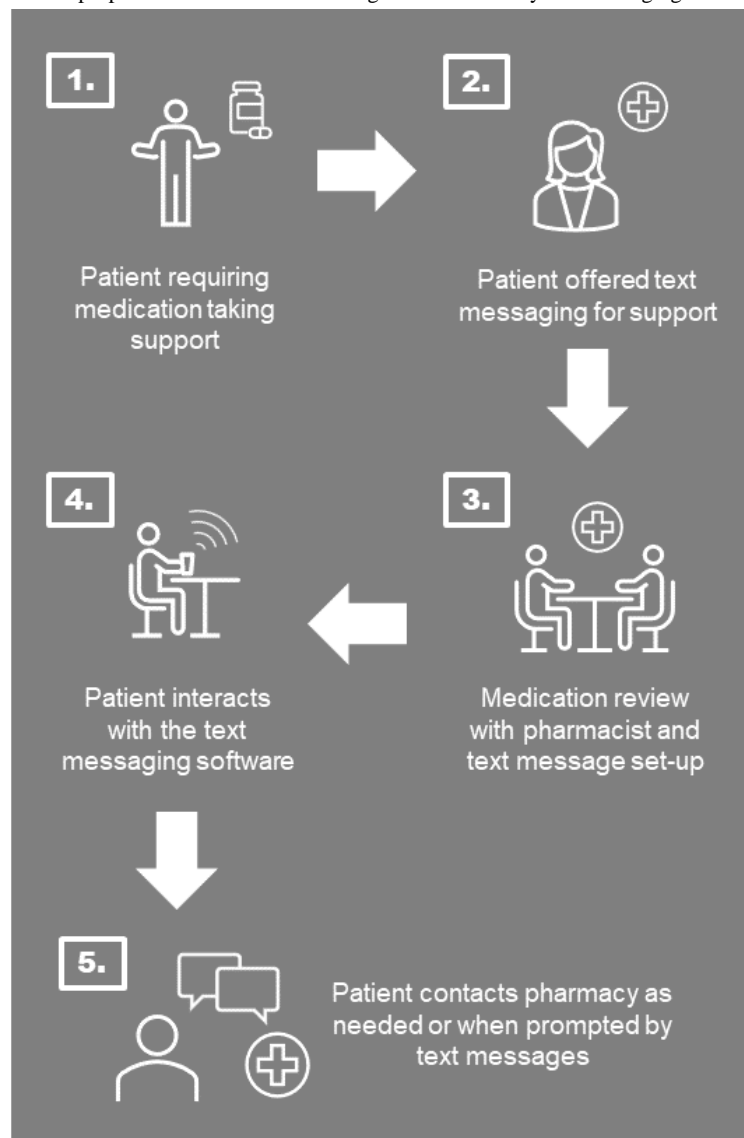
Figure 2. Patient “Journey Map” for the proposed intervention combining automated 2-way text messaging and community pharmacy support.

Table 1. Overview of the design questions, prototypes and participant feedback groups for the development study.

Step	Design question	Agreed prioritization with steering committee and rationale	Prototype used to explore	Feedback participant group
2.	What is the best way to approach patients to receive intervention in a community pharmacy setting?	High: as this is a new setting for a digital communication intervention in the United Kingdom	Video of PA ^a inviting patient to receive the intervention	Patients
2.	Who is the best person to approach them?	Medium: need to check if this needs to be a pharmacist or could be support staff	Video of PA inviting patient to receive the intervention	Patients
2.	What would encourage patients to find out more about the intervention?	High: need to ask about information needs to help patients decide if intervention will be helpful for them	Video of PA inviting patient to receive the intervention	Patients
3.	How should the pharmacist consultation be structured?	High: explore content and delivery for initial acceptability	Video of pharmacist consultation	Patients and HCPs ^b
3.	How would barriers to medication adherence be assessed?	High: explore useability of the tailoring tool, ease of completion	Tailoring questionnaire	Patients
3.	What information will patient need before setting up the text messaging?	High: identify key questions patients have about the text messaging	PIL ^c for intervention	Patients
3.	What should the information for patients look like?	Medium: explore length and presentation of the information materials	PIL for intervention	Patients
4.	What information should text messages contain?	Medium: information will be taken from other studies but will require a sense check	Document describing the TM ^d content tailoring process	HCPs
4.	Which messages should we ask patients to respond to?	Medium: information will be taken from other studies but will require a sense check	Document describing the TM content tailoring process	HCPs
4.	What information will we ask patients to send back?	Medium: information will be taken from other studies but will require a sense check	Document describing the TM content tailoring process	HCPs
5.	What happens if the pharmacy needs to refer the patient to another health care professional?	High: need to explore how communication between community pharmacies and general practice might work	Communication diagram	HCPs

^aPA: pharmacy assistant.

^bHCP: health care professional.

^cPIL: patient information leaflet.

^dTM: text message.

Feedback Using Modified NGT

Overview

Five nominal group meetings based on a modified NGT [23] were arranged to gather and analyze the feedback gathered using the intervention concept prototypes. The participants were provided with prompts to generate ideas about what they liked about the prototype and suggestions for changes. Ideas were initially generated silently, followed by sharing and discussions. Following qualitative analysis, ranking statements were generated and a web-based questionnaire was administered to rank these statements. Only 1 round of ranking was performed.

Participants

The participants included patients, community pharmacists, GPs, and practice nurses. Inclusion criteria for patients included active use of a mobile phone and ability to self-manage at least one long-term condition. As this was formative research, we did not wish to initially limit our sampling frame by dictating that participants had more than one long-term condition and

felt that those with only one condition would still provide useful feedback at this stage in the development process. HCPs were required to be currently providing patient-facing care.

Convenience sampling was used in this study. Patient participants were recruited through a patient, public, and caregiver involvement network hosted by the University of Sunderland. This network is a collection of people involved in health care teaching and research. HCPs were recruited via professional networks of the research team. Each participant was provided with an invitation letter, participant information sheet, and consent form in advance and were completed before data collection.

Ranking Statement Generation

Ranking statements were generated based on a qualitative analysis of the data generated during face-to-face nominal group meetings. The prototypes were presented individually. Elements that participants liked about the prototypes, and what they thought needed to be changed, were captured in the verbal discussion and on participants' written notes. Topic guides for

these discussions are available in [Multimedia Appendix 3](#). The nominal group meetings were audio recorded and transcribed verbatim before qualitative analysis. The meetings were facilitated by GD with contemporaneous notes taken by NH. Meetings were arranged separately for patients and HCPs, as shown in [Table 1](#). Five nominal group meetings were conducted in autumn 2018.

Transcripts and notes were analyzed using the framework approach [23] to identify the statements for the ranking exercise. Initially, an analytic framework was deductively applied. Data were coded for the prototype to which they related, and whether the data were related to a suggested change or aspect that

participants liked. Within these categories, individual suggestions were coded inductively to generate ranking statements. Examples of these processes are presented in [Table 2](#). Although the prototypes differed for patients and HCPs, there was some overlapping content. For example, the tailoring questions in the patient prototype also appeared in the HCP prototype which described how the tailoring questionnaire would determine text message content. Where feedback was for a different prototype to that examined directly, we decided to include it in the ranking exercise for the relevant participant group. The analysis was performed using NVivo 11 (QSR International) [37].

Table 2. Examples of how ranking statements were generated from qualitative analysis coding.

Qualitative extract from nominal group meetings	Deductive framework application	Inductively coded ranking statement
<ul style="list-style-type: none"> “I think it was explained well by the pharmacist about what was going to be involved in the scheme” [Patient, Focus Group] 2 “I think [the pharmacy assistant] gave [the patient] plenty of information” [Patient, Focus Group 3] 	Prototype: video of pharmacy assistant inviting patient to receive the intervention like statement	Right information given to allow the patient to make a decision
<ul style="list-style-type: none"> “...because of the way I’ve answered the questionnaire I won’t get a reminder.” [Patient, Focus Group 3] “A lot of people forget their statins at night. It wasn’t checked that he was taking his statin at night, but if he was forgetting at night, he might want to have it at night to remind him to take his statin” [General practitioner, Focus Group 5] 	Prototype: tailoring questionnaire suggested change	Ask whether medication reminders is something the patient would benefit from
<ul style="list-style-type: none"> “I really liked the MUR... I think having that conversation at the start is really good” [Pharmacist, Focus Group 1] “I think if the pharmacist looked at it and thought hang on a minute, why are they taking that in the morning and it’s definitely something they should be taking at night. It raises maybe a bit more care in the review” [Patient, Focus Group 2] 	Prototype: pharmacist consultation video like statement	Including a medication review as part of the set-up
<ul style="list-style-type: none"> “[The text messages put] the buck on them in a way that they’re going to have to be more responsible and I quite like that.” [Practice Nurse, Focus Group 4] “Minimal impact on clinician burden (patient ownership and responsibility placed on them)” [Extract from notes, Pharmacist, Focus Group 1] 	Prototype: document describing how tailoring questionnaire determines text messaging content like statement	The patient self-care emphasis which encourages patients to take responsibility
<ul style="list-style-type: none"> “So your target can be slightly different. Also depending which medications they’re on, because sometimes you just have to accept that level.” [Practice Nurse, Focus Group 4] 	Prototype: diagram suggesting community pharmacy and general practice collaboration suggested change	Confirm individual monitoring targets for patients with GP ^a practice before using home monitoring (eg, blood pressure targets for patients using home blood pressure monitoring)

^aGP: general practitioner.

Statement Ranking

The ranking statements were transferred to a web-based Qualtrics [38] questionnaire. Two versions of the questionnaire were created, one for patient participants and another for HCP participants, reflecting the prototypes examined in the nominal group meetings. Copies of these are available in [Multimedia Appendix 4](#). Participants were asked to rank 5 statements relating to the elements that they liked and then the suggested

change statements for each prototype individually. A ranking questionnaire was sent to all participants at nominal group meetings. The rank for the selected statements was then converted into a weighted score, with statements ranked first given a score of 5, second a score of 4, and so on. These scores were then summed to create a score across all the participants who provided feedback on the prototype.

Using the scores from the NGT ranking, decisions were then made about changes to the intervention design. Changes were

made when statements were either ranked in the top 3 most important changes or if the suggested changes were small amendments to documents or processes requiring minimal resource change to the overall intervention.

Ethics Approval

This study was approved by the London Riverside Research Ethics Committee (reference 18/LO/1201), Health Research Authority (IRAS ID:238,875), and University of Sunderland (reference number 002718). The participants were provided with a £20 (US \$24.58) gift voucher. No incentives were provided to HCP participants.

Results

Overview

A total of 9 patients participated in 2 meetings, and all but 1 had MLTCs. There were 21 HCPs across the 3 meetings. HCPs included pharmacists (n=7), practice nurses (n=5), and GPs (n=9). The average length of the nominal group meeting was 1 hour 21 minutes, with the fifth meeting intentionally shorter (59 minutes) to replace a meeting at a general practice site. The response rate to the ranking questionnaire was 57% (17/30; 6 patients and 11 HCPs). This means that the ranking score had a theoretical maximum of 30 for an individual statement evaluated by patients, 55 for HCPs, and 85 for statements ranked by both patients and HCPs. For most prototypes, there were more than 5 statements to rank for each prototype; therefore, the minimum ranking score could be 0, where it was not ranked as important by any participant.

The following results are organised by prototype. The change statements from the analysis are provided alongside their rank scores and whether changes were made by us to the intervention following feedback. Across all prototypes there were 51 proposed changes to the intervention generated by the analysis. Of these 51 changes, 27 (53%) were made at this point in the design process.

Video of a Community Pharmacy Assistant Inviting a Patient to Receive the New Intervention (Patient Feedback Only)

This video prototype showed a pharmacy assistant offering the new intervention to a patient waiting to collect their prescription from a community pharmacy. Following an initial expression of interest, the video showed that the patient was provided with the tailoring questionnaire and asked to complete it before a consultation with the pharmacist. The NGT statements for the aspects participants liked about this idea and be found in [Table 3](#), and the suggested changes can be found in [Table 4](#).

Feedback from the patient participants about the proposed design was generally positive. Patients liked the informal approach shown, and most felt that there was enough information provided for patients to decide whether they wanted to find out more. Patients also liked that there was no requirement for the patient to be identified as nonadherent to their medicines, and there was no pressure placed on the patient to sign up.

Most of the suggested changes were incorporated into a reiterated design; however, as the invitation to receive the intervention specified that it involved text messaging, we felt that actively asking patients if they had a mobile phone was an unnecessary change. The suggested change to offer the intervention to patients when problems were identified in a medication review was sensible, but as the intervention was designed for a scenario with no clinical issues to address, this represented a significant change to the context of the intervention and was therefore beyond the scope of the current intervention.

Overall, we were able to answer our original design questions regarding how patients might be invited to receive the proposed intervention. The use of a pharmacy assistant was not raised as something to change, and patients liked that the invitation used the preexisting relationship between the patient and pharmacy assistant. Using wording, which was nonjudgmental and focused on the intervention, increasing motivation to take medicines, was seen as a good way to encourage patients to find out more.

Table 3. Summary of Nominal Group Technique statements and scores for aspects that the participants liked about the video of pharmacy assistant inviting patients to the intervention.

Like statements	Ranking score (maximum 30)	Retained in design?
Right information given to allow the patient to make a decision	17	Yes
The informal approach	16	Yes
No pressure was put on the patient to sign up	15	Yes
The introduction was very general, not targeted at a specific patient based on a judgment of their previous compliance	13	Yes
There was an open amount of time given to complete the questionnaire	12	Yes
That it was built on an existing relationship between the patient and the pharmacy assistant	11	Yes

Table 4. Summary of Nominal Group Technique statements and scores for suggested changes to the video of pharmacy assistant inviting patients to the intervention.

Change statements	Ranking score (maximum 30)	Design changed?
Patient should be offered help to complete the questionnaire if they need it	28	Yes
The patient information leaflet should be offered before the patient is asked to complete the questionnaire	19	Yes
Patient should be offered the option to complete the questionnaire in the consultation room or at home and bring in later	16	Yes
Communication should be at the same level (eg, both sitting down or both standing)	15	Yes
The pharmacy assistant should ask the patient if they have a mobile phone before introducing them to the service	13	No
There needs to be a way of offering the service to patients who may have medicines delivered or who are housebound	11	Yes
Pharmacists should also offer the service if issues are identified as part of a medication review	5	No

Tailoring Questionnaire (Patient Feedback Only)

The tailored questionnaire prototype was designed to assess patient suitability for the proposed intervention and to inform the selection of text message content. This included an assessment of medication perceptions using the Beliefs about Medicines Questionnaire [25], the Automaticity subscale of the Self-Reported Habit Index [39], and questions about perceived medicine effectiveness adapted from Phillips et al [26] (see [Multimedia Appendix 2](#) for further details). Patient participants were asked to complete the questionnaire before providing feedback. The feedback statements and ranking scores are presented in [Tables 5](#) and [6](#).

Patients felt that the questionnaire was clear and easy to complete. During the meeting discussion, patients requested more information about how the questionnaire responses would be used to select the text message content. Further information was provided based on the suggestions in the prototype for principles of intervention tailoring. After receiving this information, patients felt they should be able to choose reminder

text messages rather than this being decided by an algorithm, and this was then the highest-ranked statement for change.

Another suggested change was to remove the “neither agree nor disagree” option from the responses. However, as these responses are components of validated tools, their removal was not felt to be appropriate. Participants reported that a question about caregivers would be helpful in understanding medication use. However, as carers would be a different end user group, this was beyond the scope of our intervention.

As the proposed intervention was designed for delivery in MLTCs, whether the questionnaire felt appropriate for patients in this context was a key question to answer using this prototype. Although it did not receive a high score in the ranking exercise, the feedback provided reassurance that the tailoring questionnaire was suitable for use in this group. One key change suggested by participants relating to MLTCs was to discuss long-term conditions verbally with the pharmacist and for pharmacists to complete this section of the questionnaire liaising with the patients’ GPs where needed, instead of patients completing this section using tick-boxes as suggested in the questionnaire prototype.

Table 5. Summary of Nominal Group Technique statements and scores for aspects that participants liked about the intervention tailoring questionnaire.

Like statements	Ranking score (maximum 30)	Retained in design?
Easy to read and understand	22	Yes
Clear layout	22	Yes
Use of tick-boxes for most of the questions	21	Yes
Questions did not feel too intrusive	17	Yes
Felt that my responses would identify any problems to address	2	Yes

Table 6. Summary of Nominal Group Technique statements and scores for suggested changes to the intervention tailoring questionnaire.

Change statements	Ranking score (maximum 30)	Design changed?
Ask whether medication reminders is something the patient would benefit from	18	Yes
Remove “neither agree nor disagree” option in the questionnaire responses so that people have to answer positively or negatively	15	No
Add in a space for the phone number to be given	14	Yes
Add a question asking if the patient has regular carers	12	No
Pharmacist completes long-term conditions, liaising with the GP ^a surgery instead of the patient completing this on the form	11	Yes
Add an additional statement in the questionnaire about medicines taking routine (eg, I have a routine for taking my medicines)	11	No
Add in a question to ask about who looks after the phone contract (eg, son or daughter)	9	Yes
Ask whether people would like information about text to voice functions available on their phone	0	No

^aGP: general practitioner.

Patient Information Leaflet (Patient Feedback Only)

The patient information leaflet aimed to provide information to prospective users of the intervention and support interaction with automated text messaging. Feedback from the NGT exercise revealed that the leaflet was easy to read and understand. The highest-ranked suggested changes included adding more information on what to expect during text messaging, such as the time for the system to respond and what would happen if patients made a typographical error when

replying to text messages. The statements and their ranks are listed in [Tables 7 and 8](#).

This feedback successfully answered the design questions regarding patients' information needs. Issues for delivering an intervention for MLTCs were raised again with this prototype, as participants requested more examples of text message content beyond the 4 examples included. However, it is unlikely that a full spectrum of examples would be feasible to include, so we chose not to make this change, but this would be subject to further testing using the experience from a future design phase.

Table 7. Summary of Nominal Group Technique statements and scores for the aspects participants liked about the intervention patient information leaflet.

Like statements	Ranking score (maximum 30)	Retained in design?
Easy to read and understand	29	Yes
Clear layout	21	Yes
Real examples of text messages the patient might receive	18	Yes
Covered most of the information the patient would need	14	Yes
Comments from other people who have used the service	8	Yes

Table 8. Summary of Nominal Group Technique statements and scores for suggested changes to the intervention patient information leaflet.

Change statements	Ranking score (maximum 30)	Design changed?
Add information on how long it will take Flo to respond	27	Yes
Include information on what happens if patient uses an error (eg, typo) in the message	26	Yes
Use real photos rather than graphics (eg, ClipArt)	23	Yes
Add space for a pharmacy stamp with name and contact details	20	Yes
Add in information about NHS 111	19	No
Include more general message examples (eg, not specific to high blood pressure)	17	No
Make emergency information more prominent	15	Yes
Change references to “SMS” to “text message”	6	Yes
Change “Flo says hello” to something more formal	0	No

Video of Pharmacist Consultation (Patient and HCP Feedback)

The video of the pharmacist consultation included the pharmacist reviewing the tailoring questionnaire and adding the patient to the software system, including receipt and sending of a confirmation text message by the patient. A summary of the aspects that participants liked is presented in Table 9 with the suggested changes in Table 10. As feedback was gathered on this prototype from both patients and HCPs, scores were presented for both groups along with the total score.

There was positive feedback from most participants on the pharmacist consultation structure proposed in the video prototype. Participants liked the Medicine Use Review format and felt that the medication review provided an opportunity to address medication-related issues that could not be addressed by text messaging. The text messaging set-up was also felt to work well, with participants liking a clear explanation of the service being offered and that this included personalizing the times that text messages were sent. Using face-to-face consultation was also highly valued by the patients and HCP participants.

Adding in a written consent process as part of the consultation was the highest-ranked change by HCPs, mainly due to high ranking by community pharmacists. Written consent was

prevalent in this setting at the time of data collection. However, these processes have since been removed, in part because of the COVID-19 pandemic; therefore, these processes have not been added for future implementation. Adding verbal information about data protection was a suggested change by HCPs but was a low priority for patients and therefore not prioritized for change in the next stage of intervention design as part of the consultation. This may be because information about this was provided in the patient information leaflet which was reviewed by patients, but not HCPs.

Patient participants' highest-ranked change was to include a verbal explanation that *Flo*, the persona used for the text messaging interaction, was not a real person. This information was provided in the patient information leaflet, but the patient participants felt that it was sufficiently important that it should also appear in the pharmacist consultation. Patient participants also suggested ensuring that long-term conditions and medication timing were captured and checked as part of the consultation rather than in the tailoring questionnaire. This was suggested so that any clinical issues could be identified, and support the selection of timing for text message reminders. These changes were included in the future iterations of the design. This prototype was successful in exploring our design questions regarding consultation content and delivery for our proposed intervention.

Table 9. Summary of Nominal Group Technique statements and scores for aspects participants liked about the pharmacist consultation video.

Like statements	HCP ^a rank score (maximum 55)	Patients rank score (maximum 30)	Total rank score (maximum 85)	Retained in design?
A clear explanation of the service being offered	36	17	53	Yes
Using a face-to-face method of communication	25	19	44	Yes
Ability of patients to choose the times messages were sent	22	5	27	Yes
Including a medication review as part of the set-up	18	9	27	Yes
Checking if the patient is experiencing any side effects from medication	10	9	19	Yes
Clear communication that the patient can opt out of receiving messages at any time	7	8	15	Yes
Providing a patient information leaflet	7	6	13	Yes
The opportunity to address adherence problems not covered by text messages	10	2	12	Yes
The use of Flo as a persona to communicate with	9	3	12	Yes
Taking place in a private consultation room	7	5	12	Yes
Explanation about the costs of participating to the patient	1	7	8	Yes
Setting up the service with a message in the consultation	7	0	7	Yes
Use of home monitoring equipment and sending in readings	6	0	6	Yes

^aHCP: health care professional.

Table 10. Summary of Nominal Group Technique statements and scores for suggested changes to intervention pharmacist consultation based on video prototype.

Change statements	HCP ^a rank score (maximum 55)	Patients rank score (maximum 30)	Total rank score (maximum 85)	Design changed?
Add in a more formal written consent process (eg, sign a consent form)	24	7	31	No
Make sure that timing of medication taking is captured and checked	17	14	31	Yes
Check patient knows how to correctly use home monitoring equipment in the consultation before use (eg, peak flowmeter)	22	7	29	Yes
Include a verbal explanation that Flo is not a real person	13	15	28	Yes
Cover data protection and regulation in the verbal consent process	18	6	24	No
Talk about the expected benefits of using text messages to support medicines taking	12	8	20	No
Option for consultation to be done in patients' home	11	9	20	No
Ensure that home blood pressure monitoring equipment is accurate (calibrated) prior to use	13	4	17	No
Confirm long-term conditions as part of the consultation	2	13	15	Yes
Add a question to assess adherence (eg, how many doses have you missed in the last 7 days)	9	4	13	No
Add in verbal instructions on how to cancel text messages	5	3	8	No
Provide an estimation of how many text messages the patient is likely to receive	8	0	8	No

^aHCP: health care professional.

Text Message Content Tailoring Document (HCP Feedback Only)

The feedback from HCP participants on the document describing how text message content would be tailored can be found in [Tables 11](#) and [12](#). The suggested tailoring process and the proposed text message content were well received, with participants liking the emphasis on self-care. Participants felt that the inclusion of medication reminders was a valuable component and liked that the suggested feedback and monitoring of medication taking allowed for “imperfect” adherence.

However, some of the example text messages were found to be inappropriate for some patients, especially those linked to the more extreme consequences of uncontrolled diseases. This included messages about the risk of amputation and excess health care costs associated with uncontrolled diabetes. It was agreed that these messages would be best maintained for future testing with patients.

A suggestion to provide home monitoring devices was also not changed, as this was felt to be unrealistic in the NHS given the potentially large number of patients who could receive the intervention. Another suggestion was to add a question about routine medication use, which would be unvalidated; therefore, it was also not included in the next design phase iteration. Using the intervention to support side effect monitoring was also suggested; however, we felt that the proposed pharmacist consultation was a better method to assess this.

Our design questions about what information should be requested from patients by the text messaging intervention were only partially answered by this feedback. The participants indicated that they liked the 2-way interaction, but there was little detail in the feedback on specific examples. However, there were no suggested changes in the feedback on the examples provided in the prototype. This suggests that the examples in the prototype were acceptable; however, this requires further investigation.

Table 11. Summary of Nominal Group Technique statements and scores for aspects participants liked about the document showing text message content selection based on tailoring questionnaire.

Like statements	Ranking score (maximum 55)	Retained in design?
The patient self-care emphasis which encourages patients to take responsibility	33	Yes
The tailoring of content to individual patients	30	Yes
Realistic targets which allow “imperfect” adherence	20	Yes
Providing information in smaller “chunks” which may be easier for the patient to digest	17	Yes
The simple language used in the messages	16	Yes
The inclusion of prompts or cues to support medicines taking	14	Yes
Messages tailored to patients’ beliefs about medication	14	Yes
Messages encouraging patients to get feedback on medicines taking (eg, blood pressure)	8	Yes
Two-way communication between the patient and Flo	7	Yes
Prioritization of concerns, then necessity, then experience, then habit	4	Yes
That the intervention is automated	1	Yes
The use of habit as a model for the messages	1	Yes

Table 12. Summary of Nominal Group Technique statements and scores for suggested changes to the document showing text message content selection based on tailoring questionnaire.

Change statements	Ranking score (maximum 55)	Design changed?
Create layers of messages, with more dramatic messages (eg, amputation being reserved for those with persistent nonadherence)	30	No
Reword the behavior experimentation message to seek approval from a health care professional before stopping medication to notice any impact	29	Yes
Provide home monitoring devices (eg, blood pressure monitor where messages are indicated but patients do not have the equipment)	23	No
Add an additional statement in the questionnaire about medicines taking routine (eg, I have a routine for taking my medicines)	22	No
Remove “neither agree nor disagree” option in the questionnaire responses so that people have to answer positively or negatively	21	No
Add in side effect monitoring as part of the intervention	17	No
Remove requirement to input keywords in responses such as “MEDS” or “DAYS”	10	No

Community Pharmacy and General Practice Collaboration Diagram (HCP Feedback Only)

The prototype showing how community pharmacy and general practice teams might work together to support patients receiving the new intervention is outlined in a document showing a series of flow diagrams. The proposed text messaging software allows all HCPs to access the same patient data (with patient permission), which has been highlighted as a method to support communication. Feedback from participants in the suggested collaboration model is provided in [Tables 13](#) and [14](#). Participants liked the community pharmacy-led design of the proposed intervention, and the data were still accessible to all HCPs through the software provider. Community pharmacist participants also liked the proposed use of pharmacy support staff to distribute the workload associated with the intervention.

The highest-ranked suggested change was to confirm patients’ individual monitoring targets (such as blood pressure) with general practices. However, this could add significantly to the intervention set-up. In addition, as blood pressure targets are

standardized in clinical guidance, we felt that the intervention should reflect these targets. However, this should be explored with patients using the system in the next stage of design development.

Our design question to answer using this prototype was to explore how pharmacies would coordinate with general practices when patients required clinical input or changes to medicines. Participants provided useful insights and suggestions on how best to achieve this based on current health care processes. This included a suggestion for general practice to add notifications indicating patient participation in the intervention into patient records. The participants also suggested that pharmacists should coordinate patient support arising from the intervention rather than patients directly contacting their general practice, as indicated in the prototype. This approach reflects other community pharmacy-led interventions such as the New Medicines Service.

There was some uncertainty in the feedback on how much information general practice staff needed about the intervention.

The suggestion of using a website to provide more detailed information is desirable. However, the suggested changes included both that practices only needed to be informed as courtesy and that notifications should include more detail. These would seem to be opposing pieces of feedback and, therefore, further exploration is needed about what level of information general practices might want and at what point.

The link between the intervention and the delivery of regular care by community pharmacies was also highlighted in the feedback on this prototype. There was a suggested change in that only a patient's nominated pharmacy could provide the intervention. The nominated pharmacy automatically receives

patient prescriptions from the NHS Electronic Prescription Service. This information is easily visible in general practice and is therefore seen as an important prerequisite for patients receiving the intervention. Pharmacies also felt that follow-up with patients shortly after initiating the intervention would be helpful to check that patients seemed to be receiving the correct text messages and any initial issues could be addressed. Participants also highlighted the need to ensure that patients could access support when they received a text message telling them to contact an HCP. This led to a change in the intervention to ensure that text messages are only sent to patients from Monday to Thursday.

Table 13. Summary of Nominal Group Technique statements and scores for aspects participants liked about the suggested collaboration model in community pharmacies and general practices.

Like statements	Ranking score (maximum 55)	Retained in design?
Community pharmacy-led service	41	Yes
That data is accessible to all health care professionals	40	Yes
Process is clear and makes sense	33	Yes
Makes good use of pharmacy support staff	20	Yes
Use of PharmOutcomes (a software platform for community pharmacy teams)	16	Yes
A website can act as a portal for more detailed information about specific content where needed	15	Yes

Table 14. Summary of Nominal Group Technique statements and scores for suggested changes to the collaboration model in community pharmacies and general practices.

Change statements	Ranking score (maximum 55)	Design changed?
Confirm individual monitoring targets for patients with GP ^a practice prior to using home monitoring (eg, blood pressure targets for patients using home blood pressure monitoring)	28	No
GP practices should add notification of patient using Flo to GP record, to ensure any medication changes are communicated to the pharmacy	23	Yes
Community pharmacies should contact the GP practice on behalf of patients initially where queries arise	22	Yes
Notification to practices should include which protocols have been set up for patients	20	Yes
General practice should receive notification of set up for information only	17	Yes
Add in a message to ask if the patient is happy with the messages so far shortly after initiation of intervention	17	Yes
The nominated pharmacy should be the only one able to provide the service	16	Yes
Messages should only be sent Monday to Thursday to allow quick access to health care professionals where there are queries	13	Yes

^aGP: general practitioner.

Discussion

Principal Findings

This paper includes a detailed description of our intervention, methods, and HCD approach used in its design and development. It provides a novel worked example of the application of HCD in developing prototypes for a complex health care intervention. Our HCD approach combined with modified NGT allowed us to successfully obtain and prioritize feedback from key stakeholders on our intervention concept for a community pharmacy-delivered text messaging intervention to support medication taking. The process facilitated the development of

our proposed intervention and informed changes to the intervention materials, including the tailoring questionnaire, patient information leaflet, and pharmacist consultation. The proposed implementation of the intervention was found to be acceptable to patients, community pharmacists, and general practice staff. Suggestions for changes in how the intervention will be implemented will be carried forward in the next iteration of the design.

Our findings provide confidence that our research-informed intervention will be acceptable and implementable to all key stakeholders. The open nature of the feedback gathering, aligned with the principles of HCD, led to some suggestions in the data

that were either outside the current design scope or contradicted the evidence base used to create the initial prototypes. In these cases, we labeled these as changes that were not implemented, although feedback could inform future alternative designs. The experience of the steering group was also used to make design change decisions alongside feedback from participants in this study, which also led to some suggestions not being taken forward. However, the generation and ranking of the feedback from key stakeholders in this study ensured that these opinions were heard and considered equally as part of the design process and reflected the intention of a co-design process similar to that suggested elsewhere [40].

Utility of HCD and Prototypes for the Proposed Intervention Design

Digital health care interventions do not exist in isolation; they are integrated into existing clinical pathways, so designing and gaining feedback on the whole patient journey as per HCD principles is important. For example, the discussions in our nominal group meetings were not always limited to one prototype at a time. Participants reflected on the relationship between the prototypes as the discussion progressed, leading to new feedback on the previously discussed prototypes. These ideas would not have been captured if feedback had been collected on individual components of the intervention at separate times. Our approach also allowed participants to see the relationships between the digital and nondigital components of the complex intervention, which can be overlooked in digital health care intervention design.

As the number of individuals involved in a co-design process is often limited, there is a potential risk that design outputs can ignore important evidence-based ideas. The process of creating prototypes allowed the incorporation of evidence and a theoretically informed approach at the start of the design process, but with the opportunity for the approach to be “sense checked” by key stakeholders. The unknowns associated with translating the research evidence into a novel setting could also be explicitly explored through our process, such as the acceptability of delivering text messaging from community pharmacies.

By using prototypes of the intervention, we were also able to explore a range of design questions without the need to build and test the intervention in the “real world.” This approach has the potential to minimize research waste, as feedback can be gathered on ideas early in the design process, preventing time spent on undesirable intervention components, or even stopping intervention design altogether where ideas are not acceptable to important stakeholders. Our approach may be particularly helpful when the design questions are more focused because evidence exists for most aspects of the intervention, but smaller changes, such as delivery context (as was the case in this scenario), require exploration.

Using NGT to Gather Feedback on Proposed Intervention Design

Using NGT to structure data collection and analysis increases the robustness of feedback gathering. Collating feedback statements across meetings for ranking allowed everyone to

consider all the generated statements. This was particularly useful for integrating feedback from both patients and HCPs in the pharmacist consultation prototype. Here, the elements that participants liked were ranked similarly, but the priority for changes differed between the participant groups. Overall, the questions we as designers thought were important (as articulated in our design questions) often did not rank highly as important by our participants. This provides evidence for the differing priorities between stakeholders and intervention designers, and the need for multiperspective input into complex health care intervention design.

The prioritization of feedback using NGT also helped us discern which aspects of the intervention and changes participants felt most strongly about. This was useful for suggestions that dominated discussions in meetings, which, if only analyzed qualitatively, may have indicated greater importance than the NGT scores revealed. However, as we did not have a high response rate, this could also be due to participants raising these issues in meetings, but not completing the ranking questionnaire.

We feel that using NGT offers a potential solution to some of the challenges associated with co-design processes, particularly the need to integrate feedback and ideas from a range of perspectives. The NGT also seemed to complement the use of prototypes, as there was consistency in how ideas were communicated to participants and, therefore, confidence that all feedback was linked to the same design ideas.

Comparison With Development Methods for Similar Interventions

We chose an HCD process as we felt it could accommodate both the digital and nondigital aspects of our proposed intervention and to ensure that we were creating a solution that met patients’ needs around medication taking. Guidance on developing digital-only health care interventions has been published by Abrams et al [40] and has been used for other digital communication interventions currently in development in the United Kingdom to support medication use [41,42]. Abrams et al [40] recommended the use of many of the features presented in this study, including a theoretically informed behavioral approach and the design of a delivery framework.

Other United Kingdom-based digital communication interventions to improve medication adherence have been designed for delivery from the general practice setting [41,42] and have used different approaches to ours. Interventions by Bartlett et al [42] for patients with diabetes and by Kassavou et al [43,44] for patients with hypertension and type 2 diabetes both seemed to focus initially on the digital communication content rather than implementation during the developmental process, which were explored simultaneously in our study.

Limitations

The discussions within nominal group meetings were not as multidisciplinary as we hoped to achieve. One nominal group meeting contained only pharmacists, and another was predominantly practice nurses. One meeting was multidisciplinary but was based on a single general practice. Combined, the data seem to offer a good range of feedback on

the initial design concept for the new intervention but may have benefited from a more multidisciplinary discussion.

Research using this type of methodology can be limited by the self-selection of participants in the studies in which they are interested. Therefore, broader perspectives may have been omitted from our results. Recruiting patient participants from a university-based network within a pharmacy school may have also increased the acceptability of the intervention. Our inclusion criteria also did not specify that patient participants needed to have MLTCs; however, most of our patients did, and many also drew on their experiences as caregivers for those who did. As these findings are just a starting point for further exploration and iteration, we believe that these limitations did not significantly affect the intervention development process.

Our modified version of the NGT included only 1 round of ranking that was not shared to inform a second round of discussion and voting, as per a normal NGT process. This means that we captured the individual perspectives of our participants but may have achieved a stronger consensus if a second round of discussion and ranking was conducted. However, given that we had a low response rate for the initial ranking questionnaire, we suspect that engagement in any further rounds of feedback and ranking would likely be poor. As we also considered all suggestions from the NGT data collection as part of a co-design process, we believe that generating a stronger consensus as defined by larger ranking scores across a smaller number of

statements would be unlikely to change the final design decisions we made.

Further Development and Research

The next step in intervention design is to develop a library of text messages for MLTCs that can be delivered using the approach described here. This will be used in a future study to further explore the delivery model using the IDEO.org “Live Prototyping” process, incorporating the changes from our results. This is important for gathering patient feedback on text message content, which was not included at this stage in the design process. Future studies should design training programs for community pharmacies and communication tools to be used in general practice. Each of these are intended to lead to an evaluation of the intervention in a real-world setting, which will also enable the testing of our intervention program theory.

Conclusions

Combining HCD with NGT allowed us to create a research-informed design for a text messaging intervention for medication adherence, gather feedback from key stakeholders, and reiterate for future testing. Although HCD has been proposed as a potential strategy for developing complex interventions in recent MRC guidance, examples of using such an approach are lacking. Our work can serve as a model for developing complex health care interventions in the future, especially where they combine digital and nondigital components.

Acknowledgments

GD acted as the chief investigator for the study, collected the study data, conducted data analysis, and prepared the initial manuscript. NH cofacilitated nominal group meetings and reviewed drafts of the manuscript. SW, FS, and JL provided input into the study design and feedback on the manuscript drafts. Gemma Donovan was funded by a National Institute for Health Research Doctoral Research Fellowship (2016-09-163). The views expressed are those of the authors and not necessarily those of the National Institute for Health and Care Research or Department of Health and Social Care.

Data Availability

We did not seek consent from participants or approval for data sharing beyond the study team; however, data requests may be made to the corresponding author for consideration on a case-by-case basis.

Conflicts of Interest

GD is employed by Generated Health, the company which owns and operate Florence Intelligent Health Messaging software; however, this work was undertaken before her employment. Generated Health have not influenced the design, data collection, analysis, interpretation, or write-up of this publication.

Multimedia Appendix 1

Guidance for reporting intervention development studies in health research checklist.

[[PDF File \(Adobe PDF File\), 222 KB - formative_v6i12e41735_app1.pdf](#)]

Multimedia Appendix 2

Detailed descriptions of prototype development.

[[DOCX File , 1365 KB - formative_v6i12e41735_app2.docx](#)]

Multimedia Appendix 3

Nominal group topic guides.

[[DOCX File , 2456 KB - formative_v6i12e41735_app3.docx](#)]

Multimedia Appendix 4

Ranking questionnaires.

[\[DOCX File , 2950 KB - formative_v6i12e41735_app4.docx \]](#)**References**

1. Organisation mondiale de la santé, World Health Organization. Adherence to Long-term Therapies Evidence for Action. Geneva: World Health Organization; 2003.
2. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. *BMJ Open* 2018 Jan 21;8(1):e016982 [FREE Full text] [doi: [10.1136/bmjopen-2017-016982](https://doi.org/10.1136/bmjopen-2017-016982)] [Medline: [29358417](https://pubmed.ncbi.nlm.nih.gov/29358417/)]
3. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012 Jul 07;380(9836):37-43 [FREE Full text] [doi: [10.1016/S0140-6736\(12\)60240-2](https://doi.org/10.1016/S0140-6736(12)60240-2)] [Medline: [22579043](https://pubmed.ncbi.nlm.nih.gov/22579043/)]
4. Vetrano DL, Bianchini E, Onder G, Cricelli I, Cricelli C, Bernabei R, et al. Poor adherence to chronic obstructive pulmonary disease medications in primary care: role of age, disease burden and polypharmacy. *Geriatr Gerontol Int* 2017 Dec 28;17(12):2500-2506. [doi: [10.1111/ggi.13115](https://doi.org/10.1111/ggi.13115)] [Medline: [28656664](https://pubmed.ncbi.nlm.nih.gov/28656664/)]
5. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keenanasseril A, et al. Interventions for enhancing medication adherence. *Cochrane Database Syst Rev* 2014 Nov 20(11):CD000011 [FREE Full text] [doi: [10.1002/14651858.CD000011.pub4](https://doi.org/10.1002/14651858.CD000011.pub4)] [Medline: [25412402](https://pubmed.ncbi.nlm.nih.gov/25412402/)]
6. Donovan G, Hall N, Ling J, Smith F, Wilkes S. Influencing medication taking behaviors using automated two-way digital communication: a narrative synthesis systematic review informed by the Behavior Change Wheel. *Br J Health Psychol* 2022 Sep 26;27(3):861-890 [FREE Full text] [doi: [10.1111/bjhp.12580](https://doi.org/10.1111/bjhp.12580)] [Medline: [35080811](https://pubmed.ncbi.nlm.nih.gov/35080811/)]
7. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open* 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: [10.1136/bmjopen-2019-029954](https://doi.org/10.1136/bmjopen-2019-029954)] [Medline: [31420394](https://pubmed.ncbi.nlm.nih.gov/31420394/)]
8. Fenerty SD, West C, Davis SA, Kaplan SG, Feldman SR. The effect of reminder systems on patients' adherence to treatment. *Patient Prefer Adherence* 2012;6:127-135 [FREE Full text] [doi: [10.2147/PPA.S26314](https://doi.org/10.2147/PPA.S26314)] [Medline: [22379363](https://pubmed.ncbi.nlm.nih.gov/22379363/)]
9. Mistry N, Keenanasseril A, Wilczynski NL, Nieuwlaat R, Ravall M, Haynes RB, Patient Adherence Review Team. Technology-mediated interventions for enhancing medication adherence. *J Am Med Inform Assoc* 2015 Apr;22(e1):e177-e193 [FREE Full text] [doi: [10.1093/jamia/ocu047](https://doi.org/10.1093/jamia/ocu047)] [Medline: [25726568](https://pubmed.ncbi.nlm.nih.gov/25726568/)]
10. Ciciriello S, Johnston RV, Osborne RH, Wicks I, deKroo T, Clerehan R, et al. Multimedia educational interventions for consumers about prescribed and over-the-counter medications. *Cochrane Database Syst Rev* 2013 Apr 30(4):CD008416. [doi: [10.1002/14651858.CD008416.pub2](https://doi.org/10.1002/14651858.CD008416.pub2)] [Medline: [23633355](https://pubmed.ncbi.nlm.nih.gov/23633355/)]
11. Registers. General Pharmaceutical Council. URL: <https://www.pharmacyregulation.org/registers> [accessed 2022-03-18]
12. General pharmaceutical services in England 2015/16 - 2020/21. NHS Business Services Authority. 2021 Oct 28. URL: <https://www.nhs.uk/statistical-collections/general-pharmaceutical-services-england/general-pharmaceutical-services-england-201516-202021> [accessed 2022-03-18]
13. New medicine service (NMS). Pharmaceutical Services Negotiating Committee. 2013 Jun 18. URL: <https://psnc.org.uk/national-pharmacy-services/advanced-services/nms/> [accessed 2022-03-18]
14. Elliott RA, Boyd MJ, Tanajewski L, Barber N, Gkoutouras G, Avery AJ, et al. 'New Medicine Service': supporting adherence in people starting a new medication for a long-term condition: 26-week follow-up of a pragmatic randomised controlled trial. *BMJ Qual Saf* 2020 Apr;29(4):286-295 [FREE Full text] [doi: [10.1136/bmjqs-2018-009177](https://doi.org/10.1136/bmjqs-2018-009177)] [Medline: [31732700](https://pubmed.ncbi.nlm.nih.gov/31732700/)]
15. Latif A, Pollock K, Boardman HF. The contribution of the Medicines Use Review (MUR) consultation to counseling practice in community pharmacies. *Patient Educ Couns* 2011 Jun;83(3):336-344 [FREE Full text] [doi: [10.1016/j.pec.2011.05.007](https://doi.org/10.1016/j.pec.2011.05.007)] [Medline: [21621943](https://pubmed.ncbi.nlm.nih.gov/21621943/)]
16. National Pharmacy Services. Pharmaceutical Services Negotiating Committee. 2013 Apr 25. URL: <https://psnc.org.uk/national-pharmacy-services/> [accessed 2022-03-18]
17. 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](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
18. Kok G, Gottlieb NH, Peters GY, Mullen PD, Parcel GS, Ruiters RA, et al. A taxonomy of behaviour change methods: an intervention mapping approach. *Health Psychol Rev* 2016 Sep;10(3):297-312 [FREE Full text] [doi: [10.1080/17437199.2015.1077155](https://doi.org/10.1080/17437199.2015.1077155)] [Medline: [26262912](https://pubmed.ncbi.nlm.nih.gov/26262912/)]
19. Jackson C, Eliasson L, Barber N, Weinman J. Applying COM-B to medication adherence. *Eur Health Psychol* 2014;16(1).
20. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. *Pilot Feasibility Stud* 2019 Mar 12;5(1):41 [FREE Full text] [doi: [10.1186/s40814-019-0425-6](https://doi.org/10.1186/s40814-019-0425-6)] [Medline: [30923626](https://pubmed.ncbi.nlm.nih.gov/30923626/)]

21. IDEO (Firm). The Field Guide to Human-centered Design Design Kit. San Francisco, California, United States: IDEO; 2015.
22. Duncan E, O’Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. *BMJ Open* 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: [10.1136/bmjopen-2019-033516](https://doi.org/10.1136/bmjopen-2019-033516)] [Medline: [32273313](https://pubmed.ncbi.nlm.nih.gov/32273313/)]
23. Van de Ven AH, Delbecq AL. The nominal group as a research instrument for exploratory health studies. *Am J Public Health* 1972 Mar;62(3):337-342. [doi: [10.2105/ajph.62.3.337](https://doi.org/10.2105/ajph.62.3.337)] [Medline: [5011164](https://pubmed.ncbi.nlm.nih.gov/5011164/)]
24. Florence. Generated Health. URL: <https://getflorence.net/> [accessed 2022-08-06]
25. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychol Health* 1999 Jan 15;14(1):1-24. [doi: [10.1080/08870449908407311](https://doi.org/10.1080/08870449908407311)]
26. Phillips LA, Cohen J, Burns E, Abrams J, Renninger S. Self-management of chronic illness: the role of 'habit' versus reflective factors in exercise and medication adherence. *J Behav Med* 2016 Dec;39(6):1076-1091. [doi: [10.1007/s10865-016-9732-z](https://doi.org/10.1007/s10865-016-9732-z)] [Medline: [26980098](https://pubmed.ncbi.nlm.nih.gov/26980098/)]
27. Your Introduction to the Florence Telehealth System. Simple Telehealth Member’s Resources. URL: <https://docs.google.com/viewer?a=v&pid=sites&srcid=c2ltcGxILnVrLm5ldHxjb21tdW5pdHI8Z3g6NGU2NzA3OWYxMTc4Y2M5OQ> [accessed 2022-10-02]
28. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. *J Appl Social Psychol* 2003 Jun;33(6):1313-1330. [doi: [10.1111/j.1559-1816.2003.tb01951.x](https://doi.org/10.1111/j.1559-1816.2003.tb01951.x)]
29. Gardner B, Abraham C, Lally P, de Bruijn G. Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *Int J Behav Nutr Phys Act* 2012 Aug 30;9(1):102 [FREE Full text] [doi: [10.1186/1479-5868-9-102](https://doi.org/10.1186/1479-5868-9-102)] [Medline: [22935297](https://pubmed.ncbi.nlm.nih.gov/22935297/)]
30. Horne R. Representations of medication and treatment: advances in theory and measurement. In: *Perceptions of Health & Illness*. London, England, United Kingdom: Psychology Press; 1997.
31. Cooper V, Buick D, Horne R, Lambert N, Gellaitry G, Leake H, et al. Perceptions of HAART among gay men who declined a treatment offer: preliminary results from an interview-based study. *AIDS Care* 2002 Jun 27;14(3):319-328. [doi: [10.1080/09540120220123694](https://doi.org/10.1080/09540120220123694)] [Medline: [12042077](https://pubmed.ncbi.nlm.nih.gov/12042077/)]
32. Horne R. Treatment perceptions and self-regulation. In: *The Self-regulation of Health and Illness Behaviour*. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2003.
33. PharmOutcomes homepage. PharmOutcomes. URL: <https://pharmoutcomes.org/pharmoutcomes/> [accessed 2022-11-25]
34. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017 Jan 26;17(1):88 [FREE Full text] [doi: [10.1186/s12913-017-2031-8](https://doi.org/10.1186/s12913-017-2031-8)] [Medline: [28126032](https://pubmed.ncbi.nlm.nih.gov/28126032/)]
35. Latif A, Waring J, Watmough D, Barber N, Chuter A, Davies J, et al. Examination of England’s New Medicine Service (NMS) of complex health care interventions in community pharmacy. *Res Social Adm Pharm* 2016 Nov;12(6):966-989. [doi: [10.1016/j.sapharm.2015.12.007](https://doi.org/10.1016/j.sapharm.2015.12.007)] [Medline: [26806858](https://pubmed.ncbi.nlm.nih.gov/26806858/)]
36. Watson M, Silver K, Watkins R. "What counts can't always be measured": a qualitative exploration of general practitioners' conceptualisation of quality for community pharmacy services. *BMC Fam Pract* 2020 Nov 28;21(1):244 [FREE Full text] [doi: [10.1186/s12875-020-01319-2](https://doi.org/10.1186/s12875-020-01319-2)] [Medline: [33248452](https://pubmed.ncbi.nlm.nih.gov/33248452/)]
37. NVivo homepage. NVivo. URL: <https://www.qsrinternational.com/nvivo/home> [accessed 2022-11-25]
38. Qualtrics XM. Qualtrics. URL: <https://www.qualtrics.com> [accessed 2022-11-25]
39. Aarts H, Verplanken B, Knippenberg A. Predicting behavior from actions in the past: repeated decision making or a matter of habit? *J Appl Social Psychol* 1998 Aug;28(15):1355-1374 [FREE Full text] [doi: [10.1111/j.1559-1816.1998.tb01681.x](https://doi.org/10.1111/j.1559-1816.1998.tb01681.x)]
40. Abrams 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](https://doi.org/10.2196/mhealth.4917)] [Medline: [26690917](https://pubmed.ncbi.nlm.nih.gov/26690917/)]
41. Berman A, Uribe-Rodríguez AF, Pérez-Rivero PF, Prieto-Merino D, Cáceres Rivera DI, Guio E, et al. Evaluation of the efficacy and safety of text messages targeting adherence to cardiovascular medications in secondary prevention: the txt2heart Colombia randomised controlled trial protocol. *BMJ Open* 2019 Dec 08;9(12):e028017. [doi: [10.1136/bmjopen-2018-028017](https://doi.org/10.1136/bmjopen-2018-028017)]
42. Bartlett YK, Farmer A, Rea R, French DP. Use of brief messages based on behavior change techniques to encourage medication adherence in people with type 2 diabetes: developmental studies. *J Med Internet Res* 2020 May 13;22(5):e15989 [FREE Full text] [doi: [10.2196/15989](https://doi.org/10.2196/15989)] [Medline: [32401214](https://pubmed.ncbi.nlm.nih.gov/32401214/)]
43. Kassavou A, Houghton V, Edwards S, Brimicombe J, Sutton S. Development and piloting of a highly tailored digital intervention to support adherence to antihypertensive medications as an adjunct to primary care consultations. *BMJ Open* 2019 Jan 06;9(1):e024121 [FREE Full text] [doi: [10.1136/bmjopen-2018-024121](https://doi.org/10.1136/bmjopen-2018-024121)] [Medline: [30613027](https://pubmed.ncbi.nlm.nih.gov/30613027/)]
44. Kassavou A, Mirzaei V, Brimicombe J, Edwards S, Massou E, Prevost AT, et al. A highly tailored text and voice messaging intervention to improve medication adherence in patients with either or both hypertension and type 2 diabetes in a UK primary care setting: feasibility randomized controlled trial of clinical effectiveness. *J Med Internet Res* 2020 May 19;22(5):e16629 [FREE Full text] [doi: [10.2196/16629](https://doi.org/10.2196/16629)] [Medline: [32427113](https://pubmed.ncbi.nlm.nih.gov/32427113/)]

Abbreviations

COM-B: Capability, Opportunity, and Motivation model

GP: general practitioner

HCD: human-centered design

HCP: health care professional

MLTC: multiple long-term condition

MRC: Medical Research Council

NGT: Nominal Group Technique

NHS: National Health Service

Edited by A Mavragani; submitted 06.08.22; peer-reviewed by R Brar Prayaga, M Östbring; comments to author 06.10.22; revised version received 22.10.22; accepted 24.10.22; published 21.12.22.

Please cite as:

Donovan G, Hall N, Smith F, Ling J, Wilkes S

Two-way Automated Text Messaging Support From Community Pharmacies for Medication Taking in Multiple Long-term Conditions: Human-Centered Design With Nominal Group Technique Development Study

JMIR Form Res 2022;6(12):e41735

URL: <https://formative.jmir.org/2022/12/e41735>

doi: [10.2196/41735](https://doi.org/10.2196/41735)

PMID: [36542458](https://pubmed.ncbi.nlm.nih.gov/36542458/)

©Gemma Donovan, Nicola Hall, Felicity Smith, Jonathan Ling, Scott Wilkes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Digital Medicine System in Veterans With Severe Mental Illness: Feasibility and Acceptability Study

Sarah Gonzales^{1,2}, BA; Olaoluwa O Okusaga^{3,4}, MD, MscPHR; J Corey Reuteman-Fowler⁵, PhD; Megan M Oakes^{1,2}, MS; Jamie N Brown⁶, PharmD; Scott Moore^{7,8}, MD, PhD; Allison A Lewinski^{1,9}, PhD; Cristin Rodriguez^{3,4}, BS; Norma Moncayo^{3,4}, MD; Valerie A Smith^{1,2,10}, DrPH; Shauna Malone^{1,2}; Justine List^{3,4}, BAsC; Raymond Y Cho^{3,4}, MD; Amy S Jeffreys¹, MStat; Hayden B Bosworth^{1,2,8,10}, PhD

¹Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham Veterans Affairs Medical Center, Durham, NC, United States

²Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, United States

³Mental Health Care Line, Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX, United States

⁴Department of Psychiatry and Behavioral Health Sciences, Baylor College of Medicine, Houston, TX, United States

⁵Global Clinical Development, Otsuka Pharmaceutical Development and Commercialization Inc., Princeton, NJ, United States

⁶Pharmacy Service, Durham Veterans Affairs Health Care System, Durham, NC, United States

⁷Durham Veterans Affairs Medical Center, Durham, NC, United States

⁸Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, United States

⁹School of Nursing, Duke University, Durham, NC, United States

¹⁰Division of General Internal Medicine, Department of Medicine, Duke University, Durham, NC, United States

Corresponding Author:

Hayden B Bosworth, PhD

Durham Center of Innovation to Accelerate Discovery and Practice Transformation

Durham Veterans Affairs Medical Center

411 W. Chapel Hill Street

Suite 600

Durham, NC, 27701

United States

Phone: 1 919 286 0411 ext 7101

Email: hayden.bosworth@duke.edu

Abstract

Background: Suboptimal medication adherence is a significant problem for patients with serious mental illness. Measuring medication adherence through subjective and objective measures can be challenging, time-consuming, and inaccurate.

Objective: The primary purpose of this feasibility and acceptability study was to evaluate the impact of a digital medicine system (DMS) among Veterans (patients) with serious mental illness as compared with treatment as usual (TAU) on medication adherence.

Methods: This open-label, 2-site, provider-randomized trial assessed aripiprazole refill adherence in Veterans with schizophrenia, schizoaffective disorder, bipolar disorder, or major depressive disorder. We randomized 26 providers such that their patients either received TAU or DMS for a period of 90 days. Semistructured interviews with patients and providers were used to examine the feasibility and acceptability of using the DMS.

Results: We enrolled 46 patients across 2 Veterans Health Administration sites: 21 (46%) in DMS and 25 (54%) in TAU. There was no difference in the proportion of days covered by medication refill over 3 and 6 months (0.82, SD 0.24 and 0.75, SD 0.26 in DMS vs 0.86, SD 0.19 and 0.82, SD 0.21 in TAU, respectively). The DMS arm had 0.85 (SD 0.20) proportion of days covered during the period they were engaged with the DMS (mean 144, SD 100 days). Interviews with patients (n=14) and providers (n=5) elicited themes salient to using the DMS. Patient findings described the positive impact of the DMS on medication adherence, challenges with the DMS patch connectivity and skin irritation, and challenges with the DMS app that affected overall use. Providers described an overall interest in using a DMS as an objective measure to support medication adherence in their patients. However, providers described challenges with the DMS dashboard and integrating DMS data into their workflow, which decreased the usability of the DMS for providers.

Conclusions: There was no observed difference in refill rates. Among those who engaged in the DMS arm, the proportion of days covered by refills were relatively high (mean 0.85, SD 0.20). The qualitative analyses highlighted areas for further refinement of the DMS.

Trial Registration: ClinicalTrials.gov NCT03881449; <https://clinicaltrials.gov/ct2/show/NCT03881449>

(*JMIR Form Res* 2022;6(12):e34893) doi:[10.2196/34893](https://doi.org/10.2196/34893)

KEYWORDS

ABILIFY MYCITE; digital medicine; adherence; aripiprazole; Veterans; qualitative methods; mental illness; mental health; medication; mobile phone

Introduction

Background

Suboptimal medication adherence is a significant problem for patients with serious mental illness (SMI), including those with schizophrenia, bipolar disorder, posttraumatic stress disorder, and major depressive disorder. Suboptimal adherence among these individuals may lead to symptom exacerbation, relapse, and hospital readmissions [1]. Moreover, suboptimal adherence to prescribed psychotropic medication is associated with increased mortality [1,2]. However, measuring medication adherence in the clinical setting is challenging and primarily subjective (ie, patient self-report) [3,4]. While objective methods such as pill counts and pharmacy refill data are helpful, they are time-consuming to calculate and often do not provide an accurate assessment of actual medication ingestion [5].

A digital medicine system (DMS), consisting of a drug-device combination, is a way to obtain objective treatment adherence data. DMS may enable patients with SMI to measure and report ingestion of atypical antipsychotic medications, most of which have broad therapeutic indications [6]. The collection of objective real-time data using a DMS enables providers to address nonadherence to medications, as well as facilitate interactions among patients and providers to promote and support medication adherence. Thus, adherence data obtained from a DMS may assist in understanding potential barriers to improving outcomes and more informed shared decision-making regarding a patient's treatment for individuals with SMI [6].

The ABILIFY MYCITE System

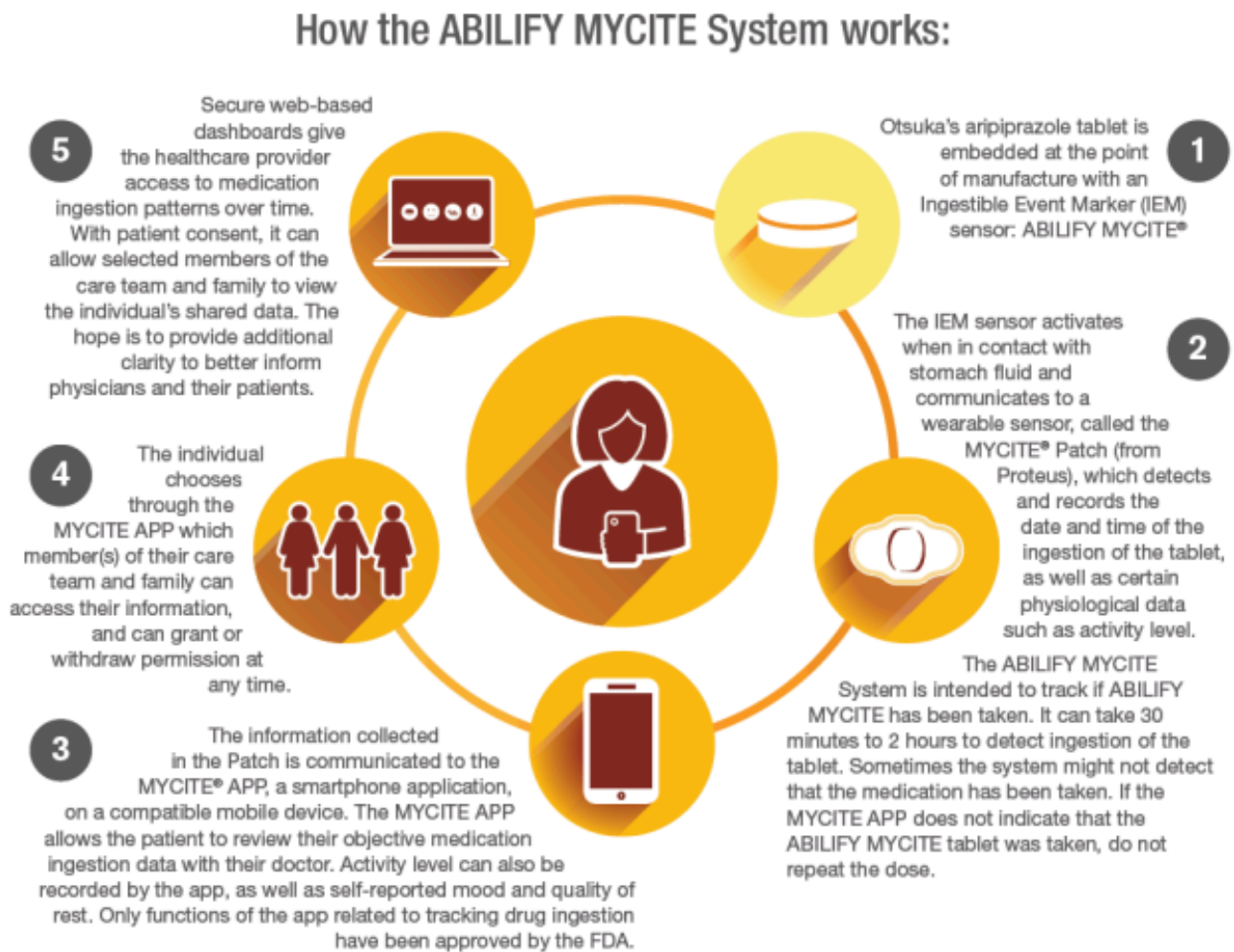
The ABILIFY MYCITE System is an example of a DMS developed to track adherence to oral aripiprazole, an atypical antipsychotic. This system enables patients and their mental health providers the opportunity to view real-time adherence data. Specifically, the ABILIFY MYCITE System is a drug-device combination product (aripiprazole tablets with a sensor) that comprises 4 separate components that enable the monitoring of treatment adherence by a patient and the patient's provider (Table 1). The four components include (1) the ABILIFY MYCITE tablet (DMS tablet), an aripiprazole tablet embedded with an ingestible event marker sensor; (2) MYCITE System patch (DMS patch); (3) MYCITE System smartphone app (DMS app); and (4) MYCITE System dashboard (DMS dashboard).

Once a participant swallows the tablet, the ingestible sensor transmits an electrical signal that is detected and then recorded by software within the patch that is worn by the participant on the left rib cage. Using Bluetooth, the patch then transmits the aripiprazole ingestion data to the participant's smartphone, which is then saved to the secure, cloud-based MYCITE System dashboard. Participants have the ability to view this medication data each day on their smartphone, while the patient's provider, study team, and selected caregivers are able to view the data on the MYCITE System dashboard via the cloud-based server (Figure 1).

Table 1. Description of the 4 components of the digital medicine system (DMS)—Abilify MYCITE System.

Components	Description
ABILIFY MYCITE tablet (DMS tablet)	Aripiprazole tablet embedded with an ingestible event marker sensor
MYCITE System patch (DMS patch)	Wearable sensor patch that detects the signal from the ingestible event marker sensor after ingestion and transmits data to a smartphone
MYCITE System app (DMS app)	Smartphone app used to display medication ingestion information for the patient
MYCITE System dashboard (DMS dashboard)	Two separate web-based portals, one for health care providers and one for family and friends who care for the patient

Figure 1. Schematic representation of digital medicine intervention. FDA: Food and Drug Administration; IEM: ingestible event marker. Image used with permission from Otsuka Development and Commercialization Inc.



Purpose

Minimal research exists on the real-world comparison between a DMS and treatment as usual (TAU) with regard to medication adherence for individuals with SMI. Therefore, before testing a DMS in a large randomized controlled trial, we wanted to examine the feasibility and acceptability of a DMS among patients with SMI. As the Veterans Health Administration (VHA) is the largest health care provider for individuals with SMI in the United States [7,8], it provides a conducive setting for a clinical trial comparing DMS and TAU. Thus, the primary purpose of this study was to evaluate the impact of a DMS among Veterans (patients) with SMI as compared with TAU on medication adherence. The secondary purpose was to obtain patient and provider perspectives on the feasibility and acceptability of using a specific DMS, the ABILIFY MYCITE System. Notably, we use the abbreviation DMS for the remainder of the manuscript to represent the overall ABILIFY MYCITE System; however, we refer to specific components of the DMS (eg, tablet, patch, app, or dashboard) as needed.

Methods

Study Design

This was an open-label, 2-site, provider-randomized, prospective, 2-arm (DMS vs TAU) clinical trial. Patients assigned to the intervention group by provider randomization were enrolled in the DMS for a period of 90 days with the option to continue use for up to 9 additional months. Participants assigned to the control arm (TAU) continued to receive care as recommended by their mental health provider, which included their continued use of aripiprazole. Study duration for both groups was up to 12 months or study closeout, whichever came first. At the study conclusion, we used a descriptive qualitative analysis design and rapid qualitative analysis procedures to examine the feasibility and acceptability of using the DMS. All study procedures were approved by both sites' respective institutional review boards and research and development committees. The trial is registered at ClinicalTrials.gov (NCT03881449), and the sponsor requested for the study to be concluded prematurely.

Setting and Participants

Study participants were recruited from 2 VHA Medical Centers in Durham, North Carolina, and Houston, Texas. Eligible patients were aged ≥ 18 years and met the Diagnostic and

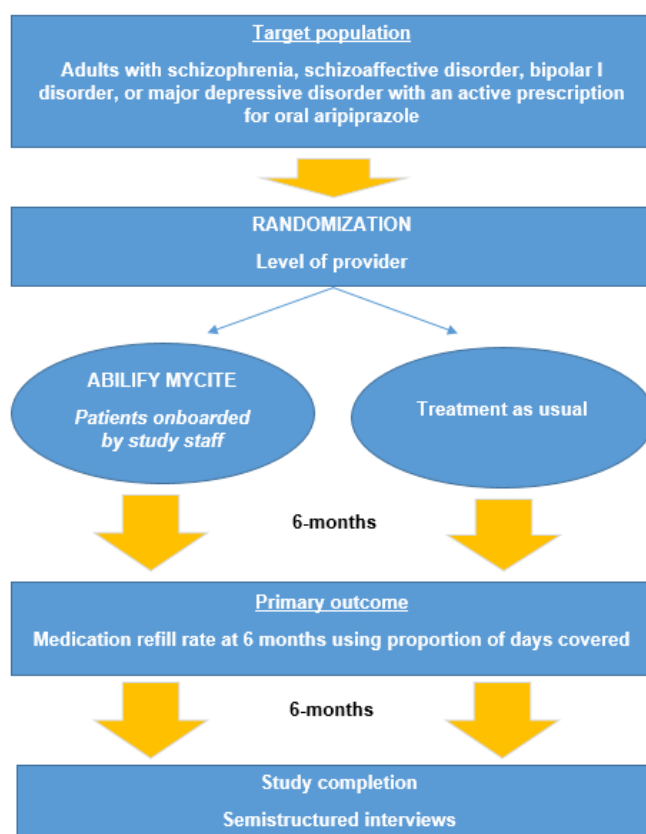
Statistical Manual of Mental Disorders, Fifth Edition, criteria for schizophrenia, schizoaffective disorder, bipolar I disorder, or major depressive disorder. Additional study eligibility included (1) an active prescription for oral aripiprazole and (2) approval to participate in the study from their mental health provider. Exclusion criteria included (1) a current neurocognitive disorder that would affect the patient's ability to complete the trial (eg, dementia); (2) the patient's mental health provider determining that the patient was not fit to participate; (3) the patient being currently enrolled in an investigational drug trial, a medication management study, or program or participation in an investigational drug trial 30 days before trial enrollment; (4) the patient being pregnant, planning on becoming pregnant during the trial, or breastfeeding; (5) the patient failing an initial cognitive screener; (6) the patient having a known allergy to adhesive tape or any pertinent components of the DMS patch; (7) the patient not having skin on the anterior chest just above the lower edge of the rib cage, having dermatologic conditions, such as dermatitis or open wounds, in the location where the patch would be placed, or unwilling to refrain from the use of topical products on the skin patch sites; and (8) the patient

having <20% proportion of days covered (PDC) with aripiprazole in the 6 months before enrollment.

Screening and Recruitment

Potential participants were identified initially by a data pull from the electronic medical record. From this data pull, only patients of those providers who agreed to be involved in the study were screened further. Qualifying patients were sent an introductory letter in mail, describing the study and inviting them to contact the study team for more information and further eligibility screening by telephone. The study used an opt-out recruitment strategy that entailed contacting participants approximately 7 to 10 business days after a recruitment letter was sent, unless participants contacted the study team to indicate that they were not interested in participating. Once patients were confirmed eligible and were interested in participating, they were seen at a scheduled in-person baseline appointment. At the baseline study visit, written informed consent was obtained, smartphone compatibility was verified or a study-owned smartphone was provided if needed, and baseline assessments and surveys were completed (Figure 2).

Figure 2. Schematic of enrollment procedures.



Randomization

This was a provider-randomized clinical trial. Providers who agreed to have their patients approached and potentially enrolled were randomized in a 1:1 ratio using stratified block randomization. Provider randomization was stratified by site.

Treatment Arms

DMS Arm

Once participants were enrolled in the trial, they were onboarded by a trained study team member. The onboarding process included obtaining DMS tablets from the site's pharmacy, successfully placing the DMS patch on the skin at the proper location, and pairing the patch with the DMS app. The

participant was then provided with additional training materials and contact information for the study team and the DMS product's call center. Participants were encouraged to reach out to the company's call center (DMS support) for technical assistance related to the DMS.

As part of the 12-month trial, participants in the DMS arm used the DMS tablet for 90 days. Participants then decided whether to continue beyond the initial 90 days. If participants discontinued the DMS tablet, they restarted their oral aripiprazole as prescribed by their provider. We continued to follow these individuals and obtained pharmacy refill data. Follow-up visits for both the DMS and TAU arms occurred at 3, 6, and 12 months. Early termination visits were completed for DMS arm participants only, per the study sponsor.

TAU Arm

Participants randomized to the TAU arm continued to receive usual care as provided by their mental health provider. Participants assigned to the TAU arm completed all required study visits and data collection surveys.

Measures

Quantitative Measures

All study measures were collected by trained study staff at both sites.

Demographic and Clinical Data

A research assistant collected demographic data (eg, race, sex, age, and comorbidities); clinical data (eg, clinical diagnoses and medications); and data regarding the use of mobile health devices. Diagnoses were obtained from medical records and the following International Classification of Diseases codes were used: F33.0 (major depressive disorder), F31.0 (bipolar I disorder), F25.0 (schizoaffective disorder), and F20.0 (schizophrenia).

Medication Refill Adherence

The primary outcome was medication refill based upon the number of days covered from baseline to 6 months. Using 2 approaches, we measured adherence using PDC [9], a leading method used to calculate medication adherence at a population level. The first set of PDCs were calculated for 3 and 6 months independent of whether an individual was recommended to stop using the DMS (intention to treat). The second PDC measure was calculated as the number of days covered until there was documentation that a patient was recommended to stop using DMS or until the patient reported a problem with the intervention (eg, skin irritation) for the DMS group.

Qualitative Measures

Guided by rapid qualitative analysis procedures [10], we completed semistructured interviews to examine the feasibility and acceptability of the DMS to support medication adherence. We used a convenience sampling plan to identify participants enrolled in the DMS arm up to their 12-month participation in the study or the end of study activities, whichever came first. Providers whose patients were enrolled in the DMS arm were invited to complete an interview up to the date of the end of study activities. Research assistants contacted intervention

patients currently enrolled in the study and invited participants and providers to complete interviews.

Interview questions inquired about the feasibility of and facilitators of and barriers to the DMS. Questions for the patients focused on medication experience (pre-enrollment), onboarding to DMS, system usability, satisfaction with support, and feedback. Questions for the providers inquired about the prescriber experience (prestudy), DMS dashboard account setup, system usability, satisfaction with support, and feedback. We used probes (eg, "Please describe your experience in greater detail" and "What do you mean?") to obtain greater detail and clarify responses. Interviews were completed by a trained research assistant and included a notetaker who recorded responses via a structured note form. After the interview, the research assistant and notetaker debriefed and reviewed interview responses in the context of other interviews. Interviews were conducted via the telephone and were recorded but not transcribed. Patient and provider interviews lasted for 43 (SD 12) minutes on average.

Analytic Strategy

Quantitative Measures

Oral aripiprazole refill was measured by the number of days covered from baseline to 3 and 6 months using PDC in both intention-to-treat and DMS-engaged analyses. Owing to the small number of participants, we conducted descriptive analyses rather than a model-based approach. The intention-to-treat analyses used all data from baseline to 3 or 6 months, depending on the outcome, while the DMS-engaged analysis censored participants in the DMS arm at system discontinuation.

Qualitative Measures

We followed rapid analysis procedures for data analysis and Microsoft Excel (version 2002) to support coding and analysis. Two authors (AAL and SG) reviewed all notes and debrief notes taken during interviews with patients and providers. These authors used thematic analysis [11] and the matrix method [12] to analyze and identify salient themes across all interviews. We established rigor and validity by independently coding and summarizing all data, discussing emerging codes and thematic groupings during meetings, and reviewing findings with the larger study team.

Ethics Approval

This study was approved by the Durham Veterans Affairs Medical Center Institutional Review Board (ID number 02188) on January 19, 2019.

Results

Sample

A total of 26 providers were randomized for this trial (Durham: 22/26, 85%; Houston: 4/26, 15%). Of the eligible patients from participating providers, a total of 46 patients consented and enrolled in the trial (Durham: 28/46, 61%; Houston: 18/46, 39%), with 21 (46%) participating in the DMS arm and 25 (54%) participating in the TAU arm (Figure 3). We issued 6 study-owned smartphones to patients to facilitate enrollment in

the study. The sample was on average aged 53 (SD 13.3) years and mostly male (33/46, 72%), 52% (24/46) self-reported as Black, and 15% (7/46) had a high school education or less. The clinical diagnoses breakdown for enrolled patients included 54% (25/46) with major depressive disorder, 7% (3/46) with schizophrenia, 11% (5/46) with schizoaffective disorder, and

28% (13/46) with bipolar I disorder. Before study enrollment, half (23/46, 50%) of the enrolled patients had downloaded a health app onto their mobile phone, and approximately one-third (14/46, 30%) of the participants had used a wearable tech device such as a fitness tracker or smartwatch (Table 2).

Figure 3. Enrollment. N/A: not applicable.

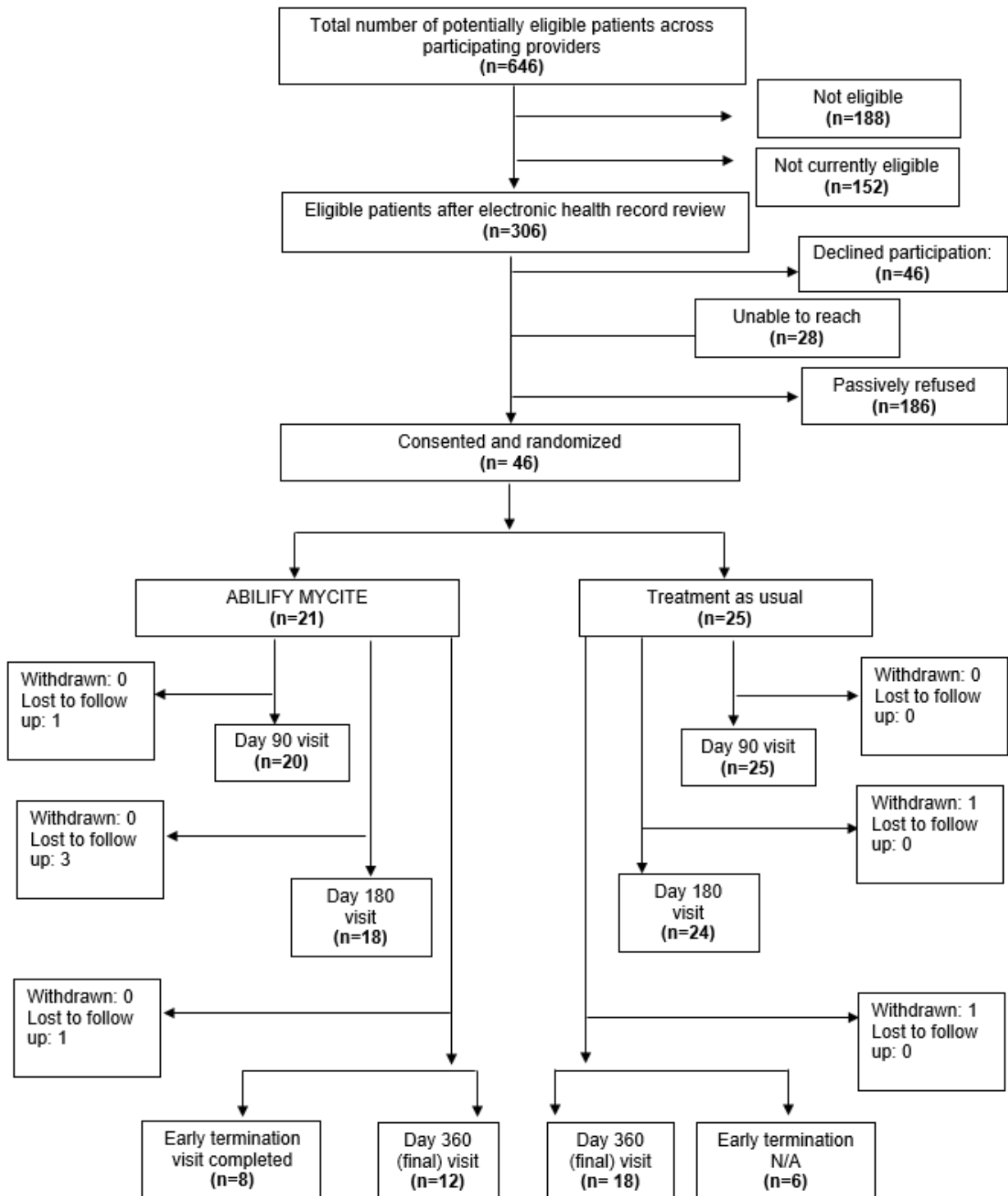


Table 2. Participant demographics (N=46).

	DMS ^a (n=21)	TAU ^b (n=25)
Age (years), mean (SD)	54.67 (12.73)	51.64 (13.40)
Sex, n (%)		
Male	17 (81)	16 (64)
Female	4 (19)	9 (36)
Intersex	0 (0)	0 (0)
Race, n (%)		
Black	12 (57)	12 (48)
White	8 (38)	13 (52)
Asian	1 (5)	0 (0)
Education, n (%)		
High school or less	5 (24)	2 (8)
Any college	14 (67)	21 (84)
Graduate school	2 (10)	2 (8)
Clinical diagnosis, n (%)		
Major depressive disorder	14 (67)	11 (44)
Bipolar I disorder	3 (14)	10 (40)
Schizoaffective disorder	2 (10)	3 (12)
Schizophrenia	2 (10)	1 (4)
Prestudy mobile device use, n (%)		
Health app	9 (43)	14 (56)
Wearable technology	8 (38)	6 (24)

^aDMS: digital medicine system.

^bTAU: treatment as usual.

Quantitative Outcomes

In the intention-to-treat analyses, PDC over 3 and 6 months was 0.82 (SD 0.24) and 0.75 (SD 0.26) in the DMS arm and 0.86 (SD 0.19) and 0.82 (SD 0.21) in the TAU arm, respectively. In the DMS-engaged analysis, the DMS arm had 0.85 (SD 0.20) PDC over the period (Table 3).

Patients in the DMS arm stayed engaged with the DMS for 144 (SD 100; median 147; range 0-376) days on average. Among the participants in the DMS arm, 5 (24%) stopped using the system by 3 months because of skin irritation adverse events from the use of the DMS patch.

Table 3. Proportion of days covered (PDC) for the drug of interest.

PDC	DMS ^a (n=21), mean (SD)	TAU ^b (n=25), mean (SD)
3 months	0.82 (0.24)	0.86 (0.19)
6 months	0.75 (0.26)	0.82 (0.21)
DMS engaged	0.85 (0.20)	N/A ^c

^aDMS: digital medicine system.

^bTAU: treatment as usual.

^cN/A: not applicable.

Qualitative Outcomes

Overview

Qualitative interviews were completed with 14 patients and 4 providers, with half (7/14, 50% and 2/4, 50%, respectively) of each group from each of the 2 sites. Patient respondents were

mostly male 86% (12/14), while 64% (9/14) were Black, 29% (4/14) were White, and 7% (1/14) were Asian. Overall, 86% (12/14) of patient respondents used the DMS for ≥ 90 days; 43% (6/14) used the DMS for ≥ 180 days. Of the providers, 50% (2/4) identified as male, 75% (3/4) were providers, and 25% (1/4) were prescribing pharmacist.

We identified 5 themes when analyzing the patient data and 5 themes for the providers' data that described the feasibility and acceptability of using the DMS. Patient themes included pre-enrollment adherence strategies and interest in the DMS, positive impact on medication adherence, system usability challenges, support needs, and suggested improvements to system design and functionality. Provider themes included prestudy concerns for patient medication adherence and interest in the DMS, concerns with the DMS (prestudy), DMS dashboard usability issues and support, challenges in impact of the DMS, and suggestions to increase provider use.

Semistructured Interviews—Patient Responses

Pre-enrollment Strategies and Interest in the DMS

Patients described the following pre-enrollment medication adherence strategies: maintaining a routine and setting up reminders through environmental cues (eg, seeing pills on their counter, a reminder from their spouse, or from reminders on their phone or calendar). Regarding their pre-enrollment medication adherence strategy, a patient stated that the “hardest part for me is sometimes I can't remember if I took my meds that day, so that's why I try to take them first thing in the morning when I wake up...” Patients were interested in using the DMS because they felt that the DMS would help them with their medication adherence. A patient shared the following statement:

...I have issues with taking medication on time, not remembering to take it, skipping doses, so the way that the [DMS] would remind you...that would be helpful to someone.

Many patients expressed interest in the technology used in the DMS, including a patient who stated the following:

I was interested because of the new technology...having the [DMS] app so you can see what your daily activities are.

Positive Impact on Medication Adherence

Patients reported the DMS made them more mindful of taking their medication. For example, a patient shared the following statement:

[The DMS] made me more cognizant of what time my dosages were and definitely to make sure that I took [my medications] daily...when I opened my phone, I would see the DMS app and remember, oh, I have to take my medicine in the morning.

The visual reminder of seeing their DMS patch and the DMS app on their phone acted as a prompt to take their medication. In addition, the DMS app's medication notifications to log missed doses into the DMS app reinforced the habit of taking their medication consistently. A patient stated the following:

The [DMS app] itself, it asks you a series of questions why you forgot to take it and so when you're going through that it just feels like it's telling you to not forget it again. I definitely thought it was helpful...it reinforced that good habit of remembering to take your medication every day.

System Usability Challenges

Many patients shared that they experienced problems with DMS patch connectivity and skin irritation. A patient described the challenges they experienced as follows:

If the [DMS] patch wasn't paired and I wasn't paying attention to the [DMS] app, I would take my medication and it wouldn't register and that was very frustrating...I [also] had problems with the DMS patch adhering and when it wouldn't adhere, it would unpair [with the DMS app]

A patient who experienced skin irritation from the DMS patch stated the following:

It seemed the more that I changed the DMS patch, the more irritated my skin got with the DMS patch...no matter where I put it.

Patients also expressed that the DMS patch was uncomfortable to wear, as the DMS patch caught on things when working in tight spaces and fell off when the patient sweated because of weather or physical activity.

Patients experienced challenges with the DMS app during DMS tablet registration (eg, DMS app would freeze and would need to be rebooted, technical assistance needed, or unable to log missed doses). A patient stated the following:

I had a lot of issues with the [DMS] app also, the [DMS] app was freezing up on me and I would have to call [DMS support] and we would have to walk through it, we'd have to uninstall it and reinstall it to get the [DMS] app to not be frozen anymore so that it would download the [DMS] patch and pair the [DMS] patch.

Overall, despite assistance from study staff, patients felt the DMS was complicated to learn because of the numerous steps needed to complete each process, and many patients shared that they discontinued use because of recurring challenges.

Support Needs

Patients expressed that their onboarding to the DMS was helpful because of the in-person, one-on-one training with their local study staff members. Patients who called DMS support felt that DMS support was professional and knowledgeable in describing step-by-step solutions. However, patients reported that DMS support was not always able to resolve the patient's issues and provide the patient with a long-term fix for their challenges with using the DMS. A patient described this as follows:

I continued to have the same issue and that's why I discontinued using the DMS...[DMS support]'s recommendation to change the patch and re-pair it wasn't a long-term fix.

Patients preferred in-person support for resolving issues with the DMS because study staff members could see their smartphone and DMS app in real time. When asked about support for the DMS, a patient stated the following:

Since the study team was local, I was more inclined to call them if I needed help with something...just being there and having them explain to me

face-to-face and answer my questions right away, it was helpful...and also that they were able to see my DMS app and see what was going on.

Suggested Improvements to System Design and Functionality

Patient feedback for the DMS included two main suggestions: (1) developing a smaller DMS patch or an alternative way to track ingestion that did not involve wearing a DMS patch and (2) improving the usability and functionality of the DMS app as well as the reliability and accuracy of the DMS as a whole.

Regarding patch improvement suggestions, a female patient shared the following:

If I could have placed the DMS patch somewhere else on my body, I might have continued...if it could be worn somewhere discreetly on the body that would be ideal...where it had to be at the top of my torso, that was problematic for me...I don't think it would be problematic for a man or someone with a smaller chest.

Several patients described how the DMS patch was cumbersome to wear, wearing the DMS patch for long periods led to skin irritation, or that they experienced issues with the DMS patch sticking to their skin. Another patient shared the following:

I did not like the DMS patch...it gave me a rash...if we could figure out some other way of doing it without the DMS patch involved that would be wonderful.

A patient who experienced DMS patch adherence issues stated the following:

I think having to have it on your torso is a problem...how active I am, it just couldn't stay on me. Maybe if it was moved to an extremity...even if it was a fitness tracker, that would have made that part of the whole process much easier.

Regarding suggested improvements to the DMS app usability and functionality, a patient said the following:

When you log into the DMS app, it logs you off rather quickly...the login information wasn't saved, so it was time consuming to access the DMS app.

In addition, another patient suggested the following:

If your old DMS patch doesn't upload information [to the DMS app]...having a mechanism to bypass that and record it and send it to DMS support, so you can continue going through the pairing process with the new DMS patch [would be helpful]

Finally, some patients suggested that DMS support be more proactive (eg, calling patients periodically to check in rather than patients needing to call DMS support to obtain technical assistance). A patient stated the following:

For people like us that's not as savvy with technology as others are...if [DMS support] have called me and checked, maybe [using the DMS] would have been easier and simpler.

In addition, patients described an interest in receiving further training and material on strategies for addressing potential

challenges to the DMS, including technical issues with the DMS app, the DMS, DMS patch connectivity, and DMS patch contact. A patient expressed that they would have liked to receive written material, as this patient wanted to refer back to information during the study.

Semistructured Interviews—Provider Responses

Prestudy Concerns for Patient Medication Adherence and Interest in the DMS

Providers shared a common concern for patient compliance with medication and acknowledged several challenges to medication adherence (eg, disease specific, health care system related, and side effects of medications). A provider explained, "It's hard to know, by patient report, how consistently they've been taking the medication..." A provider described challenges patients have in refilling medications as follows:

It's very daunting to get the refills and if the refills are done, they often times...get them late or they forget to order it...it's not like they don't want to intentionally not take their medicine.

Overall, providers were interested in the DMS, as this system could provide an objective measure of adherence instead of self-report. A provider verbalized as follows:

I'm always looking for ways to help my patients...take ownership of their own care and to improve quality of care. Those are the two things that drew me to the DMS.

Providers also shared that a common issue can be reconciling what a patient says compared with medication data in the electronic health record. A provider said the following:

I was drawn to something like technology that would help both on the provider's end and the patient's end to overcome that kind of barrier.

Concerns with the DMS (Prestudy)

Providers expressed concerns with the implementation of DMS in their patient populations, specifically those who experience paranoia and would be apprehensive of using the DMS to track medication adherence. A provider said the following:

My initial reaction was paranoid, psychotic patients might be a bit concerned that they're taking a pill that has a sensor...I was thinking that they might not be willing to take such a medication where one would know [provider or family member] whether they are taking it or not.

An additional concern was around the use of the DMS patch because of the provider's experience with skin sensitivity to adhesives.

DMS Dashboard Usability

Providers experienced frustration with the multistep process to log into the DMS dashboard via the notification email. This frustration led them to not check notification emails or log into the DMS dashboard. Improvement suggestions included streamlining the log-in process for the DMS dashboard (eg, the ability to save password and not needing to re-enter it each time)

and including additional details (eg, missed dose or multiple doses taken at one time) in the notification email to encourage them to log into the DMS dashboard more frequently. Notably, none of the providers interviewed called DMS support for assistance with the DMS dashboard, rather these providers called the study staff because of their accessibility.

Challenges in the Impact of the DMS

Providers stated that the objective medication adherence data for each patient in the DMS was helpful. However, owing to numerous usability challenges (eg, data inaccuracy and multistep log-in process), these providers did not use the DMS dashboard, rather they relied on the study team for information on their patient's medication data.

Suggestions to Increase Provider Use

Provider feedback for DMS centered on (1) improving accessibility to the DMS dashboard and (2) change in DMS data management to improve workflow. Providers suggested a streamlined process for receiving notifications such as including additional information in the notification email and embedding the DMS log-in into the electronic health record (eg, an embedded link taking them to the DMS dashboard) to make it easier to integrate into their regular workflow. A provider stated the following:

If there was some way we could incorporate [ingestion data] in our templated notes...while we are writing, documenting the notes, we could click the link to the DMS dashboard...that might work.

Finally, many providers recommended that the DMS dashboard and data should instead be managed by an individual in the clinic (eg, nurse or clinical coordinator), as this individual can summarize patient data for the provider. A provider said the following:

...If the DMS could be created in a way that there will be a go-between, an intermediary between the provider and the patient on the DMS, who would be monitoring, more closely, the ingestion data...[and summarize] this is number of days of adherence, number of days of not adhering...so that the prescriber has that information right in front of them, even before seeing the patient.

Discussion

Principal Findings

This open-label, 2-site, 12-month, provider-randomized trial assessed aripiprazole refill adherence in patients with SMI and examined patient and provider perspectives on the feasibility and acceptability of DMS for this population in the VHA health care system. Our study showed that there was no notable difference in the refill rates between the DMS and TAU arms in intention-to-treat analyses. Among the users of the DMS, the PDC by refill rates were 0.85 (SD 0.20) for those patients when engaged in the DMS. Qualitative findings indicated several challenges to the feasibility and acceptability of the DMS for patients and providers. These challenges included technical issues and contact issues with the DMS patch, including skin

irritation and adherence on the skin, and affected the length of user participation as well as overall confidence and interest in using the DMS as a reliable and accurate method of tracking medication ingestion. Both patients and providers discussed recommendations to improve the patient-provider experience and overall satisfaction with the DMS. Notably, patients and providers commented on how to increase confidence in using the DMS for patients with SMI, particularly with patients who may not have extensive experience with smartphones, Bluetooth technology, and health app use.

Our results indicated extended patient use of the DMS to manage their SMI. However, our results should be interpreted with caution, as the DMS did not appear to outperform TAU in this study, and enrollment was only a fraction of our original goal. Patients were asked to use the DMS for 90 days, at which time they, along with their provider, could decide whether to continue using the DMS beyond 90 days and up to the full 12 months of study participation. Compared with previous studies using a DMS with an ingestible sensor pill for a period of 8 [13] or 12 [14] weeks, the use of the DMS lasted up to 12 months, which provided insight into DMS use that was 90 days or longer (12/21, 57%), with 29% (6/21) of respondents opting to use the DMS for 180 days or more.

In another trial using the same DMS [13] as in our study, adherence was measured by the proportion of days with good DMS patch coverage (ingestible event marker registration). However, in our study, medication adherence was assessed by the PDC on aripiprazole using pharmacy record data. Owing to the prevalence of patient-reported technical difficulties with DMS tablet registration for the DMS and provider concern for data accuracy that was described during the qualitative interviews, PDC was a preferred measure of medication adherence in this study. Compared with a medication adherence rate of $\geq 80\%$ reported over an 8-week period (when good DMS patch coverage was reported) in the previous study [13], our study saw 82% adherence in our intention-to-treat analysis over a 3-month (approximately 12 weeks) period, which was comparable. However, this rate changed to 75% at 6 months in our study, which reveals an area of further study into the potential cause of lower adherence rates after 90 days.

Our study's qualitative findings highlighted concerns regarding DMS patch use and contact issues (skin irritation and adherence to skin). In a recent study that evaluated patient responses to using a comparable digital medicine program with an ingestible sensor coencapsulated with antiretroviral therapy medication, patients reported similar issues as seen in our qualitative findings, including issues with patch adherence to skin and overall frustration with using the digital medicine program patch [15]. As detailed earlier and in our findings, DMS patch issues affected the length of user participation; consideration should be given to making improvements to the usability of the DMS patch, particularly long-term use, to improve the patient experience and adherence to the DMS.

Limitations

Several considerations should be acknowledged. First, enrollment and data collection for the study ended early per the study sponsor's request, presumably before more null or

potentially negative findings could emerge. Second, technology issues may have affected the continued use of the DMS and engagement in the intervention. Third, a limitation for consideration with regard to adherence is the potential of the Hawthorne effect related to the observation of both participant groups (DMS and TAU) in a trial. This may have affected adherence rates as measured by PDC.

In addition, while qualitative interviews were conducted with 14 VHA patients who used the DMS during their study participation, only 4 providers were interviewed. Overall, provider engagement with the DMS dashboard was limited—some providers were unable to share feedback on their experience with the DMS dashboard and the impact that the use of the DMS had on their patients and on patient-provider communication. Providers who consented to participate in the qualitative interviews shared their level of interaction with the DMS dashboard, and interview questions were designed to capture potential barriers to use and to gather feedback on how to increase provider use. Individuals enrolled had a relatively high rate of refill at baseline; future studies using DMS may want to focus on individuals who are having greater challenges with refill adherence. Finally, some patients were interviewed

about their DMS experience several months or sometimes up to 1 year after enrollment, which could have affected how well the patients remembered specific details about their experience.

Strengths

Despite the aforementioned limitations, we collected data on the impact of DMS use in a specific patient population, while also gathering detailed feedback on patient and provider experiences, and suggested modifications and considerations for DMS improvements. Evaluation of this novel DMS in the VHA health care system provided insight into real-world use for increased use in community-based, private practice, and public health care settings.

Conclusions

Our study of a DMS in patients with SMI did not demonstrate a detectable improvement in medication adherence. Our findings indicate critical issues to consider in improving the feasibility and acceptability of a DMS for patients and providers. Specifically, our study highlights the importance of conducting a feasibility and acceptability trial and collecting quantitative and qualitative data to further refine and improve the DMS for long-term adherence.

Acknowledgments

The research reported in this publication was supported by Otsuka Pharmaceutical Development and Commercialization, Inc. The content is solely the responsibility of the authors and does not reflect the position or policy of Duke University, the US Department of Veterans Affairs, the United States government, or Otsuka Pharmaceutical Development and Commercialization, Inc.

Conflicts of Interest

HBB reports receiving research funds from Sanofi, Otsuka, Improved Patient Outcomes, Novo Nordisk, PhRMA Foundation, and Boehringer Ingelheim as well as consulting funds from Sanofi, Otsuka, Abbott, and Novartis. AAL reports receiving funds from PhRMA Foundation and Otsuka. VAS and S Moore report receiving funds from Otsuka. JCRF is an employee of Otsuka.

References

1. Warriach ZI, Sanchez-Gonzalez MA, Ferrer GF. Suicidal behavior and medication adherence in schizophrenic patients. *Cureus* 2021 Jan 04;13(1):e12473 [FREE Full text] [doi: [10.7759/cureus.12473](https://doi.org/10.7759/cureus.12473)] [Medline: [33552789](https://pubmed.ncbi.nlm.nih.gov/33552789/)]
2. Krivoy A, Balicer RD, Feldman B, Hoshen M, Zalsman G, Weizman A, et al. Adherence to antidepressants is associated with lower mortality: a 4-year population-based cohort study. *J Clin Psychiatry* 2016 May;77(5):e566-e572. [doi: [10.4088/JCP.14m09531](https://doi.org/10.4088/JCP.14m09531)] [Medline: [27136668](https://pubmed.ncbi.nlm.nih.gov/27136668/)]
3. Stirratt MJ, Dunbar-Jacob J, Crane HM, Simoni JM, Czajkowski S, Hilliard ME, et al. Self-report measures of medication adherence behavior: recommendations on optimal use. *Transl Behav Med* 2015 Dec;5(4):470-482 [FREE Full text] [doi: [10.1007/s13142-015-0315-2](https://doi.org/10.1007/s13142-015-0315-2)] [Medline: [26622919](https://pubmed.ncbi.nlm.nih.gov/26622919/)]
4. Freudenreich O. What is the role of digital medicine for adherence monitoring in patients with serious mental illness? *J Clin Psychiatry* 2022 Apr 11;83(3):22com14443 [FREE Full text] [doi: [10.4088/JCP.22com14443](https://doi.org/10.4088/JCP.22com14443)] [Medline: [35421286](https://pubmed.ncbi.nlm.nih.gov/35421286/)]
5. Sangeda RZ, Mosha F, Prospero M, Aboud S, Vercauteren J, Camacho RJ, et al. Pharmacy refill adherence outperforms self-reported methods in predicting HIV therapy outcome in resource-limited settings. *BMC Public Health* 2014 Oct 04;14:1035 [FREE Full text] [doi: [10.1186/1471-2458-14-1035](https://doi.org/10.1186/1471-2458-14-1035)] [Medline: [25280535](https://pubmed.ncbi.nlm.nih.gov/25280535/)]
6. Batra S, Baker RA, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. *Med Devices (Auckl)* 2017 Oct 4;10:237-251 [FREE Full text] [doi: [10.2147/MDER.S144158](https://doi.org/10.2147/MDER.S144158)] [Medline: [29042823](https://pubmed.ncbi.nlm.nih.gov/29042823/)]
7. Deckro J, Phillips T, Davis A, Hehr AT, Ochylski S. Big data in the veterans health administration: a nursing informatics perspective. *J Nurs Scholarsh* 2021 May;53(3):288-295. [doi: [10.1111/jnu.12631](https://doi.org/10.1111/jnu.12631)] [Medline: [33689232](https://pubmed.ncbi.nlm.nih.gov/33689232/)]
8. Trivedi RB, Post EP, Sun H, Pomerantz A, Saxon AJ, Piette JD, et al. Prevalence, comorbidity, and prognosis of mental health among US veterans. *Am J Public Health* 2015 Dec;105(12):2564-2569. [doi: [10.2105/AJPH.2015.302836](https://doi.org/10.2105/AJPH.2015.302836)] [Medline: [26474009](https://pubmed.ncbi.nlm.nih.gov/26474009/)]

9. PQA Adherence Measures. Pharmacy Quality Alliance. 2021. URL: <https://www.pqaalliance.org/adherence-measures> [accessed 2021-10-25]
10. Lewinski AA, Crowley MJ, Miller C, Bosworth HB, Jackson GL, Steinhauer K, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. *Med Care* 2021 Jun 01;59(Suppl 3):S242-S251 [FREE Full text] [doi: [10.1097/MLR.0000000000001553](https://doi.org/10.1097/MLR.0000000000001553)] [Medline: [33976073](https://pubmed.ncbi.nlm.nih.gov/33976073/)]
11. Guest G, MacQueen KM, Namey EE. *Applied Thematic Analysis*. Thousand Oaks, CA, USA: Sage Publications; 2012.
12. Averill JB. Matrix analysis as a complementary analytic strategy in qualitative inquiry. *Qual Health Res* 2002 Jul;12(6):855-866. [doi: [10.1177/104973230201200611](https://doi.org/10.1177/104973230201200611)] [Medline: [12109729](https://pubmed.ncbi.nlm.nih.gov/12109729/)]
13. Fowler JC, Cope N, Knights J, Fang H, Skubiak T, Shergill SS, et al. Hummingbird study: results from an exploratory trial assessing the performance and acceptance of a digital medicine system in adults with schizophrenia, schizoaffective disorder, or first-episode psychosis. *Neuropsychiatr Dis Treat* 2021 Feb 12;17:483-492 [FREE Full text] [doi: [10.2147/NDT.S290793](https://doi.org/10.2147/NDT.S290793)] [Medline: [33603385](https://pubmed.ncbi.nlm.nih.gov/33603385/)]
14. Frias J, Virdi N, Raja P, Kim Y, Savage G, Osterberg L. Effectiveness of digital medicines to improve clinical outcomes in patients with uncontrolled hypertension and type 2 diabetes: prospective, open-label, cluster-randomized pilot clinical trial. *J Med Internet Res* 2017 Jul 11;19(7):e246 [FREE Full text] [doi: [10.2196/jmir.7833](https://doi.org/10.2196/jmir.7833)] [Medline: [28698169](https://pubmed.ncbi.nlm.nih.gov/28698169/)]
15. Kamal S, Rosen MI, Lazar C, Siqueiros L, Wang Y, Daar ES, et al. Perceptions of people living with HIV and HIV healthcare providers on real-time measuring and monitoring of antiretroviral adherence using ingestible sensors: a qualitative study. *AIDS Res Treat* 2020 May 27;2020:1098109 [FREE Full text] [doi: [10.1155/2020/1098109](https://doi.org/10.1155/2020/1098109)] [Medline: [32566285](https://pubmed.ncbi.nlm.nih.gov/32566285/)]

Abbreviations

- DMS:** digital medicine system
PDC: proportion of days covered
SMI: serious mental illness
TAU: treatment as usual
VHA: Veterans Health Administration

Edited by A Mavragani; submitted 09.12.21; peer-reviewed by C Miller, A Kazdin; comments to author 11.04.22; revised version received 04.08.22; accepted 24.08.22; published 22.12.22.

Please cite as:

Gonzales S, Okusaga OO, Reuteman-Fowler JC, Oakes MM, Brown JN, Moore S, Lewinski AA, Rodriguez C, Moncayo N, Smith VA, Malone S, List J, Cho RY, Jeffreys AS, Bosworth HB

Digital Medicine System in Veterans With Severe Mental Illness: Feasibility and Acceptability Study

JMIR Form Res 2022;6(12):e34893

URL: <https://formative.jmir.org/2022/12/e34893>

doi: [10.2196/34893](https://doi.org/10.2196/34893)

PMID: [36548028](https://pubmed.ncbi.nlm.nih.gov/36548028/)

©Sarah Gonzales, Olaoluwa O Okusaga, J Corey Reuteman-Fowler, Megan M Oakes, Jamie N Brown, Scott Moore, Allison A Lewinski, Cristin Rodriguez, Norma Moncayo, Valerie A Smith, Shauna Malone, Justine List, Raymond Y Cho, Amy S Jeffreys, Hayden B Bosworth. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 22.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Evaluating User Preferences, Comprehension, and Trust in Apps for Environmental Health Hazards: Qualitative Case Study

Annabelle Workman¹, PhD; Fay H Johnston^{1,2}, PhD; Sharon L Campbell^{1,2}, PhD; Grant J Williamson³, PhD; Chris Lucani³, PhD; David M J S Bowman³, DSc; Nick Cooling⁴, MEd; Penelope J Jones¹, PhD

¹Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia

²Public Health Services, Tasmanian Department of Health, Hobart, Australia

³School of Natural Sciences, University of Tasmania, Hobart, Australia

⁴School of Medicine, University of Tasmania, Hobart, Australia

Corresponding Author:

Penelope J Jones, PhD

Menzies Institute for Medical Research, University of Tasmania

17 Liverpool Street

Hobart, 7000

Australia

Phone: 61 362267726

Email: penelope.jones@utas.edu.au

Abstract

Background: Climate change is projected to increase environmental health hazard risks through fire-related air pollution and increased airborne pollen levels. To protect vulnerable populations, it is imperative that evidence-based and accessible interventions are available. The environmental health app, AirRater, was developed in 2015 in Australia to provide information on multiple atmospheric health hazards in near real time. The app allows users to view local environmental conditions, and input and track their personal symptoms to enable behaviors that protect health in response to environmental hazards.

Objective: This study aimed to develop insights into users' perceptions of engagement, comprehension, and trust in AirRater to inform the future development of environmental health apps. Specifically, this study explored which AirRater features users engaged with, what additional features or functionality needs users felt they required, users' self-perception of understanding app information, and their level of trust in the information provided.

Methods: A total of 42 adult AirRater users were recruited from 3 locations in Australia to participate in semistructured interviews to capture location- or context-specific experiences. Participants were notified of the recruitment opportunity through multiple avenues including newsletter articles and social media. Informed consent was obtained before participation, and the participants were remunerated for their time and perspectives. A preinterview questionnaire collected data including age range, any preexisting conditions, and location (postcode). All participant data were deidentified. Interviews were recorded, transcribed, and analyzed using thematic analysis in NVivo 12 (QSR International).

Results: Participants discussed app features and functionality, as well as their understanding of, and trust in, the information provided by the app. Most (26/42, 62%) participants used and valued visual environmental hazard features, especially maps, location settings, and hazard alerts. Most (33/42, 78%) found information in the app easy to understand and support their needs, irrespective of their self-reported literacy levels. Many (21/42, 50%) users reported that they did not question the accuracy of the data presented in the app. Suggested enhancements include the provision of meteorological information (eg, wind speed or direction, air pressure, UV rating, and humidity), functionality enhancements (eg, forecasting, additional alerts, and the inclusion of health advice), and clarification of existing information (eg, symptom triggers), including the capacity to download personal summary data for a specified period.

Conclusions: Participants' perspectives can inform the future development of environmental health apps. Specifically, participants' insights support the identification of key elements for the optimal development of environmental health app design, including streamlining, capacity for users to customize, use of real time data, visual cues, credibility, and accuracy of data. The results also suggest that, in the future, iterative collaboration between developers, environmental agencies, and users will likely promote better functional design, user trust in the data, and ultimately better population health outcomes.

KEYWORDS

health app; evaluation; air pollution; pollen; temperature; mobile phone

Introduction

Background

Climate change is predicted to increase health risks from environmental hazards worldwide. For example, climate projections suggest that many regions will experience an increased risk of air pollution from landscape fires [1], whereas exposure to heat-related health risks is projected to increase [2]. Adverse health outcomes associated with exposure to air pollution include lung and heart diseases, lung cancer, diabetes, neurological conditions, poor pregnancy outcomes, and premature death [3]. A changing climate is also likely to increase the health risks from airborne pollen owing to changes in pollen loads, allergenicity, and pollen season length [4]. The human and economic costs of such increases are likely to be substantial [5], and individuals living with pre-existing chronic conditions or from lower socioeconomic circumstances have been identified as the most vulnerable [6].

In this context, it is imperative to provide evidence-based, cost-effective interventions to reduce the impacts of environmental hazards on human health. Smartphone apps may offer part of the solution, for example, by providing individuals, including vulnerable populations, with easy and timely access to environmental hazard information or by providing tools to support the diagnosis and management of health outcomes triggered by environmental conditions [7-9]. Smartphone apps that collect symptom data can also provide a mechanism to track population-level health impacts of environmental hazards in real time [10]. For example, access to aggregated data on respiratory symptoms such as shortness of breath or heat-related symptoms such as light-headedness, can help public health departments monitor population health outcomes associated with environmental hazards and respond quickly to spikes in outcomes by releasing health alerts and targeted health campaigns [11].

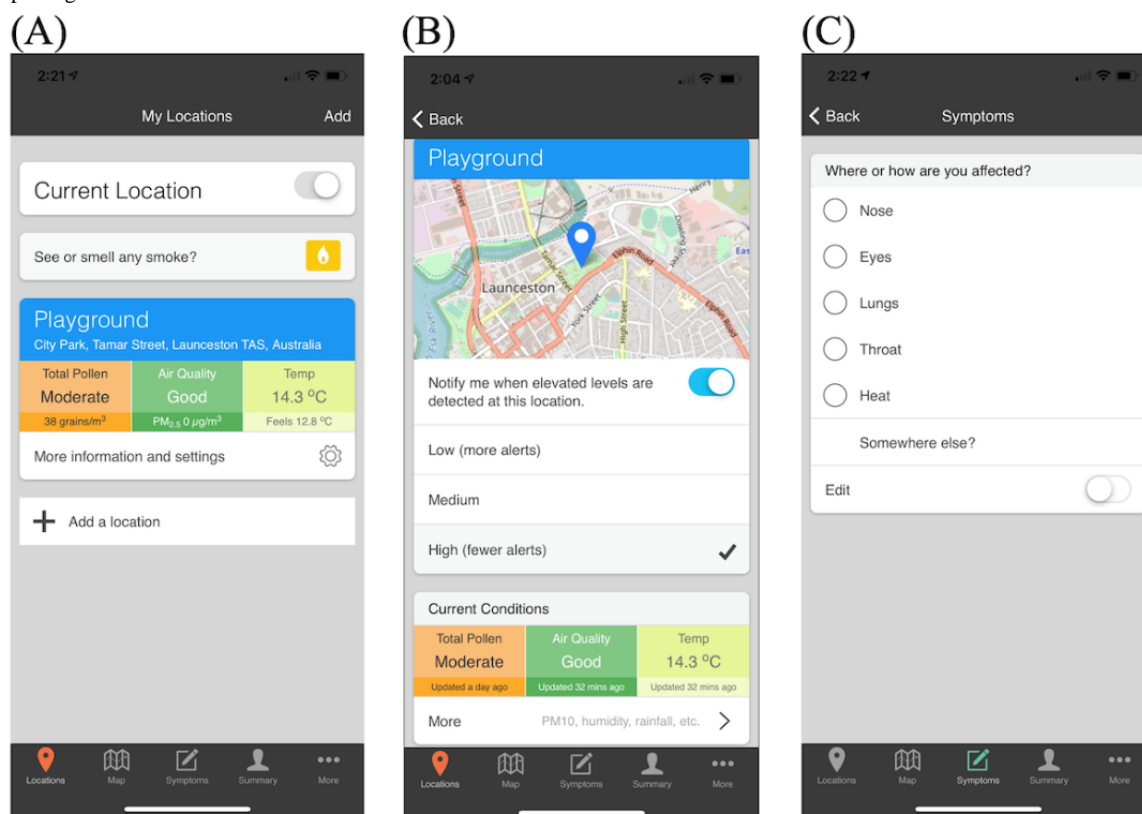
Apps related to environmental health, such as air quality and pollen count apps, are rapidly proliferating and targeting different user groups, including children with asthma [12,13]. Although user perspectives and preferences regarding apps

designed for other health conditions have been studied [14-16], there remains a paucity of evidence regarding how and why such apps do (or do not) support users in understanding and responding appropriately to atmospheric health hazard information. Some studies have investigated the effectiveness of environmental health apps with respect to messaging strategies [17] and behavior change support [18-20]; however, very few have tested user preferences related to app design or the extent to which users comprehend and trust the information presented. Given the plethora of possibilities available with respect to functionality, risk communication, and the design of both the overall app and individual features, it is imperative that we address this research gap to underpin effective and evidence-based environmental health app design in the future.

Protecting Health From Environmental Hazard Risks

The AirRater app offers an ideal opportunity to explore these critical questions. AirRater was developed in Australia in 2015 with the aim of protecting individuals from 3 key atmospheric health hazards: pollen, particulate pollution, and extreme heat [8]. The app was co-designed by a consortium of multidisciplinary researchers and environmental and health government agency representatives and is now freely available across Australia. The functionality of the app has been described in detail elsewhere [8,19]; however, its core features include (1) the provision of near real time information on air quality, temperature, and pollen counts; (2) notifications when atmospheric conditions are poor; and (3) the capacity for individuals to log their symptoms and learn about their personal sensitivities (Figure 1). Users can input and track symptoms related to the nose (eg, itchy), eyes (eg, watery), lungs (eg, wheezy), throat, and heat (eg, light-headed). Users can also input custom symptoms related to other health outcomes and self-report the severity of the symptoms experienced (mild, moderate, and severe). The symptom tracking functionality also enables the app to support public health surveillance and ongoing epidemiological research [21,22]. For a full outline of the app's functionality and features, please see [Multimedia Appendix 1](#). The app is supported by a website that provides users with information about data streams and sources.

Figure 1. Key features and functionality of the AirRater app. Panel (A) shows the home screen where users can see up-to-date information on environmental conditions at their current or saved locations. Panel (B) shows how users can opt into alert notifications. Panel (C) shows part of the symptom reporting interface.



Evaluating User Preferences, Comprehension, and Trust

Several factors make AirRater an ideal case study to explore how users perceive and react to different features of environmental health apps and the extent to which users understand and trust the information presented. First, AirRater represents an unusually holistic environmental health app because it provides information on multiple environmental health hazards (pollen, air pollution, and temperature) in multiple ways (via maps and by saved location) and couples environmental hazard information provision with features designed to help individuals understand and manage their health conditions (eg, symptom tracking, symptom modeling, and personalized alerts). In addition, AirRater has been well-evaluated, and detailed information on who uses the app and their motivations for doing so has been previously published [19,23]. Importantly, these prior evaluations have also demonstrated that the app is successful in supporting a diverse range of users to make decisions and implement behaviors to protect their health, both in the context of severe air pollution episodes and more routine conditions [8,19,23].

This study leverages these characteristics and understandings of AirRater to explore how users perceive and respond to various features of environmental health app design and how such design features do or do not support user comprehension and trust in the information presented. Using qualitative methods, we explore user experiences with AirRater from this perspective, aiming to inform the development of effective environmental health apps. We specifically address 4 research questions.

1. Which features of the AirRater app do users engage with?
2. What are users' additional information and functionality needs?
3. How do users perceive their comprehension of the information provided in the app?
4. What level of trust do users have in the information the app provides?

Methods

Overview

Numerous tools and frameworks exist to assess app integrity and quality, some of which have been validated [24-34]. Such tools are often prescriptive. Qualitative research methods provide the capacity to explore participant experiences in greater detail by enabling researchers to seek elaboration or clarification of initial participant responses. Accordingly, we developed a customized mixed methods evaluation framework based on the guide by the World Health Organization (WHO) to monitor and evaluate a digital health intervention [35] and the mobile app rating scale [34]. The WHO guide provides a practical framework and flexibility to explore specific questions of interest. The mobile app rating scale framework informed the thematic structure of the qualitative evaluation. On this basis, we developed a semistructured interview schedule that provided the capacity to prompt users for further details on their experiences based on their initial responses [36,37].

This study was conducted as part of a larger mixed methods study of AirRater users that covered multiple themes. The results

with respect to usability and behavior change have been published elsewhere [19], and this paper specifically reports the results with respect to user perceptions of app design, comprehension, and trust.

In the interest of an independent evaluation, a qualitative researcher (AW) with no affiliation to the stakeholders involved in the app design and development was used to conduct the evaluation, undertake all interviews, and analyze all data. A qualitative researcher at the University of Tasmania, who was not affiliated with the app, reviewed the framework and protocol as an independent reviewer before its submission to ethics.

Recruitment

As reported elsewhere [19], adults (18 years and older) were selected for recruitment if they resided in one of the following three locations in Australia: (1) Tasmania, (2) the Australian Capital Territory (ACT), or (3) Port Macquarie, New South Wales. Users were recruited from multiple locations to capture any location- or context-specific experiences, which were strategically chosen based on known episodes of high pollen days (Tasmania) and prolonged poor air quality because of the 2019-2020 wildfire events (ACT and Port Macquarie). Users were recruited through calls for participation via multiple media platforms, including a targeted email to registered AirRater users, an article in the AirRater newsletter, social media posts (Facebook and Twitter), and requests for participation during local radio interviews with AirRater team members. In total, 42 users were recruited to participate in the qualitative evaluation, with numbers almost evenly divided between Tasmania (20/42, 48%) and ACT (21/42, 50%). Despite multiple attempts to secure more participants, only 2% (1/42) of participants were recruited from Port Macquarie given that both locations experienced substantial wildfire smoke from 2019-20 wildfire events, ACT and Port Macquarie data were aggregated to protect the anonymity of the Port Macquarie participants.

Ethical Considerations

An ethics application for the evaluation was submitted to and approved by the University of Tasmania Health and Human Research Ethics Committee (ID: H0015006). As per the ethics approval, participants who contacted the team expressing an interest in participating were sent an information sheet and consent form before confirming their eligibility. The information sheet included information on the purpose of the study, the study's funding arrangements, details about what participants would be asked to do during the study, and the benefits and risks of participation. The participants were also informed that they were free to withdraw from the evaluation at any time. The evaluation and data collection methods were discussed with potential participants, who then returned their consent form if they were still happy to proceed with participation. Participants were notified that deidentified data would be stored at the University of Tasmania for a minimum of 5 years in accordance with the Australian National Health and Medical Research Council guidelines and would only be accessible to a subset of investigators involved in the evaluation (AW, FHJ, SLC, NC, and PJJ). Consent forms, as well as deidentified electronic transcripts and audio recordings, were stored in separate folders on a secure server at the Menzies Institute for Medical Research,

Tasmania. All data were password-protected and accessible only to a subset of investigators involved in the evaluation (AW, FHJ, SLC, NC, and PJJ). Informed consent was obtained from each participant before their involvement in the study. Participants were remunerated for their time and contributions with an AUD \$20 (US \$14.50) gift card following their participation in the interview.

Data Collection

Data were collected using 2 methods. First, a preinterview questionnaire delivered through SurveyMonkey collected key demographic details from the participants, including age range, any pre-existing conditions, and location (postcode, [Multimedia Appendix 1](#)). The questionnaire also collected data from participants on their personal descriptions of the app's purpose, period of use, primary motivation for downloading the app, and perceptions of use over time. Second, semistructured interviews were then conducted. The full interview schedule was published in the study by Workman et al [19]; the interview schedule components relevant to the results presented in this paper are provided in [Multimedia Appendix 1](#). All interviews were conducted via telephone or a web-based conferencing platform because of COVID-19 restrictions. With participants' permission, all interviews were recorded to verify the accuracy of the transcripts.

Data Analysis

Demographic data collected by the preinterview questionnaire were aggregated for descriptive analysis. The qualitative analysis software NVivo 12 (QSR International) was used to support thematic coding and analysis of the interview data [38]. All 42 interviews were transcribed by 2 research team members and verified by 1 before being uploaded to NVivo. Starting with the themes underpinning the interview schedule as an initial coding framework [39], all interview transcripts were thematically analyzed by the qualitative researcher leading the evaluation (AW). The overarching themes supported the identification of subthemes that emerged from the data. The results of the analysis were initially discussed with 2 researchers affiliated with the design and development of the app (PJJ and SLC) and who were involved in previous AirRater evaluations of user surveys.

Results

Questionnaire Data

The majority of the 42 participants (38/42, 90%) completed the preinterview questionnaire. Detailed results have been reported elsewhere [19]; however, the key results for contextualizing this study are presented in [Table 1](#). Notably, most users had one or more respiratory conditions (asthma, allergic rhinitis, other lung conditions, or a combination), and most primarily used the app during seasons in which particular environmental triggers were present. Open-text responses from users indicated that they were most likely to use the app during the wildfire season (summer) or the pollen season (spring and summer), although some participants indicated that use was sporadic outside of these times.

Table 1. Summary of participant characteristics derived from the preinterview questionnaire^a.

Characteristics	Responses, n (%)
Sex	
Female	30 (79)
Male	8 (21)
Age range (years)	
21-30	3 (8)
31-40	8 (21)
41-50	4 (10)
51-60	8 (21)
61-70	9 (24)
>70	6 (16)
Pre-existing health conditions	
Asthma	17 (45)
Lung condition other than asthma	9 (24)
Allergic rhinitis	24 (63)
Heart condition	3 (8)
Stroke	0 (0)
Diabetes	2 (5)
Pregnancy	0 (0)
Other	12 (32)
Time since app download	
<6 months	8 (21)
6-12 months	13 (34)
1-2 years	6 (16)
2-3 years	5 (13)
3-4 years	5 (13)
4-5 years	1 (3)

^aAdapted from Workman et al [19].

Interview Data

Overview

The interview results are reported below and stratified into four sections aligned with our research questions: (1) engagement with different features of AirRater, (2) additional information

and functionality needs, (3) comprehension of AirRater, and (4) trust in AirRater. Additional interview data covering different themes have been published elsewhere [19]. This paper focuses only on the data relevant to the themes outlined above. The key findings relevant to these 4 themes are summarized in Table 2, and the remainder of this section presents detailed results from each theme.

Table 2. Summary of key findings from interview data, stratified by research question.

Research question	Summary of key findings
Which features of the app do users engage with?	<ul style="list-style-type: none"> The visual map and location features are used most frequently. Receipt of general alerts for elevated environmental hazard levels prompted app use. Participants used the symptom reporting feature to varying extents. A few participants indicated that the provision of real time information made AirRater their preferred information source.
What are users' additional information or functionality needs?	<ul style="list-style-type: none"> Participants made a diverse range of suggestions and requests for additional data, features and functionality to enhance the app (see Multimedia Appendix 1 for a detailed summary). Common requests included a more detailed breakdown of pollen data, automatic notification of nearest monitoring station, the inclusion of wind speed or direction, and the capacity to download personal summary data for a specified period.
How do users perceive their comprehension of the information in the app?	<ul style="list-style-type: none"> Participants indicated they brought varying levels of scientific or air quality literacy to their interactions with the app. The interface design, such as the color code, 1-word rating system, helped participants and their family members understand environmental hazard data. Some participants indicated they were confused by the use of different metrics across different information sources.
Do users trust the information the app provides?	<ul style="list-style-type: none"> Many participants did not question the accuracy of environmental hazard data presented in the app. Some participants felt the app's affiliation with a university provided credibility. Some participants expressed mistrust in AirRater data accuracy given their distance from a monitoring station. The extent to which app ratings aligned with personal symptoms and visual cues (such as visible smoke) influenced trust in AirRater data. Some participants questioned the conclusions of their personal health profile based on inputs to the symptom reporting feature.

Which Features of the App Do Users Engage With?

The participants were asked which features they used. Many participants (26/42, 62%) indicated using location ([Figure 1](#) and [Multimedia Appendix 1](#)) and maps ([Figure 1](#) and [Multimedia Appendix 1](#)) features most frequently. Some participants (4/42, 9%) specifically noted the value of the location function in supporting their decisions and monitoring the locations where they had a family:

I loved how it could have the multiple locations which for us over summer was really helpful cos we had family moving around... [ID_16, ACT]

...one of the features I like was that I could set up... location zones and monitor not just how we were going but them (family) and use AirRater to determine whether they were likely to be having a good day or bad day and whether to give them a call or, you know if they would be OK today. [ID_10, ACT]

I like it how you can... set your own locations and then it gives you... accurate data for the pollen level, air quality, and the temperature. And... the map feature so... you can look at sort of the air pollution levels on the map. I actually looked at that a lot during the bushfire season [ID_19, TAS]

Importantly, the close-to-real time nature of the information available in the map and location functionality was particularly useful to people. A few (3/42, 7%) users indicated that up-to-date information was the core reason why they engaged not only with these specific features but with the app overall:

This one (AirRater) was far more useful from a day-to-day point of view than what we were getting information-wise from other sources...It didn't help me knowing that we'd had the worst air in the world over the last 24 hours. What I needed to know was right now. [ID_34, ACT]

...I ended up finding that AirRater seemed to have the most useful, up-to-date information... there was a lot of discussion amongst people and... I think that it ended up being that people started relying on AirRater being the most up-to-date, useful information. [ID_39, ACT]

Some (7/42, 17%) participants indicated that the receipt of alerts for elevated environmental hazard levels ([Multimedia Appendix 1](#)) and reminders for symptom reporting were effective in prompting their engagement with the app:

I really like the notifications when it gave the particulate or the pollen rating cos then I double-check my medications and make sure I have done all the right things. Every time I was sent an alert, I would do a report...I liked the frequency, because I probably reported twice a week at least. [ID_24, TAS]

...I would generally use it (symptom reporting) when I got a notification cos it would remind me to do it, and unless I'm reminded to do it then I tend to forget... [ID_28, TAS]

There were varying levels of engagement with, and perspectives on, the symptom reporting feature more generally (for an illustration of the symptom reporting feature, see [Multimedia](#)

Appendix 1). Some (4/42, 9%) participants liked the feature and found it easy to use:

I liked it actually. I just clicked on mild, moderate...and I thought, oh god, I'm a bit of whinger...then I thought, no, this data will be good for somebody, and I've just found now the summary, which is really good [ID_38, ACT]

...the reporting symptoms feature works really well... prompting you to report symptoms and report medication taken...[ID_41, ACT]

However, some (9/42, 21%) participants identified technical and practical issues with the symptom-reporting feature:

...there was no feedback to say, you know, your symptoms have been constant for the last week, you should go and see a doctor or anything like that. It was literally...just answering the questions...do you have itchy eyes, what have you done, have you put eyedrops in or whatever, so I'm just ticking the things that I have to tick and then that was it, now I've done my homework (laughs) That's how I saw it. [ID_35, ACT]

...Most of the time I have symptoms, and most of the time my symptoms are controlled by medication. The way this works, I have to tell them I've got a runny nose...and it's mild. Why is it mild? Because I'm on medication...when you're doing this, you know, four or five times a day you do get sick of that kind of stuff [ID_13, TAS]

What Are Users' Additional Information and Functionality Needs?

Participants were asked whether there were any additional features or information they wanted AirRater to provide. Most (38/42, 90%) participants recommended at least one enhancement to the app to supplement existing features or requested the provision of additional information. Given the diversity of participants, their specific health conditions, and their unique information needs, the responses were wide-ranging. However, common requests included a more detailed breakdown of the pollen count data, automatic notification of the nearest monitoring station, inclusion of wind speed or direction, and the capacity to download personal summary data for a specified period. A full summary is provided in [Multimedia Appendix 1](#); examples of responses included the following:

...the total pollen count's irrelevant...why should I care...if there's a lot of gum tree pollen around? ...I'm not allergic to gum tree pollen...I want to know specific levels of specific pollens which apply to me [ID_13, TAS]

I understand now why the PM2.5 is the most obvious particle size that's there (for air quality) so I sort of understand the other one (PM10) being a bit subliminal... I have found that I've had to hunt for, I've had to discover it for myself rather than it all being presented there on the front page... [ID_27, TAS]

...it would be good if they'd (AirRater) just tell me where the closest station is... getting the information from... so the user knows how accurate it is [ID_2, TAS]

I find windy conditions exacerbate my asthma...so even just having a wind speed...I feel like wind speed is more important than temperature... [ID_7, TAS]

What I would like to see is the capacity to be able to print it out... over a period of time, it would be perhaps better to look at it that way... and find out what the changes were. [ID_37, ACT]

How Do Users Perceive Their Comprehension of the Information in the App?

Participants were asked to describe their understanding of the information the app provided, how the interface does or does not support their comprehension, and whether additional information is needed to assist them in understanding the data. Participants indicated varying levels of perceived general air quality literacy; however, irrespective of this, most (33/42, 78%) participants indicated that they were able to sufficiently understand the information in the app to support their needs:

I have a very good understanding of the...information. I have a science degree... so I don't have any trouble understanding the information and I know exactly what the particulate issues are and all the rest of it. [ID_4, ACT]

...every now and then I thought, that's interesting that that's the metric for good and that's the metric for poor...because the poor kicked in at quite a low numeric value...and more just a curiosity cos I didn't know enough about...the scale that's used for air quality. ...if I got really seriously curious I could have Googled what the Australian air quality standards meant, but...I was happy to take, that's the standard, that's the number, I can make a decision on that. [ID_5, ACT]

I knew nothing about air quality data before the bushfires...I know that there's...a website associated with it (AirRater) where the data is translated from and...I did look at that a few times to sort of try and...get an idea of...what does this PM2.5 thing mean...I got a fairly basic but comprehensive understanding of the rating system in terms of what it meant when it said 'good' or 'poor'...and what that might mean for my health, but I certainly wouldn't ever be able to like tell somebody else it in a comprehensive manner. I just learnt enough to know that I felt comfortable that what was happening in the app was a reflection of comprehensive and good data that I could trust. [ID_14, ACT]

I had to Google what PM squared meant, in air quality, and I still didn't quite get it, but um, then I just look at the word good now, or poor, or whatever. I didn't understand that much, um, and then I thought, oh well, that's fairly scientific actually which is good, but yeah, I didn't understand, and I s'pose I don't

really need to understand that, I just Googled it out of interest. [ID_38, ACT]

Numerous (9/42, 21%) participants indicated that the interface helped them and their family members comprehend the information.

...some of my family members...they're from non-scientific backgrounds...they didn't seem to understand it...as well as me, but they could still get a sense of...what the information was showing, so I think it's good that, like you have color codes for the different levels. [ID_19, TAS]

I think the interface is fairly user-friendly for people who don't necessarily have...the science and health background to...interpret it, because of the color change and the...good, fair, poor, that made it really quick and simple to understand. [ID_34, ACT]

...I really liked especially for my kids being able to explain the coloring. ...my eldest is like, oh it's red, we can't go outside. Oh, it's purple, no, don't even open the window... [ID_16, ACT]

...it became a bit of a family game...we had different levels of how many trees we could see outside (laughs) ...we actually started being able to have a good gauge of...what the numbers, levels were like, and...connecting those to our visual observations, so...it became quite an interest factor as well for the whole family. [ID_34, ACT]

Several (6/42, 14%) participants discussed how the use of different air quality metrics on different apps and websites affected their own or others' ability to understand air quality information (in AirRater and beyond). This was particularly relevant from the beginning of 2019 to 2020 wildfire events because of the diversity of air quality metrics used by various sources:

...there was a wide disconnect in the greater community...because different people were using different pieces of information. So, the ones who were using 24-hour rolling...worked on a different poor/fair scale than the AirRater scale... [ID_34, ACT]

...people were using...a bunch of different apps that all had really different information...and people were really confused and they're like, well, this is telling me...everything's fine and that I can go outside, and I'm like, no, it's cos that's an average... [ID_39, ACT]

Do Users Trust the Information the App Provides?

Participants were asked whether they had ever questioned the data provided by AirRater. Participants indicated varying levels of trust in the information provided by AirRater. Many (21/42, 50%) participants did not question the accuracy of the data, citing reasons such as their inclusion on credible government agency websites and the comprehensiveness of the supplementary information.

I did read like, where it was made, and who was running it and all those bits, and it was also helpful,

like, that it was on the ACT Health website as well, I think. [ID_31, ACT]

...I had read a fair amount about it (air quality)...just out of interest because I do try and keep up with what's going on out there...When it came out on the (ACT government) health website...I thought, well...that's good enough. [ID_37, ACT]

I found that the one (app) that I implicitly trusted the most was just the AirRater one, because the information on the website was so comprehensive about, like this is how we are measuring it and this is why it's a good way of measuring it, and...it wouldn't be inaccessible to people without that sort of background and knowledge, but it was also like, it was also this is good data because of this reason, so it was enough information to go, yep this is what I trust, this is what I am okay with despite these debates happening in these social media spaces that I was involved in. [ID_14, ACT]

Some (3/42, 7%) participants acknowledged that their trust in the information—and in the case of one participant, trust in sharing personal information—was a result of their affiliation with a university, with references to the importance of data protection and ethical integrity:

...because it was run by a university...I implicitly trust...with data protection approaches and stuff...a university, particularly in Australia, with the really stringent ethics requirements that we have...I do feel comfortable sharing my data and sharing my information...that was a huge...reason why I even downloaded the app in the first place. [ID_14, ACT]

...if it's associated with a university it's more likely to be neutral, less likely to be contaminated by advertising dollars... generally it just seemed to me that it's likely that whichever algorithm you use or whatever to deduce it, it's probably likely to be quite accurate...[ID_15, TAS]

In contrast, several (10/42, 24%) participants indicated that they had questioned the information presented, in many cases attributing this lack of trust to a lack of data sensitivity or accuracy away from the monitoring stations:

I felt like the location, the air quality...wasn't necessarily sensitive enough, I think also because I live in a valley, so...I might be...experiencing something at my home that wasn't showing up on...the map at all... [ID_1, TAS]

...the air monitors don't actually accurately monitor where I am...we're just in a bit of a black hole...it's really limited, and I sort of stopped using it (AirRater) a lot because it was just, the data isn't there for what I was most needing it for. [ID_23, TAS]

...one of the issues even across an area the size of Canberra is the number of monitoring stations, so we're always taking AirRater and even the ACT air quality index with a grain of salt compared to your actual local situation. [ID_10, ACT]

I think I noticed that our closest one was Belconnen, which is probably not too far away... 10 or 15 kilometers away, so I guess you just sort of like, take that a bit into consideration. (ID_20, ACT).

The importance of personal experience with symptoms, as well as (the absence of) visual cues, also proved a fundamental factor influencing user trust, as some (3/42, 7%) participants acknowledged that they were at times skeptical or surprised by AirRater information:

Sometimes I'm feeling really rubbish and I look...it says that the pollen count's low, and I'm like, oh well then, maybe I'm not well, but if you go and look at a grass count specifically it says that it's quite high and like, oh that's contradictory... [ID_21, ACT]

you'd look outside and think, oh wow, I can see sort of the haze in the background but, I've...got close to normal...visual range, it doesn't look bad, it doesn't smell bad, but also realising that...harmful air quality kicks in the level below what the normal human can sense so, I tended to say, OK, yeah, if it's telling me that (air quality is poor), it just means I can't see the bad stuff, or smell the bad stuff. [ID_5, ACT]

A couple of times... I've felt a bit tight in the chest and I've been surprised that the air quality's been good... but... I put it down to there must be something else going on in my body that was causing that tightness, and it wasn't the air quality. [ID_6, TAS]

Some (3/42, 7%) participants also queried the accuracy of the conclusions in their personal health profile based on inputs into the symptom reporting feature.

AirRater was saying...your problem is with animals, because you've (been) exposed to animals and you're having symptoms, and I was like, well, actually it's not that simple, it's not that clear cut, and I've had animals for longer than I've had symptoms, so it's not necessarily the case... [ID_1, TAS]

...sometimes...I had a look at the graph when I did my symptoms and...it was all mornings and I'm like, pretty sure I said 24 hours for some of those, so I'm not sure if it's captured that as well as it could've. [ID_31, ACT]

Discussion

Principal Findings

Overview

Through semistructured interviews with 42 users, this study gained valuable insights into what features are useful, what facilitates user comprehension, and what influences users' levels of trust in the information provided by AirRater. Specifically, we found that location and map functions are the most useful features of AirRater for users. Furthermore, most participants were able to suggest app enhancements based on personal needs or preferences (Multimedia Appendix 1). We also found that, irrespective of self-reported literacy levels, most participants reported that the information provided in the app was easy to

understand and supported their needs. Finally, we found that many participants did not question the accuracy of the data presented in the app. As discussed below, this new knowledge extends our understanding of effective design principles in environmental health apps and highlights key considerations for future app design (Multimedia Appendix 1). Furthermore, our results confirm that AirRater is used by a diverse group of individuals with unique health conditions, personal preferences, and information needs, highlighting the need for environmental health apps to manage the core design tension between meeting diverse user needs and streamlining for ease of use and comprehension. In the 4 sections below, we discuss our key findings in further detail and compare them with those of previous research. We also discuss implications for the development of future environmental health apps.

Engagement With Features

The map and location features of AirRater were highly valued because they supported users in monitoring the movement of hazards or environmental conditions in other locations where family members were based. This finding is pivotal for informing the future design of environmental health apps, as the ability to access hazard information quickly and easily at multiple locations supports app users not only to assess, plan, and mitigate their own risk but also that of family members located elsewhere. Furthermore, numerous participants reported that the provision of near real time data via the map and location functions was highly valued, particularly in the context of rapid changes in air quality during the Australian 2019-2020 wildfire seasons. Other studies have determined the importance of accurate and timely information. For example, Gooze et al [40] found that consumer satisfaction and use increased following the introduction of real time information for a transportation information tool. For environmental hazard information in particular, the presentation of near real time information is critical for users to adequately support decision-making.

We also found that alerts for hazard levels and symptom reporting were effective in prompting users to look at and engage with AirRater. This finding is congruent with earlier research findings of increased engagement and higher-frequency app visits from push notifications [18,41]. The role of alerts in hazard communication is particularly important for "invisible" threats, such as pollen and, at times, air quality, as it can raise awareness and literacy of such hazards. Finally, AirRater participant perspectives on the symptom reporting feature varied; a few participants indicated that it was useful, while other participants indicated that they had experienced both technical and practical difficulties with the feature. Difficulties reported included the inability to modify symptom reports and the repetitiveness of symptom reporting. Previous research findings have confirmed that repetitive and multistep processes can discourage users from engaging with a particular app or feature. In their analysis, Cho et al [42] examined cognitive factors, including eHealth literacy and health app use efficacy, in app use among a sample of 765 app users from South Korea. They found that health app use efficacy is correlated with continued app use and hypothesized that app features that require multiple steps to access information require users to invest time and energy, which may impact their perceived health app use

efficacy [42]. Our findings underline the importance of simplified and streamlined symptom reporting processes to maximize the engagement and utility of such features in environmental health apps.

Additional Information and Functionality Needs

As detailed in the [Multimedia Appendix 1](#), participants suggested numerous enhancements to the app, including the incorporation of additional meteorological information (eg, wind speed or direction, air pressure, UV rating, humidity), functionality enhancements (eg, forecasting functionality, additional alerts, the inclusion of health advice), and clarification of existing information (eg, symptom triggers). Four common requests were presented by the participants: (1) a more detailed breakdown of pollen data, (2) automatic notification of the nearest monitoring station, (3) the inclusion of local wind speed or direction, and (4) the capacity to download personal summary data for a specified period. Our results on additional information and functionality highlight that AirRater users are diverse in their personal information and functionality needs. Although this diversity demonstrates the capacity of an environmental health app to effectively engage various populations, it also implies the fundamental importance of app personalization and customization to meet individual user needs [43,44], particularly when balancing the simultaneous need for simplicity and streamlining to enhance usability and comprehension.

Comprehension

First, despite acknowledging that they brought varying levels of air quality literacy to their engagement with AirRater, participants indicated that the current AirRater design, particularly the use of colors, facilitated their understanding of the data. The relationship between interface design and information comprehension has been reported by other researchers. For example, Caburnay et al [45] assessed a random sample of 110 diabetes-related apps for health literate design. They found that most apps studied used enabling elements, such as everyday language (88/110, 80%), bold and contrasting colors (89/110, 80.9%), and included visual, customizable content (77/110, 70%). Similarly, in their study exploring stakeholder perspectives for improving storm surge risk communication, Morrow et al [46] found that stakeholders preferred maps that used multiple colors to convey different levels of storm surge risk. Our results reaffirm the importance of color as a part of a strategic design approach for environmental health app development. Another key result was that some participants indicated that the use of different indices and metrics across information sources was confusing during the wildfire season (eg, 24-hour average vs hourly or real time levels). In the absence of standardized information, our findings highlight the importance of providing clear and accessible information on how to interpret the specific metrics presented for a given setting and the value of supplementary information via multiple media to address confusion when it arises. Finally, Stonbraker et al [47] found a group of end users with low health literacy preferred simple bar graphs with emojis, again emphasizing the need for simple visualizations that convey the health implications of data. Stonbraker, Porras, and Schnall [47] also

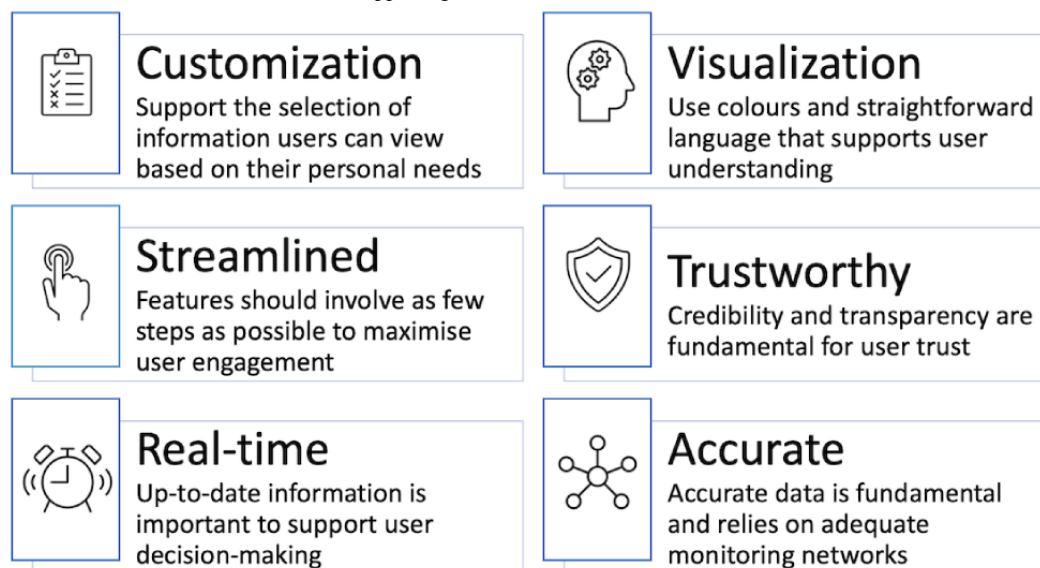
emphasized the importance of testing visualizations with target groups, a targeted approach that could be used in future iterations of AirRater to assess preferences and verify comprehension.

Trust

Many participants reported general trust in the information provided by the app, without questioning the accuracy of the information provided, with some reporting AirRater's affiliation with a university as a pivotal factor. Our findings are congruent with results reported elsewhere [48]. For example, in their scoping review on trust in digital health interventions, Adjekum et al [49] identified factors such as ease of use, self-efficacy, customizability, credibility, and stakeholder engagement as enabling trust in digital health interventions. Participants who expressed mistrust in the information provided by AirRater reported distance from an air quality monitor or the app's conclusions about their personal triggers as their reasons to question the data, although this did not seem to impact their overall willingness to use the app. This finding underscores the importance of data integrity in environmental health apps for protecting individuals from unnecessary or inadvertent exposure to environmental hazards. The need for a comprehensive and accurate air quality and pollen monitoring network is pivotal to ensuring that environmental health apps can effectively support populations. It also suggests opportunities exist for indoor air quality monitoring and the potential for environmental health apps to connect with devices, such as indoor air quality monitors or air purifiers, to provide supplementary data on potential triggers present in an indoor environment.

Implications for Future App Design and Future Research

[Figure 2](#) synthesizes the core implications of our study with respect to the environmental health app design. First, our study clearly showed that users value a level of customization with the capacity to access and select information based on their individual needs. However, streamlined design is important: users prefer features that are visual and involve minimal steps. In addition, users value real time information with transparency around data sources. Our results also highlight the importance of using visual cues to support comprehension; for example, the use of a one-word color rating system in AirRater supports users in understanding scientific information irrespective of their perceived level of health or air quality literacy and to support their family members, including children. The credibility of the app developer also proved to be important for influencing user trust as well as the perceived accuracy of data, in this case, provided by established air quality and pollen monitoring networks. Accordingly, environmental health apps must be supported by sufficient data inputs from adequate and robust monitoring networks. To ensure that key elements are integrated into app design, iterative collaboration among developers, environmental agencies, and users is likely to support better app functionality, enhance user trust in the data presented, and support the ultimate goal of improving population health outcomes.

Figure 2. Implications for future environmental health app design.

With respect to future research, this evaluation of AirRater as an example of an environmental health app suggests several future research needs. First, as environmental health issues can be highly localized, there is merit in pursuing similar evaluations of app utility and efficacy in other jurisdictions and countries. Although AirRater is currently only available in Australia, the implementation and evaluation of environmental health apps, such as AirRater, in other settings would extend the conclusions outlined here. Second, there is a strong need to investigate usability and comprehension among a more diverse range of user groups, including caregivers, children, and the elderly, across environmental health apps, to establish specific preferences and needs and to optimize health outcomes. At last, with respect to trust, it is important to investigate whether levels of trust in app data change because of likely changes to the air quality and pollen monitoring network in the coming years, either through the expansion of the existing network or the introduction of low-cost air quality sensors for individual residential use.

Limitations

Several limitations impact the strength of the conclusions presented in this study. First, in sampling a cohort of current AirRater users, results presented here are most useful for the purposes, as intended, of an in-depth evaluation of AirRater. However, the sample is not likely to be representative of all AirRater users or other populations who may choose to use environmental health apps. In this context, it is not possible to determine the extent to which the results presented here are generalizable to other groups. None of the participants in this study were culturally and linguistically diverse. Our choice of a “within person” study design has been vindicated by recent digital health reviews however greater diversity of participants would have been useful [50]. Another recognized limitation is the lack of representation of older children and young adults. Similarly, while a small number of respondents indicated having

children aged under 15 years, conversations focused on the personal use of and experience with the app. An opportunity exists to evaluate AirRater further by targeting older children (ie, 10–17-year-olds) and young adults (ie, 18–24-year-olds) to determine whether any age-specific customization is required to support further uptake of the app in these age groups. This is particularly important, given the app’s educational potential, as indicated by some participants. Finally, this evaluation only captures insights from AirRater app users, who are both engaged with the app and agree to participate. The opinions and responses of historical and current AirRater app users disengaged from the app were not accounted for in this study. Actively seeking the perspectives of users who no longer engage with the app would prove valuable for gaining unique insights into potential barriers to app use, as well as additional suggested app enhancements.

Conclusions

This qualitative evaluation of the free smartphone health app AirRater explored user preferences, comprehension, and trust. The participants’ perspectives can inform the future development of environmental health apps. Accordingly, the perspectives presented in this paper contribute to identifying key considerations for successful environmental health app design, including customization, streamlining, use of real time data, visual cues, credibility, and accuracy of the data. These considerations are more likely to be prioritized when apps are designed in collaboration with developers, environmental agencies, and users. In the future, environmental health apps are likely to play a pivotal role in supporting populations facing increasingly severe and frequent global environmental changes and extreme events. Ensuring that environmental health apps are fit for purpose, effective, accurate, comprehensible, and trustworthy will assist in maximizing the health outcomes of the population.

Acknowledgments

The authors thank Emerson Easley for her valuable administrative contributions to this research. The authors would also like to thank the 42 AirRater users who generously provided their time and perspective during the study. All the authors have read and agreed to the published version of the manuscript. This study was funded by the Menzies Philanthropic Fund. AirRater is currently funded by the Department of Health, Tasmania, Australian Capital Territory Health, Northern Territory Health, and the Menzies Institute for Medical Research, with previous funding from the Northern Territory Environment Protection Authority. AirRater collaborators include Environment Protection Authority Tasmania, Commonwealth Scientific and Industrial Research Organisation, Australian National University, Charles Darwin University, Tasmania Fire Service, the Bureau of Meteorology, and Asthma Australia.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

AW, PJJ, and FHJ conceptualized the manuscript; AW collected and analyzed the data; AW, PJJ, SLC, GJW, CL, DMJSB, NC, and FHJ all contributed to writing and editing the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A summary of AirRater features and functionality, interview questions and suggested app enhancements.

[[PDF File \(Adobe PDF File\), 416 KB - formative_v6i12e38471_app1.pdf](#)]

References

1. Bowman DM, Williamson GJ, Abatzoglou JT, Kolden CA, Cochrane MA, Smith AM. Human exposure and sensitivity to globally extreme wildfire events. *Nat Ecol Evol* 2017 Feb 06;1(3):58. [doi: [10.1038/s41559-016-0058](https://doi.org/10.1038/s41559-016-0058)] [Medline: [28812737](https://pubmed.ncbi.nlm.nih.gov/28812737/)]
2. Smith KR, Woodward A, Campbell-Lendrum D, Chadee DD, Honda Y, Liu Q, et al. Human health: impacts, adaptation, and co-benefits. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, et al, editors. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press; 2014:709-754.
3. Romanello M, Di Napoli C, Drummond P, Green C, Kennard H, Lampard P, et al. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *Lancet* 2022 Nov 05;400(10363):1619-1654. [doi: [10.1016/S0140-6736\(22\)01540-9](https://doi.org/10.1016/S0140-6736(22)01540-9)] [Medline: [36306815](https://pubmed.ncbi.nlm.nih.gov/36306815/)]
4. Barnes CS. Impact of climate change on pollen and respiratory disease. *Curr Allergy Asthma Rep* 2018 Sep 20;18(11):59. [doi: [10.1007/s11882-018-0813-7](https://doi.org/10.1007/s11882-018-0813-7)] [Medline: [30238321](https://pubmed.ncbi.nlm.nih.gov/30238321/)]
5. Ebi KL, Hess JJ, Watkiss P. Health risks and costs of climate variability and change. In: Mock CN, Nugent R, Kobusingye O, Smith KR, editors. *Injury Prevention and Environmental Health*. 3rd edition. Washington, DC, USA: The International Bank for Reconstruction and Development / The World Bank; 2017.
6. Pörtner HO, Roberts DC, Poloczanska E, Mintenbeck K, Tignor M, Alegría A, et al. Summary for policymakers. In: Pörtner HO, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegría A, et al, editors. *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press; 2022:3-33.
7. Che W, Frey HC, Fung JC, Ning Z, Qu H, Lo HK, et al. PRAISE-HK: a personalized real-time air quality informatics system for citizen participation in exposure and health risk management. *Sustain Cities Soc* 2020 Mar;54:101986. [doi: [10.1016/j.scs.2019.101986](https://doi.org/10.1016/j.scs.2019.101986)]
8. Johnston FH, Wheeler AJ, Williamson GJ, Campbell SL, Jones PJ, Koolhof IS, et al. Using smartphone technology to reduce health impacts from atmospheric environmental hazards. *Environ Res Lett* 2018 Mar 29;13(4):044019. [doi: [10.1088/1748-9326/aab1e6](https://doi.org/10.1088/1748-9326/aab1e6)]
9. Perez-Ramos JG, McIntosh S, Barrett ES, Velez Vega CM, Dye TD. Attitudes toward the environment and use of information and communication technologies to address environmental health risks in marginalized communities: prospective cohort study. *J Med Internet Res* 2021 Sep 23;23(9):e24671 [FREE Full text] [doi: [10.2196/24671](https://doi.org/10.2196/24671)] [Medline: [34554103](https://pubmed.ncbi.nlm.nih.gov/34554103/)]
10. Larkin A, Hystad P. Towards personal exposures: how technology is changing air pollution and health research. *Curr Environ Health Rep* 2017 Dec;4(4):463-471 [FREE Full text] [doi: [10.1007/s40572-017-0163-y](https://doi.org/10.1007/s40572-017-0163-y)] [Medline: [28983874](https://pubmed.ncbi.nlm.nih.gov/28983874/)]
11. Diamond CC, Mostashari F, Shirky C. Collecting and sharing data for population health: a new paradigm. *Health Aff (Millwood)* 2009;28(2):454-466. [doi: [10.1377/hlthaff.28.2.454](https://doi.org/10.1377/hlthaff.28.2.454)] [Medline: [19276005](https://pubmed.ncbi.nlm.nih.gov/19276005/)]

12. Kim S, Stanton K, Park Y, Thomas S. A mobile app for children with asthma to monitor indoor air quality (AirBuddy): development and usability study. *JMIR Form Res* 2022 May 23;6(5):e37118 [[FREE Full text](#)] [doi: [10.2196/37118](https://doi.org/10.2196/37118)] [Medline: [35604753](https://pubmed.ncbi.nlm.nih.gov/35604753/)]
13. Sousa-Pinto B, Anto A, Berger M, Dramburg S, Pfaar O, Klimek L, et al. Real-world data using mHealth apps in rhinitis, rhinosinusitis and their multimorbidities. *Clin Transl Allergy* 2022 Nov;12(11):e12208. [doi: [10.1002/ct2.12208](https://doi.org/10.1002/ct2.12208)] [Medline: [36434742](https://pubmed.ncbi.nlm.nih.gov/36434742/)]
14. Knitza J, Simon D, Lambrecht A, Raab C, Tascilar K, Hagen M, et al. Mobile health usage, preferences, barriers, and eHealth literacy in rheumatology: patient survey study. *JMIR Mhealth Uhealth* 2020 Aug 12;8(8):e19661 [[FREE Full text](#)] [doi: [10.2196/19661](https://doi.org/10.2196/19661)] [Medline: [32678796](https://pubmed.ncbi.nlm.nih.gov/32678796/)]
15. Mushquash AR, Pearson ES, Waddington K, MacIsaac A, Mohammed S, Grassia E, et al. User perspectives on a resilience-building app (JoyPop): qualitative study. *JMIR Mhealth Uhealth* 2021 Jul 08;9(7):e28677 [[FREE Full text](#)] [doi: [10.2196/28677](https://doi.org/10.2196/28677)] [Medline: [34255696](https://pubmed.ncbi.nlm.nih.gov/34255696/)]
16. Schmaderer M, Miller JN, Mollard E. Experiences of using a self-management mobile app among individuals with heart failure: qualitative study. *JMIR Nurs* 2021 Aug 09;4(3):e28139 [[FREE Full text](#)] [doi: [10.2196/28139](https://doi.org/10.2196/28139)] [Medline: [34406966](https://pubmed.ncbi.nlm.nih.gov/34406966/)]
17. Delmas MA, Kohli A. Engagement with air quality information: stated versus revealed preferences. *Organ Environ* 2019 Aug 31;34(3):413-434. [doi: [10.1177/1086026619837690](https://doi.org/10.1177/1086026619837690)]
18. Delmas MA, Kohli A. Can apps make air pollution visible? Learning about health impacts through engagement with air quality information. *J Bus Ethics* 2019 Jun 13;161(2):279-302. [doi: [10.1007/s10551-019-04215-7](https://doi.org/10.1007/s10551-019-04215-7)]
19. Workman A, Jones PJ, Wheeler AJ, Campbell SL, Williamson GJ, Lucani C, et al. Environmental hazards and behavior change: User perspectives on the usability and effectiveness of the AirRater smartphone app. *Int J Environ Res Public Health* 2021 Mar 30;18(7):3591 [[FREE Full text](#)] [doi: [10.3390/ijerph18073591](https://doi.org/10.3390/ijerph18073591)] [Medline: [33808395](https://pubmed.ncbi.nlm.nih.gov/33808395/)]
20. Rappold AG, Hano MC, Prince S, Wei L, Huang SM, Baghdikian C, et al. Smoke sense initiative leverages citizen science to address the growing wildfire-related public health problem. *Geohealth* 2019 Dec;3(12):443-457 [[FREE Full text](#)] [doi: [10.1029/2019GH000199](https://doi.org/10.1029/2019GH000199)] [Medline: [32159029](https://pubmed.ncbi.nlm.nih.gov/32159029/)]
21. Jones PJ, Koolhof IS, Wheeler AJ, Williamson GJ, Lucani C, Campbell SL, et al. Can smartphone data identify the local environmental drivers of respiratory disease? *Environ Res* 2020 Mar;182:109118. [doi: [10.1016/j.envres.2020.109118](https://doi.org/10.1016/j.envres.2020.109118)] [Medline: [32069747](https://pubmed.ncbi.nlm.nih.gov/32069747/)]
22. Jones PJ, Koolhof IS, Wheeler AJ, Williamson GJ, Lucani C, Campbell SL, et al. Characterising non-linear associations between airborne pollen counts and respiratory symptoms from the AirRater smartphone app in Tasmania, Australia: a case time series approach. *Environ Res* 2021 Sep;200:111484 [[FREE Full text](#)] [doi: [10.1016/j.envres.2021.111484](https://doi.org/10.1016/j.envres.2021.111484)] [Medline: [34116012](https://pubmed.ncbi.nlm.nih.gov/34116012/)]
23. Campbell SL, Jones PJ, Williamson GJ, Wheeler AJ, Lucani C, Bowman DM, et al. Using digital technology to protect health in prolonged poor air quality episodes: a case study of the AirRater app during the Australian 2019–20 fires. *Fire* 2020 Aug 04;3(3):40. [doi: [10.3390/fire3030040](https://doi.org/10.3390/fire3030040)]
24. Glattacker M, Boeker M, Anger R, Reichenbach F, Tassoni A, Bredenkamp R, et al. Evaluation of a mobile phone app for patients with pollen-related allergic rhinitis: prospective longitudinal field study. *JMIR Mhealth Uhealth* 2020 Apr 17;8(4):e15514 [[FREE Full text](#)] [doi: [10.2196/15514](https://doi.org/10.2196/15514)] [Medline: [32301735](https://pubmed.ncbi.nlm.nih.gov/32301735/)]
25. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. *Lancet Digit Health* 2019 Jun;1(2):e52-e54 [[FREE Full text](#)] [doi: [10.1016/S2589-7500\(19\)30013-5](https://doi.org/10.1016/S2589-7500(19)30013-5)] [Medline: [33323229](https://pubmed.ncbi.nlm.nih.gov/33323229/)]
26. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR Mhealth Uhealth* 2017 Dec 18;5(12):e190 [[FREE Full text](#)] [doi: [10.2196/mhealth.8758](https://doi.org/10.2196/mhealth.8758)] [Medline: [29254914](https://pubmed.ncbi.nlm.nih.gov/29254914/)]
27. Kowatsch T, Otto L, Harperink S, Cotti A, Schlieter H. A design and evaluation framework for digital health interventions. *Inf Technol* 2019;61(5-6):253-263. [doi: [10.1515/itit-2019-0019](https://doi.org/10.1515/itit-2019-0019)]
28. Lagan S, Aquino P, Emerson MR, Fortuna K, Walker R, Torous J. Actionable health app evaluation: translating expert frameworks into objective metrics. *NPJ Digit Med* 2020 Jul 30;3:100 [[FREE Full text](#)] [doi: [10.1038/s41746-020-00312-4](https://doi.org/10.1038/s41746-020-00312-4)] [Medline: [32821855](https://pubmed.ncbi.nlm.nih.gov/32821855/)]
29. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. *J Telemed Telecare* 2018 Jan;24(1):22-30. [doi: [10.1177/1357633X16673538](https://doi.org/10.1177/1357633X16673538)] [Medline: [27760883](https://pubmed.ncbi.nlm.nih.gov/27760883/)]
30. Moshi MR, Tooher R, Merlin T. Suitability of current evaluation frameworks for use in the health technology assessment of mobile medical applications: a systematic review. *Int J Technol Assess Health Care* 2018 Jan;34(5):464-475. [doi: [10.1017/S026646231800051X](https://doi.org/10.1017/S026646231800051X)] [Medline: [30201060](https://pubmed.ncbi.nlm.nih.gov/30201060/)]
31. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
32. Nouri R, R Niakan Kalhori S, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. *J Am Med Inform Assoc* 2018 Aug 01;25(8):1089-1098 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy050](https://doi.org/10.1093/jamia/ocy050)] [Medline: [29788283](https://pubmed.ncbi.nlm.nih.gov/29788283/)]

33. Paglialonga A, Lugo A, Santoro E. An overview on the emerging area of identification, characterization, and assessment of health apps. *J Biomed Inform* 2018 Jul;83:97-102 [FREE Full text] [doi: [10.1016/j.jbi.2018.05.017](https://doi.org/10.1016/j.jbi.2018.05.017)] [Medline: [29852318](https://pubmed.ncbi.nlm.nih.gov/29852318/)]
34. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
35. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. World Health Organization. Geneva, Switzerland: World Health Organization; 2016. URL: <https://apps.who.int/iris/handle/10665/252183> [accessed 2022-12-09]
36. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
37. Bryman A. Interviewing in qualitative research. In: Bryman S, editor. *Social Research Methods*. 4th edition. Oxford, UK: Oxford University Press; 2012:468-499.
38. NVivo Version 12. QSR International. 2018. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/support-services/nvivo-downloads> [accessed 2020-03-12]
39. Bryman A. Qualitative data analysis. In: Bryman A, editor. *Social Research Methods*. 4th edition. Oxford, UK: Oxford University Press; 2012:564-589.
40. Gooze A, Watkins K, Borning A. Benefits of real-time transit information and impacts of data accuracy on rider experience. *Transport Res Record* 2013 Jan 01;2351(1):95-103. [doi: [10.3141/2351-11](https://doi.org/10.3141/2351-11)]
41. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. *JMIR Mhealth Uhealth* 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: [10.2196/10123](https://doi.org/10.2196/10123)] [Medline: [30497999](https://pubmed.ncbi.nlm.nih.gov/30497999/)]
42. Cho J, Park D, Lee HE. Cognitive factors of using health apps: systematic analysis of relationships among health consciousness, health information orientation, eHealth literacy, and health app use efficacy. *J Med Internet Res* 2014 May 09;16(5):e125 [FREE Full text] [doi: [10.2196/jmir.3283](https://doi.org/10.2196/jmir.3283)] [Medline: [24824062](https://pubmed.ncbi.nlm.nih.gov/24824062/)]
43. Birkhoff SD, Smeltzer SC. Perceptions of smartphone user-centered mobile health tracking apps across various chronic illness populations: an integrative review. *J Nurs Scholarsh* 2017 Jul;49(4):371-378. [doi: [10.1111/jnu.12298](https://doi.org/10.1111/jnu.12298)] [Medline: [28605151](https://pubmed.ncbi.nlm.nih.gov/28605151/)]
44. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. *Med 20* 2015 Mar 02;4(1):e1 [FREE Full text] [doi: [10.2196/med20.3912](https://doi.org/10.2196/med20.3912)] [Medline: [25733033](https://pubmed.ncbi.nlm.nih.gov/25733033/)]
45. Caburnay CA, Graff K, Harris JK, McQueen A, Smith M, Fairchild M, et al. Evaluating diabetes mobile applications for health literate designs and functionality, 2014. *Prev Chronic Dis* 2015 May 07;12:E61 [FREE Full text] [doi: [10.5888/pcd12.140433](https://doi.org/10.5888/pcd12.140433)] [Medline: [25950568](https://pubmed.ncbi.nlm.nih.gov/25950568/)]
46. Morrow BH, Lazo JK, Rhome J, Feyen J. Improving storm surge risk communication: stakeholder perspectives. *Bull Am Meteorol Soc* 2015;96(1):35-48. [doi: [10.1175/bams-d-13-00197.1](https://doi.org/10.1175/bams-d-13-00197.1)]
47. Stonbraker S, Porras T, Schnall R. Patient preferences for visualization of longitudinal patient-reported outcomes data. *J Am Med Inform Assoc* 2020 Feb 01;27(2):212-224 [FREE Full text] [doi: [10.1093/jamia/ocz189](https://doi.org/10.1093/jamia/ocz189)] [Medline: [31670816](https://pubmed.ncbi.nlm.nih.gov/31670816/)]
48. Vo V, Auroy L, Sarradon-Eck A. Patients' perceptions of mHealth apps: meta-ethnographic review of qualitative studies. *JMIR Mhealth Uhealth* 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: [10.2196/13817](https://doi.org/10.2196/13817)] [Medline: [31293246](https://pubmed.ncbi.nlm.nih.gov/31293246/)]
49. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. *J Med Internet Res* 2018 Dec 13;20(12):e11254 [FREE Full text] [doi: [10.2196/11254](https://doi.org/10.2196/11254)] [Medline: [30545807](https://pubmed.ncbi.nlm.nih.gov/30545807/)]
50. Kwasnicka D, Keller J, Perski O, Potthoff S, Ten Hoor GA, Ainsworth B, et al. White Paper: Open Digital Health - accelerating transparent and scalable health promotion and treatment. *Health Psychol Rev* 2022 Dec;16(4):475-491 [FREE Full text] [doi: [10.1080/17437199.2022.2046482](https://doi.org/10.1080/17437199.2022.2046482)] [Medline: [35240931](https://pubmed.ncbi.nlm.nih.gov/35240931/)]

Abbreviations

ACT: Australian Capital Territory

WHO: World Health Organization

Edited by A Mavragani; submitted 04.04.22; peer-reviewed by K Druce, S Faruqui, S Dramburg; comments to author 01.11.22; revised version received 21.11.22; accepted 22.11.22; published 22.12.22.

Please cite as:

*Workman A, Johnston FH, Campbell SL, Williamson GJ, Lucani C, Bowman DMJS, Cooling N, Jones PJ
Evaluating User Preferences, Comprehension, and Trust in Apps for Environmental Health Hazards: Qualitative Case Study
JMIR Form Res 2022;6(12):e38471*

URL: <https://formative.jmir.org/2022/12/e38471>

doi: [10.2196/38471](https://doi.org/10.2196/38471)

PMID: [36548030](https://pubmed.ncbi.nlm.nih.gov/36548030/)

©Annabelle Workman, Fay H Johnston, Sharon L Campbell, Grant J Williamson, Chris Lucani, David M J S Bowman, Nick Cooling, Penelope J Jones. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 22.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Understanding Emergency Room Visits for Nontraumatic Oral Health Conditions in a Hospital Serving Rural Appalachia: Dental Informatics Study

Raj K Khanna¹, MD, DMD; Alfred A Cecchetti², MSci, MScIS, PhD; Niharika Bhardwaj², MBBS, MS; Bobbi Steele Muto¹, MPH, RDH; Usha Murughiyan², MBBS

¹Department of Dentistry, Oral & Maxillofacial Surgery, Joan C Edwards School of Medicine, Marshall University, Huntington, WV, United States

²Department of Clinical and Translational Science, Joan C Edwards School of Medicine, Marshall University, Huntington, WV, United States

Corresponding Author:

Alfred A Cecchetti, MSci, MScIS, PhD
Department of Clinical and Translational Science
Joan C Edwards School of Medicine
Marshall University
1600 Medical Center Drive
Huntington, WV, 25701
United States
Phone: 1 3046911585
Email: cecchetti@marshall.edu

Abstract

Background: In the Appalachian region, a variety of factors will impact the ability of patients to maintain good oral health, which is essential for overall health and well-being. Oral health issues have led to high costs within the Appalachian hospital system. Dental informatics examines preventable dental conditions to understand the problem and suggest cost containment.

Objective: We aimed to demonstrate the value of dental informatics in dental health care in rural Appalachia by presenting a research study that measured emergency room (ER) use for nontraumatic dental conditions (NTDCs) and the associated economic impact in a hospital system that primarily serves rural Appalachia.

Methods: The Appalachian Clinical and Translational Science Institute's oral health data mart with relevant data on patients (n=8372) with ER encounters for NTDC between 2010 and 2018 was created using Appalachian Clinical and Translational Science Institute's research data warehouse. Exploratory analysis was then performed by developing an interactive Tableau dashboard. Dental Informatics provided the platform whereby the overall burden of these encounters, along with disparities in burden by age groups, gender, and primary payer, was assessed.

Results: Dental informatics was essential in understanding the overall problem and provided an interactive and easily comprehensible visualization of the situation. We found that ER visits for NTDCs declined by 40% from 2010 to 2018, but a higher percentage of visits required inpatient care and surgical intervention.

Conclusions: Dental informatics can provide the necessary tools and support to health care systems and state health departments across Appalachia to address serious dental problems. In this case, informatics helped identify that although inappropriate ER use for NTDCs diminished due to ER diversion efforts, they remain a significant burden. Through its visualization and data extraction techniques, dental informatics can help produce policy changes by promoting models that improve access to preventive care.

(*JMIR Form Res* 2022;6(12):e31433) doi:[10.2196/31433](https://doi.org/10.2196/31433)

KEYWORDS

dental informatics; visualization; nontraumatic dental care; emergency room; cost; utilization; economic impact

Introduction

Overview

Oral health is critical to overall health and well-being [1]. However, various factors impact the ability to maintain oral health, such as the availability of excellent and accessible preventive dental care, socioeconomic status, age, location, race, ethnicity, and health behaviors [1]. Poor oral health or the absence of affordable regular dental care negatively affects health (associated with diabetes, heart disease, stroke, increased admission risk, and adverse pregnancy outcomes) and increases health care costs [2-7]. An emerging field, dental informatics, which involves the application of informatics to dental science, is increasingly used to solve problems in dental practice, education, training, and research [8-10].

Appalachia, a region with a largely rural population, spreads over 420 counties in 13 states. Rural Appalachia is medically underserved to the extent that there is an average of only 4 dentists per 100,000 individuals, compared to the United States' average of nearly 61 dentists per 100,000 [11]. West Virginia (WV) is the only state of the 13 states entirely in Rural Appalachia that is underserved and hit hard economically. WV leads the nation in edentulism, which afflicts 36% of adults, and reports indicate that only 61% of adults visited a dentist in the past year [12]. However, due to economic constraints, most health systems in Appalachia do not have the technological resources needed to investigate and solve the oral health issues of this population.

There are several indicators of oral health in a community. One such indicator is a visit to an emergency room (ER) for a preventable dental condition. When people don't have access to preventive dental care for problems like gum disease and tooth decay, treatable dental issues become a much bigger problem, often causing excruciating pain, leading people to seek care in an emergency department. Data show that ERs are the first and last resort for many low-income adults nationwide to obtain emergency care for preventable dental conditions. ERs are a significant part of the health care system that should be used for emergency health care needs and should not be a place for routine dental care. Further, these settings are ineffective for treating dental problems [13]. Over the past few years, dental care has increasingly shifted from dental offices to emergency departments, with a more significant portion of dental visits occurring in EDs [14]. ER use for dental problems is on the rise, and most visits are for nontraumatic dental conditions (NTDCs) [15,16]. It is harder to get good dental care in Appalachia, especially in WV, which encompasses some of the country's poorest and most remote communities. Compounding decreased availability of dental care and lower socioeconomic status is illiteracy. These factors contribute to the increased number of visits to the ER.

The use and financial impact of dental-related trips to medical settings, such as ERs and urgent care sites in the Huntington hospital system's catchment area in WV, has not been documented. We used dental informatics to develop a synopsis of dental visits to ER for preventable dental conditions in southern WV.

This paper highlights how dental informatics was used to examine preventable dental conditions in the Appalachian region of WV and suggests cost containment.

Literature Review

Dental informatics, a comparatively juvenile field [17], can be used to understand a wide variety of problems within dentistry but is not widely used [18] and historically may not be widely understood [19]. One of the difficulties with global acceptance is that dental practitioners may lack computer knowledge, creating a hurdle for introducing dental informatics [20]. e-Communities within dental societies may overcome these difficulties by bringing dental professionals together at a resource hub and allowing them to share information [21].

Dental informatics can assist in patient care [22] with new methods grounded in machine learning that can be used in the detection of teeth, caries, restored teeth, dental crowns, dental implants, prostheses for dentofacial deformities [23], oral pathology [24], oral cancer screening [25], and endodontic treatment [26]. Dental informatics has phenomenal potential in evaluating public health initiatives, and it could be used for the assessment of health goals achievements [26] as well as in education [27]; however, it has functional limitations resulting from medical and dental patient care not being stored in the same system. Dental practices typically silo their health care delivery systems, hindering easy access to institutional analysis [10].

Integrating dental informatics into health information systems provides an effective service. These combined systems will collect data from population-based surveys, disease surveillance systems, hospital information systems, and family health surveys and provide access to whole-patient patterns that can improve patient dental services [28]. Providing examples of how integrated dental informatics can focus on problems and provide solutions is expected to lead to more adaptive, patient-focused, and efficient dental care with educational advantages in training [29].

To understand the more enormous challenges within dentistry, especially concerning health in rural areas [30], the dental informatics system will require extensive collections of patient information [31], with both dental and medical information. Dental informatics can benefit the Appalachian region, but institutional costs are always of concern [32]. Hence, examples of how dental informatics could identify and initiate cost containment are of high value.

Methods

Procedures

The Appalachian Clinical and Translational Science Institute (ACTSI) Division of Clinical Informatics has a functional multi-institutional clinical research data warehouse (CRDW) containing more than 12 years of billing and electronic medical record data. The CRDW consists of relational tables, dimensions, and fact tables (Online Analytical Processing cube) that store multi-institutional medical information and provide data for operational and analytical model development (machine learning). It contains structured electronic health record data

(eg, vitals, medication, procedure, and diagnosis), non-electronic health record survey data, and unstructured (text) information received from Marshall Health practice plan, Cabell-Huntington Hospital, and Marshall University Joan C Edwards School of Medicine's Edwards Comprehensive Cancer Center. It uses the technological tools of information science (eg, computer workstations, mobile smartphones, interactive visualization, programming, and machine learning) to build a platform that can gather, analyze, and present information to address, in this case, oral health needs [33].

For this retrospective longitudinal study, we used the oral health data mart that was developed using the ACTSI's CRDW to understand and improve the dental health of the population in the area. Relevant clinical and financial data from the data mart over 9 years (2010-2018) were extracted, verified, and analyzed. ER encounters for preventable dental conditions were identified using the primary diagnosis codes (International Classification of Diseases, Ninth Revision, Clinical Modification; and International Classification of Diseases, Tenth Revision, Clinical Modification). We calculated the overall burden of dental-related

ER visits for avoidable conditions. Disparities in their burden regarding demographic variables, such as patient's age, gender, and primary insurance for the visit, were also examined. Further, patterns and trends in ER use for such visits and the associated charges were studied.

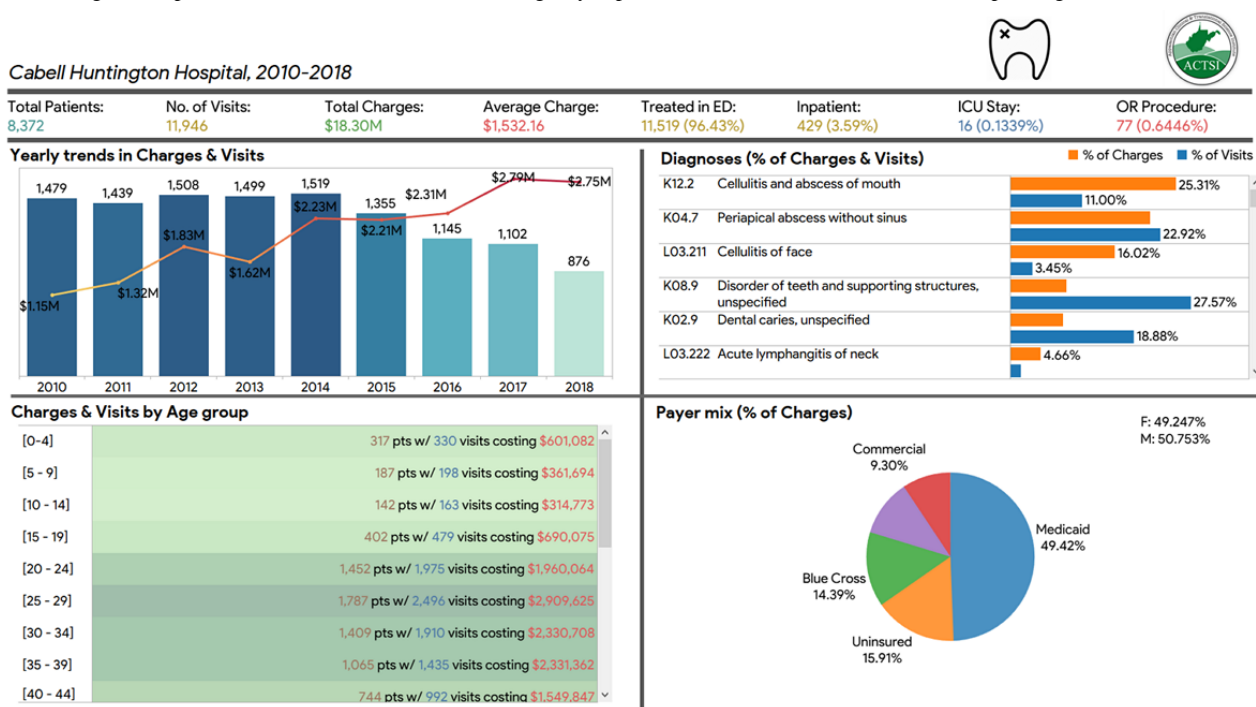
Ethical Considerations

The retrospective research study was conducted with the best intent and information obtained from the CRDW. This study was approved by the institutional review board of Marshall University, Huntington, WV (1069363-1).

Results

Transact-SQL coding found that 8372 patients made 11,946 visits to the ER for a nontraumatic dental condition and generated US \$18,303,173 worth of hospital charges over 9 years (Figure 1). Of these 8372 patients, 4123 (49.25%) were female, and 4249 (50.75%) were male, with mean ages of 32.43 (SD 0.20) and 32.37 (SD 0.198) years, respectively.

Figure 1. Tableau dashboard displaying patterns and trends in charges and visits for nontraumatic dental conditions to the emergency department at Cabell Huntington Hospital between 2010 and 2018. ED: emergency department; ICU: intensive care unit; OR: operating room.



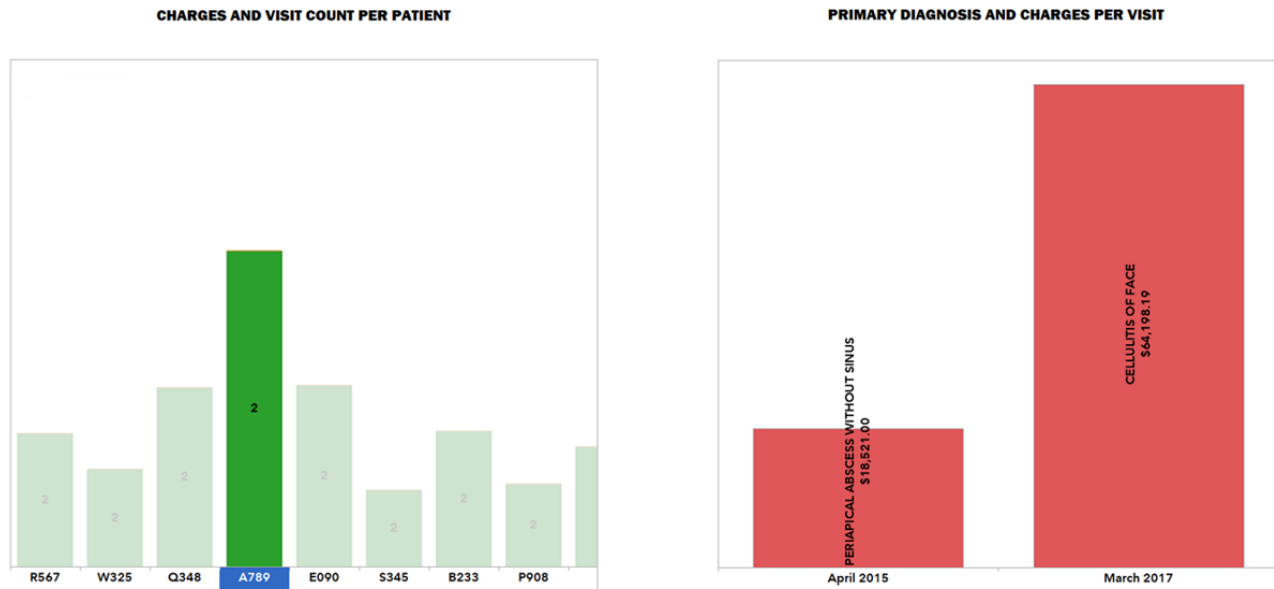
Using the oral health data mart, we found that although the number of visits decreased yearly, the charges quadrupled, with average costs per visit increasing from US \$776.64 in 2010 to US \$3136.79 in 2018. Additionally, the percentage of visits resulting in an inpatient admission or requiring a surgical intervention rose yearly from 2010 to 2018. Most of these visits were for teeth and supporting structure disorders and periapical abscesses without sinus. Meanwhile, the visits for cellulitis and mouth abscesses were the most expensive. This information was presented using a Tableau dashboard, which provided an interactive display that was informative but also simple to understand (Figure 1).

Medicaid was the primary payer for most of these visits in 2018 compared to a third of the payments in 2010. Of all the age groups, adults between the ages of 25 and 29 years had the highest visits and charges. It has also been noted that there was no distinct gender predisposition in the number of visits and charges accrued. However, it has been observed that female patients' ER visits were less for NTDCs compared to ER visits by male patients. Other predictors of an ER visit for NTDCs could be the age factor and the individuals' insurance status. It has been observed that individuals between the ages of 25 and 29 years with no insurance coverage had more ER visits. A similar trend was observed regarding the insurance status, as ER visits by uninsured patients were higher than those by the

insured patients. The type of insurance also played a role in the ER visits, as it has been noted that Medicaid-insured patients were more likely to be in ER for dental problems compared to Medicare-insured patients.

Using interactive visualization tools, we were able to drill down to the patient level, which showed that patients returned to the ER for NTDCs multiple times. An example of such visits by one such patient, along with the primary diagnosis and charges for the visits, is shown in Tableau-derived Figure 2.

Figure 2. Tableau drill-down. Patient-centered view showing all ER visits with corresponding primary diagnosis and charges for the selected patient.



Discussion

Principal Findings

Dental informatics combines technological tools with information science to hasten improvements in dental practice, research, education, and management [21,24,34]. In this paper, we examined how dental informatics was used to identify the vast resources spent on dental visits to the ER for NTDCs and how it was used to help formulate a plan to channel these resources toward preventive care, such as setting up mobile dental clinics or free oral health care checkups for the homeless to reduce ER dental visits.

Using dental informatics, we found that ER dental visits for preventable conditions posed a significant burden of more than US \$18.3 million. Further, despite a decelerating trend in the number of visits from 2010 to 2018, the average charges per patient per visit have increased. An accelerating trend has been traced in the percentage of visits that required hospital admissions and dental procedures for the same tenure. Dental informatics was instrumental in demonstrating that patients who visited the ER for NTDCs tended to be sicker every year, accruing more charges per visit. We speculate that this occurred due to the institution of services, such as a mobile dental clinic and residency programs, leading to diverting patients with less severe dental conditions from the ER settings.

Comparisons

In the past, several programs promoting these types of services have successfully reduced ER use for NTDCs [35]. Despite these measures, ER visits for NTDCs continue to be a substantial strain on the health care system, with emergency services provided in the ER being the only dental coverage available to

many patients in WV and surrounding regions [36]. Further, ERs are not well suited for treating dental conditions and promoting oral hygiene habits to prevent the problem from recurring; patients tended to come multiple times for similar dental issues, which we noticed in the patient-centered view of our interactive dashboard. We continue to use dental informatics to explore these areas of high interest and concern.

Limitations

This retrospective study used available data from 2010 to 2018 and consisted of those patients seen within the Marshall Health practice plan, Cabell-Huntington Hospital, and Edwards Comprehensive Cancer Center. This study did not include other outside hospital systems within the local area.

Conclusions

Dental informatics tools and approaches improve the dental practice's understanding and help assess dental services' economic burden [8]. Using dental informatics, we found that emergency department dental visits remain a significant and costly public health problem for vulnerable individuals. Institution of ER dental diversion services may lead to lower costs, which is more effective and can provide more appropriate care. However, they do not entirely address the issue, as there is a need to expand existing medical coverage, including coverage for preventive or restorative dental care. This will require buy-in from payers, an area where visualization tools can play an important part. But providing dental coverage alone might not be enough to reduce dental ER visits if patients do not have access to dental providers [37]. Thus, policy changes that direct efforts enabling dental coverage and access to preventive oral care, resulting in care quality and cost benefits, are necessary. These changes could save millions annually (US

\$2.75 million in avoidable charges in 2018) in the southern WV region alone. Dental informatics played an especially significant role in this Appalachia area study by providing accurate information in easy-to-understand presentations. This is particularly important in poorer rural areas, such as Appalachia, where multiple medical issues compete for scarce funding and attention.

Appalachian hospitals are especially concerned about costs with affordability, equity, and nonhealth benefits, factoring into decisions about health spending [38]. It is essential to present examples of how an Appalachian hospital's dental group, using dental informatics, could find and suggest solutions to a problem and guide decisions about where to spend limited resources that respond to the population's most significant health needs.

Acknowledgments

The authors would like to thank Gina Sharps, Bhavya Achen, and Wendy Moesteller for their assistance, with a special mention of Kim Tiemann, the program director of the Benedum Foundation. This publication was supported by the Claude Worthington Benedum Foundation and the National Institute of General Medical Sciences of the National Institutes of Health (NIH) under award number P20GM121299-01A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

This project was supported by the Marshall University School of Medicine Appalachian Clinical and Translational Science Institute (ACTSI). The content is solely the responsibility of the authors and does not necessarily represent the official views of the ACTSI.

Conflicts of Interest

None declared.

References

1. Oral Health in America: A Report of the Surgeon General.: US Public Health Service; 2000. URL: <https://www.nidcr.nih.gov/sites/default/files/2017-10/hck1ocv.%40www.surgeon.fullrpt.pdf> [accessed 2022-03-02]
2. Humphrey LL, Fu R, Buckley DI, Freeman M, Helfand M. Periodontal disease and coronary heart disease incidence: a systematic review and meta-analysis. *J Gen Intern Med* 2008 Dec;23(12):2079-2086 [FREE Full text] [doi: [10.1007/s11606-008-0787-6](https://doi.org/10.1007/s11606-008-0787-6)] [Medline: [18807098](https://pubmed.ncbi.nlm.nih.gov/18807098/)]
3. Pillai R, Iyer K, Spin-Neto R, Kothari S, Nielsen J, Kothari M. Oral health and brain injury: causal or casual relation? *Cerebrovasc Dis Extra* 2018 Jan 9;8(1):1-15 [FREE Full text] [doi: [10.1159/000484989](https://doi.org/10.1159/000484989)] [Medline: [29402871](https://pubmed.ncbi.nlm.nih.gov/29402871/)]
4. Ihezor-Ejiofor Z, Middleton P, Esposito M, Glennly A. Treating periodontal disease for preventing adverse birth outcomes in pregnant women. *Cochrane Database Syst Rev* 2017 Jun 12;6:CD005297 [FREE Full text] [doi: [10.1002/14651858.CD005297.pub3](https://doi.org/10.1002/14651858.CD005297.pub3)] [Medline: [28605006](https://pubmed.ncbi.nlm.nih.gov/28605006/)]
5. Verhulst MJL, Loos BG, Gerdes VEA, Teeuw WJ. Evaluating all potential oral complications of diabetes mellitus. *Front Endocrinol (Lausanne)* 2019 Feb 18;10:56 [FREE Full text] [doi: [10.3389/fendo.2019.00056](https://doi.org/10.3389/fendo.2019.00056)] [Medline: [30962800](https://pubmed.ncbi.nlm.nih.gov/30962800/)]
6. Laurence B, Mould-Millman NK, Scannapieco FA, Abron A. Hospital admissions for pneumonia more likely with concomitant dental infections. *Clin Oral Investig* 2015 Jul 31;19(6):1261-1268. [doi: [10.1007/s00784-014-1342-y](https://doi.org/10.1007/s00784-014-1342-y)] [Medline: [25359325](https://pubmed.ncbi.nlm.nih.gov/25359325/)]
7. Salomon D, Heidel RE, Kolokythas A, Miloro M, Schlieve T. Does restriction of public health care dental benefits affect the volume, severity, or cost of dental-related hospital visits? *J Oral Maxillofac Surg* 2017 Mar;75(3):467-474. [doi: [10.1016/j.joms.2016.10.019](https://doi.org/10.1016/j.joms.2016.10.019)] [Medline: [27875708](https://pubmed.ncbi.nlm.nih.gov/27875708/)]
8. Priya S, Ayeni M, Kalaivani S. Dental informatics in disaster management chapter. In: *Public Health Informatics—The Future Perspective*. Tamil Nadu, India: Shanlax Publications; Aug 31, 2017:49.
9. Madhu P, Kumar P, Prashant G, Sushanth V, Imranulla M, Nair A. Dental informatics: a click to the future. *J Oral Health Comm Dent* 2017:38-43 [FREE Full text] [doi: [10.5005/jp-journals-10062-0009](https://doi.org/10.5005/jp-journals-10062-0009)]
10. Webb A, Sommers S. Medical and dental data integration: how can this become a reality? *Denta Economics*. 2021 Feb 01. URL: <https://www.proquest.com/docview/2509706613?pq-origsite=gscholar&fromopenview=true> [accessed 2022-11-21]
11. Casto BC, Sharma S, Fisher JL, Knobloch TJ, Agrawal A, Weghorst CM. Oral cancer in Appalachia. *J Health Care Poor Underserved* 2009 Feb;20(1):274-285. [doi: [10.1353/hpu.0.0097](https://doi.org/10.1353/hpu.0.0097)] [Medline: [19202262](https://pubmed.ncbi.nlm.nih.gov/19202262/)]
12. Wiener RC. Dental fear and delayed dental care in Appalachia-West Virginia. *J Dent Hyg* 2015 Aug;89(4):274-281 [FREE Full text] [Medline: [26304952](https://pubmed.ncbi.nlm.nih.gov/26304952/)]
13. Figueiredo R, Dempster L, Quiñonez C, Hwang SW. Emergency department use for dental problems among homeless individuals: a population-based cohort study. *J Health Care Poor Underserved* 2016;27(2):860-868 [FREE Full text] [doi: [10.1353/hpu.2016.0081](https://doi.org/10.1353/hpu.2016.0081)] [Medline: [27180713](https://pubmed.ncbi.nlm.nih.gov/27180713/)]
14. Wall, T, Nasseh, K. Dental-related emergency department visits on the increase in the United States. Wall T, Nasseh K. American Dental Association: Health Policy Institute; 2013 May 01. URL: <https://txohc.org/wp-content/uploads/2018/10/dental-related-emergency-department-visits-on-the-increase-in-the-united-states.pdf> [accessed 2021-06-18]

15. EDs grapple with surging demand from patients with dental problems. *ED Manag* 2012 May;24(5):56-59. [Medline: [22545341](#)]
16. Zaborowski M, Dawson R. *EMOJ* 2016 Mar;4(1):23-32 [[FREE Full text](#)] [doi: [10.4236/ojem.2016.41004](#)]
17. Chhabra KG, Mulla SH, Deolia SG, Chhabra C, Singh J, Marwaha BS. Dental informatics in India: time to embrace the change. *J Clin Diagn Res* 2016 Mar;10(3):ZE12-ZE15 [[FREE Full text](#)] [doi: [10.7860/JCDR/2016/16970.7453](#)] [Medline: [27135022](#)]
18. Schleyer T, Spallek H. Dental informatics. A cornerstone of dental practice. *J Am Dent Assoc* 2001 May;132(5):605-613 [[FREE Full text](#)] [doi: [10.14219/jada.archive.2001.0237](#)] [Medline: [11367964](#)]
19. Schleyer T. Dental informatics: an emerging biomedical informatics discipline. *Adv Dent Res* 2016 Dec 01;17(1):4-8 [[FREE Full text](#)] [doi: [10.1177/154407370301700103](#)]
20. Montalvo Núñez K, Valenzuela Ramos M, Valenzuela Muñoz A, Scipión Castro R, Mendoza Murillo P. Management and administration of dental health services. *Llamkasun* 2021 Mar 15;2(1):97-104 [[FREE Full text](#)] [doi: [10.47797/llamkasun.v2i1.34](#)]
21. Spallek H, Irwin JY, Schleyer T, Butler BS, Weiss PM. Supporting the emergence of dental informatics with an online community. *Int J Comput Dent* 2007 Jul;10(3):247-264 [[FREE Full text](#)] [Medline: [18271498](#)]
22. Ghatole K. Bioinformatics: a new arsenal in dental sciences. *Int J Appl Dent Sci* 2020;6(3):610-614 [[FREE Full text](#)] [doi: [10.22271/oral.2020.v6.i3i.1015](#)]
23. Kim S, Lee J, Kim S, Lee D. Dental informatics to characterize patients with dentofacial deformities. *PLoS One* 2013;8(8):e67862 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0067862](#)] [Medline: [23940512](#)]
24. Singaraju S, Prasad H, Singaraju M. Evolution of dental informatics as a major research tool in oral pathology. *J Oral Maxillofac Pathol* 2012;16(1):83. [doi: [10.4103/0973-029x.92979](#)]
25. Sresumatchai V, Klongnoi B, Nityasuddhi D, Klaewkla J, Akko M. The development of dental informatics of oral cancer screening for prevention in community level: a prototype study In Thailand. *Asian Pac J Cancer Prev* 2021 Dec 01;22(12):3865-3873 [[FREE Full text](#)] [doi: [10.31557/apjcp.2021.22.12.3865](#)]
26. Coutinho D, Murthy A, Manjunath C, Shilpashree K, Thevara M, Khond M. Dental informatics: Current challenges and expanding opportunities in India. *Int J Appl Dent Sci* 2020;6(3):366-369 [[FREE Full text](#)]
27. Eisenberg EM. Dental informatics - its importance in dental education. *NYU Dentistry*. URL: <https://dental.nyu.edu/aboutus/news/nexus/millennium-issue/dental-informatics.html> [accessed 2022-11-21]
28. Garima G. Potential of health informatics towards centralized electronic health records approach. *J Electron Resour Med Libr* 2021 [[FREE Full text](#)] [doi: [10.5220/0004752402700275](#)]
29. Reynolds P, Harper J, Dunne S. Better informed in clinical practice - a brief overview of dental informatics. *Br Dent J* 2008 Mar 22;204(6):313-317 [[FREE Full text](#)] [doi: [10.1038/bdj.2008.193](#)] [Medline: [18356882](#)]
30. Hart L, Larson E, Lishner D. Rural definitions for health policy and research. *Am J Public Health* 2005 Jul;95(7):1149-1155. [doi: [10.2105/AJPH.2004.042432](#)] [Medline: [15983270](#)]
31. Sittig D, Kirshner M, Maupomé G. Grand challenges in dental informatics. *Adv Dent Res* 2003 Dec;17:16-19 [[FREE Full text](#)] [doi: [10.1177/154407370301700105](#)] [Medline: [15126200](#)]
32. Robinson C. An exploration of poverty in central Appalachia: questions of culture, industry, and technology. Tennessee State University. 2015. URL: https://digitalscholarship.tnstate.edu/cgi/viewcontent.cgi?article=1003&context=sswandurbanp_fac [accessed 2022-11-21]
33. Cecchetti A, Bhardwaj N, Murughiyan U, Kothakapu G, Sundaram U. Fueling clinical and translational research in Appalachia: informatics platform approach. *JMIR Med Inform* 2020 Oct 14;8(10):e17962 [[FREE Full text](#)] [doi: [10.2196/17962](#)] [Medline: [33052114](#)]
34. Eisner J. The future of dental informatics. *Eur J Dent Educ* 1999;3 Suppl 1:61-69. [Medline: [10865364](#)]
35. Rowland S, Leider JP, Davidson C, Brady J, Knudson A. Impact of a community dental access program on emergency dental admissions in rural Maryland. *Am J Public Health* 2016 Dec;106(12):2165-2170. [doi: [10.2105/ajph.2016.303467](#)]
36. Blackwell D, Villarroel M, Norris T. Regional variation in private dental coverage and care among dentate adults aged 18-64 in the United States, 2014-2017. Centers for Disease Control and Prevention. 2019. URL: <https://www.cdc.gov/nchs/products/databriefs/db336.htm> [accessed 2022-11-21]
37. Fingar KR, Smith MW, Davies S, McDonald KM, Stocks C, Raven MC. Medicaid dental coverage alone may not lower rates of dental emergency department visits. *Health Aff (Millwood)* 2015 Aug;34(8):1349-1357. [doi: [10.1377/hlthaff.2015.0223](#)] [Medline: [26240249](#)]
38. Naganandini S, Rekha K, Vanishree N. Economic evaluation and analysis of cost effectiveness in dental research. *J Dent Educ* 2011;3(1):67 [[FREE Full text](#)]

Abbreviations

ACTSI: Appalachian Clinical and Translational Science Institute

CRDW: clinical research data warehouse

ER: emergency room

NTDC: nontraumatic dental condition

Edited by A Mavragani; submitted 21.06.21; peer-reviewed by M Salimi, P Kanzow; comments to author 08.09.21; revised version received 31.03.22; accepted 05.11.22; published 23.12.22.

Please cite as:

Khanna RK, Cecchetti AA, Bhardwaj N, Muto BS, Murughiyan U

Understanding Emergency Room Visits for Nontraumatic Oral Health Conditions in a Hospital Serving Rural Appalachia: Dental Informatics Study

JMIR Form Res 2022;6(12):e31433

URL: <https://formative.jmir.org/2022/12/e31433>

doi: [10.2196/31433](https://doi.org/10.2196/31433)

PMID: [36563041](https://pubmed.ncbi.nlm.nih.gov/36563041/)

©Raj K Khanna, Alfred A Cecchetti, Niharika Bhardwaj, Bobbi Steele Muto, Usha Murughiyan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Perceptions of a Secure Cloud-Based Solution for Data Sharing During Acute Stroke Care: Qualitative Interview Study

Marcela Tuler de Oliveira¹, BSc, MSc; Lúcio Henrik Amorim Reis¹; Henk Marquering¹, BSc, MSc, PhD; Aeilko H Zwinderman¹, BSc, MSc, PhD; Sílvia Delgado Olabarriga¹, BSc, MSc, PhD

Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:

Marcela Tuler de Oliveira, BSc, MSc
Amsterdam University Medical Centers
University of Amsterdam
Meibergdreef 9
PO Box 22660
Amsterdam, 1100 DD
Netherlands
Phone: 31 205663273
Email: m.tuler@amsterdamumc.nl

Abstract

Background: Acute stroke care demands fast procedures performed through the collaboration of multiple professionals across multiple organizations. Cloud computing and the wide adoption of electronic medical records (EMRs) enable health care systems to improve data availability and facilitate sharing among professionals. However, designing a secure and privacy-preserving EMR cloud-based application is challenging because it must dynamically control the access to the patient's EMR according to the needs for data during treatment.

Objective: We developed a prototype of a secure EMR cloud-based application. The application explores the security features offered by the eHealth cloud-based framework created by the Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care Horizon 2020 project. This study aimed to collect impressions, challenges, and improvements for the prototype when applied to the use case of secure data sharing among acute care teams during emergency treatment in the Netherlands.

Methods: We conducted 14 semistructured interviews with medical professionals with 4 prominent roles in acute care: emergency call centers, ambulance services, emergency hospitals, and general practitioner clinics. We used in-depth interviews to capture their perspectives about the application's design and functions and its use in a simulated acute care event. We used thematic analysis of interview transcripts. Participants were recruited until the collected data reached thematic saturation.

Results: The participants' perceptions and feedback are presented as 5 themes identified from the interviews: current challenges (theme 1), quality of the shared EMR data (theme 2), integrity and auditability of the EMR data (theme 3), usefulness and functionality of the application (theme 4), and trust and acceptance of the technology (theme 5). The results reinforced the current challenges in patient data sharing during acute stroke care. Moreover, from the user point of view, we expressed the challenges of adopting the Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care Acute Stroke Care application in a real scenario and provided suggestions for improving the proposed technology's acceptability.

Conclusions: This study has endorsed a system that supports data sharing among acute care professionals with efficiency, but without compromising the security and privacy of the patient. This explorative study identified several significant barriers to and improvement opportunities for the future acceptance and adoption of the proposed system. Moreover, the study results highlight that the desired digital transformation should consider integrating the already existing systems instead of requesting migration to a new centralized system.

(*JMIR Form Res* 2022;6(12):e40061) doi:[10.2196/40061](https://doi.org/10.2196/40061)

KEYWORDS

qualitative interview study; electronic health records; cloud-based applications; acute stroke care; cross-organization data sharing; data privacy; encryption; data access control; mobile phone

Introduction

Background

A stroke is a medical condition that occurs when the blood supply to a part of the brain is suddenly interrupted, classified as ischemic, or when a blood vessel in the brain bursts, spilling blood into the spaces surrounding the brain cells, classified as hemorrhagic [1]. Fast access to information is essential in acute stroke care. During an emergency, health care professionals from different organizations need to evaluate the patient's condition, identify the type of stroke and severity, decide upon the treatment, transport the patient to the adequate care center, and perform the required intervention. Researchers have shown that the sooner the treatment is given, the better the outcomes for the patient are [2,3]. Moreover, patient transportation at the highest priority and hospital notification before patient arrival were associated with fast stroke care and better outcomes [4]. Finally, data availability through electronic medical records (EMRs) would improve decision-making and, ultimately, quality of care [5], leading to substantial reduction of unnecessary investigations and optimized communication among the acute stroke care teams involved in the treatment.

Emergency treatment of a patient usually requires cross-organizational collaboration: professionals at the emergency call centers, ambulance services, hospitals, and general practitioners' clinics. In the Netherlands, these health care organizations are independent and have different policies and systems for patient data sharing. However, from the first call to the emergency call center, all the professionals involved need to exchange information while treating the patient. Currently, this information is exchanged orally or via phone, as there is no unified EMR that all professionals can share during treatment. Such conventional information-sharing methods consume time and effort, and they are prone to errors. Therefore, the need for a system that enables acute care professionals to share patient data throughout the treatment process is evident, despite the organization in which the professionals work. Such data also represent valuable sources of evidence for later medical research.

Cloud storage services provide an environment that matches the needs for remote and ubiquitous access to the patient's EMR [6]. However, security and privacy challenges impede the widespread adoption of cloud services because they are susceptible to privacy and security threats [7]. Patients and health care organizations are afraid of losing control over the EMR when storing it on untrusted third-party clouds [8]. Finally, besides handling the privacy and security threats in cloud environments, cloud-based EMR applications must comply with the legal requirements regarding privacy and security imposed by the General Data Protection Regulation (GDPR) [9]. The GDPR attests that health care professionals and organizations are not obliged to systematically ask for patients' consent before they can use the data contained in the EMR. However, the professionals are bound by all the principles described in Article 5 of the GDPR, which ensures that the exemption from consent is proportionate and limited to what is necessary for the patient's treatment. Therefore, in the case of acute care, professionals

are allowed to access the patient's EMR only through their involvement in the treatment [10], requiring a solution that can dynamically grant and revoke access to the data.

A few solutions have been proposed to improve data availability and communication among professionals during acute care. Munich et al [11] presented a smartphone app to facilitate the tracking of the patient's location during ambulance transfer between organizations. Nam et al [12] also proposed a smartphone app based on the Cincinnati Prehospital Stroke Scale to aid self-screening and hospital decisions. However, these apps do not provide access to the patient's previous EMR.

Several studies have attempted to protect patient privacy in EMR cloud-based systems. Privacy-preserving approaches for eHealth clouds are classified as cryptographic and noncryptographic [10]. Various cryptographic approaches have been proposed to encrypt data in the cloud [13,14]. Seol et al [15] proposed a combination of approaches using attribute-based access control and encrypted files to share medical records stored in the cloud. However, these studies do not mention how to dynamically grant and revoke access to the encrypted data, which would be necessary to fully comply with GDPR.

Regarding dynamic access solutions, some systems offer *break-glass access*, which embodies the idea that, under certain conditions, a user can break the glass and explicitly override a denied access request [16]. Although some proposals use the break-glass approach to access encrypted EMR [17-20], access revocation after the emergency situation is still a problem. Thus, besides using encryption and access control to secure the data in the cloud, it is necessary to use modern techniques to adequately address all the requirements in acute care.

The Proposed Acute Stroke Care Application

Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care (ASCLEPIOS) is a project funded by the Horizon 2020 program [21]. The project developed the ASCLEPIOS eHealth cloud-based framework, which deploys several modern cryptographic and access control mechanisms for protecting corporate and personal sensitive data. The framework enables and facilitates the development of cloud-based eHealth applications that can protect the patient's privacy and prevent internal and external attacks. It combines dynamic index-based symmetric searchable encryption (DSSE) [22] and attribute-based encryption [23] to protect data in the cloud and to enable granting and revoking access to a user without interfering with the other users. These modern techniques allow dynamic management of encryption key access, therefore enabling more flexible access control that is important for acute care data sharing. Furthermore, the framework offers attribute-based access control based on flexible and configurable policies and attributes as an extra security layer to the encrypted data [24]. Only the users who hold the correct attributes can fulfill the policy and interact with the framework to access the data. Our organization participated in the ASCLEPIOS project and implemented a demonstrator exploring the framework for the acute stroke care case.

The ASCLEPIOS Acute Stroke Care demonstrator is a secure EMR cloud-based application that leverages the ASCLEPIOS

framework to share data among the acute care teams in a cross-organizational paradigm. In particular, it ensures that a team only has access to the patient's data under emergency conditions [25,26]. It relies on a unified EMR stored in the cloud in encrypted form to improve data accessibility during an emergency. Figure 1 shows the EMR data model, which follows the Fast Healthcare Interoperability Resources standard [27]. Figure 1 also shows the management entities and relations that the system uses to store necessary data, such as organizations, teams, and so on. Note that the EMR is encrypted with a unique

key for each patient, and health care professionals can obtain access to the key and encrypted data only while treating that patient.

At the beginning of the project, we collected health care and data privacy requirements from the potential stakeholders: professionals from call centers, ambulance services, and hospitals. The requirement was first published by Chomutare et al [28] and Reis et al [29]. Table 1 summarizes the requirements for the Acute Stroke Care demonstrator.

Figure 1. Electronic medical record (EMR) data model represented as entities relations of the Acute Stroke Care demonstrator, following the Health Level Seven Fast Healthcare Interoperability Resources (HL7 FHIR) standard. DSSE: dynamic index-based symmetric searchable encryption.

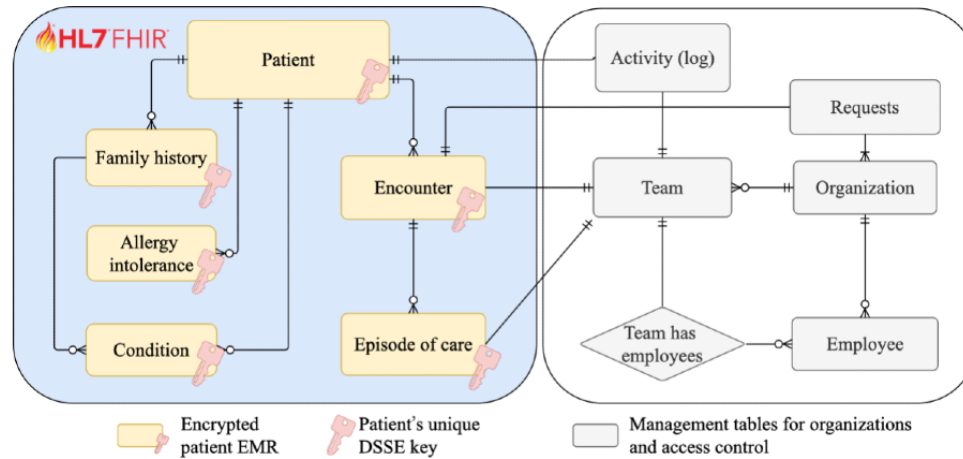


Table 1. Summary of requirements of the Acute Stroke Care demonstrator (extracted from the study by Reis et al [29]).

Requirement	Description
Availability	EMR ^a should always be available for access by legitimate users.
Confidentiality	Only authorized users should access the EMR.
Integrity	The accuracy and consistency of the EMR should be assured.
Nonrepudiation	The professional cannot deny what they have done.
Auditability	For every action, it must be possible to know who did it and what, when, where, why, and how the action occurred.

^aEMR: electronic medical record.

We implemented a web-based application to address the requirements listed in Table 1, with functionality to strengthen users' trust and comply with the GDPR. The EMR data are encrypted using a combination of DSSE to protect the data and attribute-based encryption to protect and manage the DSSE keys. The implemented attribute-based access control policies grant and revoke health care professionals' access according to their participation in the patient's acute stroke care timeline and present the EMR through the professionals' user interfaces.

Figure 2 shows a diagram of the architecture of the Acute Stroke Care demonstrator with the ASCLEPIOS framework and the stakeholders involved (patients and health care professionals). Patients and health care professionals have their own interface, through which they can interact with the system in different ways.

We implemented a specific user interface where the patients can add their medical conditions, allergies, medications, and family history; read data added by health care professionals; and visualize data access logs. Figure 3 shows an example of the patient interface with the list of organizations that treated them in a past emergency. For each organization, there are time stamps from when the organization joined, started, and completed acute care.

For each role in each organization, there is an interface through which, during an emergency, the professionals can access the patient's EMR and request other teams to join the emergency. Figure 4 shows an example of the call center interface used to treat a patient. The call center can input relevant information and request another team (eg, ambulance team), and on the right side, the EMR of the patient is presented. The interfaces for the ambulance and hospitals are similar to that shown in Figure 4.

Figure 2. Diagram of the Acute Stroke Care demonstrator architecture with the Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care (ASCLEPIOS) framework and the stakeholders involved. EMR: electronic medical record.

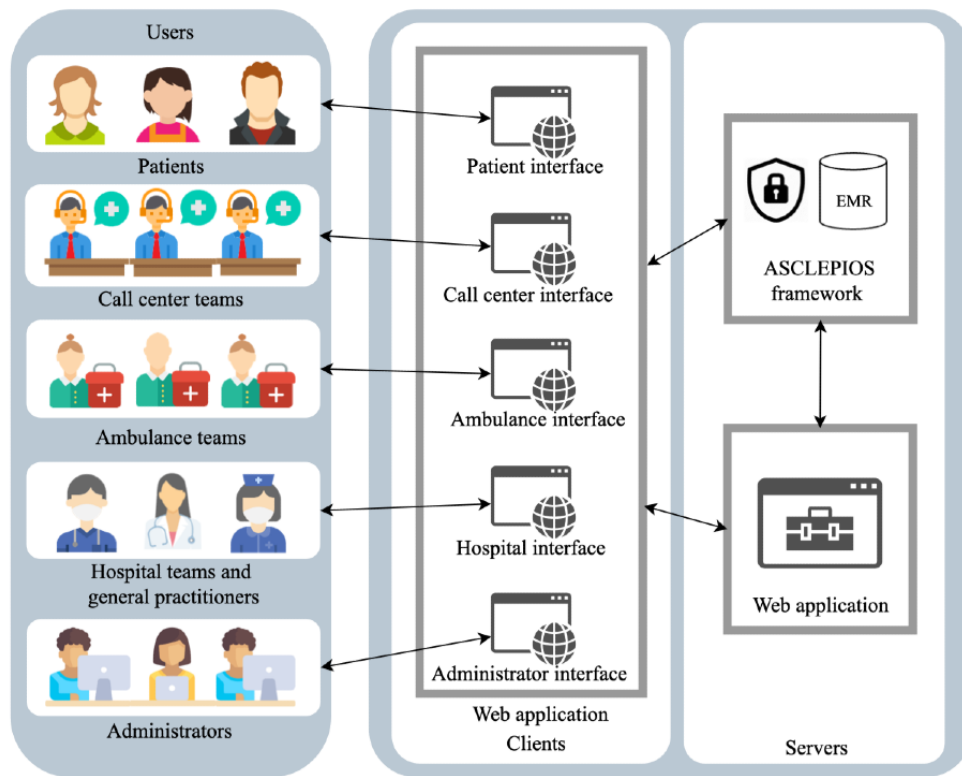
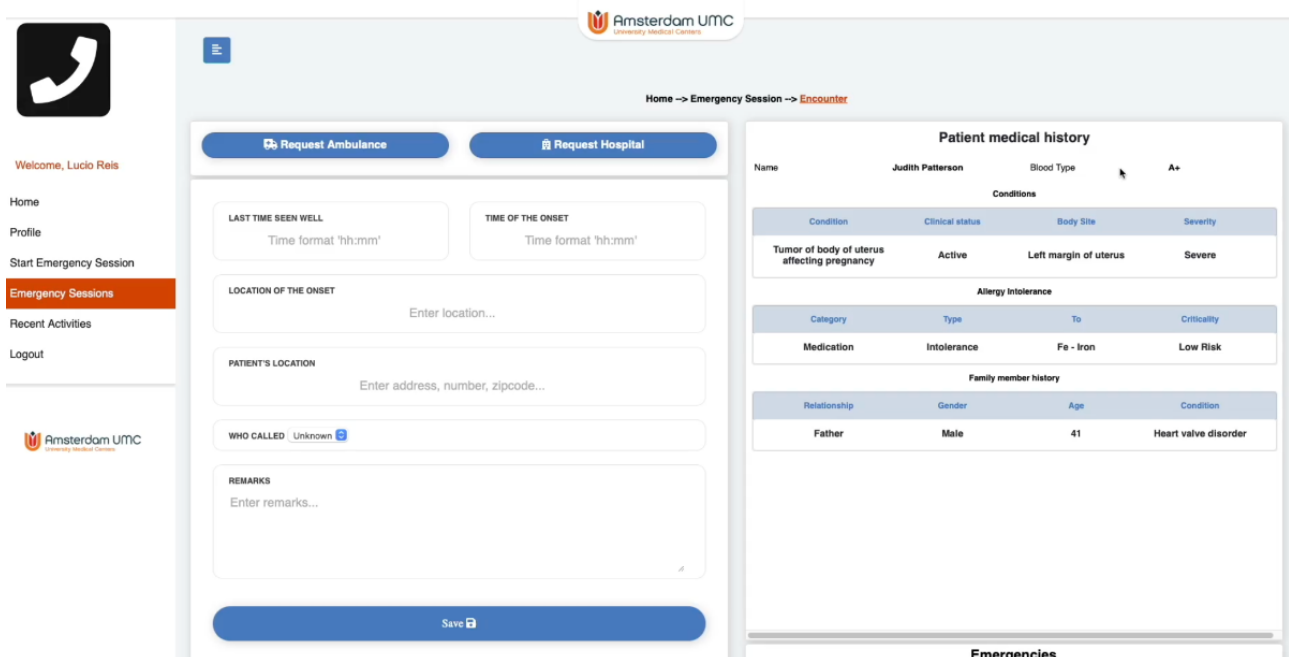


Figure 3. Example of the patient interface showing the organizations that treated them in some past emergency.



Figure 4. Example of call center interface treating a patient during an emergency.



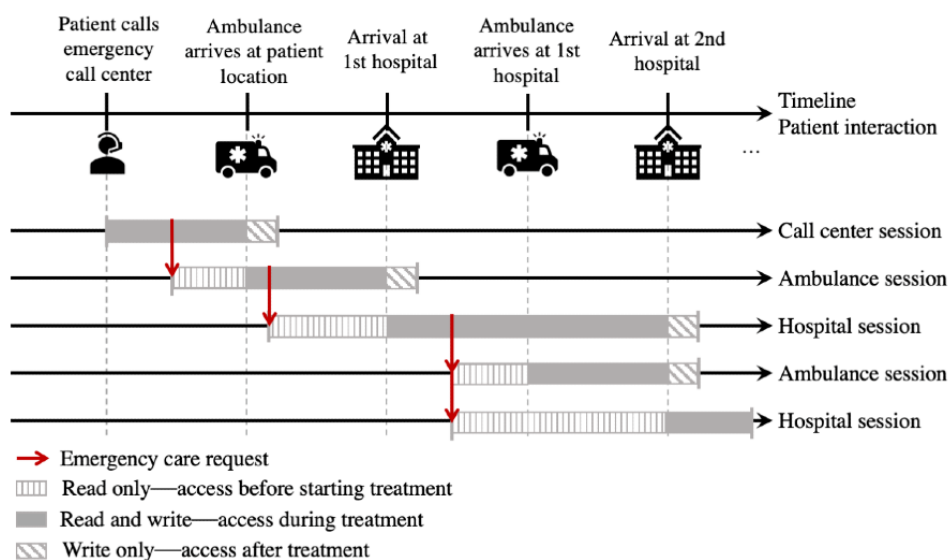
More information about the application can be obtained from our previous study [25] and videos in [Multimedia Appendices 1 and 2](#).

Figure 5 illustrates the information flow considered in the application during an emergency session, starting when the patient has a stroke until treatment completion at a hospital. An emergency session is the interval of time when all access to the patient's EMR occurs during acute care. The teams involved in the treatment become part of the emergency session for a period and leave the session when their task is completed. In this case, the patient contacts the emergency call center for help. From this moment, the call center professional searches for the patient's identification in the EMR system and starts an emergency session for this patient. Next, the call center professional requests an ambulance team to participate in this emergency session. After the ambulance arrives at the patient's location, the ambulance team performs triage and decides the hospital to which the patient must be taken for treatment. Once they know which hospital to go to, the hospital team also becomes involved in the emergency session of the patient. After arrival at the hospital, the hospital team confirms or invalidates that the patient has experienced an ischemic stroke and performs

adequate treatment. The patient is finally discharged and returns home. The same procedure occurs if the call center cannot identify the patient in the system. In such a case, a temporary identification is used to store and share the patient's data during the emergency, and later, the data are merged into the patient's EMR.

Figure 5 highlights that the health care professionals of each organization are involved only for a limited period, and access to the patient EMR must be provided only when necessary, complying with the GDPR. In an acute stroke care scenario, an involved health care team requests the participation of another team in the treatment; for example, the call center requests an ambulance to pick up the patient. Given the urgency, for adequate preparation, it would be necessary for the new team to have access to read the patient's EMR even before meeting the patient; for example, the requested ambulance team can read the patient's history during displacement. Moreover, the teams should have extra time to add data that could not be input during the treatment. Finally, access to the EMR must be revoked for any team that no longer participates actively in the patient's treatment; for example, access by the call center team is revoked after the ambulance team picks up the patient.

Figure 5. Example of an acute stroke care timeline involving multiple health care organizations.



Significance

It is essential to gain user input early in technology development to improve applications according to users' needs [30]. In this study, we presented the stakeholders with a web application designed to facilitate patient data sharing among acute care professionals using a secure cloud solution. We also explained how this application would be used during a simulated scenario of acute stroke care. This presentation served to disseminate a new vision for secure data exchange during a medical emergency, where the data are encrypted and decrypted locally in the user's device before being sent to the cloud. Moreover, access to patients' data is granted and revoked dynamically to the professionals according to their participation in the treatment. Furthermore, this study aimed to raise awareness and attract stakeholders' interest in this type of service. Finally, the

stakeholders' impressions and feedback further validated the ASCLEPIOS Acute Stroke Care application concept, thus providing valuable input for further technology development.

Objective

The goal of the interviews was 2-fold. First, the goal was to show the application's use to the main stakeholders: professionals from emergency call centers, ambulance services, and emergency hospitals and general practitioners. Second, we aimed to collect their impressions about how the application would fit into their daily acute care workflow.

Research Questions

With this study, we aimed to answer the following research questions (RQs):

- RQ1—What are the current challenges in patient data sharing during acute stroke care?
- RQ2—What are the participants' impressions about the proposed ASCLEPIOS Acute Stroke Care application?
- RQ3—What would be the challenges and suggestions for the adoption of the ASCLEPIOS Acute Stroke Care application in a real-life scenario?

Methods

Overview

We conducted an in-depth interview-based study with the main stakeholders in acute stroke care. We started recruiting participants and requesting their consent to record the interviews. The interviews were divided into 3 parts. First, we asked about the participants' familiarity with cybersecurity tools for data sharing in questionnaire part A. Second, we presented the ASCLEPIOS framework concepts and a simulation of the use of the ASCLEPIOS Acute Stroke Care application during acute stroke care and by the patient. Third, we asked about the participants' impressions regarding the use of the application in questionnaire part B. We tailored the in-depth interview according to the answers to the questionnaire, and the discussion evolved based on emerging findings. We conducted a qualitative thematic analysis of the data collected through the questionnaires and transcriptions of the interviews.

Recruitment

Participants were recruited from 4 groups, namely, representatives of emergency call centers (group 1 [G1]), ambulance services (group 2 [G2]), and emergency hospitals (group 3 [G3]) and general practitioners (group 4 [G4]). We started recruiting potential participants via email based on a contact person from the Amsterdam University Medical Center. Each message introduced the project and requested for an interview. Interviews were scheduled with those who responded

and provided informed consent to participate. After an interview, we always asked if the participants could indicate other potential participants from the 4 groups. We sent a total of 19 invitations. A follow-up email was sent to nonresponders after 1 week. When we did not get any response, we stopped any further contact with nonresponders, assuming that they had no interest in participating.

The recruitment process and interview occurred in 3 phases from September 2021 to August 2022: the first phase with 43% (6/14) of the participants, the second phase with 36% (5/14) of the participants, and the third phase with 21% (3/14) of the participants. We stopped recruitment when we reached thematic saturation and had similar representation of the 4 main stakeholders and potential users of the application. Our study's theoretical saturation refers to the point in data collection when no additional themes or insights are identified and data begin to repeat so that further data collection is redundant, signifying that an adequate sample size is reached [31]. During the second phase, we reached thematic saturation. In the third phase, we validated the saturation once the participants did not bring any new themes or suggestions in addition to those already put forward by participants in the previous phases.

Data Collection

In the study, 2 coauthors interviewed each participant individually. Of the 14 participants, 9 (64%) participants were interviewed in person and 5 (36%) were interviewed via the web. In general, the interviews lasted approximately 45 to 60 minutes. We interviewed participants from various acute care organizations in the Netherlands. During the interviews, we collected data of 2 types: the answers to the structured questionnaire (parts A and B) implemented using Google Forms (Google LLC) [32] and the audio recordings of the interviews conducted via a cell phone. All the demographic data collected are stored in a private file. Table 2 summarizes the demographic information about the interviewees.

Table 2. Demographics of the participants (N=14).

Variables	Count, n (%)
Sex	
Male	9 (64)
Female	5 (36)
Intersex	0 (0)
Role in acute care	
Emergency call center professional	3 (21)
Ambulance nurse	4 (29)
Emergency and neurologist physicians in hospital	4 (29)
General practitioner	3 (21)
Experience in acute care (years)	
0-4	2 (14)
5-9	4 (29)
10-14	1 (7)
15-19	3 (21)
20-25	1 (7)
≥25	3 (21)
Region in the Netherlands	
North Holland	9 (64)
Utrecht	3 (21)
South Holland	2 (14)

Data Management

After the interviews, we transferred the recordings via a secure private network to the *otter* service to automate the transcription process [33]. The transcriptions were treated according to the 6 steps proposed by Azevedo et al [34]. Interview transcripts, notes, and answers to the questionnaires were pseudonymized using the same identifiers and divided into 4 groups. For example, “Participant 1 from G1” is a professional from an emergency call center. The audio recordings were stored in an encrypted digital audio recorder maintained in a local machine. Only the pseudonymized transcripts were shared with other coauthors. The audio recordings will be retained for 1 year after the end of the ASCLEPIOS project (June 2023), and the transcripts and answers to the questionnaires will be retained for 5 years after the end of the project.

Data Analysis

Data were analyzed following the 4 steps from the principles of qualitative study and systematic text condensation [35]. This procedure consists of the following steps. First, we read the transcripts and the answers from the questionnaires to obtain an overall impression and identify preliminary themes as responses to the RQs of this study. The preliminary themes were directly related to the questionnaires. Second, we defined the coding that represented the themes and subthemes. Then, we read all the transcripts and answers once again and assigned themes and subthemes to the transcripts, with the support of *MAXQDA* software (VERBI GmbH) [36]. Third, we condensed

the transcripts and answers into themes and subthemes. Finally, we synthesized the descriptions of the participants’ impressions and their feedback as quotations.

Ethical Considerations

All participants were asked to provide written consent based on oral and written information about the study, and only those who provided their consent were included (14/19, 74%). The study did not collect or otherwise handle patient-related or health-related data. All the data collected using the questionnaires through Google Forms were pseudonymized and correlated to the transcripts through the time stamps. Moreover, only the authors (MTO and LHAR) had the permission to access the data in Google Forms. The ASCLEPIOS project’s ethics advisory committee and data protection officer assessed the study design and informed consent forms. They concluded that a more rigorous ethical review was unnecessary because the study did not collect any sensitive or personal data.

Results

Overview

A total of 14 participants were interviewed. They classified their roles as professionals from call centers (3/14, 21%), ambulance services (4/14, 29%), hospitals (4/14, 29%), and general practitioners’ clinics (3/14, 21%). We represent the 4 groups to show the diversity of the participants according to their roles in acute care. In general, the interviewees were very interested

in understanding the vision proposed by the application and were excited to provide feedback.

We identified 5 themes in the data analysis, namely, current challenges (theme 1), quality of the shared EMR data (theme 2), integrity and auditability of the EMR data (theme 3), usefulness and functionality of the application (theme 4), and trust and acceptance of the technology (theme 5). In the analyses phase, we did not observe any significant correlation between the groups and answers, and there was no theme that was mentioned only by a specific group. Therefore, the results are

not presented per group, and we only use the groups in the citation because it provides more context to participants' quotations.

An overview of the identified themes and subthemes is presented in [Textbox 1](#). [Table 3](#) presents the relationship among the identified subthemes, the questions from the questionnaires (parts A and B), and this study's RQs. The results presented in the following subsections use the questionnaire part and the number of the question; for example, A1 is the answer to questionnaire part A, question 1.

Textbox 1. Overview of themes and subthemes.

Theme 1

- Current challenges
- Subthemes
 - 1.1—The current systems lack standardization and structure of data
 - 1.2—Noninteroperability of systems hampers the exchange of data
 - 1.3—Achieve professionals' awareness about security and privacy of the patients' data

Theme 2

- Quality of data
- Subthemes
 - 2.1—Reliability of the data provided by the patient
 - 2.2—Reliability of the data provided by other teams

Theme 3

- Integrity and accountability
- Subthemes
 - 3.1—Prevention of data loss
 - 3.2—Accountability of the data added and edited during the treatment
 - 3.3—Duration of the extra time to add and edit data after the end of treatment
 - 3.4—How to handle unknown patients during acute care

Theme 4

- Usefulness and functionality
- Subthemes
 - 4.1—Integration of the application with other (existing) systems as data sources
 - 4.2—Granularity of access control to parts of the electronic medical record
 - 4.3—Information about the patient's condition after the treatment, for learning purposes

Theme 5

- Trust and acceptance of the technology
- Subthemes
 - 5.1—Professionals' training to use the system
 - 5.2—Extend the system to include all types of stakeholders of an electronic medical record system
 - 5.3—Merge current systems instead of proposing a new one
 - 5.4—Increase patient trust and awareness

Table 3. Questions from questionnaires (part A and part B), how they are related to the research questions of this study, and the identified subthemes.

Questionnaire part and questions	Research questions	Subthemes
A		
1. Do you use any EMR ^a system to share patient data?	1	1.1 and 1.2
2. Is the EMR system cloud-based?	1	1.1, 1.2, and 1.3
3. Is the patient data encrypted in the EMR system?	1	1.1 and 1.3
4. Would you be willing to share encrypted patient data in a cloud-based solution across multiple healthcare organisations?	1	1.1 and 1.3
5. How important is it to keep the patients' data confidential and only available to the healthcare professionals involved in their treatment?	1	1.3
6. How much would a patient data leakage affect the patient's life?	1	1.3
B		
1. How would information such as medical conditions, allergies/intolerances, and family history, as informed by the patient in the demo, be useful in case of emergency?	2	2.1
2. How much would the availability of patient data before the treatment improve the decision-making during treatment?	2	2.2
3. Do you believe that a digital system, such as the demo, could prevent data loss?	2	3.1 and 3.2
4. The demo considers accountable the professional, the team, and the organisation who added new data to the patient record during treatment. Who do you think should be accountable?	2	3.2
5. Do you think that healthcare professionals should be able to add or edit the patient's data after the treatment ends?	2	3.1, 3.2, and 3.3
6. Do you think a system like this demo could be useful in a real situation?	2	3.4
7. What would be needed to improve the usefulness of a system like this demo?	3	4.1, 4.2, and 4.3
8. Would you trust using a system like this demonstrator in your daily tasks?	3	5.1
9. What would be needed to increase your trust in a system like this demo?	3	5.1
10. How likely would your organisation be to accept adopting a system like this demo in a real situation?	3	5.2 and 5.3
11. What would be needed to improve your organisation's acceptability of a system like this demo?	3	5.2 and 5.3
12. Do you think a system like this demo could make patients feel safer about providing their data to your organisation?	3	5.4

^aEMR: electronic medical record.

Current Challenges for Patient Data Sharing During Acute Stroke Care

The first theme emerged when the participants answered questionnaire part A. All participants (14/14, 100%) told us about how they share patient data during acute care and their difficulties. Of the 14 participants, 13 (93%) said that they use EMR systems to share patient data and feel comfortable with them (A1). Overall, one-third (4/14, 29%) of the participants use cloud solutions, one-third (5/14, 36%) do not use the cloud, and the remaining one-third (5/14, 36%) do not know how the system stores the data (A2). Most participants (12/14, 86%) use different systems in different organizations, and these systems usually do not communicate directly with each other (subtheme 1.1). In the Netherlands, the call center and ambulance professionals can share data about the emergency. However, these professionals do not have access to previous medical records; they have access to data only about the ongoing acute care event. The hospitals usually do not communicate directly with the ambulance systems, and the data are generally duplicated when shared. Moreover, in North Holland, the

ambulance team can print the information collected during patient transportation and give the paper to the hospital team on arrival. A participant expressed this as follows:

...Now we are still working in such an old fashion with paper. Even after the team types the information inside the ambulance, I will receive a paper printed out or a PDF document when I receive the patient. Then I need to manually extract what I think is relevant information and insert it into another system with 10-15 words, and this is the medical report in the patient file. [Participant from G3]

The lack of interoperability was also mentioned as a big challenge because, even if they have access to other systems, they usually cannot merge the patient data into a single EMR (subtheme 1.2). The general practitioners have to merge the records manually when following up on the patient's treatment:

As a GP [general practitioner], when my patient calls and I suspect that there is a stroke, I will request an ambulance, and I will receive a notification when the patient arrives at the "hospital x" and receives the

treatment. But I can't see anything more. So I need to ask them for the treatment records, and I receive a PDF file again, and I need to insert the information again into the GP system. This is really annoying! [Participant from G4]

The participants told us about their awareness of security and privacy responsibilities regarding the patients' data (subtheme 1.3). Of the 14 participants, 10 (71%) do not know if their EMR system stores the patient data in encrypted form (A3). Nevertheless, all participants (14/14, 100%) were willing to share encrypted patient data in a cloud-based solution across multiple health care organizations (A4). In addition, they all agreed that it is important to keep patient data confidential and make them available only to the health care professionals involved in their treatment (A5). Of the 14 participants, 13 (93%) believed that patient data leakage would affect the patient's life (A6). Some of them also criticized the current data management approaches, which usually offer break-glass buttons that bypass the conventional access control mechanism of the system to any professional who has access to the system:

When I need to access some data that I usually don't have access to, a "break-glass" pop-up appears, and if I click yes, I have access to the data. [Participant from G3]

Participants' Impressions About the Proposed Application

The second theme emerged when the participants answered questionnaire part B, regarding their impressions about the application after seeing it in use.

The application enables the patient to input some information into the system, such as medical conditions, allergies, intolerances, and family history. Therefore, we asked how such information could be useful in an emergency case. Of the 14 participants, 13 (93%) believed that it would be very much useful (B1). However, all the participants (14/14, 100%) commented on the doubts about the sufficient quality and reliability of the information provided by the patient for acute care decision-making (subtheme 2.1):

Usually, when patients add medical information to their files, that is not the type of information that a doctor is looking for. For example, if patients add that they have a tumour, they cannot say the location of the tumour nor describe it as the doctor will do. Thus, the information is not that useful, but it is better than nothing. [Participant from G3]

As a doctor, I don't think that the data the patient inputs to the system is 100% reliable. I would trust it more if another doctor had added the information. [Participant from G4]

Although all participants (14/14, 100%) agreed that the availability of data before the treatment starts could improve decision-making (B2), some types of data are double-checked and input into the system again when the patient is delivered to another health care team, for example, when the ambulance delivers a patient at the hospital (subtheme 2.2):

Having access to what the teams [call center and ambulance] added about the patient can save a lot of effort and make the treatment faster. However, suppose the patient comes from another hospital and has already done some imaging. Nowadays, the next hospital team usually remakes the images exams even if they have access to the previous exam. [Participant from G4]

Well, it's great that the emergency nurses write down what's going on. As a doctor at the end of the line, I would already know the blood pressure of the patient or something. But the truth is that it is very likely that we are going to check them again. [Participant from G3]

The third theme emerged when we asked the participants' perspectives about how much a system such as our application could prevent data loss (subtheme 3.1). In theme 1, the participants mentioned that the lack of interoperability makes them rewrite essential data, and much information is lost in this process. During the interview, all participants (14/14, 100%) mentioned that using a centralized system would prevent data loss (B3):

...Prevent data loss? The central system on itself? Yes, absolutely. [Participant from G1]

... We can prevent this [data loss] when we all use one platform, and it is secure like a cloud [referring to our application]. [Participant from G2]

Moreover, we asked the participants who should be accountable for the data added to the EMR of the patient when a team treats the patient (subtheme 3.2). All of them (14/14, 100%) agreed that the person who added the data is accountable, but 71% (10/14) of the participants thought that the whole team should also be responsible and traceable for what happens to the patient, as proposed in the demonstrator (B4):

The accountability of the data is what makes the doctor remake the image exams. They do not trust that the image was made correctly in another hospital, so they need to double-check before deciding or giving a diagnostic and writing it down. [Participant from G4]

Every professional involved in the treatment should be accountable and traceable. [Participant from G1]

All the professionals who participate in the treatment should be accountable, but the professional who wrote the data must be responsible for it. [Participant from G2]

Furthermore, we asked how long access to patient data should still be provided after the treatment is over, for example, to input data that could not be added earlier owing to the urgency of the treatment or other responsibilities (subtheme 3.3). All participants (14/14, 100%) agreed that the data should be added as soon as possible to be useful to other teams involved in the acute care, but they also agreed that, sometimes, the extra time is fundamental to complete and edit all the forms. Of the 14 participants, 9 (64%) believed that a few hours are enough as extra time, whereas 5 (36%) considered a few days (B5):

At the end of our shift, my colleagues and I always go back to the reports. We write any information that we haven't added because of the hurry. So, I believe 24 hours is a good extra time, more than that is too much. [Participant from G3]

This is a difficult question. Because when I look into my practice, sometimes it happens that we arrive at the hospital, we deliver the patient. And then they call us again, and we have cardiac arrests around the corner, then we don't have time...Of course, it is not a standard procedure, but this happens quite often. So, I think if the team needs extra time, they should click the button saying that they need to keep the session open until the end of their shift and close it as soon as possible. [Participant from G2]

Because we make mistakes when we type the information, we should be able to fix them when we have time. But I think that access after the treatment is over must be logged as editing data. [Participant from G2]

We asked if the participants thought that a system such as our demonstrator would be useful in their daily tasks. All of them (14/14, 100%) responded that it would be useful, and 79% (11/14) said that it would be very useful (B6):

...The cloud solution itself will be very useful. All the [user] interfaces are not, but for the cloud solution, definitely yes. [Participant from G1]

Of the 14 participants, 5 (36%) highlighted that, sometimes, the patient cannot be rapidly identified to obtain the existing medical records in the system. They were very interested in the application's function that enables the system to store the data generated in the treatment using crypto scheme and later merge these data with the patient's EMR (subtheme 3.4):

...Sometimes when there is a tourist, for example, it takes some time to find their ID or passport or whatever. So then, it would be handy to be able to merge that [patient data] afterwards. [Participant from G2]

Challenges and Suggestions for the Adoption of the Application

The fourth theme emerged when we asked what would be needed to improve the usefulness of the system. The participants made various suggestions to enhance the usefulness and functionality of the application (B7).

The participants suggested that the application should include other types of care, such as regular physician appointments, which would require the admission of more types of users in the application and extend the access control model to cover their requests. At least, the system should be able to exchange data with other (existing) systems (subtheme 4.1):

I think one of the things that I missed is that you can push information to your base to the local EMR system. [Participant from G4]

The participants provided feedback regarding the granularity of access control to parts of the EMR (subtheme 4.1). Overall,

36% (5/14) of the participants suggested that the system should support splitting the patient's EMR into 2 parts—one part of data that is shared with the patient and another part of the data that is shared among the health care professionals. This 36% (5/14) of the participants believed that the patient should not read all the annotations that the health care professionals create. They mentioned that physicians write information about triage, which needs further investigation to remember what was done before the diagnosis. According to them, such information should only be shared among the health care professionals involved with the treatment. They affirmed that this type of information could create misunderstanding and unnecessary stress for the patients. In contrast, all participants (14/14, 100%) agreed that patients should be able to read about the diagnosis and procedures performed during treatment:

Nowadays, patients have access to part of the data. I add to the EMR only the diagnostics and measurements. I also add some notes to the patient. However, I have another place to add my comments as a doctor. For example, if I suspect that the patient has cancer, I do not add this in his report directly. First, I ask for exams, but I need to keep this note to remember the patient's case with more details. [Participant from G4]

Another 21% (3/14) of the participants said that patients should be able to read all the data about their treatment and they should be informed as much as possible:

So now [in the demonstrator], the patients can see anything I type. So now, I think I will sometimes be very careful. On the other hand, if you type it down, you can also say to the patient. If you can't say it to the patient, so maybe you shouldn't write it down. If you say, if you write down the patient is maybe faking it, you should also tell the patient that you think he is faking it. So yeah, I think anything I typed down is also something I would tell the patient. Yeah. I don't know if other doctors think otherwise. This is kind of a regulation thing. I believe. The patient has some will on this. [Participant from G3]

Of the 14 participants, 4 (29%) suggested that the application should include more data sharing opportunities for learning purposes (subtheme 4.2). These participants said that they are interested in performance measurement, such as aggregated metrics about the organizations. Others were interested to know more about what happens after they leave the patient under the care of other teams, mainly to learn whether their decision was correct:

...Can you get aggregated metrics, for example? Because this is what we need to report, some hospitals and departments, like the entry of emergency departments. Or, for instance, for ambulances, to report how fast they were for every patient with stroke because this is like a quality metric that we have to show to improve the quality of the service. [Participant from G3]

You're not a taxi when you transfer the patient in an ambulance. I believe that the professionals involved

in the treatment should see what happens with the patient even after their task is done because it is part of the learning process. [Participant from G2]

In the fifth and last theme, we analyzed the trust and acceptability of the application among the participants and the challenges regarding its adoption in a real scenario. All participants (14/14, 100%) said they would “much” and “very much” trust using the application in their daily tasks (B8). Overall, 64% (7/11) of the participants highlighted the need to train health care professionals to use a digital system such as the demonstrator (subtheme 5.1). Once the professionals understand how the system works and its security scheme, they will trust and be motivated to use it (B9):

...The point is that human errors happen pretty often because the professionals are not able to interact with the [current] system. When things go wrong in the hospital [system], that affects the patients negatively. Thus, the professionals must be trained to use the system correctly. [Participant from G3]

Of the 14 participants, 13 (93%) believed that their organization would adopt a system such as this application (B10). To improve the acceptance by health care organizations (B11), 57% (8/14) of the participants suggested that our application should include more types of users beyond the acute care teams and offer opportunities for data sharing among all of them (subtheme 5.2):

This system should be able to comprise other types of access, so we extend the security measures that you created for acute care to include the conventional and all the other types. [Participant from G4]

The feedback obtained from 71% (10/14) of the participants was to think about integrating the existing EMR systems with the ASCLEPIOS framework (subtheme 5.3). All of them (14/14, 100%) seemed to value the application, but they also reinforced that the acceptance of a new centralized national EMR system would be far-fetched. Therefore, the recommendation was to consider using the framework as an interoperability layer between the existing systems:

The organisation is very sceptical about new systems, so this can be a barrier to the organisation's acceptance. But if we prove that the system works properly and if it could be interoperable with the existing system, it would help the process. [Participant from G1]

...If you want all the acute care workers to work in the same system, that won't be easy. But if they would work in their systems and connect all those systems with web-based applications or anything else we did with this cloud solution that will be there, then there is a fair chance that it can work. [Participant from G2]

When they [acute care professionals] have to write down everything into [multiple] systems, it's too much. So they don't do it. I think the very important thing is that this system is the only one they need to work with. [Participant from G4]

Finally, all participants (14/14, 100%) answered that patients would feel safe about sharing their data (B12). However, 64% (9/14) of the participants said that most patients are not aware of the privacy risks related to EMR leakage. Therefore, 14% (2/14) of the participants suggested that health care organizations should be more transparent about the patient data processing and create awareness about privacy risks (subtheme 5.4):

I think most of the patients are not thinking at this level. Most of the patients are not thinking about their privacy risks or if their data is available in case of an emergency. They usually think about it after something happens. [Participant from G1]

It depends on the medical records of the patient. If he [patient] is applying for a job, but he had a heart problem once, maybe he will be concerned about what the company would say if they illegally already know. [Participant from G2]

Discussion

Principal Findings

The main objective of this study was to collect the current challenges for patient data sharing during acute stroke care (RQ1), the participants' impressions of the proposed ASCLEPIOS Acute Stroke Care application (RQ2), and the challenges and suggestions for adapting the ASCLEPIOS Acute Stroke Care application in a real-life scenario (RQ3). Although our study was designed in the context of a specific European Union project, the challenges of developing an EMR system that supports acute care and the collected feedback about cloud-based systems are applicable in a broad context.

From the results for RQ1, this study reinforced that the most relevant challenges for patient data sharing are the lack of interoperability and connectivity between systems from different organizations. For RQ2, this study obtained relevant feedback from every interviewee regarding the time interval for data availability, accountability, prevention of data loss, and handling of unknown patients during acute care. For RQ3, this study identified several important barriers to and improvement opportunities for the future acceptance and adoption of the proposed system.

Furthermore, this study aimed to validate the security concepts of a cloud-based medical data sharing application for acute stroke care that exploits the ASCLEPIOS framework. During the interviews with health care professionals, it became evident that they experience—daily—the lack of a properly connected and secure information infrastructure for patient data exchange across organizations. The application was well received and considered to be relevant by all participants (14/14, 100%). However, as a large number of noninteroperating systems are used in practice, replacing them with a new system—such as the developed application—did not seem realistic. An alternative path to be explored involves developing an interoperation layer for cloud-based security and trusted data exchange that could bridge legacy systems with the newly developed technology.

Another interesting finding is that the participants were excited to provide feedback when we said that we would demonstrate

the usefulness of our project in a simulation to support acute stroke care. We simulated the workflow, emphasizing that the professionals from each team could access the patient EMR only from the moment when they were invited to participate in the treatment until their tasks were completed. Thus, they could see the added value that the proposed solution could bring to facilitate data sharing among all the professionals involved. Furthermore, the received feedback validates the access control model implemented in the application.

Finally, we highlight 3 suggestions that the participants provided to increase the usefulness of the system and regarding what we could achieve using the ASCLEPIOS framework. The first suggestion was to expand the system to support all types of access to EMRs. The second suggestion was to create more granularity of access control for different types of data contained in the EMR, which would require separating the data that are sharable with the patient from those that are shared only among the health care professionals. The third suggestion was about consulting aggregated metrics from all the EMRs stored for learning purposes. All these suggestions provide valuable feedback that will be explored in future studies.

Limitations

A limitation of the study is that demonstrating the use of application interfaces can be a double-edged sword. In addition to seeing how the system would work and understanding the solution behind the screen better, the participants may also be distracted by the interfaces presented during the simulation. We anticipated this effect, and thus, we stimulated participants to provide feedback beyond the user interface. Nevertheless, we still received suggestions about interface content and design modifications, which were not relevant to this study's RQs, but they could be useful in a future application design.

Moreover, we acknowledge that collecting the perspectives of hospital administrators and technical staff is essential for accepting the new health care system. Therefore, in the future, we will design a study to collect their perceptions and feedback from management and technical perspectives.

Another limitation was related to the COVID-19 pandemic. To perform in-depth interviews, we preferred to have in-person meetings and let the participants interact with the application. However, acute care professionals are very busy, and even more so because of the pandemic; thus, it was even harder than anticipated to involve the professionals in person. Moreover, there were multiple lockdowns during the study; therefore, we had to use web-based meetings to prevent the cancellation of the already confirmed interviews. Regarding these web-based interviews, we realized that, unfortunately, the communication and interaction were limited because they could not directly visualize the application being used. Besides this limitation, the 36% (5/14) of the participants provided valuable feedback during the web-based meetings.

Comparison With Previous Studies

Researchers have successfully adopted similar sociotechnical qualitative interviews to collect stakeholders' perceptions and validate the concept of innovative technological solutions for health care.

Murry et al [37] interviewed senior managers and medical staff to explore and understand their experiences of implementing eHealth initiatives and their assessment of factors that promote the integration of eHealth initiatives. In total, 23 interviews were conducted, and they showed substantial differences in the implementation of eHealth initiatives [37]. It differed from our study because their focus was not on health professionals' perspectives. Instead, the authors interviewed the implementers, who are the staff responsible for implementing digital eHealth systems, which, according to the authors, is an under-studied group. Moreover, the implementers showed rich understanding of the barriers to and facilitators of successfully implementing such initiatives.

Georgiou et al [38] also conducted a qualitative interview study to assess the impact of introducing new health technical initiatives for medical imaging processing. They used a mixed methods study design comprising semistructured interviews with medical imaging department staff and retrospectively extracted emergency data. In the study by Georgiou et al [38], the results show that the accessibility of images and patient-related information improve the efficiency of the medical imaging department. In our study, the professionals also agreed about the potential improvement in efficiency by having the data available from other teams. Moreover, similar to the study by Georgiou et al [38], in subtheme 2.2, professionals raised concerns about the quality of the data, especially the reliability of the image data provided by other teams in acute care.

Similar to our results, the studies by Murray et al [37] and Georgiou et al [38] affirm that for the successful implementation of an eHealth system, it should be a good fit between the new technology and existing skill sets or efforts made to teach the requisite skills to users. Similarly, in our study, professionals recommended integrating the new application with other (existing) systems (subtheme 4.1) and merging current systems instead of proposing a new one (subtheme 5.3).

Azode et al [39] conducted a qualitative interview study to investigate the opportunities for and challenges of using data from wearable sensor devices in health care. In total, 16 health care, technology, business, innovation, and social sciences experts were interviewed in a qualitative, theoretically informed study. The authors concluded that current applications cannot fulfill their potential if they do not yield benefits for clinical users and integrate effectively with the existing eHealth systems. In our study, health care professionals were interested in expanding our system's application to include all types of EMR data, which could also include data from wearables.

Hasselgren et al [40] interviewed medical students and analyzed their perceptions of a blockchain-based decentralized work for maintaining professional history and credentials portfolio. The study used a qualitative approach applied with data collection through 9 semistructured interviews. The results showed that health care professionals are interested in a decentralized system in which they can control their credentials and reputation.

Brandt et al [41] interviewed patients who are overweight to identify important drivers of long-term personal lifestyle changes from a patient perspective when using a collaborative eHealth tool. Interviews were conducted 5 years after the initial

intervention and showed that all the patients still used other internet apps to benefit their health despite not having access to the eHealth tool used during the intervention.

Although the objectives of the applications used by Hasselgren et al [40] and Brandt et al [41] differ from EMR data sharing, our application has a common goal—to increase the trust on eHealth systems among patients. For this aim, our application presents to the patient a consolidated logs dashboard about how the patient data were processed by health care professionals. In the study by Hasselgren et al [40] and our study, health care professionals are not sure how aware the patients are about the digital systems and how effective these functionalities of health care transparency are, but in the study by Brandt et al [41], patients show trust and value in the use of the proposed eHealth app. This reinforces subtheme 5.4, which recommends increasing patient trust on and awareness about digital health systems and applications.

Woodward et al [42] explored the personal experiences of health care professionals using eHealth innovations for data sharing in selected postconflict situations. This study used a cross-sectional qualitative design, with 12 telephone interviews. The authors concluded that all interviewees held positive perceptions that the eHealth system can help them to access information and communicate with other health workers. However, understanding of the scope of eHealth was generally limited and often based on innovations that health workers have been introduced to by their international partners. In our study, health care professionals also raised concerns about the need for training to use eHealth applications. In the study by Woodward et al [42] and our study, the results show the importance of training so that professionals can accept and benefit from the eHealth innovation system.

Inspired by previous studies [37-42], we used similar methods and acknowledged the importance of gaining stakeholders' input for eHealth technology development, for further improvement and acceptability of new technologies.

In our previous study [29], we collected and analyzed the perspectives of medical staff regarding health care and data privacy requirements for the eHealth cloud, using a qualitative interview. At that time, we collected requirements that would guide the design of the demonstrator. Moreover, we investigated the participants' understanding of cloud services and how they envision using the ASCLEPIOS solution in their daily tasks. At that point, we did not have the Acute Stroke Care application ready to present to the clinicians.

In this study, besides validating the requirements discussed in the previous publication [29], showing the participants a working application allowed them to go deep into the matter and ask questions related to the actual usefulness and acceptance of the ASCLEPIOS solution for cross-organization acute stroke care data sharing.

Conclusions

This study validated the need for a cross-organization data sharing solution that offers the security and privacy required when patient data are processed. The participants emphasized that our cloud-based application would solve the data sharing problems, such as duplication of data, lack of information, and standardization. However, it would not be realistic to propose that all the organizations involved in acute care migrate to a unique cloud-based application. Future studies should investigate opportunities to update the system according to these inputs and further explore the ASCLEPIOS framework as a secure and interoperable layer for patient data sharing. The concept validation and feedback presented in this study incite the desire for a digital transformation in health care systems.

Acknowledgments

The authors would like to thank the participants for their willingness to participate in the interviews. This study was funded by the Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care project from the European Union's Horizon 2020 research and innovation program (grant agreement number 826093).

Conflicts of Interest

HM is a cofounder and shareholder of Nicolab and TrianecT. The authors have no further interests to declare.

Multimedia Appendix 1

Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care demonstrator for acute stroke care by the Amsterdam University Medical Center. Here, we show how various health care professionals share information about a patient who has experienced a stroke. The information is securely stored in the cloud and becomes available during acute care for the professionals in the emergency call center, ambulance service, and hospital. The fast exchange of information during acute stroke care is essential for making decisions that can have a huge impact on the correct treatment and patient recovery.

[[MP4 File \(MP4 Video\), 6668 KB - formative_v6i12e40061_app1.mp4](#)]

Multimedia Appendix 2

Presentation used during the interviews. The videos illustrate the use of the application, similar to the simulations with the participants.

[[PPTX File , 86040 KB - formative_v6i12e40061_app2.pptx](#)]

References

1. Vital M. Stroke Hope Through Research. Darby, PA, U.S.A: Diane Publishing Company; 1999.
2. Saver J. Time is brain—quantified. *Stroke* 2006 Jan;37(1):263-266. [doi: [10.1161/01.STR.0000196957.55928.ab](https://doi.org/10.1161/01.STR.0000196957.55928.ab)]
3. Lin MP. Time matters greatly in acute stroke care. *Neurol Neurochir Pol* 2020 Apr 30;54(2):104-105. [doi: [10.5603/pjnns.2020.0037](https://doi.org/10.5603/pjnns.2020.0037)]
4. Oostema JA, Nasiri M, Chassee T, Reeves MJ. The quality of prehospital ischemic stroke care: compliance with guidelines and impact on in-hospital stroke response. *J Stroke Cerebrovasc Dis* 2014 Nov;23(10):2773-2779. [doi: [10.1016/j.jstrokecerebrovasdis.2014.06.030](https://doi.org/10.1016/j.jstrokecerebrovasdis.2014.06.030)] [Medline: [25312034](https://pubmed.ncbi.nlm.nih.gov/25312034/)]
5. 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 Sep;24(5):1103-1117. [doi: [10.1377/hlthaff.24.5.1103](https://doi.org/10.1377/hlthaff.24.5.1103)] [Medline: [16162551](https://pubmed.ncbi.nlm.nih.gov/16162551/)]
6. Pierantoni G, Kiss T, Terstysanzky G, Dang H, Delgado OS, de OM, et al. A Secure Cloud-based Platform to Host Healthcare Applications. In: Proceedings of the 11th International Workshop on Science Gateways, IWSG 2019. 2019 Presented at: 11th International Workshop on Science Gateways, IWSG 2019; Jun 12-14, 2019; Ljubljana, Slovenia.
7. Ren K, Wang C, Wang Q. Security challenges for the public cloud. *IEEE Internet Comput* 2012 Jan;16(1):69-73. [doi: [10.1109/mic.2012.14](https://doi.org/10.1109/mic.2012.14)]
8. Yigzaw K, Olabarriaga S, Michalas A, Marco-Ruiz L, Hillen C, Verginadis Y, et al. Health data security and privacy: challenges and solutions for the future. In: Roadmap to Successful Digital Health Ecosystems. Cambridge, Massachusetts, United States: Academic Press; 2022.
9. General data protection regulation. Intersoft Consulting. URL: <https://gdpr-info.eu/> [accessed 2021-05-22]
10. Abbas A, Khan SU. A review on the state-of-the-art privacy-preserving approaches in the e-health clouds. *IEEE J Biomed Health Inform* 2014 Jul;18(4):1431-1441. [doi: [10.1109/jbhi.2014.2300846](https://doi.org/10.1109/jbhi.2014.2300846)]
11. Munich SA, Tan LA, Nogueira DM, Keigher KM, Chen M, Crowley RW, et al. Mobile real-time tracking of acute stroke patients and instant, secure inter-team communication - the join app. *Neurointervention* 2017 Sep;12(2):69-76 [FREE Full text] [doi: [10.5469/neuroint.2017.12.2.69](https://doi.org/10.5469/neuroint.2017.12.2.69)] [Medline: [28955508](https://pubmed.ncbi.nlm.nih.gov/28955508/)]
12. Nam HS, Heo J, Kim J, Kim YD, Song TJ, Park E, et al. Development of smartphone application that aids stroke screening and identifying nearby acute stroke care hospitals. *Yonsei Med J* 2014 Jan;55(1):25-29 [FREE Full text] [doi: [10.3349/ymj.2014.55.1.25](https://doi.org/10.3349/ymj.2014.55.1.25)] [Medline: [24339283](https://pubmed.ncbi.nlm.nih.gov/24339283/)]
13. Rahmani H, Sundararajan E, Ali ZM, Zin AM. Encryption as a service (EaaS) as a solution for cryptography in cloud. *Procedia Technol* 2013;11:1202-1210. [doi: [10.1016/j.protcy.2013.12.314](https://doi.org/10.1016/j.protcy.2013.12.314)]
14. Akinyele J, Pagano M, Green M, Lehmann C, Peterson Z, Rubin A. Securing electronic medical records using attribute-based encryption on mobile devices. In: Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices. 2011 Oct Presented at: CCS'11: the ACM Conference on Computer and Communications Security; Oct 17, 2011; Chicago Illinois USA. [doi: [10.1145/2046614.2046628](https://doi.org/10.1145/2046614.2046628)]
15. Seol K, Kim Y, Lee E, Seo Y, Baik D. Privacy-preserving attribute-based access control model for XML-based electronic health record system. *IEEE Access* 2018;6:9114-9128. [doi: [10.1109/access.2018.2800288](https://doi.org/10.1109/access.2018.2800288)]
16. Brouwers C, Merten H, Willems M, Habraken DJ, Bloemers FW, Biesheuvel TH, et al. Improving care for older patients in the acute setting: a qualitative study with healthcare providers. *Neth J Med* 2017 Oct;75(8):335-343 [FREE Full text] [Medline: [29219828](https://pubmed.ncbi.nlm.nih.gov/29219828/)]
17. Marinovic S, Craven R, Ma J, Dulay N. Rumpole: a flexible break-glass access control model. In: Proceedings of the 16th ACM symposium on Access control models and technologies. 2011 Presented at: SACMAT '11: 16th ACM Symposium on Access Control Models and Technologies; Jun 15 - 17, 2011; Innsbruck Austria. [doi: [10.1145/1998441.1998453](https://doi.org/10.1145/1998441.1998453)]
18. Li M, Yu S, Ren K, Lou W. Securing personal health records in cloud computing: patient-centric and fine-grained data access control in multi-owner settings. In: Security and Privacy in Communication Networks. Berlin, Heidelberg: Springer; 2010.
19. Brucker A, Petritsch H, Weber SG. Attribute-based encryption with break-glass. In: Information Security Theory and Practices. Security and Privacy of Pervasive Systems and Smart Devices. Berlin, Heidelberg: Springer; 2010.
20. Yang Y, Zheng X, Guo W, Liu X, Chang V. Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system. *Inform Sci* 2019 Apr;479:567-592. [doi: [10.1016/j.ins.2018.02.005](https://doi.org/10.1016/j.ins.2018.02.005)]
21. Advanced secure cloud encrypted platform for internationally orchestrated solutions in healthcare. Asclepios Project. URL: <https://www.asclepios-project.eu/> [accessed 2021-01-20]
22. Michalas A, Bakas A, Dang H, Zalizko A. MicroSCOPE: enabling access control in searchable encryption with the use of attribute-based encryption and SGX. In: Secure IT Systems. Cham: Springer; Nov 13, 2019.
23. Bethencourt J, Sahai A, Waters B. Ciphertext-policy attribute-based encryption. *IEEE Symposium Security Privacy* 2007:321-334.
24. Psarra E, Patiniotakis I, Verginadis Y, Apostolou DG, Mentzas G. Securing access to healthcare data with context-aware policies. In: Proceedings of the 11th International Conference on Information, Intelligence, Systems and Applications (IISA). 2020 Presented at: 11th International Conference on Information, Intelligence, Systems and Applications (IISA); Jul 15-17, 2020; Piraeus, Greece. [doi: [10.1109/iisa50023.2020.9284393](https://doi.org/10.1109/iisa50023.2020.9284393)]

25. de Oliveira MT, Bakas A, Frimpong E, Groot AE, Marquering HA, Michalas A, et al. A break-glass protocol based on ciphertext-policy attribute-based encryption to access medical records in the cloud. *Ann Telecommun* 2020 Mar 07;75(3-4):103-119. [doi: [10.1007/s12243-020-00759-2](https://doi.org/10.1007/s12243-020-00759-2)]
26. de Oliveira MT, Dang H, A. Reis LH, Marquering HA, D. Olabarriaga S. AC-AC: dynamic revocable access control for acute care teams to access medical records. *Smart Health* 2021 Apr;20:100190. [doi: [10.1016/j.smhl.2021.100190](https://doi.org/10.1016/j.smhl.2021.100190)]
27. Welcome to FHIR. HL7 FHIR. URL: <https://www.hl7.org/fhir/> [accessed 2022-03-31]
28. Chomutare T, Yigzaw K, Olabarriaga S, Makhlysheva A, Oliveira M, Silsand L, et al. Healthcare and data privacy requirements for e-health cloud: a qualitative analysis of clinician perspectives. In: *Proceedings of the IEEE International Conference on E-health Networking, Application & Services, HealthCom 2020*. 2021 Presented at: IEEE International Conference on E-health Networking, Application & Services, HealthCom 2020; Mar 1-2, 2021; Shenzhen, China. [doi: [10.1109/HEALTHCOM49281.2021.9399006](https://doi.org/10.1109/HEALTHCOM49281.2021.9399006)]
29. Reis L, de Oliveira MT, Mattos D, Olabarriaga S. Private data sharing in a secure cloud-based application for acute stroke care. In: *Proceedings of the IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. 2021 Presented at: IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS); Jun 7-9, 2021; Aveiro, Portugal. [doi: [10.1109/cbms52027.2021.00039](https://doi.org/10.1109/cbms52027.2021.00039)]
30. Choi YM. Utilizing end user input in early product development. *Procedia Manuf* 2015;3:2244-2250. [doi: [10.1016/j.promfg.2015.07.368](https://doi.org/10.1016/j.promfg.2015.07.368)]
31. Hennink M, Kaiser B. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. *Soc Sci Med* 2022 Jan;292:114523 [FREE Full text] [doi: [10.1016/j.socscimed.2021.114523](https://doi.org/10.1016/j.socscimed.2021.114523)] [Medline: [34785096](https://pubmed.ncbi.nlm.nih.gov/34785096/)]
32. Get insights quickly, with Google forms. Google Forms. URL: <https://www.google.com/forms/about/> [accessed 2022-04-12]
33. Capture and share insights from your meetings. Otter. URL: <https://otter.ai/> [accessed 2022-03-31]
34. Azevedo V, Carvalho M, Costa F, Mesquita S, Soares J, Teixeira F, et al. Interview transcription: conceptual issues, practical guidelines, and challenges. *Rev Enf Ref* 2017 May 22;IV Série(Nº14).
35. Malterud K. Systematic text condensation: a strategy for qualitative analysis. *Scand J Public Health* 2012 Dec 04;40(8):795-805. [doi: [10.1177/1403494812465030](https://doi.org/10.1177/1403494812465030)] [Medline: [23221918](https://pubmed.ncbi.nlm.nih.gov/23221918/)]
36. MAXQDA homepage. MAXQDA. URL: <https://www.maxqda.com> [accessed 2022-03-31]
37. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. *Implement Sci* 2011 Jan 19;6(1):6 [FREE Full text] [doi: [10.1186/1748-5908-6-6](https://doi.org/10.1186/1748-5908-6-6)] [Medline: [21244714](https://pubmed.ncbi.nlm.nih.gov/21244714/)]
38. Georgiou A, Prgomet M, Lymer S, Hordern A, Ridley L, Westbrook J. The impact of a health IT changeover on medical imaging department work processes and turnaround times. *Appl Clin Inform* 2017 Dec 19;06(03):443-453. [doi: [10.4338/aci-2015-01-ra-0014](https://doi.org/10.4338/aci-2015-01-ra-0014)]
39. Azodo I, Williams R, Sheikh A, Cresswell K. Opportunities and challenges surrounding the use of data from wearable sensor devices in health care: qualitative interview study. *J Med Internet Res* 2020 Oct 22;22(10):e19542 [FREE Full text] [doi: [10.2196/19542](https://doi.org/10.2196/19542)] [Medline: [33090107](https://pubmed.ncbi.nlm.nih.gov/33090107/)]
40. Hasselgren A, Kravlevska K, Gligoroski D, Faxvaag A. Medical students' perceptions of a blockchain-based decentralized work history and credentials portfolio: qualitative feasibility study. *JMIR Form Res* 2021 Oct 22;5(10):e33113 [FREE Full text] [doi: [10.2196/33113](https://doi.org/10.2196/33113)] [Medline: [34677137](https://pubmed.ncbi.nlm.nih.gov/34677137/)]
41. Brandt CJ, Clemensen J, Nielsen JB, Søndergaard J. Drivers for successful long-term lifestyle change, the role of e-health: a qualitative interview study. *BMJ Open* 2018 Mar 12;8(3):e017466 [FREE Full text] [doi: [10.1136/bmjopen-2017-017466](https://doi.org/10.1136/bmjopen-2017-017466)] [Medline: [29530904](https://pubmed.ncbi.nlm.nih.gov/29530904/)]
42. Woodward A, Fyfe M, Handuleh J, Patel P, Godman B, Leather A, et al. Diffusion of e-health innovations in 'post-conflict' settings: a qualitative study on the personal experiences of health workers. *Hum Resour Health* 2014 Apr 23;12(1):22 [FREE Full text] [doi: [10.1186/1478-4491-12-22](https://doi.org/10.1186/1478-4491-12-22)] [Medline: [24754997](https://pubmed.ncbi.nlm.nih.gov/24754997/)]

Abbreviations

ASCLEPIOS: Advanced Secure Cloud Encrypted Platform for Internationally Orchestrated Solutions in Health Care

DSSE: dynamic index-based symmetric searchable encryption

EMR: electronic medical record

G1: group 1

G2: group 2

G3: group 3

G4: group 4

GDPR: General Data Protection Regulation

RQ: research question

Edited by A Mavragani; submitted 03.06.22; peer-reviewed by Y Lu, AS Alaqra; comments to author 23.08.22; revised version received 13.09.22; accepted 26.09.22; published 23.12.22.

Please cite as:

*Tuler de Oliveira M, Amorim Reis LH, Marquering H, Zwinderman AH, Delgado Olabarriaga S
Perceptions of a Secure Cloud-Based Solution for Data Sharing During Acute Stroke Care: Qualitative Interview Study
JMIR Form Res 2022;6(12):e40061*

URL: <https://formative.jmir.org/2022/12/e40061>

doi: [10.2196/40061](https://doi.org/10.2196/40061)

PMID: [36563043](https://pubmed.ncbi.nlm.nih.gov/36563043/)

©Marcela Tuler de Oliveira, Lúcio Henrik Amorim Reis, Henk Marquering, Aeilko H Zwinderman, Sílvia Delgado Olabarriaga. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

State-Level COVID-19 Symptom Searches and Case Data: Quantitative Analysis of Political Affiliation as a Predictor for Lag Time Using Google Trends and Centers for Disease Control and Prevention Data

Alex Turvy¹, BA, MS

City, Culture, and Community, Department of Sociology, Tulane University, New Orleans, LA, United States

Corresponding Author:

Alex Turvy, BA, MS

City, Culture, and Community

Department of Sociology

Tulane University

6823 St Charles Ave

New Orleans, LA, 70118

United States

Phone: 1 504 865 5231

Email: aturvy@tulane.edu

Abstract

Background: Across each state, the emergence of the COVID-19 pandemic in the United States was marked by policies and rhetoric that often corresponded to the political party in power. These diverging responses have sparked broad ongoing discussion about how the political leadership of a state may affect not only the COVID-19 case numbers in a given state but also the subjective individual experience of the pandemic.

Objective: This study leverages state-level data from Google Search Trends and Centers for Disease Control and Prevention (CDC) daily case data to investigate the temporal relationship between increases in relative search volume for COVID-19 symptoms and corresponding increases in case data. I aimed to identify whether there are state-level differences in patterns of lag time across each of the 4 spikes in the data (RQ1) and whether the political climate in a given state is associated with these differences (RQ2).

Methods: Using publicly available data from Google Trends and the CDC, linear mixed modeling was utilized to account for random state-level intercepts. Lag time was operationalized as number of days between a peak (a sustained increase before a sustained decline) in symptom search data and a corresponding spike in case data and was calculated manually for each of the 4 spikes in individual states. Google offers a data set that tracks the relative search incidence of more than 400 potential COVID-19 symptoms, which is normalized on a 0-100 scale. I used the CDC's definition of the 11 most common COVID-19 symptoms and created a single construct variable that operationalizes symptom searches. To measure political climate, I considered the proportion of 2020 Trump popular votes in a state as well as a dummy variable for the political party that controls the governorship and a continuous variable measuring proportional party control of federal Congressional representatives.

Results: The strongest overall fit was for a linear mixed model that included proportion of 2020 Trump votes as the predictive variable of interest and included controls for mean daily cases and deaths as well as population. Additional political climate variables were discarded for lack of model fit. Findings indicated evidence that there are statistically significant differences in lag time by state but that no individual variable measuring political climate was a statistically significant predictor of these differences.

Conclusions: Given that there will likely be future pandemics within this political climate, it is important to understand how political leadership affects perceptions of and corresponding responses to public health crises. Although this study did not fully model this relationship, I believe that future research can build on the state-level differences that I identified by approaching the analysis with a different theoretical model, method for calculating lag time, or level of geographic modeling.

(*JMIR Form Res* 2022;6(12):e40825) doi:[10.2196/40825](https://doi.org/10.2196/40825)

KEYWORDS

COVID-19; search trends; prediction; case; political; symptom; pandemic; data; google; disease; prevention; model

Introduction

Background

Incidence of COVID-19 cases in the United States has widely varied across states and across time, as have state-level policies and some of the rhetoric heard in response. There has been ongoing investigation about how mitigation measure mandates such as mask wearing [1,2], social distancing [3], and vaccines [4] affect uptake of these measures as well as how they are associated with actual case numbers. One existing study focused on the political dimensions of mandates and cases [5] by trying to understand the broader social forces that are associated with the response to the pandemic and to mandates, but it is challenging to observe and understand the informal behaviors and tacit unreported beliefs that drive state-level differences in case numbers and response to the pandemic. There has been significant ongoing debate regarding theories for why individual behavioral responses to the pandemic have diverged so significantly over time, and many of these theories have involved analysis of political and administrative messaging. This research enters this conversation by exploring the intersection of political affiliation and perceptions of the pandemic. Increases in search traffic seem to reflect individual concerns about the pandemic as well as information seeking for individuals who are experiencing and observing symptoms. This study aimed to specifically analyze the effects that political affiliation has on this dynamic; there has been plenty of concern about the danger of public health becoming more politicized, and this research intended to add granularity to our understanding.

Both anecdotally and in popular media, there is discussion about how trends in COVID-19-related Google searches might be associated with ongoing COVID-19 case numbers as reported to the Centers for Disease Control and Prevention (CDC) [6]. Using Google Trends search data about COVID-19-related symptoms along with CDC data concerning state-level case numbers, I investigated 2 questions:

1. RQ1: Are there state-level differences in the lag time between spikes in searches for COVID-19 symptoms and spikes in reported COVID-19 cases?
2. RQ2: If these state-level differences do exist, do covariates related to political leadership contribute to state-level variance in lag time?

I hypothesized that state-level political outcomes, as a marker of the dominant or collective political identification of a state's voters, offer a route to investigating how social behavior via self-identified group affiliation explains differences in the temporal relationship between spikes in COVID-19-related searches and later spikes in total confirmed COVID-19 cases. I expected to find that political variables marking Republican identification are associated with a decrease in lag time, given what we know about the differences in compliance with mitigation measures and vaccine uptake and what this suggests about broader COVID-19 risk beliefs and self-surveillance of

symptoms. This lag time relationship may give us insight into how people think about COVID-19: Are they proactive about watching for and managing symptoms, or do they only start noticing symptoms once cases begin to increase and spike?

Literature Review

Theoretical Approach

Social cognitive theory (SCT) [7] frames learning and behavior socially, noting that there is a reciprocal relationship between an individual, their environment, and their behavior—while emphasizing the specifically social nature of this triad. That is, people tend to learn through observing the actions of those in their environment along with their own experiences. The essential lens to understand SCT in this context is how it focuses specifically on personal but environmentally contextualized agency. The social identity approach (SIA) by Abrams and Hogg [8] is a complementary perspective, adding that not only is learning and behavior social but also that knowledge of being in social groups affects how people attach emotion and value to certain behaviors and circumstances. They also emphasize the influence of one's own in-groups and out-groups as part of this individual/group relationship.

In terms of compliance with health behaviors, the research has settled around 3 major factors that tend to drive an individual's level of compliance: perceived risk to oneself, belief in behavior effectiveness, and observed risk to others. During the H1N1 influenza pandemic, a review of 26 studies found consistently strong associations between an individual's perceived susceptibility to the virus and increased compliance with recommended behaviors; this effect was strengthened when perceived severity of infection increased and was consistent across many countries and cultures [9]. Perceived risk to self is affected by factors such as perceived personal vulnerability [10], level of cultural individualism [11], fear [12,13], and anger [14].

Belief in the effectiveness of health behaviors is driven by a number of affective and epistemological factors as well; laypeople tend to create their own justifications for health behaviors by pulling from not only both establishment and nonestablishment sources [15] but also their preferred mass media sources [16]. Although this belief is reduced in all groups when they perceive recommendations to predominantly be moralistic [17], general trust in government is strongly associated with affecting perceived risk in complicated and sometimes counterintuitive ways [10,17,18]. Finally, observed risk to specific others and a more generalized community tends to be positively affected by a general sense of conscientiousness [12]. This is true when individuals feel an ethical responsibility to their community [19] but is also true on a more individual level when individuals experience the vulnerability of their close ties [20] and so avoid the perception that health concerns are "overhyped" [21].

In this paper, I do not explore the details of how particular political ideologies specifically affect compliance behaviors.

Instead, the context described in the previous paragraphs serves to highlight the power that political identities and communication have on individual compliance. From both Democratic and Republican politicians, we frequently hear justifications for COVID-19 behaviors that speak directly to these 3 factors. These justifications are often oriented around personal ethical responsibility, the meaning of a sense of community, and the epistemological justifications for agency recommendations; each of these threads is an important part of building the set of beliefs that ultimately drive behaviors. More recent research extends these claims further, demonstrating that political affiliation is associated with particular pandemic responses that cannot be explained solely via these 3 factors [5], highlighting how at least part of an individual's response is related to affinity-related dogma or that operational constraints and diverging priorities tend to be privileged over more objective risk assessments [22].

Using Search Data in Public Health Research

Public health professionals and researchers have broadly been discussing the value of search data for both detection and surveillance for quite some time, initially highlighting its value in a landmark paper that advocated for its use but cautioned that it should be predominantly used in areas with widespread internet access [23]. Given that American internet usage is now frequent and widespread across most settings, researchers have been able to turn their attention specifically to its use in early detection of emergent diseases [24].

There is also evidence that search volume is effective for ongoing monitoring and surveillance, both for active and predictive surveillance [25], and for passive or retrospective surveillance that aims to understand how factors such as the media affect the relationship between search interest and cases [26].

Search data have been used as a lens specifically for understanding COVID-19 data, but this research has had different areas of focus: searches as predictive of local metropolitan-level data [27,28], impacts on mental health [29], and more rare symptoms (anosmia and ageusia) as ineffective predictors of case incidence [30]. Eysenbach [31] took a somewhat similar approach to my own but concerning flu symptoms and incidence instead of COVID-19, finding strong correlations between clicks on sponsored flu-related links and flu diagnoses 7 days later.

There are ongoing challenges to using search data as well as other novel data streams (NDS) such as social media posts; although there is some evidence that they can help to retroactively explore associations and wield predictive power, there are still unresolved issues of how to assess reliability and validity of these data [32]. Some challenges such as lack of transparency and reproducibility [33] can be resolved by establishing accepted best practices such as sharing specific search strings and Boolean operators, but others are more related to complex sociological and psychological phenomena such as a panic-induced search increase that will likely prove to be much more difficult to solve [34]. Given the established value and unresolved challenges of using NDS such as search traffic, the use of Google Trends data should be seen as a supplemental

tool for public health researchers along with more traditional and localized practices instead of as a substitute [35].

This paper addressed these methodological limitations by using search data indirectly; although this means that some of these concerns become endogenous to the modeling, it is beneficial insofar as this captures these complex dynamics within the lag time variable. Instead of relying on search trends data as an accurate predictive or surveillance tool, I used it to highlight areas of difference across regions and explore the reasons for those differences.

Methods

Although there is discussion in popular media about using trends data in a predictive way and some push toward this in technical methods literature [36], there is not yet evidence that search data are defensible for use in a predictive way about future health trends [30] given that access to real-time or otherwise timely raw data is not possible. Instead, trends data are most useful for monitoring and evaluating relationships between events in the past, especially as one predictive element within a larger model [37,38].

For the purposes of this study, I used the publicly available COVID-19 Search Trends data set, which tracks the incidence of more than 400 symptoms associated at various levels with COVID-19. Typically, Google does not allow for large-scale downloads of granular daily search trends data except through use of their API. However, the company made this COVID-19-specific search data available specifically for researchers and journalists; the data include both daily data as well as state-level geographic data. The data set allows for what Google calls "metro areas," but these do not include shapefiles that could be used to match the search data with CDC data via geographic information system (GIS) software. Like all Google Search Trends data, it is normalized on a 0-100 scale, contextualized within the geography and time range in question, and based on a particular search string's incidence in proportion to all searches in that same geography and time frame. This study used daily trends data from each state for the time period from March 11, 2020, through April 4, 2022, and was retrieved on April 15, 2022, via Google's internally hosted GitHub.

The beginning of the study period is the day on which COVID-19 was declared a global pandemic by the World Health Organization, and the end of the period is the final day of trends data from Google's data set; so, the total number of days in the study period is 762. Although not every region experienced their first case by the beginning of the trends data time period, I was specifically interested in the lag time between increases in searches and increases in cases. Given that there was already widespread discussion of COVID-19 in popular media and so this is broadly reflected in the search data, search volume was already increasing across all regions by March 11, 2020, and this allowed me to examine state-level differences in when cases began to increase.

The CDC makes daily data detailing new cases, new hospitalizations, and new deaths associated with COVID-19 available to the public. These data are available to the county

level, but this study exclusively used state-level data in order to look at these variables alongside the state-level search trends data. These data are raw numbers, but population is controlled for in regression modeling to account for this. There are likely data gaps related to submission logistics and other issues; additionally, the data had some level of daily noise within the data as states catch up with missed submissions and correct submissions from previous days, but this has been accounted for in the process of examining the data closely in the calculations for lag time as the dependent variable.

The analytic strategy used a linear mixed model with fixed effects for all included predictors and controls and random intercepts for each state to investigate the state-level differences named in the research questions. Additionally, I included a random effect for the political predictor nested within state clusters, recognizing how SCT indicates that behavior is affected in an ongoing reciprocal way by environment—here, the state and its political climate are considered as that environment.

I called the key outcome variable “lag time,” and it measures the amount of time in days between a spike in COVID-19 symptom searches and a (typically) later corresponding spike in reported COVID-19 cases. This variable was calculated manually using the raw data and plots as a guide. Each state had 4 identifiable case peaks of varying magnitudes. After marking these and accounting for any noise or reporting gaps in the data, I turned to the search data to identify whether there was a corresponding spike in symptom searches that preceded the case spike. In nearly all cases, there was an associated spike, and this was measured in number of days.

Political variables under consideration included the proportion of Trump popular votes within a state in the 2020 election, a dummy variable indicating whether a Republican holds the Governor office, and a variable measuring the proportion of a state’s federal representatives in the House of Representatives and Senate that is Republican. The latter 2 variables did not lead to a strong model fit in any case and were discarded, so the Trump proportion variable remained as the main predictor in this model. Controls for mean daily cases, mean daily deaths, and state population were also included in the model. All predictors and controls were normalized using z scores.

Within the search data, I identified the 11 most common symptoms of COVID-19 as reported by the CDC and created

a construct to represent the collective incidence of these search terms. These symptoms included headache, nasal congestion, rhinorrhea, fever, sore throat, nausea, anosmia, ageusia, fatigue, and diarrhea. This symptom construct has a Cronbach alpha score of .812. An alpha value greater than .8 generally indicates a strong level of construct reliability. Reliability analysis showed that the alpha value would not be improved by removing any variable from the construct.

Descriptive analysis for daily new case and death data by state, initial bivariate linear regression modeling, calculations for construct reliability, lag time calculations, mixed modeling, and model comparisons were all completed in R. The primary packages used were lubridate for parsing date variables, plotly for examining ggplot2 results in more detail, lme4 and lmerTest for fitting linear mixed models, sjPlot for plotting data to test model assumptions, stargazer for table and figure creation, and all of the packages within the “tidyverse” (primarily dplyr and ggplot2) for cleaning, organizing, and preparing data for analysis and presentation.

Results

Descriptive statistics for all variables included in the final model are included in [Table 1](#). These data are the raw numbers, but predictors and controls were standardized for analysis to control for the vastly different scales for many of the variables. Given the extremely large volume of daily case and search data, this is not included in this table.

Assumptions for linear mixed model regression were checked, confirming that there was a linear relationship between the predictor and outcome variable and that the residuals were independent, uncorrelated, and normally distributed. The residual plot for homoscedasticity of residuals is in [Figure 1](#).

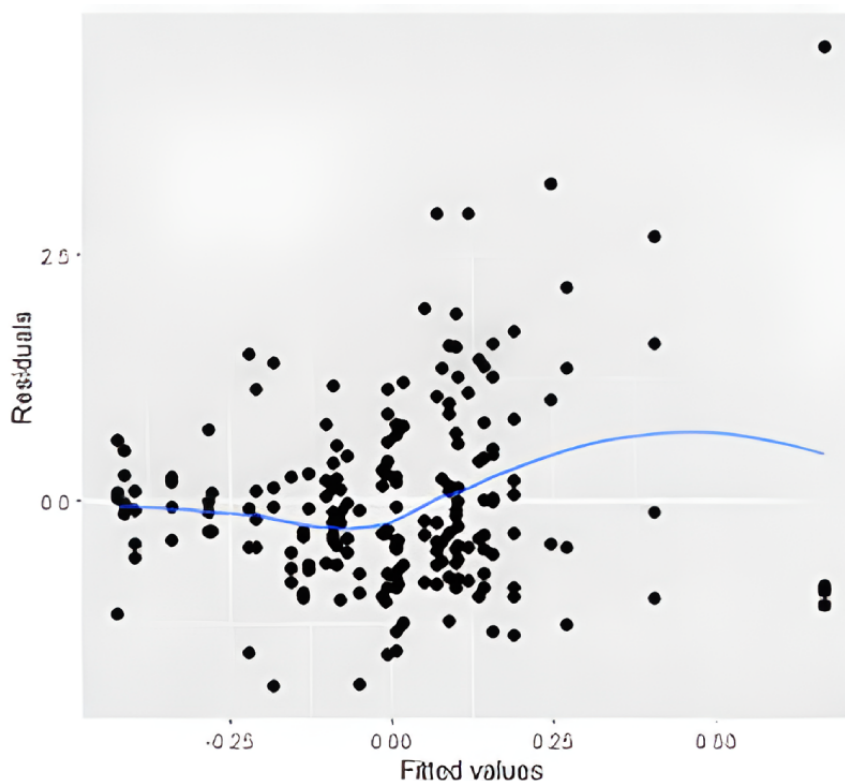
A linear mixed model was fitted using the proportion of 2020 Trump votes within a state as the primary predictive variable. This variable was also nested within state-level clusters to allow its effect to vary within each state. Models using the proposed Governor and Congressional proportion variables were discarded due to comparatively poor model fit statistics. Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores for discarded models were in the range of 600 to 800 points higher, indicating poorer fit.

Table 1. Descriptive statistics.

State	Case peaks				Search peaks				Proportion of Trump votes	GOP governor	GOP proportion of Congress	Population	Daily cases, mean	Daily deaths, mean
	1	2	3	4	1	2	3	4						
AL	185	349	584	728	157	342	581	707	0.62	1	0.83	5024803	1700.48	25.48
AK	187	319	612	729	176	273	589	714	0.53	1	0.83	732441	314.21	1.38
AZ	162	348	585	732	156	342	579	718	0.49	1	0.45	7177986	2639.60	39.21
AR	199	346	576	729	156	342	496	718	0.62	1	0.92	3012232	1092.82	14.46
CA	183	354	589	727	169	344	463	714	0.34	0	0.16	39499738	11713.8	116.06
CO	189	297	457	717	181	302	461	708	0.42	0	0.39	5784308	1775.10	15.75
CT	92	317	605	720	52	294	567	707	0.39	0	0.00	3600260	969.89	14.17
DE	104	354	598	718	57	322	567	708	0.40	0	0.00	991886	341.14	3.76
FL	172	345	583	718	157	342	567	708	0.51	1	0.58	21569932	7696.68	96.48
GA	185	353	591	713	158	343	579	707	0.49	1	0.61	10725800	3628.51	48.40
HI	204	287	586	707	174	273	537	701	0.34	0	0.00	1451911	308.27	1.80
ID	176	323	631	729	119	300	592	727	0.64	1	0.75	1847772	583.05	6.43
IL	101	297	444	720	57	293	462	708	0.41	0	0.30	12785245	4045.41	49.53
IN	97	317	598	720	57	293	579	708	0.57	1	0.82	6785644	2218.97	30.79
IA	102	297	602	724	92	293	601	714	0.53	1	0.67	3188669	990.51	12.42
KS	100	307	597	729	89	301	518	714	0.56	0	0.83	2935880	1010.45	11.14
KY	101	351	590	731	54	344	579	707	0.62	0	0.88	3503958	1532.39	22.52
LA	173	351	590	713	157	342	554	707	0.58	0	0.82	4651203	1532.39	22.52
ME	119	352	450	722	107	293	448	707	0.44	0	0.12	1362280	312.94	2.98
MD	119	318	613	719	111	317	615	707	0.32	1	0.10	6172679	1334.10	18.86
MA	94	3353	603	720	53	350	567	708	0.32	1	0.05	7022220	2246.86	31.38
MI	115	294	440	728	111	293	442	708	0.48	0	0.44	10067664	3134.87	39.68
MN	107	298	441	729	105	300	462	714	0.45	0	0.35	5707165	1869.44	16.37
MS	181	353	576	728	157	342	578	707	0.58	1	0.83	2956870	1029.39	14.62
MI	98	293	560	713	55	259	567	708	0.57	1	0.80	6154481	1841.18	22.75
MT	107	298	643	734	57	257	641	719	0.57	1	0.67	1086193	355.38	4.24
NE	107	302	603	727	57	301	574	714	0.58	1	0.80	1961455	598.27	4.97
NV	177	352	598	728	169	343	587	714	0.48	0	0.33	3114071	895.19	13.34
NH	101	348	444	728	57	259	400	714	0.45	1	0.00	1377848	399.08	3.23
NJ	73	354	604	717	57	350	567	707	0.41	0	0.07	9279743	2919.76	41.26
NM	188	303	569	728	174	302	567	707	0.43	0	0.09	3117566	680.66	9.39
NY	78	353	605	717	57	350	567	707	0.37	0	0.28	21054933	3541.09	36.21
NC	184	346	590	723	158	342	596	708	0.50	0	0.72	10457177	3452.16	30.48
ND	121	297	440	729	113	295	449	707	0.65	1	0.67	778962	314.86	2.95
OH	89	322	598	725	57	302	595	708	0.53	1	0.71	11790587	3502.58	45.09
OK	168	343	582	727	158	343	550	718	0.65	1	0.93	3962031	1358.26	17.05
OR	180	318	584	728	128	300	567	718	0.40	0	0.14	4241544	927.88	9.70
PA	79	325	448	718	57	314	462	708	0.49	0	0.50	12989625	3635.72	58.22
RI	108	317	602	716	106	258	568	709	0.39	0	0.00	1096229	454.25	4.62
SC	181	353	584	732	158	342	587	708	0.55	1	0.83	5130829	1927.38	23.20

State	Case peaks				Search peaks				Proportion of Trump votes	GOP governor	GOP proportion of Congress	Population	Daily cases, mean	Daily deaths, mean
	1	2	3	4	1	2	3	4						
SD	109	296	679	729	57	301	576	707	0.62	1	1.00	887099	310.05	3.79
TN	192	330	599	730	158	343	582	707	0.61	1	0.82	6920119	2633.41	30.82
TX	178	350	596	722	158	343	587	707	0.52	1	0.66	29217653	8525.78	113.08
UT	177	398	664	728	126	300	596	714	0.58	1	0.92	3281685	1217.06	6.19
VT	74	436	604	728	57	460	561	715	0.30	1	0.00	642495	144.45	0.78
VA	125	361	597	720	111	342	601	708	0.44	1	0.31	8632044	2202.54	25.94
WA	167	322	561	737	113	301	567	715	0.39	0	0.25	7718785	1920.45	16.46
WV	183	346	612	713	159	265	553	707	0.69	1	0.80	1789798	653.50	8.81
WI	143	297	623	727	111	293	615	707	0.49	0	0.60	5892323	2085.48	18.84
WY	189	301	589	748	159	301	603	722	0.69	1	1.00	577267	204.82	2.35

Figure 1. Homoscedasticity (constant variance of residuals). Note: amount and distance of points scattered above and below the line is equal.



The model results are in Table 2; note that the scales displayed for predictors and controls are standardized, but lag time remains measured in days. I used 762 observations for 50 states to calculate 4 intervals of lag time, which led to a sample for the model of 50 states, or units of analysis with 4 repeated measurements per state. The model is a significant improvement over the null model with a single predictor (Trump proportion), no controls, and random intercepts for states. The AIC and BIC scores for the null model were 1776.04 and 1785.93, respectively—the selected model’s scores were 588.85 and 618.535, respectively, and so were a significant improvement in terms of model fit.

However, none of the predictors within the final model were statistically significant, even those that were considered but excluded from the model (Governor party and Congressional delegation parties) because of worse overall model fit. The only significant independent variable in the selected model was mean daily deaths, which had a small negative relationship with lag time, at $P < .10$.

Random elements in the model were individual state-level intercepts and Trump proportion nested within state as a random effect. There was sufficient variance ($\sigma^2_{state} = 0.099$ and $\sigma^2_{trump} = 0.049$) in this random portion of the model to justify their inclusion. The random intercepts for state clusters captured

a portion of and thereby reduced the fixed effect residuals, but the model itself lacked predictive power. Thus, the model's fixed effect error term still captured a relatively high amount of the variance in the data. As Figure 2 shows, there was a range

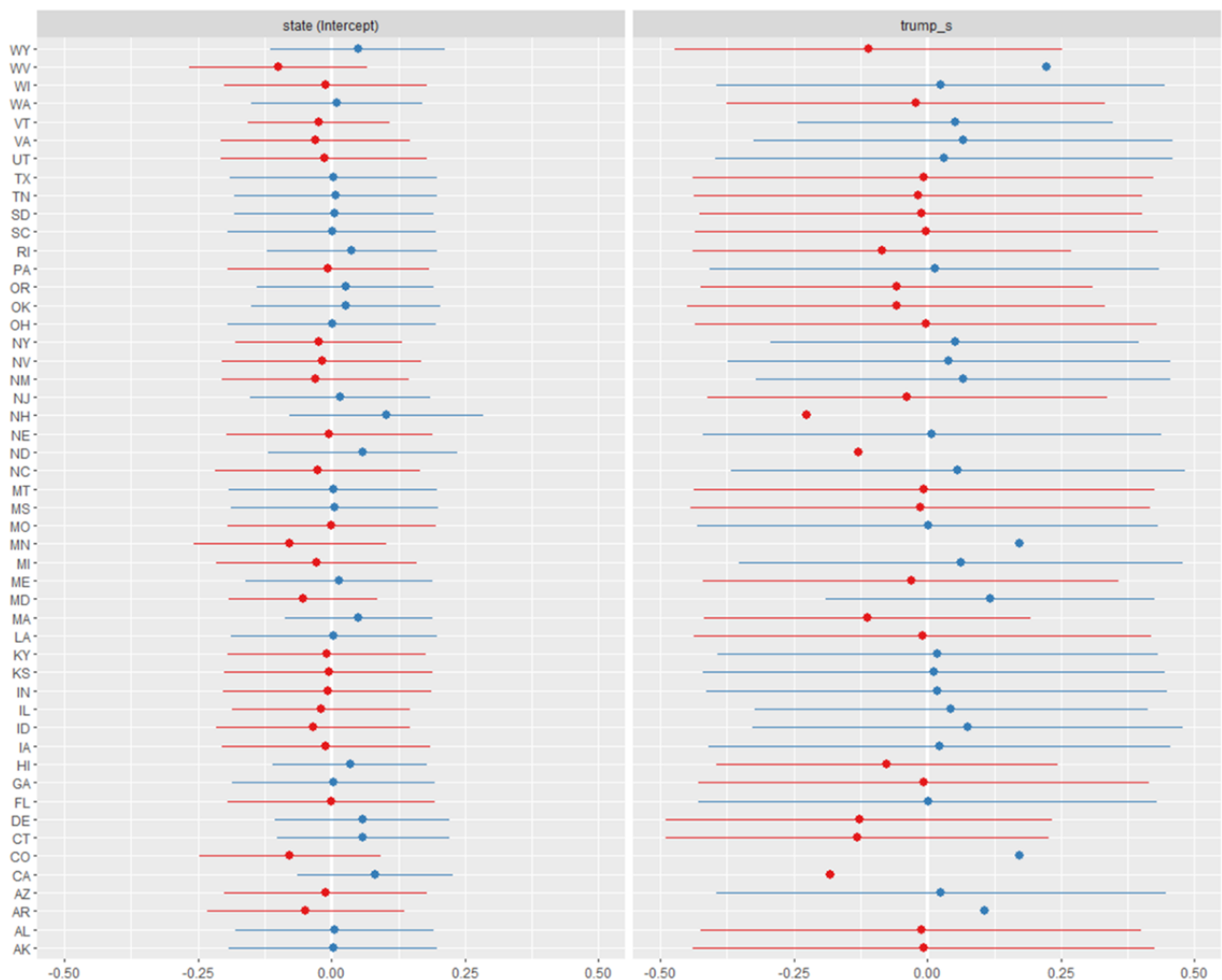
of positive and negative values for each effect, but the relatively wide confidence intervals (95% CI) were another artifact of the relatively low predictive power for this model.

Table 2. Linear mixed model results (dependent variable is lag time in days).

Variable	Statistic	P value
Proportion of Trump 2020 votes, β (SE)	0.063 (0.091)	.49
Mean daily cases, β (SE)	0.461 (0.569)	.42
Mean daily deaths, β (SE)	-0.670 (0.356)	.07
Population, β (SE)	0.168 (0.430)	.70
Constant, β (SE)	-0.012 (0.076)	.88
Overall model		
Observations	200	N/A ^a
Log likelihood	-285.425	N/A
Akaike information criterion	588.850	N/A
Bayesian information criterion	618.535	N/A

^aN/A: not applicable.

Figure 2. Random effect estimates (intercept and 95% CI) in lag time by state; red: negative; blue: positive.



Discussion

Principal Findings

I have identified evidence to support a positive answer for the first research question but not the second. Results suggest that yes, there are meaningful state-level differences in the lag time between spikes in COVID-19-related search traffic and spikes in COVID-19 cases. However, my hypothesis that political covariates would contribute to a portion of this variation in a statistically significant way was not supported. My findings indicate that including proportion of 2020 Trump voters in the final linear mixed model leads to a model fit that is overall much stronger than a null model with control variables only, but this political affiliation variable is not itself statistically significant or appropriate for predictive inference.

Although this is a partially negative finding, it challenges further work to explore how environmental factors and social group forces (both structural and interpersonal) may cause diverging responses to shared societal crises. The statistically significant differences in lag time across states demonstrate that there are meaningful differences across states that are at least partially causing these changes. I propose 3 potential causes for the mismatch between my theorized model and the results that also suggest directions for future research: level of geographic modeling, approach to calculating lag, and theoretical model mismatch.

Level of Geographic Modeling

My approach considered state as the geographic unit primarily for logistical reasons. State is an easily available unit in both the trends and cases data, meaning that they can be reliably matched for analysis. However, in doing so, I lost the opportunity for more localized nuance in my political covariates: Even using local city elections as a proxy for political identification might lead to stronger model fit and predictive power, but it may also be possible to use smaller blocks such as census tracts to map the trends data onto already narrowly geotagged CDC case data. However, the Google search data are limited by the lack of GIS shapefiles as well as the lack of city or tract-level data for areas outside of its large metro areas. An analysis with more granular geographic modeling would require the theoretical model to be reconsidered, given that it would exclude data from smaller cities and rural areas. If these localized variables were included in a hypothetical model, it would also likely be wise to include controls such as income and educational attainment, assuming that these are available for the census tracts or areas in question. On a smaller scale, these controls would likely contribute more significantly to a model than the same statistics at a broad state level.

Approach to Calculating Lag

Although I believe that the method I used to calculate lag is defensible, it is possible that another approach may uncover a significant relationship that is not present here. Pelat et al [39] provided one option within the same domain, describing a method that calculates correlations between increases in searches and incidence of a disease at predefined intervals (eg, 1 week,

1 month). These correlations were stored and used for further regression analysis along with selected predictors.

Effenberger et al [40] took another approach that accounts for lag specifically via time lag correlational analysis. Instead of calculating lag as a repeated measure in longitudinal analysis, they instead mapped multiple models as a network, examining how associations changed at predefined time intervals. This would require a significant reconfiguring of how the research questions were operationalized but may uncover relationships that were not established here.

Theoretical Model Mismatch

In short, my proposed theoretical model was that, per Bandura's SCT [7], individuals act as part of a constantly changing and reciprocal triad that is bound by personal factors, environmental influences, and past behavior. As part of this triad, the SIA also tells us that individuals sort themselves into groups by categorizing others, giving meaning to those categories, and then self-sorting into one of these groups. Given how politicized the ongoing cultural response to COVID-19 and mitigation measures has been, I suspected that the competing political understandings of the nature of the pandemic and appropriate reaction would be part of these social processes. Eventually, I hypothesized that this dynamic would affect how people managed their own symptom surveillance and perceived risk, meaning that we would discover differences in lag time between search incidence and case incidence. This theoretical model was not supported by the data here.

It is possible that another theoretical model, operationalized with a different set of predictors, would generate a significant statistical model to explain lag time variation. Barber and Pope [41] approached political identity by investigating how one's political party identification is associated with his or her individual political ideology; one element from their model that is missing here is the influence of fellow members of a political group. They also define the concepts of "party loyalist" and "policy loyalist" in the context of Trump's election, which represented an opportune time for investigating this relationship, as ideology and party often diverged. It may be possible to capture these concepts in a more localized model as described in the previous paragraphs.

Another potential theoretical model comes partially from Agadjanian and Lacy [42], who investigated how individual political leaders have a more significant influence on public opinion and behavior than party or ideology. This suggests that, if we could capture political leader characteristics and ideology with some level of granularity, this could be folded into the proposed SCT/SIA framework. As an example, coding leader rhetoric (whether manually or via sentiment analysis) would be possible under both a localized (city or tract-specific) or state-level model.

Notably, a limitation of my approach to geographic modeling is that it does not allow for investigation into whether state-level data are reliably correlated with local provincial data; future research could address this limitation via a study that specifically considers larger cities in comparison with their states or even extend this via propensity score matching at the metro and state

levels. Further, the way that I chose to operationalize political identity and affiliation is suboptimal because of the existing geographic constraint. In taking an approach that allows for more geographic granularity, future studies could also model for more granular political variables such as voting by census tract, city council representatives, and mayors.

Although there may be structural, demographic, and geographic factors that contribute to these differences, I believe that the

effects of political affiliation and its rippling effects are also closely tied to these significant differences in lag time across states. In the context of not only increasing political polarization and opportunities for political speech but also changing laws and norms around American federalism that may give states more control over shared social functions, it is more important than ever to understand how political identity and political communication affect the physical well-being of a state's residents.

Conflicts of Interest

None declared.

References

- Lyu W, Wehby GL. Community use of face masks and COVID-19: evidence from a natural experiment of state mandates in the US. *Health Aff (Millwood)* 2020 Aug;39(8):1419-1425 [FREE Full text] [doi: [10.1377/hlthaff.2020.00818](https://doi.org/10.1377/hlthaff.2020.00818)] [Medline: [32543923](https://pubmed.ncbi.nlm.nih.gov/32543923/)]
- Fischer, Adrien N, Silguero JJ, Hopper JJ, Chowdhury AI, Werler MM. Mask adherence and rate of COVID-19 across the United States. *PLoS One* 2021;16(4):e0249891 [FREE Full text] [doi: [10.1371/journal.pone.0249891](https://doi.org/10.1371/journal.pone.0249891)] [Medline: [33852626](https://pubmed.ncbi.nlm.nih.gov/33852626/)]
- Thunström L, Newbold SC, Finnoff D, Ashworth M, Shogren JF. The benefits and costs of using social distancing to flatten the curve for COVID-19. *J. Benefit Cost Anal* 2020 May 21;11(2):179-195 [FREE Full text] [doi: [10.1017/bca.2020.12](https://doi.org/10.1017/bca.2020.12)]
- Lin C, Tu P, Beitsch LM. Confidence and receptivity for COVID-19 vaccines: a rapid systematic review. *Vaccines (Basel)* 2020 Dec 30;9(1):A [FREE Full text] [doi: [10.3390/vaccines9010016](https://doi.org/10.3390/vaccines9010016)] [Medline: [33396832](https://pubmed.ncbi.nlm.nih.gov/33396832/)]
- Cowan SK, Mark N, Reich JA. COVID-19 vaccine hesitancy is the new terrain for political division among Americans. *Socius* 2021 Jun 17;7:237802312110236 [FREE Full text] [doi: [10.1177/23780231211023657](https://doi.org/10.1177/23780231211023657)]
- Stephens-Davidowitz S. Google Searches Can Help Us Find Emerging Covid-19 Outbreaks. *New York Times*. 2020 Apr 05. URL: <https://www.nytimes.com/2020/04/05/opinion/coronavirus-google-searches.html> [accessed 2022-12-07]
- Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall; 1986.
- Abrams D, Hogg MA. *Social identity theory: Constructive and critical advances*. New York, NY: Springer-Verlag Publishing; 1990.
- Bish A, Michie S. Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. *Br J Health Psychol* 2010 Nov;15(Pt 4):797-824 [FREE Full text] [doi: [10.1348/135910710X485826](https://doi.org/10.1348/135910710X485826)] [Medline: [20109274](https://pubmed.ncbi.nlm.nih.gov/20109274/)]
- Clark C, Davila A, Regis M, Kraus S. Predictors of COVID-19 voluntary compliance behaviors: An international investigation. *Glob Transit* 2020;2:76-82 [FREE Full text] [doi: [10.1016/j.glt.2020.06.003](https://doi.org/10.1016/j.glt.2020.06.003)] [Medline: [32835202](https://pubmed.ncbi.nlm.nih.gov/32835202/)]
- Cho H, Lee JS. The influence of self-efficacy, subjective norms, and risk perception on behavioral intentions related to the H1N1 flu pandemic: A comparison between Korea and the US. *Asian Journal of Social Psychology* 2015 Jun 17;18(4):311-324 [FREE Full text] [doi: [10.1111/ajsp.12104](https://doi.org/10.1111/ajsp.12104)]
- Brouard S, Vasilopoulos P, Becher M. Sociodemographic and psychological correlates of compliance with the COVID-19 public health measures in France. *Can J Pol Sci* 2020 Apr 23;53(2):253-258 [FREE Full text] [doi: [10.1017/s0008423920000335](https://doi.org/10.1017/s0008423920000335)]
- Harper CA, Satchell LP, Fido D, Litzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. *Int J Ment Health Addict* 2021;19(5):1875-1888 [FREE Full text] [doi: [10.1007/s11469-020-00281-5](https://doi.org/10.1007/s11469-020-00281-5)] [Medline: [32346359](https://pubmed.ncbi.nlm.nih.gov/32346359/)]
- Lerner JS, Keltner D. Fear, anger, and risk. *Journal of Personality and Social Psychology* 2001;81(1):146-159 [FREE Full text] [doi: [10.1037/0022-3514.81.1.146](https://doi.org/10.1037/0022-3514.81.1.146)]
- Hughner RS, Kleine SS. Views of health in the lay sector: a compilation and review of how individuals think about health. *Health (London)* 2004 Oct;8(4):395-422 [FREE Full text] [doi: [10.1177/1363459304045696](https://doi.org/10.1177/1363459304045696)] [Medline: [15358896](https://pubmed.ncbi.nlm.nih.gov/15358896/)]
- Becker MH, Maiman LA. Sociobehavioral determinants of compliance with health and medical care recommendations. *Med Care* 1975 Jan;13(1):10-24. [doi: [10.1097/00005650-197501000-00002](https://doi.org/10.1097/00005650-197501000-00002)] [Medline: [1089182](https://pubmed.ncbi.nlm.nih.gov/1089182/)]
- Nivette A, Ribeaud D, Murray A, Steinhoff A, Bechtiger L, Hepp U, et al. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study. *Soc Sci Med* 2021 Jan;268:113370 [FREE Full text] [doi: [10.1016/j.socscimed.2020.113370](https://doi.org/10.1016/j.socscimed.2020.113370)] [Medline: [32980677](https://pubmed.ncbi.nlm.nih.gov/32980677/)]
- Barari S, Caria S, Davola A, Falco O, Fetzer T, Fiorin S, et al. Evaluating COVID-19 public health messaging in Italy: self-reported compliance and growing mental health concerns. *medRxiv*. Preprint posted online April 5 2020. [FREE Full text] [doi: [10.1101/2020.03.27.20042820](https://doi.org/10.1101/2020.03.27.20042820)]
- Callahan D, Jennings B. Ethics and public health: forging a strong relationship. *Am J Public Health* 2002 Feb;92(2):169-176. [doi: [10.2105/ajph.92.2.169](https://doi.org/10.2105/ajph.92.2.169)] [Medline: [11818284](https://pubmed.ncbi.nlm.nih.gov/11818284/)]

20. Baum NM, Jacobson PD, Goold SD. "Listen to the people": public deliberation about social distancing measures in a pandemic. *Am J Bioeth* 2009 Nov 04;9(11):4-14 [[FREE Full text](#)] [doi: [10.1080/15265160903197531](https://doi.org/10.1080/15265160903197531)] [Medline: [19882444](#)]
21. Rubin G, Amlôt R, Page L, Wessely S. Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. *BMJ* 2009 Jul 02;339:b2651 [[FREE Full text](#)] [doi: [10.1136/bmj.b2651](https://doi.org/10.1136/bmj.b2651)] [Medline: [19574308](#)]
22. Callaway D, Runge J, Mullen L, Rentz L, Staley K, Stanford M, et al. Risk and the Republican National Convention: application of the novel COVID-19 operational risk assessment. *Disaster Med Public Health Prep* 2022 Aug;16(4):1612-1617 [[FREE Full text](#)] [doi: [10.1017/dmp.2020.499](https://doi.org/10.1017/dmp.2020.499)] [Medline: [33762039](#)]
23. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014 [[FREE Full text](#)] [doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634)] [Medline: [19020500](#)]
24. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection — harnessing the web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153-2157 [[FREE Full text](#)] [doi: [10.1056/nejmp0900702](https://doi.org/10.1056/nejmp0900702)]
25. Polgreen P, Chen Y, Pennock D, Nelson F. Using internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 01;47(11):1443-1448 [[FREE Full text](#)] [doi: [10.1086/593098](https://doi.org/10.1086/593098)] [Medline: [18954267](#)]
26. Cooper CP, Mallon KP, Leadbetter S, Pollack LA, Peipins LA. Cancer Internet search activity on a major search engine, United States 2001-2003. *J Med Internet Res* 2005 Jul 01;7(3):e36 [[FREE Full text](#)] [doi: [10.2196/jmir.7.3.e36](https://doi.org/10.2196/jmir.7.3.e36)] [Medline: [15998627](#)]
27. Jimenez AJ, Estevez-Reboredo RM, Santed MA, Ramos V. COVID-19 symptom-related Google searches and local COVID-19 incidence in Spain: correlational study. *J Med Internet Res* 2020 Dec 18;22(12):e23518 [[FREE Full text](#)] [doi: [10.2196/23518](https://doi.org/10.2196/23518)] [Medline: [33156803](#)]
28. Rabiolo A, Alladio E, Morales E, McNaught AI, Bandello F, Afifi AA, et al. Forecasting the COVID-19 epidemic by integrating symptom search behavior into predictive models: infoveillance study. *J Med Internet Res* 2021 Aug 11;23(8):e28876 [[FREE Full text](#)] [doi: [10.2196/28876](https://doi.org/10.2196/28876)] [Medline: [34156966](#)]
29. Hoerger M, Alonzi S, Perry LM, Voss HM, Easwar S, Gerhart JI. Impact of the COVID-19 pandemic on mental health: Real-time surveillance using Google Trends. *Psychol Trauma* 2020 Sep;12(6):567-568 [[FREE Full text](#)] [doi: [10.1037/tra0000872](https://doi.org/10.1037/tra0000872)] [Medline: [32790441](#)]
30. Asseo K, Fierro F, Slavutsky Y, Frasnelli J, Niv MY. Tracking COVID-19 using taste and smell loss Google searches is not a reliable strategy. *Sci Rep* 2020 Nov 25;10(1):20527 [[FREE Full text](#)] [doi: [10.1038/s41598-020-77316-3](https://doi.org/10.1038/s41598-020-77316-3)] [Medline: [33239650](#)]
31. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc* 2006;2006:244-248 [[FREE Full text](#)] [Medline: [17238340](#)]
32. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Sci* 2015 Oct 16;4(1):1 [[FREE Full text](#)] [doi: [10.1140/epjds/s13688-015-0054-0](https://doi.org/10.1140/epjds/s13688-015-0054-0)] [Medline: [27990325](#)]
33. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of Google Trends in health care research: a systematic review. *PLoS One* 2014 Oct 22;9(10):e109583 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0109583](https://doi.org/10.1371/journal.pone.0109583)] [Medline: [25337815](#)]
34. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 2011 May 31;5(5):e1206 [[FREE Full text](#)] [doi: [10.1371/journal.pntd.0001206](https://doi.org/10.1371/journal.pntd.0001206)] [Medline: [21647308](#)]
35. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol* 2013 Oct 17;9(10):e1003256 [[FREE Full text](#)] [doi: [10.1371/journal.pcbi.1003256](https://doi.org/10.1371/journal.pcbi.1003256)] [Medline: [24146603](#)]
36. Jun SP, Yoo HS, Choi S. Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change* 2018 May;130:69-87 [[FREE Full text](#)] [doi: [10.1016/j.techfore.2017.11.009](https://doi.org/10.1016/j.techfore.2017.11.009)]
37. Choi HY, Varian H. Predicting the present with Google Trends. *Economic Record* 2012;88(S1):2-9 [[FREE Full text](#)] [doi: [10.1111/j.1475-4932.2012.00809.x](https://doi.org/10.1111/j.1475-4932.2012.00809.x)]
38. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. *JMIR Public Health Surveill* 2019 May 29;5(2):e13439 [[FREE Full text](#)] [doi: [10.2196/13439](https://doi.org/10.2196/13439)] [Medline: [31144671](#)]
39. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron AJ. More diseases tracked by using Google Trends. *Emerg Infect Dis* 2009 Aug;15(8):1327-1328 [[FREE Full text](#)] [doi: [10.3201/eid1508.090299](https://doi.org/10.3201/eid1508.090299)] [Medline: [19751610](#)]
40. Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with internet search volumes: a Google Trends analysis. *Int J Infect Dis* 2020 Jun;95:192-197 [[FREE Full text](#)] [doi: [10.1016/j.ijid.2020.04.033](https://doi.org/10.1016/j.ijid.2020.04.033)] [Medline: [32305520](#)]
41. Barber M, Pope JC. Does party trump ideology? Disentangling party and ideology in America. *Am Polit Sci Rev* 2018 Dec 18;113(1):38-54 [[FREE Full text](#)] [doi: [10.1017/s0003055418000795](https://doi.org/10.1017/s0003055418000795)]

42. Agadjanian A, Lacy D. Changing votes, changing identities? Racial fluidity and vote switching in the 2012–2016 US Presidential Elections. *Public Opinion Quarterly* 2021;85(3):737-752 [FREE Full text] [doi: [10.1093/poq/nfab045](https://doi.org/10.1093/poq/nfab045)]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
CDC: Centers for Disease Control and Prevention
GIS: geographic information system
NDS: novel data streams
SCT: social cognitive theory
SIA: social identity approach

Edited by A Mavragani; submitted 06.07.22; peer-reviewed by JK Kumar, Y Hu; comments to author 06.10.22; revised version received 24.10.22; accepted 08.11.22; published 23.12.22.

Please cite as:

Turvy A

State-Level COVID-19 Symptom Searches and Case Data: Quantitative Analysis of Political Affiliation as a Predictor for Lag Time Using Google Trends and Centers for Disease Control and Prevention Data

JMIR Form Res 2022;6(12):e40825

URL: <https://formative.jmir.org/2022/12/e40825>

doi: [10.2196/40825](https://doi.org/10.2196/40825)

PMID: [36446048](https://pubmed.ncbi.nlm.nih.gov/36446048/)

©Alex Turvy. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Ethical Issues in the Use of Smartphone Apps for HIV Prevention in Malaysia: Focus Group Study With Men Who Have Sex With Men

Antoine Khati¹, MD; Jeffrey A Wickersham², PhD; Aviana O Rosen¹, MSc; Jeffrey Ralph B Luces³, MOHRE; Nicholas Copenhaver¹, BSc; Alma Jeri-Wahrhaftig¹, BSc; Mohd Akbar Ab Halim⁴, BA; Iskandar Azwa⁴, MD; Kamal Gautam¹, MPH; Kai Hong Ooi⁴, BSc; Roman Shrestha^{1,2}, MPH, PhD

¹Department of Allied Health Sciences, University of Connecticut, Storrs, CT, United States

²AIDS Program, Yale School of Medicine, New Haven, CT, United States

³Department of Physical Therapy, University of the Philippines, Manila, Philippines

⁴Centre of Excellence for Research in AIDS (CERiA), University of Malaya, Kuala Lumpur, Malaysia

Corresponding Author:

Roman Shrestha, MPH, PhD

Department of Allied Health Sciences

University of Connecticut

358 Mansfield Road Unit 1101

Storrs, CT, 06269

United States

Phone: 1 8604862834

Email: roman.shrestha@uconn.edu

Abstract

Background: The use of smartphone apps can improve the HIV prevention cascade for key populations such as men who have sex with men (MSM). In Malaysia, where stigma and discrimination toward MSM are high, mobile health app-based strategies have the potential to open new frontiers for HIV prevention. However, little guidance is available to inform researchers about the ethical concerns that are unique to the development and implementation of app-based HIV prevention programs.

Objective: This study aimed to fill this gap by characterizing the attitudes and concerns of Malaysian MSM regarding HIV prevention mobile apps, particularly regarding the ethical aspects surrounding their use.

Methods: We conducted web-based focus group discussions with 23 MSM between August and September 2021. Using in-depth semistructured interviews, participants were asked about the risks and ethical issues they perceived to be associated with using mobile apps for HIV prevention. Each session was digitally recorded and transcribed. Transcripts were inductively coded using the Dedoose software (SocioCultural Research Consultants) and analyzed to identify and interpret emerging themes.

Results: Although participants were highly willing to use app-based strategies for HIV prevention, they raised several ethical concerns related to their use. Prominent concerns raised by participants included privacy and confidentiality concerns, including fear of third-party access to personal health information (eg, friends or family and government agencies), issues around personal health data storage and management, equity and equitable access, informed consent, and regulation.

Conclusions: The study's findings highlight the role of ethical concerns related to the use of app-based HIV prevention programs. Given the ever-growing nature of such technological platforms that are intermixed with a complex ethical-legal landscape, mobile health platforms must be safe and secure to minimize unintended harm, safeguard user privacy and confidentiality, and obtain public trust and uptake.

(JMIR Form Res 2022;6(12):e42939) doi:[10.2196/42939](https://doi.org/10.2196/42939)

KEYWORDS

HIV; mobile health; mHealth; mobile app; HIV prevention; men who have sex with men; privacy; confidentiality; Malaysia; mobile apps; ethics; focus group; implementation; user privacy; mobile phone

Introduction

Background

Between 2020 and 2021, HIV incidence around the globe witnessed an annual decline of 3.6%, the smallest since 2016 [1]. Although recent HIV incidence and mortality trends have declined globally, some countries in Southeast Asia still bear a disproportionate HIV burden [2-4]. Among the 1.5 million new HIV cases reported globally in 2021, a total of 260,000 were from the Asia and Pacific region, where HIV infection rates are now rising, even in locations where they had been previously declining [1].

Malaysia has one of the highest HIV prevalence in the Asia-Pacific region and remains one of the only a few countries globally where HIV-related mortality has been steadily increasing [5]. In addition, it remains one of the HIV hot spots in that region, with alarming increases in HIV infections [1] and 5500 new cases in 2021 among adults aged ≥ 15 years [6]. Malaysia's HIV epidemic is rapidly expanding, with recent evidence suggesting accelerated sexual transmission, especially in men who have sex with men (MSM) [7,8]. Several factors may potentiate HIV transmission among Malaysian MSM, including condomless sex, sexually transmitted infections, and comorbid psychiatric or substance use disorders [9-14].

Same-sex behavior is illegal in Malaysia under the Sharia law and section 377 of the Malaysian Penal Code. Under section 377, "carnal intercourse against the order of nature" is punishable with imprisonment that may last for 20 years and include whipping [15,16]. Similarly, the Islamic Sharia law criminalizes anal sex among Muslims and carries similar punishments with imprisonment, fines, and whipping. Instances of Malaysian authorities raiding venues regularly visited by MSM to enforce the law have also been reported, along with cases of physical and psychological abuse during incarceration among MSM and other men perceived to be MSM who were arrested under section 377 [15].

In this hostile environment that criminalizes and punishes same-sex sexual activity, high levels of stigma and discrimination against Malaysian MSM ensue, including in health care settings [17-19]. The resulting fear of legal repercussions related to the disclosure of same-sex behavior and judgment from health care providers [17,20-22] hinder access to and use of HIV prevention and treatment services, contesting the role of in-person venues for HIV prevention services, such as HIV testing and pre-exposure prophylaxis (PrEP).

The recent advances in wireless technology, along with the change in the traditional health care service delivery model, have led to the development of mobile health (mHealth), which offers an unparalleled opportunity to deliver internet-based health care services. These technologies have increasingly been used in health care management and delivery to address various health conditions and diseases, including diabetes and hypertension management, smoking cessation, weight loss, increased physical activity, and sexually transmitted infections [23-28]. Such technologies have many potential health care

benefits, such as monitoring users' health status remotely and continuously, improving access to health care services, lowering health care costs, increasing patients' awareness of health status, and improving patient-provider communications.

mHealth interventions can particularly benefit HIV prevention efforts in Malaysia, where smartphone ownership among MSM is nearly universal. Existing data indicate that over 97% of MSM in Malaysia have access to a smartphone, making mHealth a highly feasible platform for this subpopulation [29]. Furthermore, findings from recent studies demonstrate a strong preference among MSM for smartphone apps over other modalities (eg, text, phone calls, and emails) to engage in mHealth HIV prevention tools, thus supporting the development and deployment of smartphone apps for HIV prevention [29-31]. mHealth could serve as an innovative platform to improve access to and use of HIV testing, linkage to PrEP and antiretroviral therapy, and other support services (eg, mental health and substance use).

Despite the numerous benefits of app-based platforms, several ethical challenges accompany their use and must be considered to safeguard user safety, privacy, and rights. These include the protection of privacy and confidentiality, informed consent (including transparency with users about potential risks), data management and careful communication of data, mHealth product regulation and evaluation before public dissemination and uptake, and equity and equitable access [32-34]. Researchers, clinicians, and app developers must be knowledgeable about the ethical issues surrounding the use of HIV prevention apps to improve the perceived usefulness, interpretability, navigability, feasibility, and acceptability of such platforms and tailor them to the specific needs of their target population.

Objectives

Although ethical considerations for telehealth have been well addressed in the literature, particularly in high-income countries [35,36], patient feedback on ethical issues around using smartphone apps for HIV prevention is sparse, especially in low- and middle-income countries such as Malaysia [31]. Therefore, this study aims to fill this gap by characterizing the attitudes and concerns of Malaysian MSM regarding HIV prevention mobile apps, particularly regarding the ethical aspects surrounding their use.

Methods

Study Design

The reporting of this manuscript was guided by the Standards for Reporting Qualitative Research [37]. This study used a qualitative exploratory design using focus group discussions (FGDs) as an interview technique. The authors first developed an interview guide, which was consolidated according to the available literature on key ethical issues related to mHealth [32,33,38].

Existing Ethical Principles and Interview Guide Development

To generate a comprehensive list of ethical principles related to HIV prevention mHealth app use, we first identified the general ethical framework's 8 principles for biomedical research [39]. These included collaborative partnership, social value, scientific validity, fair subject selection, favorable risk-benefit ratio, independent review, and informed consent. We then explored ethical issues in the context of mHealth research among people living with HIV [32], which further stressed the importance of associated physical, social, behavioral, and psychological risks and privacy and confidentiality risks. In addition, we identified 2 mHealth ethics frameworks [33,38] that elaborated on the accessibility of mHealth platforms, informed consent, and regulation of mHealth products.

To develop the interview guide, we modified the general ethical framework for biomedical research principles by adapting it to mHealth-related research. For example, the social value and scientific validity principles were combined [32]. The collaborative partnership and independent review ethical principles were not addressed in the mHealth-specific ethics frameworks and were removed from the guide. Furthermore, Carter et al [33] described issues with access to mHealth, such as socioeconomic status and physical or mental impairments. These were subsumed under fair subject selection. Carter et al [33] and Fisher et al [38] discussed privacy and consent in more detail, which helped to populate probing questions. Finally, Shrestha et al [33] discussed regulation, which was not addressed in other frameworks and was thus added to the topic list. Finally, feedback from authors was taken until a consensus was reached. This was followed by 4 iterative rounds of revisions, rearrangements, and merging of topics between coauthors that ultimately generated the final version of the guide used in all FGDs.

Study Setting and Recruitment

We recruited a convenience sample of 23 MSM between August and September 2021. The eligibility criteria were as follows: (1) a self-reported negative or unknown HIV status, (2) aged ≥ 18 years, (3) identifying as male, and (4) the ability to read and understand English or Bahasa Malaysia. Participants were recruited using advertisements on geosocial networking (GSN) apps for MSM (ie, Hornet) and Facebook (a popular social networking website for the general population). The GSN apps pushed the advertisement as a message to the chat inboxes of all users in Malaysia. Targeted banner advertisements were used on Facebook. These banners appeared either as a static advertisement on the right-hand pane of the website or an advertisement that resembled a standard post that users could encounter while scrolling through their feed; clicking on the advertisements directed interested persons to another page to provide their contact information. A research staff member then contacted eligible individuals, assessed their willingness to join the study, and shared more details with eligible and willing participants.

Study Procedures and Measures

Owing to movement restrictions related to the COVID-19 pandemic in Malaysia, participant recruitment and FGDs were conducted via the internet. FGDs were conducted using a videoconferencing platform (ie, WebEx), and each session lasted approximately 90 minutes. Three FGDs were conducted until theoretical saturation was reached [40]. The first 2 FGDs included 8 participants, whereas the last session included 7. A trained facilitator led the FGD session, whereas a cofacilitator took notes, recorded nonverbal cues, and collected chat entries.

Each session began with a brief introduction in which a description of the purpose of the FGD was communicated to participants; time was dedicated to answering participants' questions and concerns, and participants were reminded that the session would be recorded. Participants were informed that they were not required to use their real names and that they could keep their cameras off during the session to ensure privacy. Participants were encouraged to share their feedback verbally or through the chat function, which was frequently reviewed to ensure that all participant responses were collected. Each session was audio recorded and transcribed. Two authors reviewed the transcripts for accuracy.

Following informed consent, participants completed a brief web-based survey on Qualtrics that included sociodemographic information (eg, age, ethnicity, and sexual orientation), access to communication devices (eg, mobile phones, tablets, and laptops), and awareness of and previous use of HIV prevention services (eg, HIV testing and PrEP).

Before the discussion, participants were briefly provided with information regarding the use of smartphone apps for HIV prevention, including the common features and functions that apps for HIV prevention comprise to support HIV prevention efforts and care. Some key app features included receiving reminders to take PrEP; scheduling appointments with health care providers; ordering PrEP and HIV self-testing kits; and reviewing test results, medication intake, and other health-related information through the app.

Participants were then asked to provide insights into perceived individual- and community-level ethical challenges and concerns about using an HIV prevention app with such features. The FGDs started with a general question to participants regarding their impression and thoughts on the use of a mobile app that contains the features and functions that were just presented ("What is your general impression regarding the use of such an app with these features embedded?"). A more specific question followed, pertaining to the risks of using such apps for HIV prevention ("What do you think are some of the risks of using such an app for HIV prevention efforts?"). Participants were then probed on each of the 5 ethical principles (previously described) that were the focus of the FGD sessions through multiple questions on each ethical aspect.

The first ethical concern was privacy, specifically regarding anonymity and data deidentification. The subconstructs included app visibility (eg, "how would you feel about the app being visible on your phone?"), photos or avatars (eg, "how do you feel about using photo avatars and aliases in setting up your

profile in the app?”), and reminders and notifications (eg, “what concerns might you have regarding PrEP reminders?”). The second ethical principle was the third-party use of data. Therefore, questions regarding potential confidentiality breaches were asked (eg, “do you have any concerns about your data being accessible to others?”). Questions regarding the storage and transmission of data were then addressed, including questions about server locations and types (eg, “if the data were stored on a server in Malaysia or in a foreign country, how would you feel about your data security/privacy?”), governmental interception of stored data (eg, “are you concerned about the government accessing your personal information from the app?”), and sharing personal health-related information (eg, “how comfortable do you feel about sharing personal health information via the app?”). Finally, questions about access to mHealth technology (eg, “do you think there are smaller groups in your community who will have difficulty using the app or getting access to a smartphone?”) and regulation of mHealth products in Malaysia (eg, “how do you feel about using an app that does or does not meet certain minimum standards for security, privacy or quality that leading companies or industry groups set?”) were raised. In addition, while each ethical issue was being discussed, participants were encouraged to provide recommendations or suggestions to help mitigate or solve the concerns they brought forth. For example, during the discussion on third-party use of data, participants were asked the following question: “What steps could app developers take to address the security and confidentiality breach concerns that you have?”

Data Analysis

Descriptive statistics for variables collected via a brief Qualtrics survey were computed, including frequencies and percentages for categorical variables, using SAS (version 9.4; SAS Institute, Inc). The transcripts were analyzed using abductive thematic analysis to inductively identify and interpret the concepts and themes that emerged from the interview transcripts. This method involves multiple readings of transcripts and interview notes and analytic induction via open and axial data coding using

Dedoose software (SocioCultural Research Consultants, Los Angeles, CA, US) to organize transcripts thematically. Transcripts were checked for any inconsistencies or mistakes before coding was initiated. A codebook was developed with mutually agreed-upon codes derived from the interview transcripts, and coding was completed independently by 2 researchers (including a senior coder). To ensure reliability, codes were constantly compared for agreement and discussed between the 2 coders, and the senior coder cross-checked all codes [41,42]. The Cohen κ coefficient for agreement was estimated to assess the interrater concordance. Open coding, which involved assigning conceptual codes to small sections of words, phrases, and sentences in transcripts, was followed by axial coding, whereby relationships among similar concepts and categories were identified and combined into themes.

Ethics Approval

Participants provided verbal consent before starting the FGDs and were informed that participation in the study was voluntary. Participants were compensated with RM 45 (approximately US \$10) per person for their participation. The study protocol was approved by the Institutional Review Board at the University of Connecticut (L21-0007). All FGD transcripts were deidentified before the analysis, and the web-based survey data were anonymous.

Results

Demographics

Table 1 provides information on participant characteristics. The mean age of the participants was 33.4 (SD 12.0) years. Most (13/23, 57%) participants were Chinese, identified as being gay (21/23, 91%), had daily access to a smartphone with internet (22/23, 96%), and had daily access to the internet (23/23, 100%). In addition, the vast majority (22/23, 96%) had taken an HIV test at least once and had heard of PrEP previously (22/23, 96%). Finally, 61% (14/23) of participants had already taken PrEP in the past.

Table 1. Characteristics of participants (N=23).

Variables	Values
Age (years), mean (SD)	33.4 (12.0)
Ethnicity^a, n (%)	
Chinese	13 (57)
Malaya	8 (35)
Indian	2 (9)
Sexual orientation, n (%)	
Gay	21 (91)
Bisexual	2 (9)
Access to communication technology^b, n (%)	
Landline	1 (4)
Mobile phone with internet access (ie, smartphone)	22 (96)
Mobile phone without internet access	2 (9)
Tablet	13 (57)
Laptop	19 (83)
PC	7 (30)
Had daily access to internet, n (%)	23 (100)
Primary device to access the internet, n (%)	
Smartphone	17 (74)
Tablet	1 (4)
Laptop	3 (13)
PC	2 (9)
Ever tested for HIV, n (%)	
Yes	22 (96)
No	1 (4)
Ever heard of PrEP^c, n (%)	
Yes	22 (96)
No	1 (4)
Ever taken PrEP, n (%)	
Yes	14 (61)
No	9 (39)

^aPercentages may not add to 100% owing to rounding.

^bPercentages may not add to 100% because answers are nonexclusive.

^cPrEP: pre-exposure prophylaxis.

Ethical Issues and Concerns Around the Use of HIV Prevention Smartphone Apps

Overview

Several themes were identified during FGDs regarding ethical concerns around using mobile apps for HIV prevention. Participants were generally concerned about privacy and confidentiality issues, including uploading personal information to mobile apps and how apps would be visible to others on mobile home screens. Another emerging theme was storage and data ownership, whereby participants raised concerns about

data management and storage, including server types and locations and governmental access to data. Moreover, concerns about informed consent, access to web-based communication technology (eg, smartphones), and regulation were raised. Participants noted several subpopulations in Malaysia that might have trouble using or accessing mobile apps and demonstrated preferences for endorsing agencies, quality control checks, and locations of app developers and their affiliations. The major themes (privacy and confidentiality, storage and data ownership, access to mHealth technology, informed consent, and regulation) are discussed below in order of overall importance. Cohen κ

coefficient showed strong agreement between the coders (Cohen $\kappa=0.851$).

Privacy and Confidentiality

The participants highlighted concerns regarding the privacy and confidentiality of the data collected by the app (Table 2). Participants were concerned about third-party access to the data collected via the app (or knowledge of their HIV app use)

through incidental discovery, for instance, by someone accessing the phone (eg, family or friends):

Some people have a keen eye, and they might read the app's name, for example, and guess that, it is related to HIV. Some people would discriminate against others based on their medications, like antiretroviral medications.

Table 2. Ethical concerns brought up in focus group discussions regarding privacy and confidentiality (N=23).

Ethical category themes and subthemes	Mentions, n
Uploading personal information	12
Information for identity verification	2
Uploading HIV test result	10
Weary of uploading positive results	3
Not concerned with uploading information	6
Setting up account	3
Concerns over disclosing personal information	3
Doctors upholding privacy	2
App visibility	14
App icon	8
PrEP ^a stigma within MSM ^b community	1
Visibility is not an issue	5
Notifications	7
Discreet or silenced notifications preferred	1
Notification customization options	4
Pop-up display or message preview	3
Sounds or ringtone	5
Photos and avatars	5
Avatars are an excellent option for discretion	2
Customization options are preferred	3
Photos are not an issue for younger MSM	2
Not worried about privacy	5
Level of security	5
Breach via theft or hacking	4
Access via log-in, password, or OTP ^c	2
Not a concern	2

^aPrEP: pre-exposure prophylaxis.

^bMSM: men who have sex with men.

^cOTP: one-time password.

Although this was a relatively uncommon concern, some participants highlighted the need to ensure that information regarding the identity of the app developers or owners (eg, academic institutions, health care organizations, nongovernmental organizations, pharmaceutical companies, and government agencies) is transparent. For example, one individual mentioned:

I want to know "who owns this company" or whatnot. Suppose I feel it is associated with an academic institution or at least with one established in the health industry. In that case, I think it'll be safer to upload my personal info on it.

Possible solutions to address privacy and confidentiality issues included a log-in screen on opening the app and multifactor authentication or fingerprint or facial recognition methods:

I wonder, sending out like a multi-factor authentication is quite good.

Privacy concerns regarding app visibility and tailoring notifications and avatars were also discussed. For example, 1 participant pointed out that app visibility should be minimized, suggesting that this “would be great for people who are still not comfortable with their sexuality.” One recommendation to address this issue was to incorporate a discreet app icon to make the app inconspicuous when notifications appear, thereby reducing the risk of accidental disclosure to family members or friends. Another suggestion was to deliver encrypted messages to the app users:

Maybe you could customize [the notifications] so that the app sends a message that says, “It’s 8 am, time to brush your teeth,” and you know, that doesn’t really mean brush your teeth, it means “take your PrEP.”

Another opportunity for discretion within the app, as suggested by the participants, involved the customization of profile photos and avatars. A participant suggested:

I think having the option to have avatars and stuff like that will help them to have the confidence to use the app, so I think it’s a great idea to have a few avatars as an option. For me personally, I wouldn’t be uploading my picture on your app, but I don’t think it really matters to have a personal picture on it.

In general, participants seemed to value discretion within the app, leaving no opportunity for accidental disclosure incidents.

Storage and Data Ownership

Another prominent concern that the participants raised was storing and sharing data collected via the app (Table 3). Participants expressly referred to the management of collected data and the data storage methods, including the entities in charge of data storage and individuals and organizations with access to it:

The communication itself is not a problem, but the custodian of the data is. I believe the chat will be safe. I mean, it will be encrypted, right? But it’s just a matter of how the data is being kept.

Table 3. Ethical issues and concerns brought up in focus group discussions regarding data storage and ownership (N=23).

Ethical category themes and subthemes	Mentions, n
Data ownership	3
Government access to stored data	3
Data management concerns	34
How will data be used	1
Data storage	14
Location of the server	2
Reputable operator	3
Who stores or manages data	3
Location of the server is not a concern	5
Data transmission risks	4
Avoid public Wi-Fi	1
Not a concern	2

Many participants expressed concerns that their personal health-related information may become publicly available and accessible by government agencies or that their sexual orientation may be disclosed to others, including health care providers:

I’m concerned about the custodian of the data that’s being collected by this app. Because it is not just a normal app, you will also be collecting data regarding health.

Participants further pointed out the threat to privacy through third-party access of the app data through government interception or hacking of information sent over the internet. For many participants, the underlying concern in this topic involved the likelihood of Malaysian government officials obtaining access to app data.

Interestingly, most participants who raised concerns about data storage servers were not too concerned about the server’s location. The need for reputable server hosting companies was discussed in several instances. Moreover, data transmission risks were not a major concern. Avoiding public Wi-Fi (or other telecommunication networks) was infrequently mentioned, and no other specific transmission risks were raised.

Access to mHealth Technology

Participants stressed that there is an ethical imperative to not exclude any individuals at risk for HIV (eg, discreet or hidden MSM subgroups, those with low income, those who reside in rural areas, and those who cannot afford technology) from benefiting from this platform (Table 4):

But then you also need to understand that the living standard itself, you know, people who are from the lower class might not be able to afford all this.

Table 4. Ethical issues and concerns brought up in focus group discussions regarding mobile health (mHealth) access and regulation (N=23).

Ethical category themes and subthemes	Mentions, n
Access to mHealth^a technology	N/A ^b
MSM^c subpopulations	10
College students	3
Discreet MSM	3
Low-income MSM	2
MSM in rural areas	1
MSM who use drugs	1
Young MSM	3
Regulation	18
Endorsement by ministry or NGO ^d	3
App liability or quality control of medication	1
Assurance of HIV test kit quality	2
App developer	8
Local versus overseas	4
Private versus public	2
University versus company	4

^amHealth: mobile health.

^bN/A: not applicable.

^cMSM: men who have sex with men.

^dNGO: nongovernmental organization.

One participant shared the challenges of young individuals (eg, college students) accessing HIV prevention services and indicated the need to tailor the app-based interventions to their specific needs. One potential solution proposed to ensure equitable access in this subgroup was to target college-level LGBTQIA+ (lesbian, gay, bisexual, queer, intersex, asexual, and others) advocacy and support groups to help them access HIV prevention and other health care resources:

They [college students] are often, um, either still under their family's medical insurance and therefore have difficulty accessing PrEP or have concerns about privacy. So certainly, I think, making the app available to everybody, especially the younger, more vulnerable groups, would be valuable.

As mentioned previously, one of the MSM subpopulations identified by participants as potentially having trouble accessing or using mobile apps for HIV prevention was discreet MSM or those who are not *out* to their respective communities or social circles. It was noted that providing discreet MSM access to such apps could empower them, link them to care, and encourage them to voice their concerns. In addition, 1 participant noted that the app constitutes a way for MSM who use drugs to be potentially linked to HIV care and addiction treatment programs:

But I really hope that those people who are discreet can find an access to start off, you know, like, open up and voice out what is your concern and what is the care you need. That's why I really hope it

[referring to equitable access to MSM subpopulations] can be done.

Informed Consent

Several participants indicated the importance of informed consent (mentioned 7 times in total), whereby privacy policies and terms of services must be in place to maximize transparency and clarity between the app developers or owners and the front-end app users (Table 4). More specifically, participants suggested that this feature should be added before users sign up or create an account on the app and include details about the type of data that will be collected and the risks and benefits accompanying the use of the app. Furthermore, participants noted that these documents should ideally cover data collection and storage, whereby users would be able to consent to store their data on the mHealth platform or not:

I think at the beginning of the registration on the app, you must put an option to consent for app use, then it shouldn't be an issue anymore. So, consent should include gathering data and stuff like that.

Regulation

Participants were generally concerned about the regulation and evaluation of the app-based platforms and the services offered via the platform (Table 4). Participants expressed that one of the ways they would use to ascertain the app's efficacy was to look at the location (ie, overseas or local) of the developers or owners and the type of the institution (ie, private company or academic institution owned) that owned or endorsed the app,

with a preference for overseas entities and educational institutions. In addition, participants wanted to ensure that they could trust app developers or the app owner to conduct general quality control and ensure that the services provided through the app were of high quality and from trustworthy sources:

I am more concerned about the medication supply that I receive [through the app]. I would expect that the meds that are being supplied through the apps are legit, through the right source and as well as being approved by the regulators of Malaysia. I expect that this app will review all the medication being supplied.

Discussion

Principal Findings and Comparison With Previous Work

mHealth platforms are promising tools that can change the HIV prevention landscape [43-47]. In countries where a hostile sociopolitical environment for MSM prevails, the ensuing stigma and discrimination around homosexuality hinder health care access and have been associated with reduced HIV prevention service uptake among MSM [48,49]. mHealth platforms are well placed to fill this gap, as they provide a web-based platform to deliver HIV services, especially in countries such as Malaysia where MSM are well connected to the internet and use smartphones to seek sexual health information on the web [29]. The findings of this study are critical to assist the development of ethically sound app-based HIV prevention programs, which require the engagement of MSM early on [31].

Among our sample of MSM, concerns about privacy and confidentiality were substantial, bringing forth the issue of third-party data interception, which can occur by accidental disclosure or deliberately through governmental access and subpoena. This finding accentuates previously reported concerns among MSM subgroups regarding the unintentional disclosure of sensitive information to nearby users [31] and further highlights the need for app developers to incorporate app features that address the privacy and confidentiality concerns. Other MSM subgroups have also emphasized their fear of being inadvertently “outed” within their respective communities if they use such platforms in public [50]. Interestingly, a discrepancy between privacy concerns and privacy practices (eg, privacy settings on mobile devices and apps) has also been noted, bringing forth a form of a privacy paradox [51].

In our study, participants stressed on incorporating additional security features in the app (eg, multifactor authentication) and additional features to adjust app visibility, notifications, and avatars to make them more discreet. Similar app features and recommendations were brought up by MSM in previous studies, including clear instructions in the app that prompt users to enable phone privacy features on their mobile device and warnings to view app content in private [52]. More importantly, using encrypted messages, colloquial expressions referring to same-sex behavior, or more neutral content seems to be a recurring suggestion [52,53]. To safeguard user privacy and confidentiality, app developers must communicate with the users about instances in which deidentification of information

is not possible and the risk of third-party access to or interception of the data collected via hacking, legal interception, accidental discovery, or telecommunication companies (eg, Google) [33].

Study participants also expressed their concern that access to this platform should be equitable among members of the MSM community, emphasizing the importance of ensuring that specific subgroups, such as hidden MSM and those with low socioeconomic status, have equal access to these services. Similar concerns have been noted in the literature among people living with HIV, whereby ethical considerations such as the accommodation of literacy, infrastructure, and access to technology have been delineated and are viewed as necessary to guarantee equitable access to mHealth services and care [54]. Many established strategies can be implemented to ensure app equity. One potential solution discussed includes the need to market the app to college LGBTQIA+ advocacy groups with greater reach and aptness to share such a platform with vulnerable populations, such as younger MSM, who would be more attentive to privacy risks and, therefore, less able to access such platforms. Adopting a human-centered design approach is also crucial, whereby app developers should increase the participation of participants from underprivileged backgrounds in developing mHealth products to gain insights into their preferences and priorities. In addition to promoting inclusivity by integrating population-specific design or app features, app developers can also promote digital literacy by teaching users how to use mHealth platforms, especially in low-resource or underprivileged subpopulations and settings [55]. Although mHealth technologies are rapidly evolving and despite the impact of sociodemographic inequities on equitable access, there is still a lack of evidence on the equity implications of HIV mHealth platforms. Nevertheless, it is an ethical imperative not to exclude subgroups that lack or have limited access to mHealth, including those who cannot afford to buy a smartphone or have access to the internet and those who might have an impairment that hinders mHealth platform use [33].

A common emerging theme in FGDs was the storage and management of collected data. Participants also referred to entities who will oversee data storage and management, expressing fears regarding data dissemination to third parties (eg, governmental agencies or the public) while maintaining a higher level of trust in general for entities outside of Malaysia as opposed to local entities. MSM’s preferences for eHealth interventions listed in the literature seem to align with these findings, as they expressed fear of disclosure of their HIV status when the intervention originated in organizations or places that they did not know about in the past. In addition, participants would tend to be skeptical about the information disclosed or dismiss messages stemming from such organizations because of the lack of trust [50]. In general, app developers are required to store and transmit the least amount of data possible, which allows the app’s purpose to be fulfilled and to be sensitive to the amount and ways the data are stored on the device itself or transmitted to the clinical team [33]. This ethical consideration has been considered especially sensitive as the data collected are usually detailed and originate from multiple sources that can generate identifiable information if data streams are

combined [56]. The unauthorized use of such data can jeopardize user safety and privacy if accessed by third parties [33,56].

Participants emphasized the importance of transparency via privacy policies and terms of service that clearly and explicitly detail which personally identifiable information will be collected, why it is collected, with whom it may be shared, and how users can control their data. It is essential that the participants be given a choice to consent to specific types of data collection and sharing activities [51], which corroborates the concept of “voluntariness” in the context of informed consent, whereby participants should be given a chance to consent to specific data collection streams or app modules but not others [51,56]. However, as there are instances in which opting out of some modules interferes with overall health or risk assessments, the risks and benefits associated with opting in or out of certain app modalities should be thoroughly explained [56]. In addition, using multimodal content, such as audio, video, or infographics, and summarizing key points in laypersons’ terms is encouraged to increase the engagement and comprehension of the users [56-59]. As data protection laws related to mobile apps change over time, these policy documents must be updated, and the users must be informed of such updates accordingly [56,60].

Finally, app-based platforms, particularly when intended to function as aids or alternatives to traditional medical services, carry increased risks to the users. There is a global debate on regulating mHealth apps if they are classified as medical devices. For example, in many high-income countries (eg, the United States and Australia), there are regulatory guidelines that provide oversight to mHealth apps that fulfill the definition of a medical device to ensure safety and effectiveness. However, these do not impede app dissemination, as most mobile apps can still be downloaded by patients via app stores without regulatory filters [33]. Moreover, there are no regulations in many low- and middle-income countries, including Malaysia, that guide app developers on consent, privacy protocols, and ethical practices in delivering health services. Therefore, a dedicated government body must be in place to provide regulatory oversight or “policing” to ensure the safety and effectiveness of such platforms. In addition, mHealth app developers must maintain transparency regarding the scientific evidence (or lack of evidence) underpinning the app’s effectiveness.

Limitations and Strengths

Our study has some limitations. First, our sample of MSM excludes those who are not active on GSN apps, do not have access to the internet, and do not read or understand English or Bahasa Malaysia. Furthermore, although theoretical saturation was reached after the third session, it is likely that additional ethical perceptions that were not brought up during the FGDs exist, which could be identified, for example, by recruiting participants in ways that differ from ours (ie, GSN app advertisements and direct profile inquiries). Second, most participants in our sample were of Chinese ethnicity, had tested for HIV, and were aware of PrEP, which limits the generalizability of our findings to the broader MSM group in Malaysia. Enrolling predominantly Malay participants or MSM who were not linked to HIV prevention services could yield different perspectives on ethical issues. Nevertheless, this study is among the first to examine the ethical issues surrounding the use of smartphone apps for HIV prevention among MSM in Malaysia, which is the first step toward understanding and engaging this specific subgroup. In addition, the web-based modality of the FGDs constituted a safe, confidential, and easily accessible platform for participants, thereby creating a safe environment and fostering a higher level of truthfulness among participants.

Conclusions

In countries such as Malaysia, where homosexuality is illegal, stigma and discrimination around HIV make traditional venue-based HIV prevention services less suitable for vulnerable populations, such as MSM. The findings from this study indicate the potential implications of ethical concerns and associated risks related to using app-based HIV prevention programs in low- and middle-income countries such as Malaysia. Such platforms can be kept safe by integrating security features (eg, multifactor authentication and discreet features to adjust app or notification visibility); storing and transmitting the least amount of data possible; ensuring transparency via clear privacy policies and terms of service; giving participants a choice to consent to specific types of data being collected; and, in the case of medical devices, ensuring adequate regulation by dedicated government bodies. Our findings indicate the need for the systematic engagement of all relevant stakeholders (eg, MSM, community members, and health care providers) and the adoption of an ethical framework in the design and development phases to ensure that the platform is ethical, safe, secure, equitable, and sustainable.

Acknowledgments

The authors acknowledge the financial support in part by a career development award from the National Institute on Drug Abuse (grant K01 DA051346) and a research grant from the Fogarty International Center (grant R21 TW011665). Data collection for this study was supported by the Fordham HIV Prevention Research Ethics Training Institute through a training grant sponsored by the National Institute on Drug Abuse (grant R25-DA031608). The funders had no role in the study design, data collection, data analysis, decision to publish, or manuscript writing.

Data Availability

The data sets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. UNAIDS Global AIDS Update 2022. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2022. URL: https://www.unaids.org/sites/default/files/media_asset/2022-global-aids-update_en.pdf [accessed 2022-11-16]
2. Global AIDS Update 2016. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2016. URL: https://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf [accessed 2022-08-05]
3. Prevention Gap Report. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2016. URL: https://www.unaids.org/sites/default/files/media_asset/2016-prevention-gap-report_en.pdf [accessed 2022-08-05]
4. UNAIDS Data 2017. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2017. URL: https://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf [accessed 2022-08-05]
5. Country Progress Report on HIV/AIDS 2019. Ministry of Health Malaysia. 2019. URL: [https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Report_GAM_2019_\(Final\).pdf](https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Report_GAM_2019_(Final).pdf) [accessed 2022-08-05]
6. Country Factsheets Malaysia 2021. Joint United Nations Programme on HIV/AIDS (UNAIDS). 2022. URL: <https://www.unaids.org/en/regionscountries/countries/malaysia> [accessed 2022-11-08]
7. Global AIDS Monitoring 2020, Malaysia HIV/AIDS Progress Report 2020. Ministry of Health Malaysia. 2020. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Laporan_Global_AIDS_Monitoring_2020_new.pdf [accessed 2022-08-10]
8. Global AIDS Monitoring, Country Progress Report, Malaysia. Ministry of Health Malaysia. 2021. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/20211130_MYS_country_report_2021.pdf [accessed 2022-08-10]
9. Chooi WT, Mohd Zaharim N, Desrosiers A, Ahmad I, Yasin MA, Syed Jaapar SZ, et al. Early initiation of amphetamine-type stimulants (ATS) use associated with lowered cognitive performance among individuals with co-occurring opioid and ATS use disorders in Malaysia. *J Psychoactive Drugs* 2017;49(4):326-332 [FREE Full text] [doi: [10.1080/02791072.2017.1342152](https://doi.org/10.1080/02791072.2017.1342152)] [Medline: [28661714](https://pubmed.ncbi.nlm.nih.gov/28661714/)]
10. Desrosiers A, Chooi WT, Zaharim NM, Ahmad I, Mohd Yasin MA, Syed Jaapar SZ, et al. Emerging drug use trends in Kelantan, Malaysia. *J Psychoactive Drugs* 2016;48(3):218-226. [doi: [10.1080/02791072.2016.1185553](https://doi.org/10.1080/02791072.2016.1185553)] [Medline: [27224011](https://pubmed.ncbi.nlm.nih.gov/27224011/)]
11. Singh D, Chawarski MC, Schottenfeld R, Vicknasingam B. Substance abuse and the HIV situation in Malaysia. *J Food Drug Anal* 2013 Dec;21(4):S46-S51 [FREE Full text] [doi: [10.1016/j.jfda.2013.09.033](https://doi.org/10.1016/j.jfda.2013.09.033)] [Medline: [25278737](https://pubmed.ncbi.nlm.nih.gov/25278737/)]
12. World Drug Report: 2016. United Nations Office on Drugs and Crime. 2016. URL: https://www.unodc.org/documents/scientific/WORLD_DRUG_REPORT_2016_web.pdf [accessed 2022-08-10]
13. Kanter J, Koh C, Razali K, Tai R, Izenberg J, Rajan L, et al. Risk behaviour and HIV prevalence among men who have sex with men in a multiethnic society: a venue-based study in Kuala Lumpur, Malaysia. *Int J STD AIDS* 2011 Jan;22(1):30-37. [doi: [10.1258/ijsa.2010.010277](https://doi.org/10.1258/ijsa.2010.010277)] [Medline: [21364064](https://pubmed.ncbi.nlm.nih.gov/21364064/)]
14. Lim SH, Akbar M, Wickersham JA, Kamarulzaman A, Altice FL. The management of methamphetamine use in sexual settings among men who have sex with men in Malaysia. *Int J Drug Policy* 2018 May;55:256-262 [FREE Full text] [doi: [10.1016/j.drugpo.2018.02.019](https://doi.org/10.1016/j.drugpo.2018.02.019)] [Medline: [29605540](https://pubmed.ncbi.nlm.nih.gov/29605540/)]
15. Habib S. Islam and Homosexuality. 2 volumes. Santa Barbara, CA, USA: ABC-CLIO; 2009.
16. Itaborahy LP, Zhu J. State-Sponsored Homophobia, A world survey of laws: Criminalisation, protection and recognition of same-sex love. International Lesbian, Gay, Bisexual, Trans and Intersex Association (ILGA). 2014. URL: <https://www.refworld.org/docid/519b6c2f4.html> [accessed 2022-08-25]
17. Earnshaw VA, Jin H, Wickersham J, Kamarulzaman A, John J, Altice FL. Exploring intentions to discriminate against patients living with HIV/AIDS among future healthcare providers in Malaysia. *Trop Med Int Health* 2014 Jun;19(6):672-679 [FREE Full text] [doi: [10.1111/tmi.12306](https://doi.org/10.1111/tmi.12306)] [Medline: [24666546](https://pubmed.ncbi.nlm.nih.gov/24666546/)]
18. Earnshaw VA, Jin H, Wickersham JA, Kamarulzaman A, John J, Lim SH, et al. Stigma toward men who have sex with men among future healthcare providers in Malaysia: would more interpersonal contact reduce prejudice? *AIDS Behav* 2016 Jan;20(1):98-106 [FREE Full text] [doi: [10.1007/s10461-015-1168-x](https://doi.org/10.1007/s10461-015-1168-x)] [Medline: [26324078](https://pubmed.ncbi.nlm.nih.gov/26324078/)]
19. Jin H, Earnshaw VA, Wickersham JA, Kamarulzaman A, Desai MM, John J, et al. An assessment of health-care students' attitudes toward patients with or at high risk for HIV: implications for education and cultural competency. *AIDS Care* 2014;26(10):1223-1228 [FREE Full text] [doi: [10.1080/09540121.2014.894616](https://doi.org/10.1080/09540121.2014.894616)] [Medline: [24625279](https://pubmed.ncbi.nlm.nih.gov/24625279/)]
20. Schwartz SR, Nowak RG, Orazulike I, Keshinro B, Ake J, Kennedy S, TRUST Study Group. The immediate effect of the Same-Sex Marriage Prohibition Act on stigma, discrimination, and engagement on HIV prevention and treatment services in men who have sex with men in Nigeria: analysis of prospective data from the TRUST cohort. *Lancet HIV* 2015 Jul;2(7):e299-e306 [FREE Full text] [doi: [10.1016/S2352-3018\(15\)00078-8](https://doi.org/10.1016/S2352-3018(15)00078-8)] [Medline: [26125047](https://pubmed.ncbi.nlm.nih.gov/26125047/)]
21. Vu L, Tun W, Sheehy M, Nel D. Levels and correlates of internalized homophobia among men who have sex with men in Pretoria, South Africa. *AIDS Behav* 2012 Apr;16(3):717-723. [doi: [10.1007/s10461-011-9948-4](https://doi.org/10.1007/s10461-011-9948-4)] [Medline: [21484279](https://pubmed.ncbi.nlm.nih.gov/21484279/)]

22. Wei C, Cheung DH, Yan H, Li J, Shi LE, Raymond HF. The impact of homophobia and HIV stigma on HIV testing uptake among Chinese men who have sex with men: a mediation analysis. *J Acquir Immune Defic Syndr* 2016 Jan 01;71(1):87-93 [FREE Full text] [doi: [10.1097/QAI.0000000000000815](https://doi.org/10.1097/QAI.0000000000000815)] [Medline: [26334742](https://pubmed.ncbi.nlm.nih.gov/26334742/)]
23. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e17776 [FREE Full text] [doi: [10.2196/17776](https://doi.org/10.2196/17776)] [Medline: [32217503](https://pubmed.ncbi.nlm.nih.gov/32217503/)]
24. MacKinnon GE, Brittain EL. Mobile health technologies in cardiopulmonary disease. *Chest* 2020 Mar;157(3):654-664 [FREE Full text] [doi: [10.1016/j.chest.2019.10.015](https://doi.org/10.1016/j.chest.2019.10.015)] [Medline: [31678305](https://pubmed.ncbi.nlm.nih.gov/31678305/)]
25. Marcolino MS, Oliveira JA, D'Agostino M, Ribeiro AL, Alkmim MB, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2018 Jan 17;6(1):e23 [FREE Full text] [doi: [10.2196/mhealth.8873](https://doi.org/10.2196/mhealth.8873)] [Medline: [29343463](https://pubmed.ncbi.nlm.nih.gov/29343463/)]
26. Mönninghoff A, Kramer JN, Hess AJ, Ismailova K, Teepe GW, Tudor Car L, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2021 Apr 30;23(4):e26699 [FREE Full text] [doi: [10.2196/26699](https://doi.org/10.2196/26699)] [Medline: [33811021](https://pubmed.ncbi.nlm.nih.gov/33811021/)]
27. Sales RO, Silva RM. mHealth in the prevention of sexually transmitted infections STIs. *Cien Saude Colet* 2020 Nov;25(11):4315-4325 [FREE Full text] [doi: [10.1590/1413-812320202511.11022019](https://doi.org/10.1590/1413-812320202511.11022019)] [Medline: [33175040](https://pubmed.ncbi.nlm.nih.gov/33175040/)]
28. Wang Y, Min J, Khuri J, Xue H, Xie B, Kaminsky LA, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2020 Apr 28;8(4):e15400 [FREE Full text] [doi: [10.2196/15400](https://doi.org/10.2196/15400)] [Medline: [32343253](https://pubmed.ncbi.nlm.nih.gov/32343253/)]
29. Shrestha R, Lim SH, Altice FL, Copenhaver M, Wickersham JA, Saifi R, et al. Use of smartphone to seek sexual health information online among Malaysian men who have sex with men (MSM): implications for mHealth intervention to increase HIV testing and reduce HIV risks. *J Community Health* 2020 Feb;45(1):10-19 [FREE Full text] [doi: [10.1007/s10900-019-00713-x](https://doi.org/10.1007/s10900-019-00713-x)] [Medline: [31375976](https://pubmed.ncbi.nlm.nih.gov/31375976/)]
30. Shrestha R, Maviglia F, Altice FL, DiDomizio E, Khati A, Mistler C, et al. Mobile health technology use and the acceptability of an mHealth platform for HIV prevention among men who have sex with men in Malaysia: cross-sectional respondent-driven sampling survey. *J Med Internet Res* 2022 Jul 25;24(7):e36917 [FREE Full text] [doi: [10.2196/36917](https://doi.org/10.2196/36917)] [Medline: [35877172](https://pubmed.ncbi.nlm.nih.gov/35877172/)]
31. Shrestha R, Fisher C, Wickersham JA, Khati A, Kim R, Azwa I, et al. Privacy and confidentiality concerns related to the use of mHealth apps for HIV prevention efforts among Malaysian men who have sex with men: cross-sectional survey study. *JMIR Form Res* 2021 Dec 16;5(12):e28311 [FREE Full text] [doi: [10.2196/28311](https://doi.org/10.2196/28311)] [Medline: [34924355](https://pubmed.ncbi.nlm.nih.gov/34924355/)]
32. Labrique AB, Kirk GD, Westergaard RP, Merritt MW. Ethical issues in mHealth research involving persons living with HIV/AIDS and substance abuse. *AIDS Res Treat* 2013;2013:189645 [FREE Full text] [doi: [10.1155/2013/189645](https://doi.org/10.1155/2013/189645)] [Medline: [24171110](https://pubmed.ncbi.nlm.nih.gov/24171110/)]
33. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. *JMIR Mhealth Uhealth* 2015 Oct 16;3(4):e95 [FREE Full text] [doi: [10.2196/mhealth.4538](https://doi.org/10.2196/mhealth.4538)] [Medline: [26474545](https://pubmed.ncbi.nlm.nih.gov/26474545/)]
34. Keenan AJ, Tsourtos G, Tieman J. The value of applying ethical principles in telehealth practices: systematic review. *J Med Internet Res* 2021 Mar 30;23(3):e25698 [FREE Full text] [doi: [10.2196/25698](https://doi.org/10.2196/25698)] [Medline: [33783366](https://pubmed.ncbi.nlm.nih.gov/33783366/)]
35. Fields BG. Regulatory, legal, and ethical considerations of telemedicine. *Sleep Med Clin* 2020 Sep;15(3):409-416 [FREE Full text] [doi: [10.1016/j.jsmc.2020.06.004](https://doi.org/10.1016/j.jsmc.2020.06.004)] [Medline: [32762973](https://pubmed.ncbi.nlm.nih.gov/32762973/)]
36. Kaplan B. Revisiting health information technology ethical, legal, and social issues and evaluation: telehealth/telemedicine and COVID-19. *Int J Med Inform* 2020 Nov;143:104239 [FREE Full text] [doi: [10.1016/j.ijmedinf.2020.104239](https://doi.org/10.1016/j.ijmedinf.2020.104239)] [Medline: [33152653](https://pubmed.ncbi.nlm.nih.gov/33152653/)]
37. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)] [Medline: [24979285](https://pubmed.ncbi.nlm.nih.gov/24979285/)]
38. Fisher CB, Bragard E, Bloom R. Ethical considerations in HIV eHealth intervention research: implications for informational risk in recruitment, data maintenance, and consent procedures. *Curr HIV/AIDS Rep* 2020 Jun;17(3):180-189 [FREE Full text] [doi: [10.1007/s11904-020-00489-z](https://doi.org/10.1007/s11904-020-00489-z)] [Medline: [32358768](https://pubmed.ncbi.nlm.nih.gov/32358768/)]
39. Emanuel EJ, Wendler D, Killen J, Grady C. What makes clinical research in developing countries ethical? The benchmarks of ethical research. *J Infect Dis* 2004 Mar 01;189(5):930-937. [doi: [10.1086/381709](https://doi.org/10.1086/381709)] [Medline: [14976611](https://pubmed.ncbi.nlm.nih.gov/14976611/)]
40. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods* 2016 Jul 24;29(1):3-22 [FREE Full text] [doi: [10.1177/1525822x16639015](https://doi.org/10.1177/1525822x16639015)]
41. Creswell JW, Creswell JD. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th edition. Thousand Oaks, CA, USA: Sage Publications; 2017.
42. Horsburgh D. Evaluation of qualitative research. *J Clin Nurs* 2003 Mar;12(2):307-312. [doi: [10.1046/j.1365-2702.2003.00683.x](https://doi.org/10.1046/j.1365-2702.2003.00683.x)] [Medline: [12603565](https://pubmed.ncbi.nlm.nih.gov/12603565/)]
43. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. WHO Global Observatory for eHealth. Geneva, Switzerland: World Health Organization; 2011. URL: <https://apps.who.int/iris/handle/10665/44607> [accessed 2022-08-25]

44. 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](https://doi.org/10.1371/journal.pmed.1001363)] [Medline: [23458994](https://pubmed.ncbi.nlm.nih.gov/23458994/)]
45. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. *Lancet* 2010 Nov 27;376(9755):1838-1845. [doi: [10.1016/S0140-6736\(10\)61997-6](https://doi.org/10.1016/S0140-6736(10)61997-6)] [Medline: [21071074](https://pubmed.ncbi.nlm.nih.gov/21071074/)]
46. Milward J, Lynskey M, Strang J. Solving the problem of non-attendance in substance abuse services. *Drug Alcohol Rev* 2014 Nov;33(6):625-636. [doi: [10.1111/dar.12194](https://doi.org/10.1111/dar.12194)] [Medline: [25196817](https://pubmed.ncbi.nlm.nih.gov/25196817/)]
47. Mbuagbaw L, Mursleen S, Lytvyn L, Smieja M, Dolovich L, Thabane L. Mobile phone text messaging interventions for HIV and other chronic diseases: an overview of systematic reviews and framework for evidence transfer. *BMC Health Serv Res* 2015 Jan 22;15:33 [FREE Full text] [doi: [10.1186/s12913-014-0654-6](https://doi.org/10.1186/s12913-014-0654-6)] [Medline: [25609559](https://pubmed.ncbi.nlm.nih.gov/25609559/)]
48. Ayala G, Makofane K, Santos G, Beck J, Do TD, Hebert P, et al. Access to basic HIV-related services and PrEP acceptability among men who have sex with men worldwide: barriers, facilitators, and implications for combination prevention. *J Sex Transm Dis* 2013;2013:953123 [FREE Full text] [doi: [10.1155/2013/953123](https://doi.org/10.1155/2013/953123)] [Medline: [26316968](https://pubmed.ncbi.nlm.nih.gov/26316968/)]
49. Chesney MA, Smith AW. Critical delays in HIV testing and care: the potential role of stigma. *Am Behav Sci* 1999 Apr 1;42(7):1162-1174 [FREE Full text] [doi: [10.1177/00027649921954822](https://doi.org/10.1177/00027649921954822)]
50. Hubach RD, O'Neil A, Stowe M, Giano Z, Curtis B, Fisher CB. Perceived confidentiality risks of mobile technology-based ecologic momentary assessment to assess high-risk behaviors among rural men who have sex with men. *Arch Sex Behav* 2021 May;50(4):1641-1650 [FREE Full text] [doi: [10.1007/s10508-019-01612-x](https://doi.org/10.1007/s10508-019-01612-x)] [Medline: [32078710](https://pubmed.ncbi.nlm.nih.gov/32078710/)]
51. Rendina HJ, Mustanski B. Privacy, trust, and data sharing in web-based and mobile research: participant perspectives in a large nationwide sample of men who have sex with men in the united states. *J Med Internet Res* 2018 Jul 04;20(7):e233 [FREE Full text] [doi: [10.2196/jmir.9019](https://doi.org/10.2196/jmir.9019)] [Medline: [29973332](https://pubmed.ncbi.nlm.nih.gov/29973332/)]
52. Zhao Y, Zhu X, Pérez AE, Zhang W, Shi A, Zhang Z, et al. MHealth approach to promote Oral HIV self-testing among men who have sex with men in China: a qualitative description. *BMC Public Health* 2018 Sep 27;18(1):1146 [FREE Full text] [doi: [10.1186/s12889-018-6046-9](https://doi.org/10.1186/s12889-018-6046-9)] [Medline: [30261856](https://pubmed.ncbi.nlm.nih.gov/30261856/)]
53. Nguyen LH, Nguyen HL, Larsson M, Tran BX, L Stein M, Rocha LE, et al. An exploratory assessment of the preference for eHealth interventions to prevent HIV and sexually transmitted infections among men who have sex with men in Hanoi, Vietnam. *BMC Public Health* 2020 Sep 11;20(1):1387 [FREE Full text] [doi: [10.1186/s12889-020-09449-z](https://doi.org/10.1186/s12889-020-09449-z)] [Medline: [32917157](https://pubmed.ncbi.nlm.nih.gov/32917157/)]
54. Duthely LM, Sanchez-Covarrubias AP. Digitized HIV/AIDS treatment adherence interventions: a review of recent SMS/texting mobile health applications and implications for theory and practice. *Front Commun (Lausanne)* 2020 Nov;5:530164 [FREE Full text] [doi: [10.3389/fcomm.2020.530164](https://doi.org/10.3389/fcomm.2020.530164)] [Medline: [33644162](https://pubmed.ncbi.nlm.nih.gov/33644162/)]
55. Fortuna KL, Kadakia A, Cosco TD, Rotondi A, Nicholson J, Mois G, et al. Guidelines to establish an equitable mobile health ecosystem. *Psychiatr Serv* 2022 Nov 15;appips202200011. [doi: [10.1176/appi.ps.202200011](https://doi.org/10.1176/appi.ps.202200011)] [Medline: [36377370](https://pubmed.ncbi.nlm.nih.gov/36377370/)]
56. Jacobson NC, Bentley KH, Walton A, Wang SB, Fortgang RG, Millner AJ, et al. Ethical dilemmas posed by mobile health and machine learning in psychiatry research. *Bull World Health Organ* 2020 Apr 01;98(4):270-276 [FREE Full text] [doi: [10.2471/BLT.19.237107](https://doi.org/10.2471/BLT.19.237107)] [Medline: [32284651](https://pubmed.ncbi.nlm.nih.gov/32284651/)]
57. Opportunities and challenges of developing information technologies on behavioral and social science clinical research. National Institute of Mental Health. 2017. URL: <https://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/opportunities-and-challenges-of-developing-information-technologies-on-behavioral-and-social-science-clinical-research> [accessed 2022-09-01]
58. Martinez-Martin N, Insel TR, Dagum P, Greely HT, Cho MK. Data mining for health: staking out the ethical territory of digital phenotyping. *NPJ Digit Med* 2018;1:68 [FREE Full text] [doi: [10.1038/s41746-018-0075-8](https://doi.org/10.1038/s41746-018-0075-8)] [Medline: [31211249](https://pubmed.ncbi.nlm.nih.gov/31211249/)]
59. Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Ment Health* 2018 Apr 23;5(2):e32 [FREE Full text] [doi: [10.2196/mental.9423](https://doi.org/10.2196/mental.9423)] [Medline: [29685865](https://pubmed.ncbi.nlm.nih.gov/29685865/)]
60. O'Reilly M, Parker N, Hutchby I. Ongoing processes of managing consent: the empirical ethics of using video-recording in clinical practice and research. *Clin Ethics* 2011 Dec 05;6(4):179-185 [FREE Full text] [doi: [10.1258/ce.2011.011040](https://doi.org/10.1258/ce.2011.011040)]

Abbreviations

FGD: focus group discussion

GSN: geosocial networking

LGBTQIA+: lesbian, gay, bisexual, queer, intersex, asexual, and others

mHealth: mobile health

MSM: men who have sex with men

PrEP: pre-exposure prophylaxis

Edited by A Mavragani; submitted 24.09.22; peer-reviewed by K Macapagal, V Watzlaf; comments to author 28.10.22; revised version received 19.11.22; accepted 28.11.22; published 23.12.22.

Please cite as:

Khati A, Wickersham JA, Rosen AO, Luces JRB, Copenhaver N, Jeri-Wahrhaftig A, Ab Halim MA, Azwa I, Gautam K, Ooi KH, Shrestha R

Ethical Issues in the Use of Smartphone Apps for HIV Prevention in Malaysia: Focus Group Study With Men Who Have Sex With Men

JMIR Form Res 2022;6(12):e42939

URL: <https://formative.jmir.org/2022/12/e42939>

doi: [10.2196/42939](https://doi.org/10.2196/42939)

PMID: [36563046](https://pubmed.ncbi.nlm.nih.gov/36563046/)

©Antoine Khati, Jeffrey A Wickersham, Aviana O Rosen, Jeffrey Ralph B Luces, Nicholas Copenhaver, Alma Jeri-Wahrhaftig, Mohd Akbar Ab Halim, Iskandar Azwa, Kamal Gautam, Kai Hong Ooi, Roman Shrestha. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Barriers and Facilitators of Obtaining Social Determinants of Health of Patients With Cancer Through the Electronic Health Record Using Natural Language Processing Technology: Qualitative Feasibility Study With Stakeholder Interviews

Jordan Alpert¹, PhD; Hyehyun (Julia) Kim², MA; Cara McDonnell³, MPH; Yi Guo³, PhD; Thomas J George⁴, MD; Jiang Bian³, PhD; Yonghui Wu³, PhD

¹Cleveland Clinic, Center for Value-Based Care Research, Cleveland, OH, United States

²College of Journalism and Communications, University of Florida, Gainesville, FL, United States

³Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United States

⁴Division of Hematology and Oncology, Department of Medicine, College of Medicine, University of Florida, Gainesville, FL, United States

Corresponding Author:

Jordan Alpert, PhD

Cleveland Clinic

Center for Value-Based Care Research

9500 Euclid Ave., mail code: G10

Cleveland, OH, 44195

United States

Phone: 1 216 527 5925

Email: alpertj@ccf.org

Abstract

Background: Social determinants of health (SDoH), such as geographic neighborhoods, access to health care, education, and social structure, are important factors affecting people's health and health outcomes. The SDoH of patients are scarcely documented in a discrete format in electronic health records (EHRs) but are often available in free-text clinical narratives such as physician notes. Innovative methods like natural language processing (NLP) are being developed to identify and extract SDoH from EHRs, but it is imperative that the input of key stakeholders is included as NLP systems are designed.

Objective: This study aims to understand the feasibility, challenges, and benefits of developing an NLP system to uncover SDoH from clinical narratives by conducting interviews with key stakeholders: (1) oncologists, (2) data analysts, (3) citizen scientists, and (4) patient navigators.

Methods: Individuals who frequently work with SDoH data were invited to participate in semistructured interviews. All interviews were recorded and subsequently transcribed. After coding transcripts and developing a codebook, the constant comparative method was used to generate themes.

Results: A total of 16 participants were interviewed (5 data analysts, 4 patient navigators, 4 physicians, and 3 citizen scientists). Three main themes emerged, accompanied by subthemes. The first theme, importance and approaches to obtaining SDoH, describes how every participant (n=16, 100%) regarded SDoH as important. In particular, proximity to the hospital and income levels were frequently relied upon. Communication about SDoH typically occurs during the initial conversation with the oncologist, but more personal information is often acquired by patient navigators. The second theme, SDoH exists in numerous forms, exemplified how SDoH arises during informal communication and can be difficult to enter into the EHR. The final theme, incorporating SDoH into health services research, addresses how more informed SDoH can be collected. One strategy is to empower patients so they are aware about the importance of SDoH, as well as employing NLP techniques to make narrative data available in a discrete format, which can provide oncologists with actionable data summaries.

Conclusions: Extracting SDoH from EHRs was considered valuable and necessary, but obstacles such as narrative data format can make the process difficult. NLP can be a potential solution, but as the technology is developed, it is important to consider how key stakeholders document SDoH, apply the NLP systems, and use the extracted SDoH in health outcome studies.

(JMIR Form Res 2022;6(12):e43059) doi:[10.2196/43059](https://doi.org/10.2196/43059)

KEYWORDS

natural language processing; qualitative; social determinants of health; electronic health records; cancer; technology; education; patient; clinical; communication; data

Introduction

The World Health Organization defines social determinants of health (SDoH) as “*non-medical factors that influence health outcomes, such as where people are born, live, learn, work, worship, and age that affect health, quality-of-life, and risks*” [1]. They can broadly be categorized as (1) health care access and quality, (2) education access and quality, (3) social and community, (4) economic stability, and (5) neighborhood and built environment [2]. Within these 5 key areas, other factors, such as smoking status, substance use, homelessness, and alcohol use are the most frequently studied SDoH categories [3]. Health outcomes are impacted by SDoH in various ways. For example, data show that there is a strong correlation between socioeconomic status and diabetes [4], frequency of health care visits [5], and mental health [6]. In fact, clinical and medical care only accounts for 10%-20% of an individual’s modifiable determinants to healthy outcomes, while the other 80-90% are SDoH [7].

Although the characteristics about patients’ lifestyles and behaviors have been included in the medical record since the origin of documentation in the 1800s [8], the shift to electronic records held the promise of improving the integration of SDoH into health care delivery systems [9]. Effectively leveraging SDoH within the electronic record can yield many benefits, including improved diagnosis and treatment plan, resulting in better health outcomes [3]. However, despite the rapid expansion of SDoH documentation tools in electronic health record (EHR) systems, difficulties remain about how to effectively capture and utilize SDoH data [10,11]. For instance, a systematic review of social determinants research and data quality found that data from the EHR are often inaccurate, incomplete, and incompatible [12]. While SDoH data can be derived from structured fields, clinicians often do not use them [10], and instead enter SDoH-related data into their notes. One potential way of documenting SDoH in EHRs is to use the International Classification of Diseases, Tenth Revision, Clinical Modification Z codes (Z55-Z65), since they are intended to document patients’ SDoH related to their socioeconomic, occupational, and psychosocial circumstances [13]. We conducted a retrospective analysis of EHR data between 2015 to 2018 using a large collection of EHRs from the OneFlorida Clinical Research Consortium and found a low rate of usage for these Z codes (270.61 per 100,000 at the encounter level and 2.03% at the patient level) [13]. Clinicians often document SDoH in clinical notes; however, they were not collected in a systematic, structured format, posing further challenges and limits to their usage [3,5].

To better capture SDoH data, studies highlight natural language processing (NLP) as an effective tool for extracting insights from unstructured data [5]. NLP refers to a branch of artificial intelligence that enables computers to understand text in the same manner as humans [14]. NLP can extract SDoH data from narrative clinical notes into discrete variables [3], which can

aid in the development of screening tools, risk prediction models, and clinical decision support systems. For example, including SDoH in risk prediction models not only improves model accuracy for hospitalization and death but also produces outcomes comparable to clinical factors [15]. Moreover, NLP produced a nearly 90-fold increase in identifying patients with significant SDoH problems compared to using structured EHR data elements alone [16]. We have systematically reviewed recent NLP studies for the extraction of SDoH [3] and developed NLP systems to identify SDoH from clinical notes [17]. Further, we have also assessed the extraction rates in patients with lung cancer [18] and studied how SDoH influence disparities of treatment selections in patients with diabetes [19].

Studies from our group and others show that it is feasible to use NLP to extract SDoH from clinical notes, but it is critical to understand what type of SDoH to collect and the best way to collect such data [20] when developing NLP solutions for SDoH extraction. Further, SDoH data generated from NLP are more useful if key stakeholders could incorporate the data to study real-world health outcomes. For example, the NLP extracted SDoH needs to be normalized to clinically meaningful categories (eg, stable housing, shelter, and homeless) related to outcomes of interest for analysis. How to standardize and populate NLP extracted SDoH concepts to common data models defined by large clinical research networks such as the Patient-Centered Clinical Research Network and the Observational Medical Outcomes Partnership remains unsolved. Engaging diverse stakeholders who document and frequently use SDoH can produce many benefits. Currently, there is a dearth of literature examining stakeholders’ perceptions of SDoH generated from NLP. This gap has resulted in NLP technology being developed without considering best practices for the oncologists and analysts who are the users of these NLP systems. Thus, the objective of this study is to understand the feasibility, challenges, and benefits of developing an NLP system to uncover SDoH data in EHRs. To best understand the facilitators and barriers, qualitative interviews were conducted with four key stakeholders: (1) oncologists, (2) data analysts, (3) citizen scientists, and (4) patient navigators. The following research questions were explored: what factors facilitate obtaining SDoH data? What are the challenges to obtaining SDoH data? How can SDoH data from EHR be applied to health services research and clinical care?

Methods

Setting and Study Design

This study took place at the University of Florida Health (UFHealth) in coordination with the University of Florida Health Cancer Center in Gainesville, Florida.

Participants

A form of purposive sampling, critical case sampling, was used to identify participants as the goal of the study was to assess a

phenomenon of interest at its very early stages [21]. The 4 groups identified as critical for understanding how SDoH data can be effectively used were (1) oncologists, (2) data analysts, (3) citizen scientists, who are members of the community that engages with researchers to improve the quality of health care, and (4) patient navigators, who are typically nurses who help guide patients throughout the diagnosis and treatment processes. Inclusion criteria were being at least 18 years old, fluent in English, and willing to provide informed consent. Members of the research team identified individuals within UFHealth who had experience working with SDoH data on the basis of their job titles and referrals from the research team's network. Once a master list was formed, they were contacted via email to participate. All participation was voluntary and was done without compensation.

Data Analysis

An interview guide was developed by the research team using a grounded theory approach, in which interview questions were general to cover a wide range of experiences and also narrow to explore specific experiences [22]. Group discussions centered on existing literature formed the basis for initial questions. Modifications were made to tailor a subset of specific questions for each group. Sample questions can be found in [Textbox 1](#).

Semistructured interviews were conducted because they allow for detailed information about a phenomenon to be obtained [23], as well as for the ability to immediately ask follow-up questions for clarification [23]. All interviews were conducted by 2 of the coauthors (JA and HK) using videoconferencing technology. Upon verbatim transcription of the video recording, the lead author (JA) performed primary cycle coding by reading 2 transcripts from each group [24]. Once an initial codebook was generated using the qualitative software ATLAS.ti, another coauthor (HK) read the same transcripts as the lead author, as well as one additional transcript from each group. A list of preliminary themes was presented to all of the other coauthors, and modifications were made after discussions. The remaining transcripts were read and coded and subjected to a process of constant comparison [25], and themes were generated. Interviews and data analysis continued until thematic salience occurred through the criteria of repetition, recurrence, and forcefulness [26]. During the analysis stage, interviews continued until saturation of the data was achieved [27]. Finally, another coauthor (CM) reviewed several transcripts and coded each one as a validation strategy. To confirm the findings from our analyses, self-reflecting memos that were recorded during interviews were verified [28] and to ensure trustworthiness, in vivo quotes were included [29].

Textbox 1. Sample interview questions.

Oncologists

- Describe the types of information you regularly gather during a typical interaction with a patient.
- How could we better obtain social determinant information from patients?
- What tool within the electronic health record (EHR) system would be helpful to improve the entry of social determinant variables?

Citizen scientists

- How do you think social determinants of health information is important for your health care provider?
- How does your provider usually ask about social determinants of health?
- What suggestions do you have for providers to more effectively learn about your social determinants of health?

Data analysts

- How can social determinants of health be important for health outcomes research?
- How do you work with other researchers and oncologists using data about social determinants of health?
- What role does the EHR system play in being equipped to provide data?

Patient navigators

- Describe your typical interaction with a patient.
- In your experience, are patients comfortable sharing information about social determinants?
- What is your experience reviewing a patient's EHR with them?

Ethical Considerations

This study was approved by the University of Florida institutional review board (IRB202002156). All procedures were performed in accordance with institutional guidelines regulations and human subject protections. Informed consent was reviewed with all participants prior to interviews and it was explained that participation was voluntary and they were free to withdraw at any time.

All data were transcribed and any identifiable information was removed from the record. Data were saved on secure Health Insurance Portability and Accountability Act–approved servers, and only members of the research team had access.

Results

Participant Characteristics

From August to September 2021, a total of 16 participants agreed to be interviewed (66% recruitment rate), consisting of 5 data analysts, 4 patient navigators, 4 oncologists, and 3 citizen scientists. All participants were based in the United States; their average age was 48 years, and most of them were female (63%) and White (69%). Among oncologists, the average time from fellowship was 24 years, ranging from 3 to 35 years. Interviews averaged 26 minutes in length and 160 pages of transcribed data were generated.

Theme 1: Importance and Approaches to Obtaining SDoH

Overview

The main theme of the importance and approaches to obtaining SDoH emerged, summarized by 2 subthemes. The first subtheme, importance of SDoH, focuses on how doctors, navigators, and citizen scientists value SDoH data and why they think they are essential to be collected. The second subtheme, SDoH solicitation during patient-provider communication, addresses how SDoH are woven into discussions that occur between patients and other members of the health care team such as navigators, social workers, and nurses.

Importance of SDoH Data

Every participant (n=16) across all stakeholder groups agreed that SDoH data were very important for patient care. When asked to name a specific type of SDoH data that may not be particularly important, participants struggled to provide an example. Doctors and patient navigators expressed how SDoH played a major role in the delivery of care. For instance, one particular SDoH, proximity to the hospital, was deemed crucial because as doctor #9 said, "If the patient lives further away...they're less likely to make it to the appointments. Less likely to make it on time. More likely to miss appointments and have subsequent negative outcomes." Another SDoH, social structure, was related to helping patients get to the hospital as well as managing side effects. Doctor #9 said, "Our treatments can be physically and mentally debilitating...they see me or my colleagues for maybe 15 to 20 minutes...They spend the majority of time outside of our clinic and having someone that they can rely on to help them with their symptoms." Citizen scientist #4 recalled her experience as a patient and the importance of social support. She said, "Thank God I had the support of my family, my mom, my sister. That was a big one for me just with living with [the disease]." Doctors can use information about patients' social structure to effectively communicate with patients and family members, as well as bring clarity as to who can assist patients the most. For example, doctor #10 provided the following example:

The individual that's accompanying the patient every day to clinic, is probably not the one who's doing a lot of the heavy lifting or at least a lot of the heavy organizing. Maybe he is there for the day-to-day things but he's not he's not the person that needs to

be involved with major life transition points in her care. So that's been, that's been really, really insightful.

Other SDoH considered important were income levels and geographic location. All data analysts (n=5, 100%) mentioned that zip codes were the most frequently requested data point by health service researchers. Zip codes were powerful because they usually are connected to geography. Analyst #12 said, "Zip code is probably the most important one because you can link that back to job opportunities [and] someone's almost entire socio-economic status just from where they live." Another analyst (#13) referred to how income and geography were connected due to "air quality [and] the stresses of the environment."

SDoH Solicitation During Patient-Oncologist Communication

The process of soliciting SDoH information usually occurred during initial consultations. Traditionally, in a formal procedure conducted at the first visit, doctors ask standardized questions, prompted by smart forms in the EHR. Smart forms allow oncologists to enter data, usually in the form of a drop-down menu. However, time restraints are an obstacle during initial visits, resulting in not all SDoH being collected. Doctor #9 said, "It can be hard to get into everything in that first visit...that first visit is a pretty packed visit." Doctors found that additional information about patients surfaced once the patient-oncologist relationship was better established. For instance, doctor #10 said, "You have to be prepared to listen for it (SDoH) and then and then take that opportunity since they brought it up to let your foot in the door and pursue it a little bit more." Citizen scientists agreed with this sentiment and recognized the importance of doctors getting to know their patients. Citizen scientist #4 recommended that doctors should "make the patient feel like you're there for them. You want to know what's really going on, make them feel comfortable." Doctor #9 said the following:

Subsequent visits I'm able to have more of that kind of conversational approach about asking them about other aspects of their care and I think that is also because...with subsequent visits, I've built that relationship. Sometimes it can be difficult to ask serious and personal questions to the patient with that first visit.

Patient navigators also obtain SDoH data from patients during conversations, but the nature of the navigator-patient relationship sometimes allows for the acquisition of more personal detail. Navigators often form a close bond with patients, which differs from patient-oncologist interactions. Navigator #8 mentioned that patients confide in them because of the high level of trust that is formed. Citizen scientist #5 reiterated this notion, saying, "Can you imagine your doctor asking 'do you have enough money to buy groceries? Is your neighborhood dangerous? Is your sex life interesting?' They don't go there." Navigators either alert oncologists or enter SDoH data directly into the EHR under their own notes. Doctor #10 noted the importance of social workers discussing SDoH with patients because, in addition to helping with things like travel or insurance, they can

also be “a shoulder for [patients] to cry on.” Overall, SDoH solicitation more realistically occurs on a continuous basis during the relationship between the patient and oncologist.

Theme 2: SDoH Exists in Numerous Forms

Overview

While it is common to consider SDoH as medical information obtained during a consultation, it can transpire in other ways. The first subtheme, informal communication, describes how not all SDoH are included in the EHR, while the second subtheme, the need for SDoH expertise and knowledge, explains the ways in which SDoH can be uncovered.

Informal Communication

Although smart forms were cited as a sufficient method of capturing SDoH data, patient information can be communicated informally, placing the onus on oncologists to separately enter data that they deem significant. Doctor #9 commented, “If I do elicit something from [a] patient encounter that I think is really important, I'll put it in the note. But I think the issue is that there's no section in our typical note template.” As a result, doctors were uncertain about where to include such information in the note. Similarly, another doctor (#15) said, “I don't put it in the template...I put what I think is important to me [in] the first paragraph.” In addition to communication that occurred during examinations, SDoH can also be uncovered during alternative methods of communication. Doctor #10 said, “MyChart messaging [gave] me insights that I didn't already have in the clinic.” For instance, the doctor mentioned the following:

When there's consistently another caregiver in their social system that's engaging with us on the patient's behalf...maybe these are individuals that I haven't met...but clearly are intimately involved in the day-to-day support of the patient.

Since SDoH can come from many different sources, like messaging as well as additional notes input by navigators, it is imperative that they are properly included in the EHR. Although smart forms seem easier to enter SDoH data, doctor #10 noticed that “every discrete variable requires at least 15 clicks.” As a less time-consuming alternative, doctors will type in data, which are considered unstructured data, because the information is not necessarily included in smart forms. To get such data into the EHR, manual documentation is often required. It is common for oncologists to copy and paste previous notes, but doing so many omit new SDoH that arise during recent examinations.

Expertise and Knowledge Needed

Inputting data manually creates obstacles to extracting SDoH data. Data analyst #13 said, “Having unstructured data would be a lot more difficult for the end-user because a lot of researchers are not going to have skills...to try to actually get the data out that they need.” Deciphering unstructured SDoH data are problematic because as data analyst #12 asked, “A lot of that stuff (SDoH) is...stored in a notes section...how do we get it out of the notes and then how do we put it into a structured format?” Analysts are forced to get creative to identify SDoH that are often requested by researchers. For instance, data analyst

#12 continued, “We don't have education status. We don't have income levels or anything for individual patients...we can infer that stuff from zip code but that only goes so far.” Doctor #10 stated, “We wouldn't even consider not documenting a patient's past surgical history, but [SDoH] are not captured in the [EHR] optimally...because it requires too much manual labor to type into and put into the system.”

When NLP was proposed as a solution, analysts were intrigued, but none of the 5 analysts had the expertise to extract data that way. Data analyst #17 said, “I know what NLP does and I have played with some stuff before, but it's not something that I do in my work.” Another analyst (#14) had no experience with NLP but said, “It's an interesting area of the field, but I haven't personally worked with it.” Doctors, like participant #15, were skeptical that NLP could be a solution soon. He said, “It's not ready for prime time, but there might be beta stuff that I'm just not aware of that's ready to go.”

Theme 3: Incorporating SDoH Into Health Services Research

Overview

Given the challenges of obtaining and extracting SDoH, combined with the importance of SDoH to treat patients, it is necessary to discover methods of incorporating data into research to expand their impact. Two subthemes emerged that addressed possible solutions to improve how SDoH data can be incorporated into health services research and patient care: (1) empower patients and (2) actionable data.

Empower Patients

Truly understanding the patient as a person requires that SDoH, such as the neighborhood in which they reside, occupation, access to health, and social support, are identified and integrated into the care plan. Although oncologists solicit information from patients, there are opportunities to increase patient participation in the process. For instance, doctor #9 said, “We're always interested in getting information from the patients directly, but I think it's not a bad idea to have the patients voluntarily answer questions.” The doctor elaborated that surveys could be distributed via email up to a week before meeting with the doctor to understand more about the patient's background and environment. Doing so would “take a lot of the burden off of [oncologists].” Another doctor (#15) thought that patients should be more involved and have the ability to clarify information about themselves because “there can be innocent errors or there's a certain amount of incorrect information all the time.” However, patients have other opportunities to fill out forms with information about themselves, but because they are optional, the majority of patients choose not to. As a result, a data analyst (#16) said, “We don't see much structured data... I do know there is this one survey that has many more structured questions, but not many people have filled [it] out.” Doctor #10 reinforced this notion by saying the following:

We've done away with a paper system of patients filling out information when they're in the waiting room which is another great missed opportunity for patients to get data into the system themselves that we could then verify.

Actionable Data

To involve patients, they must be made aware of how SDoH can affect them, but they also must understand the data. Citizen scientist #3 reflected upon her own family members who have high school educations and recognized that they might not have the ability to process health information as easily as others. Inputting or viewing SDoH data is even more helpful when patients can take action. For example, doctor #9 thought it would be beneficial for patients to have someone that they could talk with about SDoH for the following reason:

I think a lot of patients aren't aware of all the resources that are available to them whether it's patient assistance programs for drugs, or whether it's support groups... [or] financial programs.

NLP innovations were welcomed as a potential solution to aggregate individual-level SDoH data and as a way to make them more prominent for both patients and oncologists. Data analysts were enthusiastic about the prospect of acquiring additional data to work with, but were curious about the process of accurately validating data. Oncologists also realized the potential efficiencies of NLP. Referring to NLP, doctor #9 commented, "It's automatic. Doesn't require anything from us. I think it would be great... You're better able to incorporate the right management and the individualized care that the patient needs with regards to those social determinants." Oncologists also considered where and how the data would be presented. For instance, doctor #10 speculated, "If the NLP system can identify the data can pull it out, what happens to it? Does it end up getting put into the discrete places in the EHR where it should have been to begin with?" Citizen scientists were cautiously optimistic about interpreting data derived from NLP. A citizen scientist (#5) provided the following example:

Textbox 2. Recommendations based on each theme.

Importance of social determinants of health (SDoH) data

- SDoH should be prioritized and as much detail as possible should be included when inputting data into the electronic health record (EHR). Smart forms should be filled out and additional information about patients, such as their lifestyle and issues they are currently confronting, should also be added.

SDoH solicitation during patient-provider communication

- Although most SDoH are input into the EHR after the initial consultation, valuable SDoH are provided by patients in subsequent visits as the patient-provider relationship grows. Richer detail about patients' lives should be continuously added to the EHR.

Informal communication

- While smart forms might guide the discussion, patients offer clues about their lifestyle throughout the visit. If a specific variable is not available via the smart form, it should be entered within the notes.

The need for SDoH expertise and knowledge

- Everyone involved in the research analysis process should be briefed and educated about technology to identify and extract SDoH.

Empower patients

- Increase patients' awareness about how sharing SDoH to providers can positively influence their care.

Actionable data

- Involve patients in the implications of SDoH and how they can directly affect decisions that are made about their care.

If I were looking at doctor's notes, and I said something about... [feeling] paranoid about my next-door neighbor...the word paranoid might appear in my record. What a misstatement that would have been for me...that could be coded into something that was not really a diagnosis for me. I worry about that process a whole lot.

Discussion

Principal Results

We interviewed 16 stakeholders with various involvements related to SDoH, including oncologists, data analysts, patient navigators, and citizen scientists. The findings revealed that no SDoH data point was deemed as unimportant. Geography, social structure, and income were most accessible and, therefore, used most often. Although the importance of SDoH on population health are known, SDoH have recently taken on a heightened level of importance, brought about by the COVID-19 pandemic. For instance, those with a lower socioeconomic status or those living with comorbidities were more vulnerable to infection [30]. Machine learning is frequently being used to uncover such SDoH [31], but it is not clear whether it can lead to improved patient experiences and more informed research [32]. Moreover, it is possible that techniques such as NLP may increase social biases [32]. Our study included the insights from multiple perspectives to help create a path for using NLP to take bias into account and can positively impact clinical encounters as well as health services research. [Textbox 2](#) pairs each theme identified from the interviews with a direct actionable recommendation.

Our results indicate that the method of obtaining SDoH from patients varied, as data arose during both formal and informal patient-provider communication and across the time line of care delivery. Physicians acquire SDoH during initial consultations, but due to time restraints, capturing all relevant SDoH is unlikely. SDoH can also be found within secure messages as well as during interactions with patient navigators. Previous studies show that patients' self-disclosure depends on the gender of physicians and their own willingness to be open and disclose information to patients [33]. Additionally, the way questions are framed can dictate patients' willingness to respond to sensitive questions about SDoH [34]. Therefore, relationship-centered communication has been suggested as a method for acquiring SDoH and sensitive information from patients [34]. This can occur by monitoring negative attitudes, displaying empathy, and honoring patients' preferences [35]. During our interviews, patient navigators mentioned that they often receive patient information that is not disclosed to the doctor because of their close relationship with the patient.

Since most navigators have access to the EHR and can enter their own notes, it is important to use unstructured data to capture SDoH. NLP has produced valid results [36], but as algorithms become more accurate, it is necessary to involve data analysts who have historically worked with SDoH to support health services research. NLP requires a different skill set than what most data analysts are accustomed to. However, we discovered that data analysts were excited about the prospect of using NLP and enthusiastic about its capabilities. NLP should not be confined to programmers and machine learning experts. Data analysts should understand the capabilities and limitations of working with NLP while the technology is being perfected.

Lastly, interviews revealed that to better incorporate SDoH data into health services research and clinical care, patients should

be empowered to understand how such data can impact their health and that valuable resources might be available upon disclosure of the information. Since time restraints are a common barrier to collecting SDoH [37], innovations that allow for patients to enter information about themselves directly into the EHR should be explored, allowing it to be entered once and be available to all members of their health care team. Patients often find errors in their health record [38], so the ability to review and edit or update the information is critical. While patient portal enrollment continues to increase, groups most affected by SDoH are often less likely to enroll in portals and have access to their health record [39].

Limitations

This study includes several limitations. First, a fairly small sample size was used. Second, all interviews were conducted at the same health system. Therefore, observations from participants may be confined to the specific procedures of the health system and not applicable to stakeholders in other facilities. Third, there was an element of selection bias, as all participants volunteered to participate after being informed about the topic. It is possible that stakeholders who did not see value in NLP or SDoH, in general, chose not to participate.

Conclusions

SDoH data are extremely valuable for patient care but can be difficult to access as a result of the way unstructured data are entered into the EHR. NLP can ease the burden on oncologists by identifying hidden SDoH data within the EHR while enabling analysts to easily extract requested data for health outcomes research. However, maintaining high levels of quality for SDoH data entered into the EHR is imperative. Processes should be developed to facilitate acquiring SDoH from patients as well as educating patients about the importance of SDoH.

Acknowledgments

This study was funded Patient-Centered Outcomes Research Institute (PCORI) Award (ME-2018C3-14754).

Data Availability

Deidentified data can be made available upon reasonable request.

Conflicts of Interest

None declared.

References

1. Social determinants of health. World Health Organization. 2022. URL: https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1 [accessed 2022-05-24]
2. Social determinants of health. Healthy people 2030. 2022. URL: <https://health.gov/healthypeople/priority-areas/social-determinants-health> [accessed 2022-09-27]
3. Patra B, Sharma MM, Vekaria V, Adekkanattu P, Patterson OV, Glicksberg B, et al. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. *J Am Med Inform Assoc* 2021 Nov 25;28(12):2716-2727 [FREE Full text] [doi: [10.1093/jamia/ocab170](https://doi.org/10.1093/jamia/ocab170)] [Medline: [34613399](https://pubmed.ncbi.nlm.nih.gov/34613399/)]
4. Walker RJ, Strom Williams J, Egede LE. Influence of race, ethnicity and social determinants of health on diabetes outcomes. *Am J Med Sci* 2016 Apr;351(4):366-373 [FREE Full text] [doi: [10.1016/j.amjms.2016.01.008](https://doi.org/10.1016/j.amjms.2016.01.008)] [Medline: [27079342](https://pubmed.ncbi.nlm.nih.gov/27079342/)]

5. Hatef E, Ma X, Rouhizadeh M, Singh G, Weiner JP, Kharrazi H. Assessing the impact of social needs and social determinants of health on health care utilization: using patient- and community-level data. *Popul Health Manag* 2021 Apr 01;24(2):222-230 [[FREE Full text](#)] [doi: [10.1089/pop.2020.0043](https://doi.org/10.1089/pop.2020.0043)] [Medline: [32598228](https://pubmed.ncbi.nlm.nih.gov/32598228/)]
6. Prokosch C, Fertig AR, Ojebuoboh AR, Trofholz AC, Baird M, Young M, et al. Exploring associations between social determinants of health and mental health outcomes in families from socioeconomically and racially and ethnically diverse households. *Prev Med* 2022 Aug;161:107150. [doi: [10.1016/j.ypmed.2022.107150](https://doi.org/10.1016/j.ypmed.2022.107150)] [Medline: [35809824](https://pubmed.ncbi.nlm.nih.gov/35809824/)]
7. Magnan S. Social determinants of health 101 for health care: five plus five. In: *NAM Perspectives*. Washington, DC: National Academy of Medicine; Oct 9, 2017.
8. Gillum RF. From papyrus to the electronic tablet: a brief history of the clinical medical record with lessons for the digital age. *Am J Med* 2013 Oct;126(10):853-857. [doi: [10.1016/j.amjmed.2013.03.024](https://doi.org/10.1016/j.amjmed.2013.03.024)] [Medline: [24054954](https://pubmed.ncbi.nlm.nih.gov/24054954/)]
9. Gottlieb LM, Tirozzi KJ, Manchanda R, Burns AR, Sandel MT. Moving electronic medical records upstream: incorporating social determinants of health. *Am J Prev Med* 2015 Feb;48(2):215-218. [doi: [10.1016/j.amepre.2014.07.009](https://doi.org/10.1016/j.amepre.2014.07.009)] [Medline: [25217095](https://pubmed.ncbi.nlm.nih.gov/25217095/)]
10. Wang M, Pantell MS, Gottlieb LM, Adler-Milstein J. Documentation and review of social determinants of health data in the EHR: measures and associated insights. *J Am Med Inform Assoc* 2021 Nov 25;28(12):2608-2616 [[FREE Full text](#)] [doi: [10.1093/jamia/ocab194](https://doi.org/10.1093/jamia/ocab194)] [Medline: [34549294](https://pubmed.ncbi.nlm.nih.gov/34549294/)]
11. Stemerman R, Arguello J, Brice J, Krishnamurthy A, Houston M, Kitzmiller R. Identification of social determinants of health using multi-label classification of electronic health record clinical notes. *JAMIA Open* 2021 Jul;4(3):ooaa069 [[FREE Full text](#)] [doi: [10.1093/jamiaopen/ooaa069](https://doi.org/10.1093/jamiaopen/ooaa069)] [Medline: [34514351](https://pubmed.ncbi.nlm.nih.gov/34514351/)]
12. Cook L, Sachs J, Weiskopf NG. The quality of social determinants data in the electronic health record: a systematic review. *J Am Med Inform Assoc* 2021 Dec 28;29(1):187-196 [[FREE Full text](#)] [doi: [10.1093/jamia/ocab199](https://doi.org/10.1093/jamia/ocab199)] [Medline: [34664641](https://pubmed.ncbi.nlm.nih.gov/34664641/)]
13. Guo Y, Chen Z, Xu K, George TJ, Wu Y, Hogan W, et al. International Classification of Diseases, tenth revision, clinical modification social determinants of health codes are poorly used in electronic health records. *Medicine* 2020 Dec 24;99(52):e23818 [[FREE Full text](#)] [doi: [10.1097/MD.00000000000023818](https://doi.org/10.1097/MD.00000000000023818)] [Medline: [33350768](https://pubmed.ncbi.nlm.nih.gov/33350768/)]
14. Natural Language Processing (NLP). IBM Cloud Education. 2020. URL: <https://www.ibm.com/cloud/learn/natural-language-processing> [accessed 2022-08-19]
15. Hammond G, Johnston K, Huang K, Joynt Maddox KE. Social determinants of health improve predictive accuracy of clinical risk models for cardiovascular hospitalization, annual cost, and death. *Circ: Cardiovascular Quality and Outcomes* 2020 Jun;13(6):e006752. [doi: [10.1161/circoutcomes.120.006752](https://doi.org/10.1161/circoutcomes.120.006752)]
16. Dorr D. Identifying patients with significant problems related to social determinants of health with natural language processing. In: *MEDINFO 2019: Health and Wellbeing e-Networks for All*. Washington, DC: IOS Press; 2019:1456-1457.
17. Yu Z, Yang X, Dang C, Wu S, Adekanattu P, Pathak J, et al. A study of social and behavioral determinants of health in lung cancer patients using transformers-based natural language processing models. *AMIA Annu Symp Proc* 2021;2021:1225-1233 [[FREE Full text](#)] [Medline: [35309014](https://pubmed.ncbi.nlm.nih.gov/35309014/)]
18. Yu Z, Yang X, Guo Y, Bian J, Wu Y. Assessing the documentation of social determinants of health for lung cancer patients in clinical narratives. *Front Public Health* 2022 Mar 28;10:778463 [[FREE Full text](#)] [doi: [10.3389/fpubh.2022.778463](https://doi.org/10.3389/fpubh.2022.778463)] [Medline: [35419333](https://pubmed.ncbi.nlm.nih.gov/35419333/)]
19. Guo J, Wu Y, Guo Y, Chen A, Yu Z, Yang X, et al. Abstract P108: natural language processing extracted social and behavioral determinants of health and newer glucose-lowering drug initiation among real-world patients with type 2 diabetes. *Circulation* 2022 Mar;145(Suppl_1):AP108. [doi: [10.1161/circ.145.suppl_1.p108](https://doi.org/10.1161/circ.145.suppl_1.p108)]
20. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1764-1773 [[FREE Full text](#)] [doi: [10.1093/jamia/ocaa143](https://doi.org/10.1093/jamia/ocaa143)] [Medline: [33202021](https://pubmed.ncbi.nlm.nih.gov/33202021/)]
21. Rai N, Thapa B. *A Study on Purposive Sampling Method in Research*. Kathmandu: Kathmandu School of Law; 2015:5.
22. Charmaz K, Belgrave L. Qualitative interviewing and grounded theory analysis. In: Gubrium JF, Holstein JA, Marvasti AB, McKinney KD, editors. *The Sage Handbook of Interview Research: The Complexity of the Craft*. Thousand Oaks, CA: SAGE; 2012:347-364.
23. Billups FD. *Qualitative Data Collection Tools: Design, Development, and Applications*. Vol. 55. Thousand Oaks, CA: Sage Publications; 2019.
24. Tracy SJ. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. Hoboken, NJ: John Wiley & Sons; 2019.
25. Glaser BG. The constant comparative method of qualitative analysis. *Social Probl* 1965 Apr;12(4):436-445. [doi: [10.2307/798843](https://doi.org/10.2307/798843)]
26. Owen WF. Interpretive themes in relational communication. *Q J Speech* 2009 Jun 05;70(3):274-287. [doi: [10.1080/00335638409383697](https://doi.org/10.1080/00335638409383697)]
27. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. *Int J Qual Meth* 2016 Nov 30;1(2):13-22. [doi: [10.1177/160940690200100202](https://doi.org/10.1177/160940690200100202)]
28. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. In: Thyer B, editor. *The Handbook of Social Work Research Methods*. Newbury Park, CA: Sage; Jan 1985.

29. Castleberry A, Nolen A. Thematic analysis of qualitative research data: is it as easy as it sounds? *Curr Pharm Teach Learn* 2018 Jun;10(6):807-815. [doi: [10.1016/j.cptl.2018.03.019](https://doi.org/10.1016/j.cptl.2018.03.019)] [Medline: [30025784](https://pubmed.ncbi.nlm.nih.gov/30025784/)]
30. Venkataramany B, Sutton JM. The importance of social determinants of health for cancer patients in the era of COVID-19. *Cureus* 2022 Aug;14(8):e27993 [FREE Full text] [doi: [10.7759/cureus.27993](https://doi.org/10.7759/cureus.27993)] [Medline: [36120243](https://pubmed.ncbi.nlm.nih.gov/36120243/)]
31. Hatef E, Ma X, Rouhizadeh M, Singh G, Weiner JP, Kharrazi H. Assessing the impact of social needs and social determinants of health on health care utilization: using patient- and community-level data. *Popul Health Manage* 2021 Apr 01;24(2):222-230. [doi: [10.1089/pop.2020.0043](https://doi.org/10.1089/pop.2020.0043)]
32. Kino S, Hsu Y, Shiba K, Chien Y, Mita C, Kawachi I, et al. A scoping review on the use of machine learning in research on social determinants of health: trends and research prospects. *SSM Popul Health* 2021 Sep;15:100836 [FREE Full text] [doi: [10.1016/j.ssmph.2021.100836](https://doi.org/10.1016/j.ssmph.2021.100836)] [Medline: [34169138](https://pubmed.ncbi.nlm.nih.gov/34169138/)]
33. Kadji K, Schmid Mast M. The effect of physician self-disclosure on patient self-disclosure and patient perceptions of the physician. *Patient Educ Couns* 2021 Sep;104(9):2224-2231 [FREE Full text] [doi: [10.1016/j.pec.2021.02.030](https://doi.org/10.1016/j.pec.2021.02.030)] [Medline: [33775499](https://pubmed.ncbi.nlm.nih.gov/33775499/)]
34. Schoenthaler A, Hassan I, Fiscella K. The time is now: fostering relationship-centered discussions about patients' social determinants of health. *Patient Educ Counsel* 2019 Apr;102(4):810-814. [doi: [10.1016/j.pec.2018.10.025](https://doi.org/10.1016/j.pec.2018.10.025)]
35. Beach MC, Inui T, Relationship-Centered Care Research Network. Relationship-centered care. A constructive reframing. *J Gen Intern Med* 2006 Jan;21 Suppl 1(Suppl 1):S3-S8 [FREE Full text] [doi: [10.1111/j.1525-1497.2006.00302.x](https://doi.org/10.1111/j.1525-1497.2006.00302.x)] [Medline: [16405707](https://pubmed.ncbi.nlm.nih.gov/16405707/)]
36. Hatef E, Rouhizadeh M, Nau C, Xie F, Rouillard C, Abu-Nasser M, et al. Development and assessment of a natural language processing model to identify residential instability in electronic health records' unstructured data: a comparison of 3 integrated healthcare delivery systems. *JAMIA Open* 2022 Apr;5(1):ooac006 [FREE Full text] [doi: [10.1093/jamiaopen/ooac006](https://doi.org/10.1093/jamiaopen/ooac006)] [Medline: [35224458](https://pubmed.ncbi.nlm.nih.gov/35224458/)]
37. Schickedanz A, Hamity C, Rogers A, Sharp AL, Jackson A. Clinician experiences and attitudes regarding screening for social determinants of health in a large integrated health system. *Med Care* 2019 Jun;57 Suppl 6 Suppl 2(Suppl 6 2):S197-S201 [FREE Full text] [doi: [10.1097/MLR.0000000000001051](https://doi.org/10.1097/MLR.0000000000001051)] [Medline: [31095061](https://pubmed.ncbi.nlm.nih.gov/31095061/)]
38. Bell SK, Delbanco T, Elmore JG, Fitzgerald PS, Fossa A, Harcourt K, et al. Frequency and types of patient-reported errors in electronic health record ambulatory care notes. *JAMA Netw Open* 2020 Jun 01;3(6):e205867 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.5867](https://doi.org/10.1001/jamanetworkopen.2020.5867)] [Medline: [32515797](https://pubmed.ncbi.nlm.nih.gov/32515797/)]
39. Sinha S, Garriga M, Naik N, McSteen BW, Odisho AY, Lin A, et al. Disparities in electronic health record patient portal enrollment among oncology patients. *JAMA Oncol* 2021 Jun 01;7(6):935-937 [FREE Full text] [doi: [10.1001/jamaoncol.2021.0540](https://doi.org/10.1001/jamaoncol.2021.0540)] [Medline: [33830178](https://pubmed.ncbi.nlm.nih.gov/33830178/)]

Abbreviations

EHR: electronic health record

NLP: natural language processing

SDoH: social determinants of health

Edited by A Mavragani; submitted 28.09.22; peer-reviewed by M Kapsetaki, S Herzig van Wees; comments to author 11.11.22; revised version received 01.12.22; accepted 05.12.22; published 27.12.22.

Please cite as:

Alpert J, Kim H, McDonnell C, Guo Y, George TJ, Bian J, Wu Y

Barriers and Facilitators of Obtaining Social Determinants of Health of Patients With Cancer Through the Electronic Health Record Using Natural Language Processing Technology: Qualitative Feasibility Study With Stakeholder Interviews

JMIR Form Res 2022;6(12):e43059

URL: <https://formative.jmir.org/2022/12/e43059>

doi: [10.2196/43059](https://doi.org/10.2196/43059)

PMID: [36574288](https://pubmed.ncbi.nlm.nih.gov/36574288/)

©Jordan Alpert, Hye Hyun (Julia) Kim, Cara McDonnell, Yi Guo, Thomas J George, Jiang Bian, Yonghui Wu. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 27.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Effectiveness of Secondary Risk–Reducing Strategies in Patients With Unilateral Breast Cancer With Pathogenic Variants of BRCA1 and BRCA2 Subjected to Breast-Conserving Surgery: Evidence-Based Simulation Study

Jelena Maksimenko¹, MD, PhD; Pedro Pereira Rodrigues², MSc, PhD; Miki Nakazawa-Miklaševiča³, MSc, PhD; David Pinto⁴, MD; Edvins Miklaševičs³, Prof Dr biol; Genadijs Trofimovičs⁵, MD, PhD; Jānis Gardovskis⁶, MD, PhD; Fatima Cardoso⁴, MD; Maria João Cardoso⁴, MD, PhD

¹Institute of Oncology, Department of Surgery, Breast Unit, Pauls Stradiņš Clinical University Hospital, Riga Stradiņš University, Riga, Latvia

²Information and Health Decision Sciences of the Faculty of Medicine, University of Porto, Porto, Portugal

³Institute of Oncology, Riga Stradiņš University, Riga, Latvia

⁴Breast Cancer Unit, Champalimaud Cancer Center, Lisbon, Portugal

⁵Faculty of Medicine, Rīga Stradiņš University, Riga, Latvia

⁶Department of Surgery, Faculty of Medicine, Pauls Stradins Clinical University Hospital, Rīga Stradiņš University, Riga, Latvia

Corresponding Author:

Jelena Maksimenko, MD, PhD

Institute of Oncology, Department of Surgery, Breast Unit

Pauls Stradiņš Clinical University Hospital

Riga Stradiņš University

Dzirčiema 16

Riga

Latvia

Phone: 371 27038577

Email: jelena.maksimenko@rsu.lv

Related Article:

This is a corrected version. See correction statement: <https://formative.jmir.org/2023/1/e45810>

Abstract

Background: Approximately 62% of patients with breast cancer with a pathogenic variant (*BRCA1* or *BRCA2*) undergo primary breast-conserving therapy.

Objective: The study aims to develop a personalized risk management decision support tool for carriers of a pathogenic variant (*BRCA1* or *BRCA2*) who underwent breast-conserving therapy for unilateral early-stage breast cancer.

Methods: We developed a Bayesian network model of a hypothetical cohort of carriers of *BRCA1* or *BRCA2* diagnosed with stage I/II unilateral breast cancer and treated with breast-conserving treatment who underwent subsequent second primary cancer risk–reducing strategies. Using event dependencies structured according to expert knowledge and conditional probabilities obtained from published evidence, we predicted the 40-year overall survival rate of different risk-reducing strategies for 144 cohorts of women defined by the type of pathogenic variants (*BRCA1* or *BRCA2*), age at primary breast cancer diagnosis, breast cancer subtype, stage of primary breast cancer, and presence or absence of adjuvant chemotherapy.

Results: Absence of adjuvant chemotherapy was the most powerful factor that was linked to a dramatic decline in survival. There was a negligible decline in the mortality in patients with triple-negative breast cancer, who received no chemotherapy and underwent any secondary risk–reducing strategy, compared with surveillance. The potential survival benefit from any risk-reducing strategy was more modest in patients with triple-negative breast cancer who received chemotherapy compared with patients with luminal breast cancer. However, most patients with triple-negative breast cancer in stage I benefited from bilateral risk-reducing mastectomy and risk-reducing salpingo-oophorectomy or just risk-reducing salpingo-oophorectomy. Most patients with luminal

stage I/II unilateral breast cancer benefited from bilateral risk-reducing mastectomy and risk-reducing salpingo-oophorectomy. The impact of risk-reducing salpingo-oophorectomy in patients with luminal breast cancer in stage I/II increased with age. Most older patients with the *BRCA1* and *BRCA2* pathogenic variants in exons 12-24/25 with luminal breast cancer may gain a similar survival benefit from other risk-reducing strategies or surveillance.

Conclusions: Our study showed that it is mandatory to consider the complex interplay between the types of *BRCA1* and *BRCA2* pathogenic variants, age at primary breast cancer diagnosis, breast cancer subtype and stage, and received systemic treatment. As no prospective study results are available at the moment, our simulation model, which will integrate a decision support system in the near future, could facilitate the conversation between the health care provider and patient and help to weigh all the options for risk-reducing strategies leading to a more balanced decision.

(*JMIR Form Res* 2022;6(12):e37144) doi:[10.2196/37144](https://doi.org/10.2196/37144)

KEYWORDS

BRCA1 and BRCA2; secondary prophylactic strategies; breast-conserving therapy; breast cancer

Introduction

Breast cancer is the most common cancer and the leading cause of cancer mortality among women in economically developed and developing countries [1]. In unselected patients with breast cancer aged 35-64 years, pathogenic variants of *BRCA1* and *BRCA2* were detected in 2.4% and 2.3%, respectively [2]. In approximately 63% of patients with breast cancer related to the pathogenic variant *BRCA1* or *BRCA2*, genetic testing was performed after surgery of the primary cancer, with 62% of these patients also undergoing a primary breast-conserving treatment (BCT) [3].

Patients with breast cancer with *BRCA1* or *BRCA2* who underwent BCT have a significantly higher risk of a second primary ipsilateral breast event that is almost exclusively a new primary breast cancer rather than a true recurrence [4-6]. In addition, 27% of carriers of the *BRCA1* pathogenic variant and 19% of carriers of the *BRCA2* pathogenic variant will develop a second primary contralateral breast cancer within 10 years after the first primary breast cancer diagnosis [7].

Current guidelines describe different cancer risk management strategies: enhanced breast cancer screening, risk-reducing bilateral mastectomy (RRBM), risk-reducing salpingo-oophorectomy (RRSO), and chemoprevention [8]. Annual breast cancer screening with mammography and magnetic resonance imaging allows one to detect breast cancer at an early stage [9], although it cannot be prevented. In carriers of the *BRCA1* and *BRCA2* pathogenic variants, prophylactic mastectomy reduces the risk of subsequent breast cancer by approximately 90% [10]. However, the prophylactic mastectomy procedure could also have a potentially damaging effect on the patient's body image and sexual well-being [11-14]. Risk-reducing bilateral salpingo-oophorectomy (RRBSO) can be offered to carriers of the *BRCA1* and *BRCA2* pathogenic variants who are more than 35 years old or who have completed childbearing [8]. RRBSO reduces the risk of ovarian cancer by approximately 80%. However, the impact of RRBSO on second primary breast cancer risk remains uncertain and research findings are inconsistent [15-19]. RRBSO may also increase patients' risk of osteopenia, osteoporosis, cardiovascular disease, and may negatively impact cognitive function and quality of life [20]. Therefore, only 70% of carriers of the *BRCA1* and *BRCA2* pathogenic variants elect for RRBSO [21].

Previous studies have revealed that different factors, such as the patient's age at first breast cancer diagnosis, the type of pathogenic variant (ie, *BRCA1* or *BRCA2*), first breast cancer subtype, adjuvant systemic treatment received, presence or absence of ovarian cancer, may influence the degree to which a particular patient will benefit from various prophylactic strategies [15,16,22-33]. Ultimately, the patient and her health care team, who have already faced the first breast cancer treatment, are confronted with complex decisions regarding the optimal prophylactic strategy for subsequent cancer risk management. In addition, there are no prospective trials comparing different cancer risk-reducing strategies in patients with breast cancer that tested positive for the *BRCA1* and *BRCA2* pathogenic variants who were treated with BCT at the first event.

The aim of this study is to develop a personalized risk management guideline for carriers of the pathogenic variants of *BRCA1* and *BRCA2* who underwent BCT for unilateral early-stage breast cancer taking into account the patient characteristics and tumor prognostic parameters as well as systemic treatment received.

Methods

Study Design: Network Model and Strategies

We have developed a temporal Bayesian network model to estimate the expected overall survival of a hypothetical cohort of *BRCA1* and *BRCA2* carriers diagnosed with stage I-II unilateral breast cancer and treated with BCT who underwent subsequent second primary cancer prevention strategies. A Bayesian network is a directed acyclic graph that represents the joint distribution of a single set of variables. Each variable is represented by a node in the graph and is dependent on the set of variables represented by its ascendant nodes. This dependence is represented by a conditional probability table that describes the probability distribution of each variable given its ascendant variables. Temporal Bayesian networks are a special type of model where each (temporal) variable is expressed in multiple linked nodes to represent events in different moments in time; for example, a 2-year model for the event "ovarian cancer" (OC) could be defined with 2 linked nodes: "OC-y1" and "OC-y2," where OC-y2 is certain if OC-y1 is true, and P(OC-y2) is given

by the yearly risk of OC if OC-y1 is false, adjusted for all their ascendant nodes.

All risk estimates were converted into yearly estimates by conditional probabilities, depending on the original metric published in the literature with needed conversions (eg, risk for years between 5 and 10 used 10-year estimates converted to actual follow-up) [34,35]. If incidence estimates were given for a certain follow-up time (eg, lifetime ovarian cancer risk), the probability for each year “i” is computed as $1 - \exp(-\text{rate}[p,o] \times c)$, where “p” is the original risk for occurrence within “o” years, and $\text{rate}(p,o) = -\ln(1 - p)/o$.

If hazard ratios were given, survival for each group was computed as ref^{hr} , where “ref” is the expected survival for the reference group.

The simulation was run for a yearly follow-up of 40 years after diagnosis, yielding a temporal Bayesian network with 40 nodes per temporal variable (eg, ipsilateral recurrence). We predicted the overall survival following different prevention strategies for 144 cohorts (Multimedia Appendix 1) of women defined by the location of the *BRCA1* and *BRCA2* pathogenic variants (*BRCA1*: exons 1-10, exon 11, and exons 12-24; *BRCA2*: exons 1-10; exon 11; exons 12-25), age at primary breast cancer diagnosis (<40 years old, 40-50 years old, and >50 years old), breast cancer subtype (luminal-like and triple negative [TN]), stage of primary breast cancer (stage I and II), and presence or absence of adjuvant chemotherapy.

Data on 1 million simulations were generated. Each subgroup combination had around 6900 patients simulated across the 9 different intervention policies: (1) surveillance; (2) contralateral risk-reducing mastectomy; (3) RRBM; (4) contralateral risk-reducing mastectomy and RRBSO; (5) RRBM-RRBSO; (6) 5-year tamoxifen therapy; (7) contralateral risk-reducing mastectomy and 5-year tamoxifen therapy; (8) RRBSO; and (9) RRBM and 5 years’ tamoxifen therapy. As a result, around 770 patients were distributed to each subgroup \times policy combination. All these intervention policies were considered with or without adjuvant chemotherapy, totaling 18 different policies.

For each patient, the first temporal node to be activated was identified, and survival computed for each patient. The overall survival of patients assigned for each subgroup \times policy combination was plotted as Kaplan-Meier curves for 40-year follow-up and compared by the log-rank test. Hazard ratios for each subgroup were computed according to the proportional hazard Cox regression. However, given the simulation nature of the data, it was not possible to analyze any *P* value or CI estimates.

A temporal Bayesian network model was constructed using R (R Foundation for Statistical Computing) statistical software packages ‘bnlearn’ [36] and ‘gRain’ [37], assigning the overall survival associated for each strategy.

Key decision variables used in baseline and sensitivity analyses were obtained from peer-reviewed English language literature published in PubMed and from publicly available databases (Multimedia Appendix 2; also see [5,19,22,23,30,31,34,38-45]). We obtained the age-specific risk of death from other causes

from Colzani et al [46]. TN breast cancer was defined as estrogen receptor (ER) <10%, progesterone receptor (PR) <10%, human epidermal growth factor receptor 2 (HER2-) 0 or 1+. Luminal phenotype breast cancer was defined as ER+, PR+, and HER2- 0 or 1+.

First Primary Breast Cancer

Both *BRCA1* and *BRCA2* pathogenic variant-related tumors rarely showed evidence of HER2 amplification or expression [47-54].

Therefore, we did not include carriers of the *BRCA1* and *BRCA2* pathogenic variants with HER2-positive breast cancer subtypes in our hypothetical cohort. We derived that the cumulative incidence of first primary TN breast cancer in *BRCA1* and *BRCA2* carriers is 69% and 15%, respectively [55], and that other breast cancers are of luminal phenotype by immunohistochemistry. According to a recently published meta-analysis of 66 studies, there is no clear evidence supporting the different prognosis for patients with breast cancer with *BRCA1* and *BRCA2* pathogenic variants compared with sporadic cases [48,50].

Therefore, we used stage-specific and breast cancer subtype-specific mortality rates adjusted for age, race/ethnicity, and socioeconomic status reported in the population-based study by *Parise and Caqqiano* [38], where 143,333 female primary first invasive breast cancer cases were included. We assumed that 52.3% of breast cancer cases were diagnosed in stage I, 43.1% in stage II, and 4% in stage III [5].

BRCA1/BRCA2 Pathogenic Variant Genotype-Phenotype Correlation and Ovarian Cancer

In the study published by *Bayraktar et al* [22], patients with exon 20 *BRCA1* pathogenic variant and patients with exons 12-25 *BRCA2* pathogenic variant had a higher risk of developing both breast and ovarian cancer compared with patients with other exon mutations. We assumed ovarian cancer lifetime incidence in patients with unilateral primary breast cancer as breast and ovarian cancer prevalence by the pathogenic variant of the *BRCA1* and *BRCA2* combined exon group [22,56]. The risk of ovarian cancer was modeled assuming an expected lifetime of 80 years. For each age stratum, we computed the initial risk of ovarian cancer for patients aged 35, 45, and 65 years, respectively.

In patients with *BRCA1*- and *BRCA2*-related breast cancer, adjuvant tamoxifen and chemotherapy showed no impact on the risk reduction of subsequent ovarian cancer [28]. Therefore, we did not evaluate the impact of systemic treatment on ovarian cancer rates in our simulation model.

We used the distribution of stage at diagnosis of ovarian cancer and 10-year survival rates for ovarian cancer reported by *Benedet et al* [39].

A recently published study [28] showed no long-term survival benefit in patients with *BRCA1*- and *BRCA2*-related ovarian cancer compared with patients with sporadic ovarian cancer.

Second Primary Contralateral Breast Cancer

For contralateral breast cancer we assumed the same breast cancer distribution as for the first primary breast cancer [38]. We summarized the stage-specific and breast cancer subtype-specific mortality rate of the first primary breast cancer with the stage-specific and breast cancer subtype-specific mortality rate of the second primary contralateral breast cancer.

Multiple studies showed that a younger age at the onset of first breast cancer is associated with a higher contralateral breast cancer risk [23-26]. We assumed the age at the onset of first breast cancer based on the cumulative, contralateral breast cancer risk estimates proposed by Graeser et al [24]. We used the lifetime breast cancer-specific mortality for ductal carcinoma in situ of 3.3% reported by Narod et al [40]. We assumed the breast cancer-specific mortality of 100% for patients with metastatic invasive contralateral breast cancer [41]. We assumed that RRBSO does not reduce the risk of contralateral breast cancer in carriers of the *BRCA1* and *BRCA2* pathogenic variants [57].

Ipsilateral Breast Cancer

According to the previously published studies, carriers of the *BRCA1* and *BRCA2* pathogenic variants who underwent BCT for the first primary breast cancer have an increased risk of ipsilateral breast events compared with carriers of the *BRCA1* and *BRCA2* pathogenic variants who underwent a mastectomy [4-6]. We assumed the local failure cumulative incidence reported by Pierce et al [6]. Previous studies demonstrated that RRBSO and adjuvant chemotherapy decrease the risk of an ipsilateral breast event. We used hazard ratios published by Valachis et al [31]. In this meta-analysis, RRBSO decreased the risk of an ipsilateral breast event by 58% and adjuvant chemotherapy decreased the risk of an ipsilateral breast event by 49%. There was no evidence to support the protective effect of tamoxifen against ipsilateral breast cancer [31]. Despite the high rate of ipsilateral events in carriers of the pathogenic variants of *BRCA1* and *BRCA2* who underwent BCT, there was no statistically significant impact on distant recurrence and disease-specific survival [5]. These findings could be explained by the limited sample size in the study's cohorts and detection of ipsilateral events in the early stage due to the close surveillance of carriers of the *BRCA1* and *BRCA2* pathogenic variants [5]. The mean time to ipsilateral events was approximately 7 years. This fact conveys the impression that most of these events were new primary breast cancers [4-6]. Therefore, for patients who developed ipsilateral breast cancer we assumed the same stage-specific and breast cancer subtype-specific mortality rates as for the first primary breast cancer [38]. We summarized the stage-specific and breast cancer subtype-specific mortality rate of the first primary breast cancer with the stage-specific and breast cancer subtype-specific mortality rate of an ipsilateral breast event. We assumed that the stage distribution and breast cancer mortality in patients with a new second primary breast cancer are the same as those for the contralateral breast cancer [38,40-42].

Prevention Strategies

Prophylactic Oophorectomy

We considered an 80% risk reduction of ovarian cancer in carriers of the *BRCA1* pathogenic variant and 79% risk reduction in carriers of the *BRCA2* pathogenic variant [15].

In our hypothetical cohort, patients underwent an RRBSO within 5 years after the first primary breast cancer diagnosis. At the moment, there is no clear evidence suggesting that hormone replacement therapy does not offset the second primary breast cancer risk induced by RRBSO in *BRCA1* and *BRCA2* carriers [18,58-60]. Therefore, we assumed that none of the patients in our cohort received hormone replacement therapy. RRBSO is associated with a higher risk of noncancer-related death. Thus, in the simulation, we applied increased noncancer mortality for patients who underwent RRBSO before the age of 45 and had an onset of first primary breast cancer before the age of 40 [43]. The mortality risk following RRBSO was assumed to be the same as that following a mastectomy.

Contralateral Prophylactic Mastectomy

We assumed contralateral risk-reducing bilateral mastectomy-adjusted risk reduction of primary mastectomy [44]. Patients who underwent contralateral mastectomy have a statistically significantly better survival rate compared with patients who underwent surveillance [42]. The survival benefit was even more pronounced in patients with primary breast cancer onset under 40 years of age, with no TN breast subtype, and not treated with chemotherapy [42].

Ipsilateral Prophylactic Mastectomy

We assumed the local failure cumulative incidence in patients who underwent a mastectomy, as reported by Pierce et al [5]. According to Pierce et al [5], or carriers of *BRCA1* and *BRCA2*, who underwent mastectomy and received chemotherapy and RRBSO, the status of the pathogenic variants of *BRCA1* and *BRCA2* had no impact on local failure rate [5]. We assumed that an isolated locoregional recurrence after mastectomy has no impact on the survival of patients with breast cancer [61].

Tamoxifen

We used data from the combined International Retrospective-Pro prospective Carriers of the Pathogenic Variants of the *BRCA1* and *BRCA2* cohort study, in which 1583 carriers of the pathogenic variant of *BRCA1* and 881 carriers of the pathogenic variant of *BRCA2* with unilateral breast cancer were included [30]. We assumed that 5-year tamoxifen administration reduces 15-year mortality by 30% in patients with luminal-like primary breast cancer [62].

We assumed that tamoxifen reduces the age-specific ER status-adjusted risk of second primary contralateral breast cancer by 56% in carriers of the *BRCA1* pathogenic variant and by 67% in carriers of the *BRCA2* pathogenic variant [30]. We assumed that chemotherapy has no additional protective effect on the contralateral breast cancer development in carriers of *BRCA1* and *BRCA2* who received tamoxifen [30]. We assumed that the use of tamoxifen for 5 years has no impact on mortality due to cardiovascular or thromboembolic disease [63].

Ethical Considerations

Our simulation model of a hypothetical cohort was based on previously published data, and therefore did not require a submission to a research ethics committee.

Results

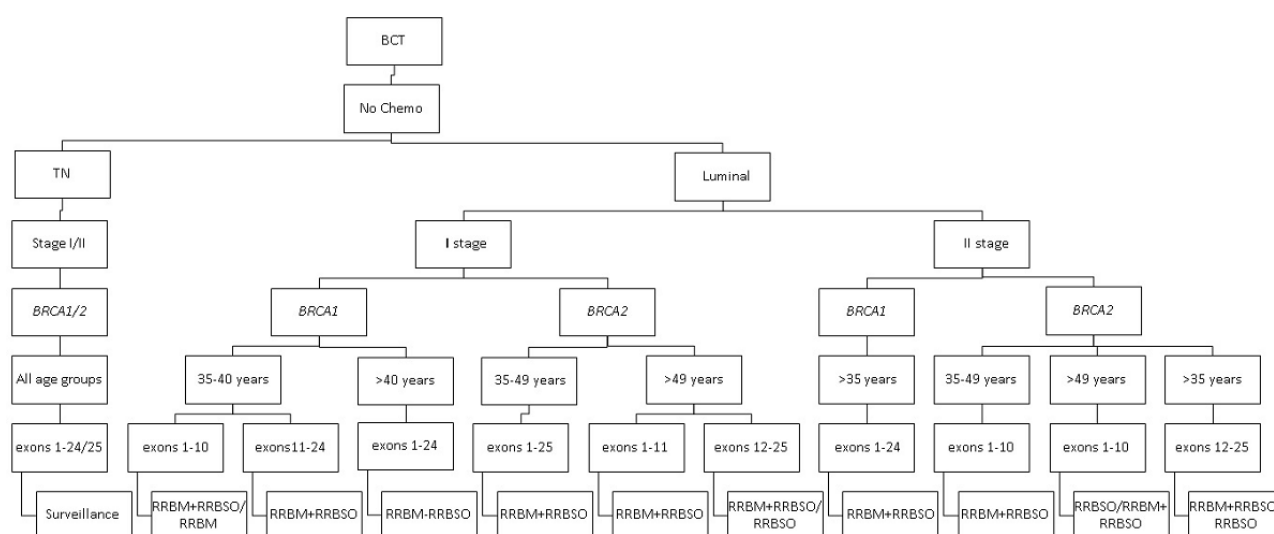
Effectiveness of Secondary Prevention Strategies

The predicted 40-year overall survival rate for carriers of *BRCA1* and *BRCA2* variants after unilateral BCT who received

secondary cancer prevention strategies or surveillance is presented in [Multimedia Appendix 3](#).

The impact of secondary prevention strategies on the survival of carriers of the *BRCA1* and *BRCA2* pathogenic variants with luminal breast cancer who received no adjuvant chemotherapy is presented in [Figure 1](#).

Figure 1. The most effective secondary prophylactic strategies in *BRCA1* and *BRCA2* carriers who received no adjuvant chemotherapy. BCT: breast-conserving treatment; RRBm: risk-reducing bilateral mastectomy; RRBSO: risk-reducing bilateral salpingo-oophorectomy; TN: triple negative.



Absence of adjuvant chemotherapy was the most powerful factor that was linked to a dramatic decline in survival for patients with breast cancer with *BRCA1* and *BRCA2* mutation. There was a negligible increase in survival for carriers of the pathogenic variants *BRCA1* and *BRCA2* with TN breast cancer and who received any secondary prevention strategy compared with surveillance.

Most carriers of *BRCA1* and *BRCA2* with luminal breast cancer in stage I benefited from RRBm-RRBSO. However, patients with breast cancer aged less than 40 years with the *BRCA1* pathogenic variant in exons 1-10 in stage I who underwent RRBm + RRBSO had an almost similar impact on survival compared with those who underwent RRBm alone. By contrast, patients with breast cancer aged over 49 years with the *BRCA2* pathogenic variant in exons 12-25 in stage I who underwent RRBm + RRBSO had an almost similar impact on survival compared with those who underwent RRBSO alone.

In patients with breast cancer with the *BRCA1* and *BRCA2* pathogenic variants with luminal breast cancer in stage II, RRBm + RRBSO had a very modest impact on survival compared with surveillance. Interestingly, RRBm-RRBSO or only RRBSO was the most effective prevention strategy in patients with luminal breast cancer aged over 35 years with the *BRCA2* pathogenic variant in exons 12-25 in stage II and in

patients aged over 49 years with the *BRCA2* pathogenic variant in exons 1-10 in stage II. By contrast, there was a negligible increase in survival among carriers of *BRCA1* and *BRCA2* aged over 35 years with luminal breast cancer in exon 11 who underwent any secondary prevention strategy compared with surveillance.

Interestingly, we noted that the impact of RRBSO in patients with the *BRCA1* and *BRCA2* pathogenic variants with luminal breast cancer in stage I/II increased with the age.

The impact on the survival of secondary risk-reducing strategies among carriers of *BRCA1* and *BRCA2* variants with luminal breast cancer after adjuvant chemotherapy is presented in [Figures 2 and 3](#).

RRBM-RRBSO was the most effective risk-reducing strategy in patients with luminal breast cancer who received adjuvant chemotherapy. The protective role of RRBSO in patients with luminal breast cancer increased with their age at diagnosis, stage of the disease, and was impacted by the type of pathogenic variant (*BRCA1* or *BRCA2*).

The impact on survival of secondary risk-reducing strategies in carriers of the pathogenic variants of *BRCA1* and *BRCA2* with TN breast cancer after adjuvant chemotherapy is shown in [Figures 4 and 5](#).

Figure 2. The most effective secondary prophylactic strategies in *BRCA1* carriers with luminal breast cancer who received adjuvant chemotherapy. BCT: breast-conserving treatment; RRBM: risk-reducing bilateral mastectomy; RRBSO: risk-reducing bilateral salpingo-oophorectomy.

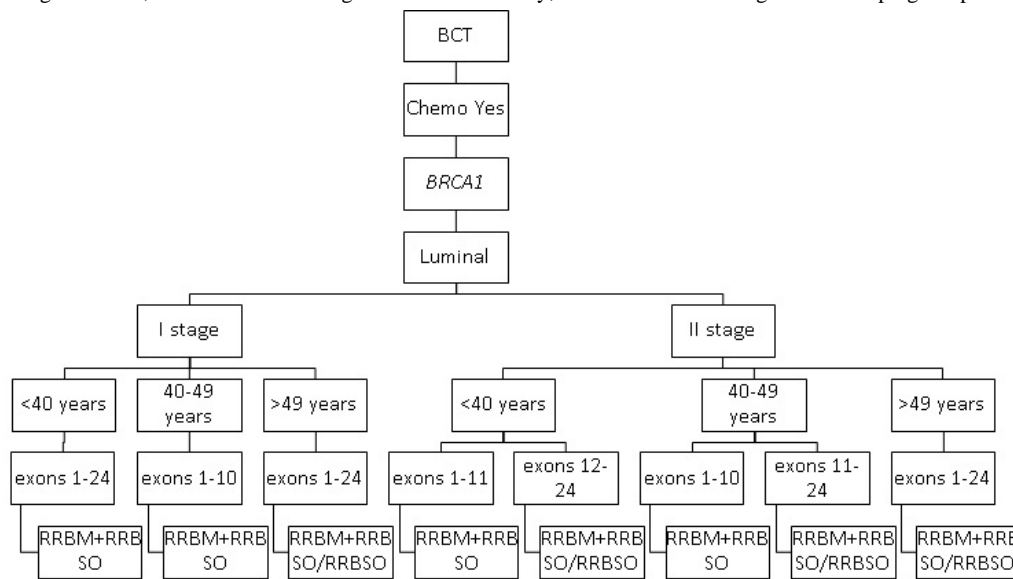


Figure 3. The most effective secondary prophylactic strategies in *BRCA2* carriers with luminal breast cancer who received adjuvant chemotherapy. BCT: breast-conserving treatment; RRBM: risk-reducing bilateral mastectomy; RRBSO: risk-reducing bilateral salpingo-oophorectomy.

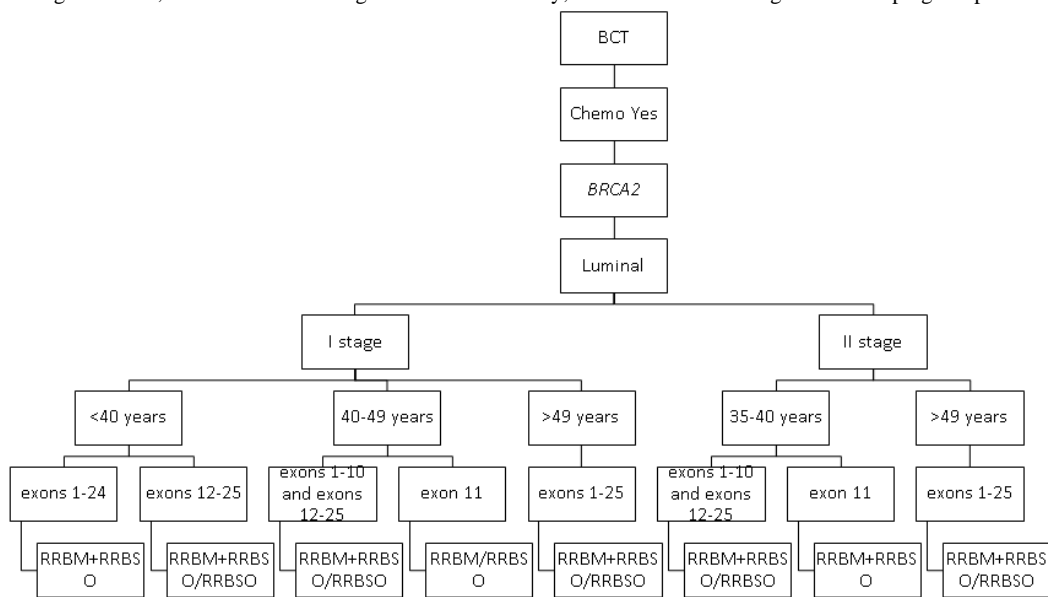


Figure 4. The most effective secondary prophylactic strategies in *BRCA1* carriers with TN breast cancer after adjuvant chemotherapy. BCT: breast-conserving treatment; RRBM: risk-reducing bilateral mastectomy; RRBSO: risk-reducing bilateral salpingo-oophorectomy; TN: triple negative.

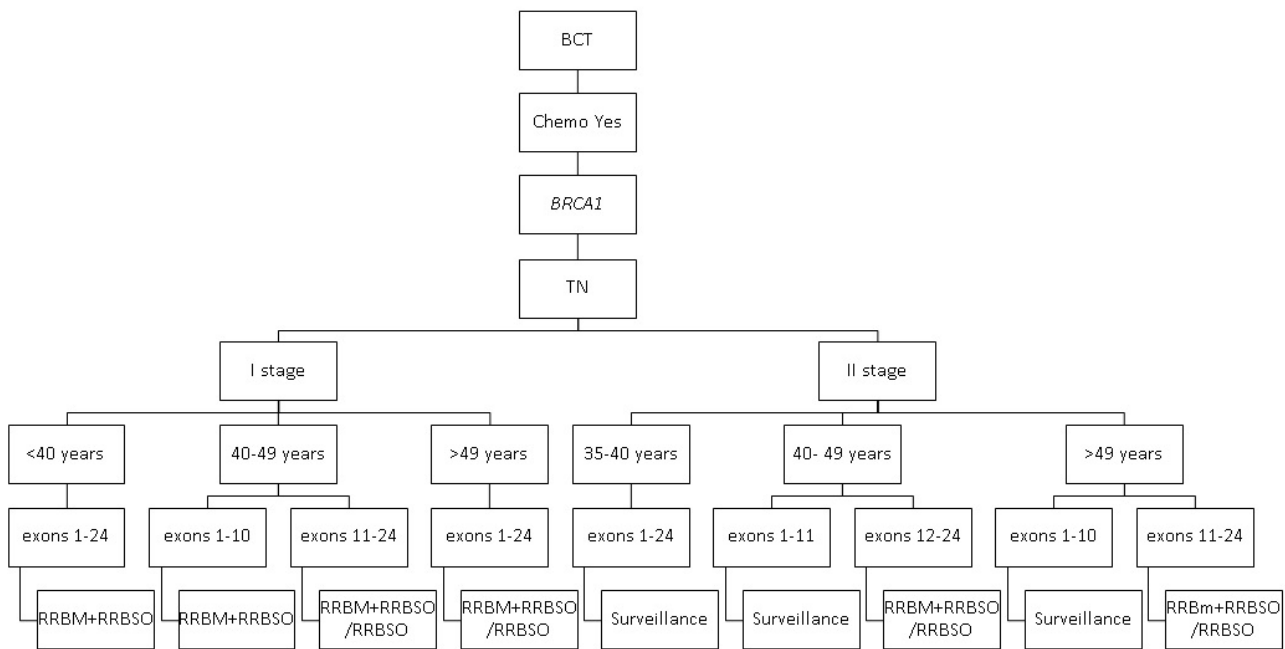
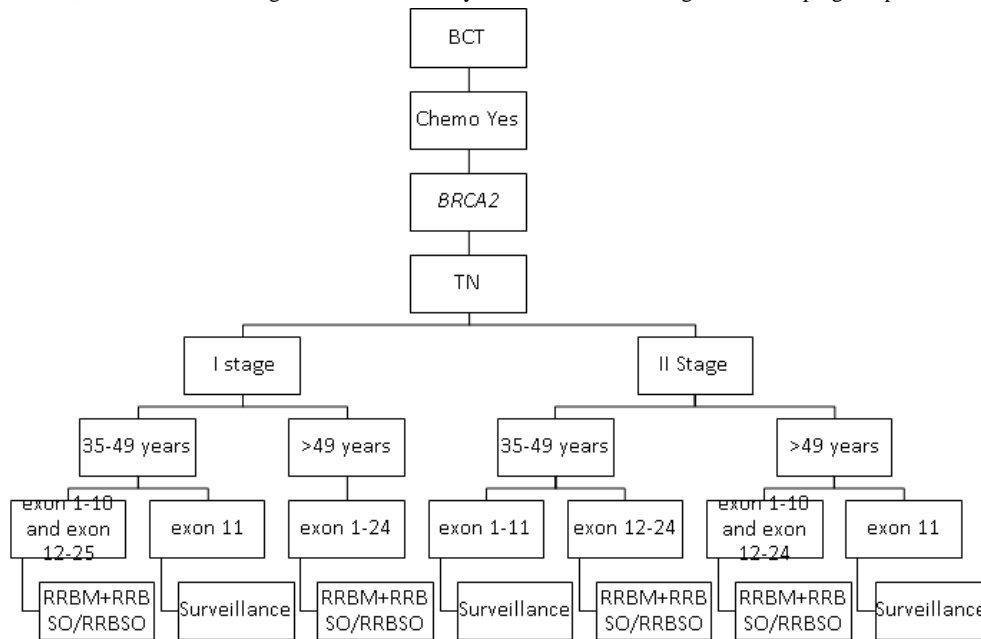


Figure 5. The most effective secondary prophylactic strategies in *BRCA2* carriers with TN breast cancer after adjuvant chemotherapy. BCT: breast-conserving treatment; RRBM: risk-reducing bilateral mastectomy; RRBSO: risk-reducing bilateral salpingo-oophorectomy; TN: triple negative.



The potential survival benefit from any risk-reducing strategy was modest in patients with TN breast cancer when compared with patients with luminal breast cancer. However, most carriers of *BRCA1* and *BRCA2* with TN breast cancer in stage I benefited from RRBM-RRBSO or just RRBSO.

RRBM-RRBSO was only the most effective risk-reducing strategy in patients with TN breast cancer under 40 years in stage I with the *BRCA1* pathogenic variant and in patients aged 40-49 years with the *BRCA1* pathogenic variant in exons 1-10. Further, patients with TN breast cancer aged 35-49 years in stage I with the *BRCA2* pathogenic variant in exon 11 had a very modest benefit from RRBM-RRBSO.

There was a negligible increase in survival in almost all carriers of the pathogenic variants of *BRCA1* and *BRCA2* with TN breast cancer in stage II who underwent any secondary risk-reducing strategy compared with surveillance.

However, RRBM-RRBSO or just RRBSO was the most effective risk-reducing strategy in patients with TN breast cancer aged over 40 years with the *BRCA1* pathogenic variant in exons 12-24 in stage II and in patients aged over 35 years with the *BRCA2* pathogenic variant in exons 12-24 in stage II. The impact of RRBSO in patients with *BRCA1* and *BRCA2* mutation with TN breast cancer in stage I/II increased with age.

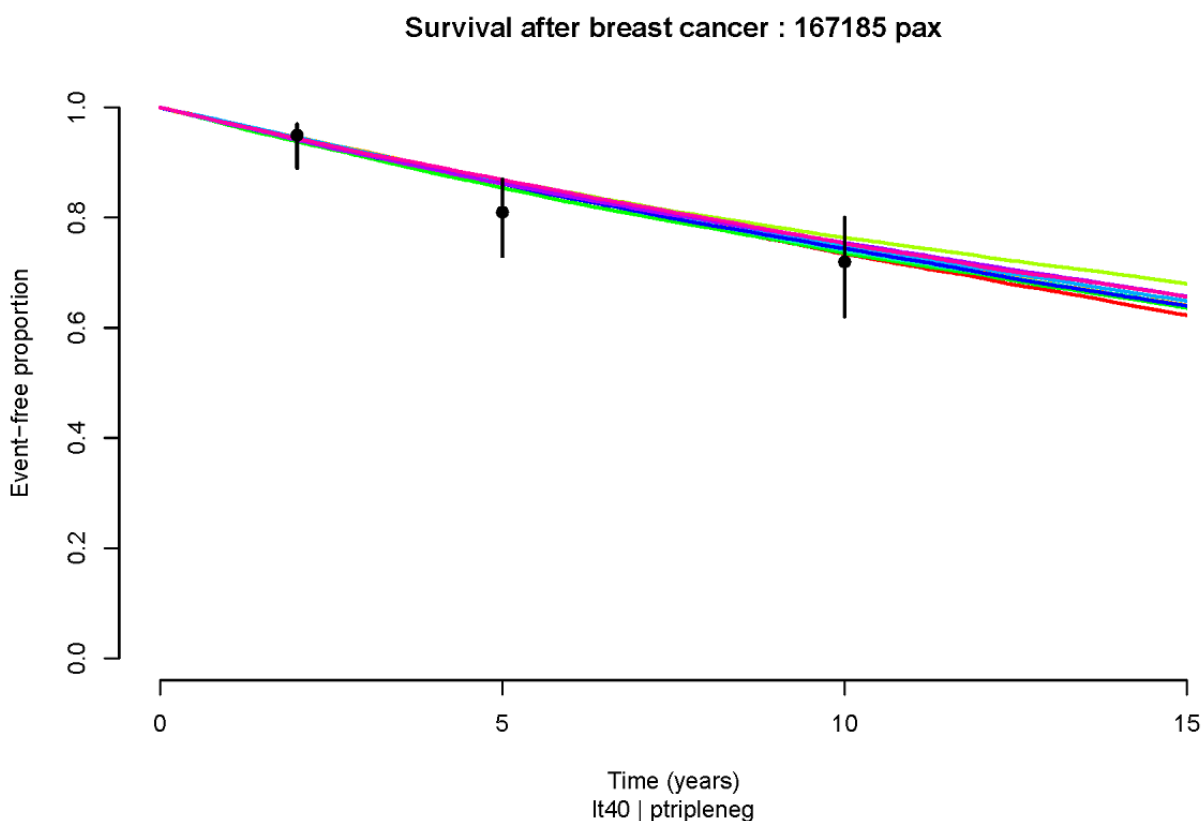
Sensitivity Analysis

There were difficulties in validating our results because of the lack of previously published studies with similar subgroups of patients that match all the detailed clinical and treatment variables.

Validation of our method was performed using TN subgroup cohort definitions from the largest published prospective study (POSH) [64]. Settings of the simulation were as follows: (1) from the 1 million patients in our simulation, we only considered those who were younger than 40 years at TN breast cancer diagnosis; (2) the distribution of *BRCA1* and *BRCA2* carriers with TN breast cancer was performed according to Copson et

al [64] (*BRCA1*: 123/136, *BRCA2*: 13/136); (3) the distribution of adjuvant chemotherapy in the TN *BRCA*-positive group was performed according to the study by Copson et al [64] (probability = 117/136); (4) the distribution of exon was uniform; (5) primary breast cancer stage distribution was used from the literature, considering the fact that all the patients were either in stage I or II (about 61.2% in stage I and nearly 38.8% in stage II). The simulation was run for 15 years after diagnosis with network steps of 2.5 years across the 9 different intervention policies. The overall survival of patients was plotted as a Kaplan-Meier curve (Figure 6) and compared with CIs from Copson et al [64] for TN *BRCA1*- and *BRCA2*-positive cases.

Figure 6. Kaplan-Meier survival plots for simulation for 15 years, performed using TN subgroup cohort definitions from the largest published prospective POSH study.



In our simulation, the overall survival was similar to the results in the POSH study. Similarly, in our simulation model patients with TN breast cancer aged under 40 years with the *BRCA1* and *BRCA2* pathogenic variants in stage I/II who received adjuvant chemotherapy had no survival benefit from RRBM. RRBM-RRBSO was the most effective risk-reducing strategy in these patients.

Discussion

Principal Findings

To our knowledge, this is the first study to simulate the expected overall survival and determine the most effective personalized management strategies for carriers *BRCA1* and *BRCA2* variants who underwent BCT for unilateral early-stage breast cancer taking into account the type of the pathogenic variant (*BRCA1* or *BRCA2*), age at primary breast cancer diagnosis, breast cancer

subtype, stage, and received systemic treatment. Absence of adjuvant chemotherapy was the most powerful factor that was linked to a dramatic decline in survival for patients with breast cancer with the pathogenic variants of *BRCA1* and *BRCA2*. There was a negligible decline in mortality among carriers of *BRCA1* and *BRCA2* with TN breast cancer who received no chemotherapy and underwent any secondary risk-reducing strategy compared with surveillance. The potential survival benefit from any risk-reducing strategy was more modest in patients with TN breast cancer who received chemotherapy compared with patients with luminal breast cancer. However, most carriers of *BRCA1* and *BRCA2* with TN breast cancer in stage I benefited from RRBM-RRBSO or just RRBSO. Most carriers of the pathogenic variant of *BRCA1* or *BRCA2* with luminal breast cancer in stage I-II (unilateral breast cancer) benefited from RRBM-RRBSO. The impact of RRBSO in patients with the *BRCA1* and *BRCA2* pathogenic variants with

luminal breast cancer in stage I/II increased with age. Most older patients with the *BRCA1* and *BRCA2* pathogenic variants in exons 12-24/25 with luminal breast cancer may gain a similar survival benefit from other risk-reducing strategies or surveillance.

Comparison With Prior Work

To date, only Schrag et al [65] have addressed secondary cancer risk-reducing strategies in *BRCA1* and *BRCA2* carriers who underwent BCT for unilateral breast cancer, using decision-analytic models. They calculated life-expectancy gains for different age groups; lymph node positive or negative status; and low, moderate, or high penetrance of the *BRCA1* and *BRCA2* pathogenic variants using a Markov model that incorporated 8 prevention strategies [65]. By contrast, we developed a temporal Bayesian network model with a total of 1 million simulated patients across 9 different intervention policies. Our study is an advancement over previous ones due to the incorporation of variables from up-to-date peer-reviewed studies considering the type of pathogenic variant (*BRCA1* or *BRCA2*), age at primary breast cancer diagnosis, breast cancer subtype and stage, status of systemic treatment received.

Our study showed that most *BRCA1* and *BRCA2* carriers benefited from RRBM + RRBSO. However, in patients with TN breast cancer who received no adjuvant chemotherapy, the impact of secondary risk-reducing strategies on overall mortality was reduced as a result of the higher risk of dying from primary breast cancer rather than from subsequent primary secondary cancer. Nevertheless, younger patients with limited disease who received adjuvant chemotherapy gained more benefit from aggressive surgical management (RRBM + RRBSO). By contrast, in older patients with more advanced disease the protective role of bilateral risk-reducing salpingo-oophorectomy was comparable to RRBM + RRBSO and was more pronounced in exons 12-24 of the *BRCA1* pathogenic variant and in exons 12-25 of the *BRCA2* pathogenic variant. In our study, we assumed that RRBSO does not reduce the risk of contralateral breast cancer in carriers of the *BRCA1* and *BRCA2* pathogenic variants. Recent studies showed no impact of RRBSO on primary and second primary breast cancer risk reduction [19,66]. As the risk of contralateral and ipsilateral breast cancer decreases with age [5,24] and the risk of ovarian cancer increases with age [67], patients with a higher probability of ovarian cancer gain a stronger protective effect from RRBSO and a low additional protective effect from RRBM.

In our model, we assumed breast cancer subtype-specific, population-based mortality and the general breast cancer population-based impact of adjuvant chemotherapy on outcomes. According to the prospective study by Clifton et al [68], for patients with TN breast cancer with the *BRCA1* and *BRCA2* pathogenic variants in stage I/II, there was no difference in overall survival for those who received neoadjuvant chemotherapy compared with those who received adjuvant chemotherapy [68].

However, a growing body of evidence indicates that *BRCA1* and *BRCA2* TN as well as luminal cancers are more chemosensitive and achieve higher pathologic complete response (pCR) rates compared with breast cancer without the *BRCA1*

or *BRCA2* pathogenic variant [69,70]. Paradoxically, in *BRCA1* and *BRCA2* breast cancer tumors it seems that pCR does not serve as a surrogate marker of better clinical outcome and a higher pCR does not translate into improved disease-free and overall survival [71,72]. Therefore, we did not include a neoadjuvant chemotherapy treatment strategy in our model.

In the systematic review published by Davey et al [73], there was no difference in 15-year mortality between BCT and mastectomy in *BRCA1* and *BRCA2* carriers with breast cancer. Most patients in both BCT and mastectomy groups had T1/2 tumors; approximately 60% of patients had ER-, N+ disease, or underwent adjuvant chemotherapy; and approximately 50% of patients underwent RRBSO [73]. In our study, for most patients with TN breast cancer in stage II who received chemotherapy, and all patients with TN breast cancer in stage I/II who did not receive chemotherapy, there was no difference in survival between any secondary prevention strategy and surveillance.

In a study published by Wan et al [74], where 8396 consecutive patients with breast cancer after surgery were included, no survival benefit was shown in carriers with *BRCA1* and *BRCA2* who underwent BCT compared with those who underwent mastectomy with or without radiotherapy. However, only 73 carriers with the *BRCA1* pathogenic variant and 106 with the *BRCA2* pathogenic variant who underwent BCT, and 104 carriers with the *BRCA1* pathogenic variant and 198 with the *BRCA2* pathogenic variant who underwent mastectomy were included in the study, with a relatively short follow-up period of 7.5 years. As the study is retrospective with a relatively small patient number in the subgroups, caution should be exercised while estimating the study's results and thus, further prospective research with a larger study population is needed.

Limitations

The main limitation of our study is that it is a computer simulation and it can misrepresent reality. Nevertheless, our model could prove to be a valuable decision support tool and we plan to validate our model on the target patient population. Risk modifiers assume adjusted risk estimates, and therefore they are additive, and some evidence is thin and dated in some of the included estimates. We assumed independence of risk factors where it was not possible to model any interaction, and this is a limitation. Nonetheless, the final validation shows concurrent results, supporting our model.

Conclusions

At present, no personalized guidelines are available for the prophylactic management of second primary breast cancer in patients with the *BRCA1* and *BRCA2* pathogenic variants with unilateral breast cancer who underwent BCT as a primary procedure. Our study showed that it is mandatory to consider the complex interplay between the type of *BRCA1* and *BRCA2* pathogenic variants, age at primary breast cancer diagnosis, breast cancer subtype and stage, and systemic treatment received. As no prospective study results are available, our simulation model could facilitate the conversation between the health care provider and patient and help to weigh all the options for risk-reducing strategies, thus leading to a more balanced

decision. However, we plan to expand and update our model and develop a computer-based clinical decision tool by including more variables from new evidence-based research

Availability of Data and Materials

The authors confirm that the data supporting the findings of this study are available within the article or its Multimedia Appendices.

Authors' Contributions

JM, DP, and MJC conceptualized this study. JM curated the data. PRP performed formal analysis of the study data. JG and FC were responsible for funding acquisition and project administration. JM and PRP proposed the study methodology and wrote the manuscript. MJC, EM, and GT were responsible for acquisition of resources and study supervision. PRP performed software/statistical analysis. JM, PRP, MN-M, MJC, and FC validated the study. JM, MN-M, MJC, EM, and GT performed data visualization. JM, MN-M, MJC, EM, and FC reviewed and edited the manuscript.

Conflicts of Interest

FC has received consultancy fees from Amgen, Astellas/Medivation, AstraZeneca, Celgene, Daiichi-Sankyo, Eisai, GE Oncology, Genentech, Gilead, GlaxoSmithKline, Iqvia, MacroGenics, Medscape, Merck-Sharp, Merus BV, Mylan, Mundipharma, Novartis, Pfizer, Pierre-Fabre, prIME Oncology, Roche, Sanofi, Samsung Bioepis, Seagen, Teva, and Touchime.

Multimedia Appendix 1

Overall survival following different prevention strategies for 144 cohorts.

[PDF File (Adobe PDF File), 4911 KB - [formative_v6i12e37144_app1.pdf](#)]

Multimedia Appendix 2

Key decision variables used in baseline and sensitivity analyses.

[DOCX File , 37 KB - [formative_v6i12e37144_app2.docx](#)]

Multimedia Appendix 3

Secondary cancer prevention strategies in *BRCA1/2* mutation carriers.

[XLSX File (Microsoft Excel File), 44 KB - [formative_v6i12e37144_app3.xlsx](#)]

References

1. Torre L, Bray F, Siegel R, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015 Mar;65(2):87-108 [FREE Full text] [doi: [10.3322/caac.21262](#)] [Medline: [25651787](#)]
2. Malone KE, Daling JR, Doody DR, Hsu L, Bernstein L, Coates RJ, et al. Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years. *Cancer Res* 2006 Aug 15;66(16):8297-8308 [FREE Full text] [doi: [10.1158/0008-5472.CAN-06-0503](#)] [Medline: [16912212](#)]
3. Trosman JR, Weldon CB, Gradishar WJ, Schink JC. Timing of genetic testing relative to breast cancer surgery. *JCO* 2010 May 20;28(15_suppl):666-666 [FREE Full text] [doi: [10.1200/jco.2010.28.15_suppl.666](#)]
4. Metcalfe K, Lynch H, Ghadirian P, Tung N, Kim-Sing C, Olopade O, et al. Risk of ipsilateral breast cancer in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res Treat* 2011 May;127(1):287-296 [FREE Full text] [doi: [10.1007/s10549-010-1336-7](#)] [Medline: [21221768](#)]
5. Pierce L, Phillips K, Griffith K, Buys S, Gaffney D, Moran M, et al. Local therapy in BRCA1 and BRCA2 mutation carriers with operable breast cancer: comparison of breast conservation and mastectomy. *Breast Cancer Res Treat* 2010 Jun;121(2):389-398 [FREE Full text] [doi: [10.1007/s10549-010-0894-z](#)] [Medline: [20411323](#)]
6. Pierce L, Levin A, Rebbeck T, Ben-David M, Friedman E, Solin L, et al. Ten-Year Multi-Institutional Results of Breast-Conserving Surgery and Radiotherapy in -Associated Stage I/II Breast Cancer. *JCO* 2006 Jun 01;24(16):2437-2443 [FREE Full text] [doi: [10.1200/jco.2005.02.7888](#)]
7. Molina-Montes E, Pérez-Nevot B, Pollán M, Sánchez-Cantalejo E, Espín J, Sánchez MJ. Cumulative risk of second primary contralateral breast cancer in BRCA1/BRCA2 mutation carriers with a first breast cancer: A systematic review and meta-analysis. *Breast* 2014 Dec;23(6):721-742 [FREE Full text] [doi: [10.1016/j.breast.2014.10.005](#)] [Medline: [25467311](#)]
8. Balmaña J, Diez O, Rubio I, Castiglione M, ESMO Guidelines Working Group. BRCA in breast cancer: ESMO Clinical Practice Guidelines. *Ann Oncol* 2010 May;21 Suppl 5:v20-v22 [FREE Full text] [doi: [10.1093/annonc/mdq161](#)] [Medline: [20555082](#)]
9. Evans D, Gareth ED, Kesavan N, Nisha K, Lim Y, Yit L, MARIBS Group, et al. MRI breast screening in high-risk women: cancer detection and survival analysis. *Breast Cancer Res Treat* 2014 Jun 9;145(3):663-672 [FREE Full text] [doi: [10.1007/s10549-014-2931-9](#)] [Medline: [24687378](#)]

10. Rebbeck T, Friebel T, Lynch H, Neuhausen S, van 't Veer L, Garber J, et al. Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. *J Clin Oncol* 2004 Mar 15;22(6):1055-1062 [[FREE Full text](#)] [doi: [10.1200/JCO.2004.04.188](https://doi.org/10.1200/JCO.2004.04.188)] [Medline: [14981104](https://pubmed.ncbi.nlm.nih.gov/14981104/)]
11. Frost MH, Hoskin TL, Hartmann LC, Degnim AC, Johnson JL, Boughey JC. Contralateral prophylactic mastectomy: long-term consistency of satisfaction and adverse effects and the significance of informed decision-making, quality of life, and personality traits. *Ann Surg Oncol* 2011 Oct;18(11):3110-3116 [[FREE Full text](#)] [doi: [10.1245/s10434-011-1917-7](https://doi.org/10.1245/s10434-011-1917-7)] [Medline: [21947589](https://pubmed.ncbi.nlm.nih.gov/21947589/)]
12. Altschuler A, Nekhlyudov L, Rolnick S, Greene S, Elmore J, West C, et al. Positive, negative, and disparate--women's differing long-term psychosocial experiences of bilateral or contralateral prophylactic mastectomy. *Breast J* 2008;14(1):25-32 [[FREE Full text](#)] [doi: [10.1111/j.1524-4741.2007.00521.x](https://doi.org/10.1111/j.1524-4741.2007.00521.x)] [Medline: [18186862](https://pubmed.ncbi.nlm.nih.gov/18186862/)]
13. Frost M, Slezak J, Tran N, Williams C, Johnson J, Woods J, et al. Satisfaction after contralateral prophylactic mastectomy: the significance of mastectomy type, reconstructive complications, and body appearance. *J Clin Oncol* 2005 Nov 01;23(31):7849-7856 [[FREE Full text](#)] [doi: [10.1200/JCO.2005.09.233](https://doi.org/10.1200/JCO.2005.09.233)] [Medline: [16204003](https://pubmed.ncbi.nlm.nih.gov/16204003/)]
14. Isern A, Tengrup I, Loman N, Olsson H, Ringberg A. Aesthetic outcome, patient satisfaction, and health-related quality of life in women at high risk undergoing prophylactic mastectomy and immediate breast reconstruction. *J Plast Reconstr Aesthet Surg* 2008 Oct;61(10):1177-1187. [doi: [10.1016/j.bjps.2007.08.006](https://doi.org/10.1016/j.bjps.2007.08.006)] [Medline: [17938010](https://pubmed.ncbi.nlm.nih.gov/17938010/)]
15. Marchetti C, De Felice F, Palaia I, Perniola G, Musella A, Musio D, et al. Risk-reducing salpingo-oophorectomy: a meta-analysis on impact on ovarian cancer risk and all cause mortality in BRCA 1 and BRCA 2 mutation carriers. *BMC Womens Health* 2014 Dec 12;14(1):150 [[FREE Full text](#)] [doi: [10.1186/s12905-014-0150-5](https://doi.org/10.1186/s12905-014-0150-5)] [Medline: [25494812](https://pubmed.ncbi.nlm.nih.gov/25494812/)]
16. Metcalfe K, Lynch H, Foulkes W, Tung N, Kim-Sing C, Olopade O, et al. Effect of Oophorectomy on Survival After Breast Cancer in BRCA1 and BRCA2 Mutation Carriers. *JAMA Oncol* 2015 Jun;1(3):306-313 [[FREE Full text](#)] [doi: [10.1001/jamaoncol.2015.0658](https://doi.org/10.1001/jamaoncol.2015.0658)] [Medline: [26181175](https://pubmed.ncbi.nlm.nih.gov/26181175/)]
17. Mavaddat N, Peock S, Frost D, Ellis S, Platte R, Fineberg E, EMBRACE. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. *J Natl Cancer Inst* 2013 Jun 05;105(11):812-822 [[FREE Full text](#)] [doi: [10.1093/jnci/djt095](https://doi.org/10.1093/jnci/djt095)] [Medline: [23628597](https://pubmed.ncbi.nlm.nih.gov/23628597/)]
18. Rebbeck T, Kauff N, Domchek S. Meta-analysis of risk reduction estimates associated with risk-reducing salpingo-oophorectomy in BRCA1 or BRCA2 mutation carriers. *J Natl Cancer Inst* 2009 Jan 21;101(2):80-87 [[FREE Full text](#)] [doi: [10.1093/jnci/djn442](https://doi.org/10.1093/jnci/djn442)] [Medline: [19141781](https://pubmed.ncbi.nlm.nih.gov/19141781/)]
19. Heemskerk-Gerritsen B, Seynaeve C, van Asperen CJ, Ausems M, Collée JM, van Doorn HC, Hereditary BreastOvarian Cancer Research Group Netherlands. Breast cancer risk after salpingo-oophorectomy in healthy BRCA1/2 mutation carriers: revisiting the evidence for risk reduction. *J Natl Cancer Inst* 2015 May;107(5):djh033. [doi: [10.1093/jnci/djh033](https://doi.org/10.1093/jnci/djh033)] [Medline: [25788320](https://pubmed.ncbi.nlm.nih.gov/25788320/)]
20. Finch A, Evans G, Narod S. BRCA carriers, prophylactic salpingo-oophorectomy and menopause: clinical management considerations and recommendations. *Womens Health (Lond)* 2012 Sep;8(5):543-555 [[FREE Full text](#)] [doi: [10.2217/whe.12.41](https://doi.org/10.2217/whe.12.41)] [Medline: [22934728](https://pubmed.ncbi.nlm.nih.gov/22934728/)]
21. Bradbury A, Ibe C, Dignam JJ, Cummings S, Verp M, White M, et al. Uptake and timing of bilateral prophylactic salpingo-oophorectomy among BRCA1 and BRCA2 mutation carriers. *Genet Med* 2008 Mar;10(3):161-166 [[FREE Full text](#)] [doi: [10.1097/GIM.0b013e318163487d](https://doi.org/10.1097/GIM.0b013e318163487d)] [Medline: [18344704](https://pubmed.ncbi.nlm.nih.gov/18344704/)]
22. Bayraktar S, Jackson M, Gutierrez-Barrera AM, Liu D, Meric-Bernstam F, Brandt A, et al. Genotype-Phenotype Correlations by Ethnicity and Mutation Location in BRCA Mutation Carriers. *Breast J* 2015;21(3):260-267 [[FREE Full text](#)] [doi: [10.1111/tbj.12392](https://doi.org/10.1111/tbj.12392)] [Medline: [25789811](https://pubmed.ncbi.nlm.nih.gov/25789811/)]
23. Rhiem K, Engel C, Graeser M, Zachariae S, Kast K, Kiechle M, et al. The risk of contralateral breast cancer in patients from BRCA1/2 negative high risk families as compared to patients from BRCA1 or BRCA2 positive families: a retrospective cohort study. *Breast Cancer Res* 2012 Dec 07;14(6):R156 [[FREE Full text](#)] [doi: [10.1186/bcr3369](https://doi.org/10.1186/bcr3369)] [Medline: [23216834](https://pubmed.ncbi.nlm.nih.gov/23216834/)]
24. Graeser MK, Engel C, Rhiem K, Gadzicki D, Bick U, Kast K, et al. Contralateral breast cancer risk in BRCA1 and BRCA2 mutation carriers. *J Clin Oncol* 2009 Dec 10;27(35):5887-5892 [[FREE Full text](#)] [doi: [10.1200/JCO.2008.19.9430](https://doi.org/10.1200/JCO.2008.19.9430)] [Medline: [19858402](https://pubmed.ncbi.nlm.nih.gov/19858402/)]
25. Verhoog LC, Brekelmans CT, Seynaeve C, Meijers-Heijboer EJ, Klijn JG. Contralateral breast cancer risk is influenced by the age at onset in BRCA1-associated breast cancer. *Br J Cancer* 2000 Aug;83(3):384-386 [[FREE Full text](#)] [doi: [10.1054/bjoc.2000.1239](https://doi.org/10.1054/bjoc.2000.1239)] [Medline: [10917555](https://pubmed.ncbi.nlm.nih.gov/10917555/)]
26. Metcalfe K, Gershman S, Lynch H, Ghadirian P, Tung N, Kim-Sing C, et al. Predictors of contralateral breast cancer in BRCA1 and BRCA2 mutation carriers. *Br J Cancer* 2011 Apr 26;104(9):1384-1392 [[FREE Full text](#)] [doi: [10.1038/bjc.2011.120](https://doi.org/10.1038/bjc.2011.120)] [Medline: [21487411](https://pubmed.ncbi.nlm.nih.gov/21487411/)]
27. Malone K, Begg C, Haile R, Borg A, Concannon P, Tellhed L, et al. Population-based study of the risk of second primary contralateral breast cancer associated with carrying a mutation in BRCA1 or BRCA2. *J Clin Oncol* 2010 May 10;28(14):2404-2410 [[FREE Full text](#)] [doi: [10.1200/JCO.2009.24.2495](https://doi.org/10.1200/JCO.2009.24.2495)] [Medline: [20368571](https://pubmed.ncbi.nlm.nih.gov/20368571/)]
28. McLaughlin J, Rosen B, Moody J, Pal T, Fan I, Shaw P, et al. Long-term ovarian cancer survival associated with mutation in BRCA1 or BRCA2. *J Natl Cancer Inst* 2013 Jan 16;105(2):141-148 [[FREE Full text](#)] [doi: [10.1093/jnci/djs494](https://doi.org/10.1093/jnci/djs494)] [Medline: [23257159](https://pubmed.ncbi.nlm.nih.gov/23257159/)]

29. Metcalfe KA, Lynch HT, Ghadirian P, Tung N, Olivotto IA, Foulkes WD, et al. The risk of ovarian cancer after breast cancer in BRCA1 and BRCA2 carriers. *Gynecol Oncol* 2005 Jan;96(1):222-226. [doi: [10.1016/j.ygyno.2004.09.039](https://doi.org/10.1016/j.ygyno.2004.09.039)] [Medline: [15589605](https://pubmed.ncbi.nlm.nih.gov/15589605/)]
30. Phillips KA, Milne RL, Rookus MA, Daly MB, Antoniou AC, Peock S, et al. Tamoxifen and risk of contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *J Clin Oncol* 2013 Sep 01;31(25):3091-3099 [FREE Full text] [doi: [10.1200/JCO.2012.47.8313](https://doi.org/10.1200/JCO.2012.47.8313)] [Medline: [23918944](https://pubmed.ncbi.nlm.nih.gov/23918944/)]
31. Valachis A, Nearchou A, Lind P. Surgical management of breast cancer in BRCA-mutation carriers: a systematic review and meta-analysis. *Breast Cancer Res Treat* 2014 Apr;144(3):443-455 [FREE Full text] [doi: [10.1007/s10549-014-2890-1](https://doi.org/10.1007/s10549-014-2890-1)] [Medline: [24567198](https://pubmed.ncbi.nlm.nih.gov/24567198/)]
32. Goodwin P, Phillips K, West D, Ennis M, Hopper J, John E, et al. Breast cancer prognosis in BRCA1 and BRCA2 mutation carriers: an International Prospective Breast Cancer Family Registry population-based cohort study. *J Clin Oncol* 2012 Jan 01;30(1):19-26 [FREE Full text] [doi: [10.1200/JCO.2010.33.0068](https://doi.org/10.1200/JCO.2010.33.0068)] [Medline: [22147742](https://pubmed.ncbi.nlm.nih.gov/22147742/)]
33. Robson ME, Chappuis PO, Satagopan J, Wong N, Boyd J, Goffin JR, et al. A combined analysis of outcome following breast cancer: differences in survival based on BRCA1/BRCA2 mutation status and administration of adjuvant treatment. *Breast Cancer Res* 2004 Oct 24;6(1):R8-R17 [FREE Full text] [doi: [10.1186/bcr658](https://doi.org/10.1186/bcr658)] [Medline: [14680495](https://pubmed.ncbi.nlm.nih.gov/14680495/)]
34. Miller D, Homan S. Determining transition probabilities: confusion and suggestions. *Med Decis Making* 1994;14(1):52-58 [FREE Full text] [doi: [10.1177/0272989X9401400107](https://doi.org/10.1177/0272989X9401400107)] [Medline: [8152357](https://pubmed.ncbi.nlm.nih.gov/8152357/)]
35. Naglie G, Krahn M, Naimark D, Redelmeier D, Detsky A. Primer on medical decision analysis: Part 3--Estimating probabilities and utilities. *Med Decis Making* 1997;17(2):136-141 [FREE Full text] [doi: [10.1177/0272989X9701700203](https://doi.org/10.1177/0272989X9701700203)] [Medline: [9107608](https://pubmed.ncbi.nlm.nih.gov/9107608/)]
36. Scutari M. Learning Bayesian Networks with the Package. *J. Stat. Soft* 2010;35(3):22 [FREE Full text] [doi: [10.18637/jss.v035.i03](https://doi.org/10.18637/jss.v035.i03)]
37. Højsgaard S. Graphical Independence Networks with the Package for R. *J. Stat. Soft* 2012;46(10):1-26 [FREE Full text] [doi: [10.18637/jss.v046.i10](https://doi.org/10.18637/jss.v046.i10)]
38. Parise CA, Caggiano V. Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers. *J Cancer Epidemiol* 2014;2014:469251-469298 [FREE Full text] [doi: [10.1155/2014/469251](https://doi.org/10.1155/2014/469251)] [Medline: [24955090](https://pubmed.ncbi.nlm.nih.gov/24955090/)]
39. Benedet J, Bender H, Jones H, Ngan H, Pecorelli S. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. *Int J Gynaecol Obstet* 2000 Aug;70(2):209-262. [Medline: [11041682](https://pubmed.ncbi.nlm.nih.gov/11041682/)]
40. Narod S, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. *JAMA Oncol* 2015 Oct;1(7):888-896 [FREE Full text] [doi: [10.1001/jamaoncol.2015.2510](https://doi.org/10.1001/jamaoncol.2015.2510)] [Medline: [26291673](https://pubmed.ncbi.nlm.nih.gov/26291673/)]
41. Bayraktar S, Gutierrez-Barrera AM, Lin H, Elsayegh N, Tasbas T, Litton JK, et al. Outcome of metastatic breast cancer in selected women with or without deleterious BRCA mutations. *Clin Exp Metastasis* 2013 Jun;30(5):631-642 [FREE Full text] [doi: [10.1007/s10585-013-9567-8](https://doi.org/10.1007/s10585-013-9567-8)] [Medline: [23370825](https://pubmed.ncbi.nlm.nih.gov/23370825/)]
42. Heemskerck-Gerritsen B, Rookus M, Aalfs C, Ausems M, Collée JM, Jansen L, HEBON, et al. Improved overall survival after contralateral risk-reducing mastectomy in BRCA1/2 mutation carriers with a history of unilateral breast cancer: a prospective analysis. *Int J Cancer* 2015 Feb 01;136(3):668-677 [FREE Full text] [doi: [10.1002/ijc.29032](https://doi.org/10.1002/ijc.29032)] [Medline: [24947112](https://pubmed.ncbi.nlm.nih.gov/24947112/)]
43. Rocca W, Grossardt B, de Andrade M, Malkasian G, Melton L. Survival patterns after oophorectomy in premenopausal women: a population-based cohort study. *The Lancet Oncology* 2006 Oct;7(10):821-828 [FREE Full text] [doi: [10.1016/S1470-2045\(06\)70869-5](https://doi.org/10.1016/S1470-2045(06)70869-5)]
44. van Sprundel TC, Schmidt M, Rookus M, Brohet R, van Asperen CJ, Rutgers EJT, et al. Risk reduction of contralateral breast cancer and survival after contralateral prophylactic mastectomy in BRCA1 or BRCA2 mutation carriers. *Br J Cancer* 2005 Aug 08;93(3):287-292 [FREE Full text] [doi: [10.1038/sj.bjc.6602703](https://doi.org/10.1038/sj.bjc.6602703)] [Medline: [16052221](https://pubmed.ncbi.nlm.nih.gov/16052221/)]
45. El-Tamer MB, Ward BM, Schiffner T, Neumayer L, Khuri S, Henderson W. Morbidity and mortality following breast cancer surgery in women: national benchmarks for standards of care. *Ann Surg* 2007 May;245(5):665-671. [doi: [10.1097/01.sla.0000245833.48399.9a](https://doi.org/10.1097/01.sla.0000245833.48399.9a)] [Medline: [17457156](https://pubmed.ncbi.nlm.nih.gov/17457156/)]
46. Colzani E, Liljegren A, Johansson AL, Adolfsson J, Hellborg H, Hall PF, et al. Prognosis of Patients With Breast Cancer: Causes of Death and Effects of Time Since Diagnosis, Age, and Tumor Characteristics. *JCO* 2011 Oct 20;29(30):4014-4021 [FREE Full text] [doi: [10.1200/jco.2010.32.6462](https://doi.org/10.1200/jco.2010.32.6462)]
47. Lakhani S, Van De Vijver MJ, Jacquemier J, Anderson TJ, Osin P, McGuffog L, et al. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* 2002 May 01;20(9):2310-2318 [FREE Full text] [doi: [10.1200/JCO.2002.09.023](https://doi.org/10.1200/JCO.2002.09.023)] [Medline: [11981002](https://pubmed.ncbi.nlm.nih.gov/11981002/)]
48. Rennert G, Bisland-Naggan S, Barnett-Griness O, Bar-Joseph N, Zhang S, Rennert H, et al. Clinical Outcomes of Breast Cancer in Carriers of and Mutations. *N Engl J Med* 2007 Jul 12;357(2):115-123 [FREE Full text] [doi: [10.1056/nejmoa070608](https://doi.org/10.1056/nejmoa070608)]

49. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001 Feb 22;344(8):539-548 [FREE Full text] [doi: [10.1056/NEJM200102233440801](https://doi.org/10.1056/NEJM200102233440801)] [Medline: [11207349](https://pubmed.ncbi.nlm.nih.gov/11207349/)]
50. van den Broek AJ, Schmidt MK, van 't Veer LJ, Tollenaar RAEM, van Leeuwen FE. Worse breast cancer prognosis of BRCA1/BRCA2 mutation carriers: what's the evidence? A systematic review with meta-analysis. *PLoS One* 2015;10(3):e0120189 [FREE Full text] [doi: [10.1371/journal.pone.0120189](https://doi.org/10.1371/journal.pone.0120189)] [Medline: [25816289](https://pubmed.ncbi.nlm.nih.gov/25816289/)]
51. Foulkes W. Germline BRCA1 Mutations and a Basal Epithelial Phenotype in Breast Cancer. *CancerSpectrum Knowledge Environment* 2003 Oct 01;95(19):1482-1485 [FREE Full text] [doi: [10.1093/jnci/djg050](https://doi.org/10.1093/jnci/djg050)] [Medline: [14519755](https://pubmed.ncbi.nlm.nih.gov/14519755/)]
52. Garber JE. BRCA1/2-associated and sporadic breast cancers: fellow travelers or not? *Cancer Prev Res (Phila)* 2009 Feb;2(2):100-103 [FREE Full text] [doi: [10.1158/1940-6207.CAPR-08-0243](https://doi.org/10.1158/1940-6207.CAPR-08-0243)] [Medline: [19174575](https://pubmed.ncbi.nlm.nih.gov/19174575/)]
53. Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, et al. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast Cancer Res* 2014 Dec 05;16(6):475 [FREE Full text] [doi: [10.1186/s13058-014-0475-x](https://doi.org/10.1186/s13058-014-0475-x)] [Medline: [25475740](https://pubmed.ncbi.nlm.nih.gov/25475740/)]
54. Bayraktar S, Amendola L, Gutierrez-Barrera AM, Hashmi SS, Amos C, Gambello M, et al. Clinicopathologic characteristics of breast cancer in BRCA-carriers and non-carriers in women 35 years of age or less. *Breast* 2014 Dec;23(6):770-774 [FREE Full text] [doi: [10.1016/j.breast.2014.08.010](https://doi.org/10.1016/j.breast.2014.08.010)] [Medline: [25231195](https://pubmed.ncbi.nlm.nih.gov/25231195/)]
55. Kuchenbaecker KB, Neuhausen SL, Robson M, Barrowdale D, McGuffog L, Mulligan AM, Breast Cancer Family Registry, EMBRACE Study, GEMO Study Collaborators, HEBON, KConFab Investigators, CIMBA. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res* 2014 Dec 31;16(6):3416-3487 [FREE Full text] [doi: [10.1186/s13058-014-0492-9](https://doi.org/10.1186/s13058-014-0492-9)] [Medline: [25919761](https://pubmed.ncbi.nlm.nih.gov/25919761/)]
56. Thompson D, Easton D, Breast Cancer Linkage Consortium. Cancer Incidence in BRCA1 mutation carriers. *J Natl Cancer Inst* 2002 Sep 18;94(18):1358-1365 [FREE Full text] [doi: [10.1093/jnci/94.18.1358](https://doi.org/10.1093/jnci/94.18.1358)] [Medline: [12237281](https://pubmed.ncbi.nlm.nih.gov/12237281/)]
57. Kotsopoulos J, Lubinski J, Lynch H, Tung N, Armel S, Senter L, Hereditary Breast Cancer Clinical Study Group. Oophorectomy and risk of contralateral breast cancer among BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res Treat* 2019 Jun;175(2):443-449 [FREE Full text] [doi: [10.1007/s10549-019-05162-7](https://doi.org/10.1007/s10549-019-05162-7)] [Medline: [30756284](https://pubmed.ncbi.nlm.nih.gov/30756284/)]
58. Rebbeck TR, Friebel T, Wagner T, Lynch HT, Garber JE, Daly MB, PROSE Study Group. Effect of short-term hormone replacement therapy on breast cancer risk reduction after bilateral prophylactic oophorectomy in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. *J Clin Oncol* 2005 Nov 01;23(31):7804-7810 [FREE Full text] [doi: [10.1200/JCO.2004.00.8151](https://doi.org/10.1200/JCO.2004.00.8151)] [Medline: [16219936](https://pubmed.ncbi.nlm.nih.gov/16219936/)]
59. Marchetti C, Iadarola R, Palaia I, di DV, Perniola G, Muzii L. Hormone therapy in oophorectomized BRCA1 and BRCA2 mutation carriers. *Menopause* 2014;21(7):763-768 [FREE Full text] [doi: [10.1097/gme.0000000000000126](https://doi.org/10.1097/gme.0000000000000126)]
60. Figueiredo JC, Haile RW, Bernstein L, Malone KE, Largent J, Langholz B, et al. Oral contraceptives and postmenopausal hormones and risk of contralateral breast cancer among BRCA1 and BRCA2 mutation carriers and noncarriers: the WECARE Study. *Breast Cancer Res Treat* 2010 Feb 14;120(1):175-183 [FREE Full text] [doi: [10.1007/s10549-009-0455-5](https://doi.org/10.1007/s10549-009-0455-5)] [Medline: [19597986](https://pubmed.ncbi.nlm.nih.gov/19597986/)]
61. Willner J, Kiricuta I, Kölbl O. Locoregional recurrence of breast cancer following mastectomy: Always a fatal event? Results of univariate and multivariate analysis. *International Journal of Radiation Oncology*Biophysics*Physics* 1997 Mar;37(4):853-863 [FREE Full text] [doi: [10.1016/s0360-3016\(96\)00556-1](https://doi.org/10.1016/s0360-3016(96)00556-1)]
62. Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *The Lancet* 2011 Aug;378(9793):771-784 [FREE Full text] [doi: [10.1016/S0140-6736\(11\)60993-8](https://doi.org/10.1016/S0140-6736(11)60993-8)]
63. Iqbal J, Ginsburg OM, Wijeratne TD, Howell A, Evans G, Sestak I, et al. Endometrial cancer and venous thromboembolism in women under age 50 who take tamoxifen for prevention of breast cancer: a systematic review. *Cancer Treat Rev* 2012 Jun;38(4):318-328 [FREE Full text] [doi: [10.1016/j.ctrv.2011.06.009](https://doi.org/10.1016/j.ctrv.2011.06.009)] [Medline: [21775065](https://pubmed.ncbi.nlm.nih.gov/21775065/)]
64. Copson E, Maishman T, Tapper W, Cutress R, Greville-Heygate S, Altman D, et al. Germline BRCA mutation and outcome in young-onset breast cancer (POSH): a prospective cohort study. *The Lancet Oncology* 2018 Feb;19(2):169-180 [FREE Full text] [doi: [10.1016/S1470-2045\(17\)30891-4](https://doi.org/10.1016/S1470-2045(17)30891-4)]
65. Schrag D, Kuntz K, Garber J, Weeks JC. Life expectancy gains from cancer prevention strategies for women with breast cancer and BRCA1 or BRCA2 mutations. *JAMA* 2000 Feb 02;283(5):617-624 [FREE Full text] [doi: [10.1001/jama.283.5.617](https://doi.org/10.1001/jama.283.5.617)] [Medline: [10665701](https://pubmed.ncbi.nlm.nih.gov/10665701/)]
66. Kotsopoulos J, Huzarski T, Gronwald J, Singer C, Moller P, Lynch H, Hereditary Breast Cancer Clinical Study Group. Bilateral Oophorectomy and Breast Cancer Risk in BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst* 2017 Jan;109(1):djw177 [FREE Full text] [doi: [10.1093/jnci/djw177](https://doi.org/10.1093/jnci/djw177)] [Medline: [27601060](https://pubmed.ncbi.nlm.nih.gov/27601060/)]
67. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ, BRCA1BRCA2 Cohort Consortium, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA* 2017 Jun 20;317(23):2402-2416 [FREE Full text] [doi: [10.1001/jama.2017.7112](https://doi.org/10.1001/jama.2017.7112)] [Medline: [28632866](https://pubmed.ncbi.nlm.nih.gov/28632866/)]
68. Clifton K, Gutierrez-Barrera A, Ma J, Bassett JR, Litton J, Kuerer H, et al. Adjuvant versus neoadjuvant chemotherapy in triple-negative breast cancer patients with BRCA mutations. *Breast Cancer Res Treat* 2018 Jul;170(1):101-109 [FREE Full text] [doi: [10.1007/s10549-018-4727-9](https://doi.org/10.1007/s10549-018-4727-9)] [Medline: [29470805](https://pubmed.ncbi.nlm.nih.gov/29470805/)]

69. Halpern N, Sonnenblick A, Uziely B, Divinsky L, Goldberg Y, Hamburger T, et al. Oncotype Dx recurrence score among BRCA1/2 germline mutation carriers with hormone receptors positive breast cancer. *Int J Cancer* 2017 May 01;140(9):2145-2149 [FREE Full text] [doi: [10.1002/ijc.30616](https://doi.org/10.1002/ijc.30616)] [Medline: [28120435](https://pubmed.ncbi.nlm.nih.gov/28120435/)]
70. Nicolas E, Bertucci F, Sabatier R, Gonçalves A. Targeting Deficiency in Breast Cancer: What are the Clinical Evidences and the Next Perspectives? *Cancers (Basel)* 2018 Dec 11;10(12):506 [FREE Full text] [doi: [10.3390/cancers10120506](https://doi.org/10.3390/cancers10120506)] [Medline: [30544963](https://pubmed.ncbi.nlm.nih.gov/30544963/)]
71. Paluch-Shimon S, Friedman E, Berger R, Papa M, Dadiani M, Friedman N, et al. Neo-adjuvant doxorubicin and cyclophosphamide followed by paclitaxel in triple-negative breast cancer among BRCA1 mutation carriers and non-carriers. *Breast Cancer Res Treat* 2016 May;157(1):157-165 [FREE Full text] [doi: [10.1007/s10549-016-3800-5](https://doi.org/10.1007/s10549-016-3800-5)] [Medline: [27113739](https://pubmed.ncbi.nlm.nih.gov/27113739/)]
72. Fasching P, Loibl S, Hu C, Hart S, Shimelis H, Moore R, et al. BRCA1/2 Mutations and Bevacizumab in the Neoadjuvant Treatment of Breast Cancer: Response and Prognosis Results in Patients With Triple-Negative Breast Cancer From the GeparQuinto Study. *J Clin Oncol* 2018 Aug 01;36(22):2281-2287 [FREE Full text] [doi: [10.1200/JCO.2017.77.2285](https://doi.org/10.1200/JCO.2017.77.2285)] [Medline: [29791287](https://pubmed.ncbi.nlm.nih.gov/29791287/)]
73. Davey M, Davey C, Ryan E, Lowery A, Kerin M. Combined breast conservation therapy versus mastectomy for BRCA mutation carriers - A systematic review and meta-analysis. *Breast* 2021 Apr;56:26-34 [FREE Full text] [doi: [10.1016/j.breast.2021.02.001](https://doi.org/10.1016/j.breast.2021.02.001)] [Medline: [33582622](https://pubmed.ncbi.nlm.nih.gov/33582622/)]
74. Wan Q, Su L, Ouyang T, Li J, Wang T, Fan Z, et al. Comparison of Survival After Breast-Conserving Therapy vs Mastectomy Among Patients With or Without the BRCA1/2 Variant in a Large Series of Unselected Chinese Patients With Breast Cancer. *JAMA Netw Open* 2021 Apr 01;4(4):e216259 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.6259](https://doi.org/10.1001/jamanetworkopen.2021.6259)] [Medline: [33890992](https://pubmed.ncbi.nlm.nih.gov/33890992/)]

Abbreviations

BCT: breast-conserving treatment
ER: estrogen receptor
FIGO: International Federation of Gynaecology and Obstetrics
HER2: human epidermal growth factor receptor 2
pCR: pathologic complete response
PR: progesterone receptor
RRBM: risk-reducing bilateral mastectomy
RRBSO: risk-reducing bilateral salpingo-oophorectomy
TN: triple negative

Edited by A Mavragani; submitted 18.02.22; peer-reviewed by ER Khalilian, I M, J Zhu; comments to author 07.09.22; revised version received 30.10.22; accepted 01.11.22; published 29.12.22.

Please cite as:

Maksimenko J, Rodrigues PP, Nakazawa-Miklaševića M, Pinto D, Miklaševićs E, Trofimovičs G, Gardovskis J, Cardoso F, Cardoso MJ

Effectiveness of Secondary Risk-Reducing Strategies in Patients With Unilateral Breast Cancer With Pathogenic Variants of BRCA1 and BRCA2 Subjected to Breast-Conserving Surgery: Evidence-Based Simulation Study

JMIR Form Res 2022;6(12):e37144

URL: <https://formative.jmir.org/2022/12/e37144>

doi: [10.2196/37144](https://doi.org/10.2196/37144)

PMID: [36580360](https://pubmed.ncbi.nlm.nih.gov/36580360/)

©Jelena Maksimenko, Pedro Pereira Rodrigues, Miki Nakazawa-Miklaševića, David Pinto, Edvins Miklaševićs, Genadijs Trofimovičs, Jānis Gardovskis, Fatima Cardoso, Maria João Cardoso. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 29.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Acceptance, Usability, and Utility of a Web Portal for Back Pain as Recommended by Primary Care Physicians: Qualitative Interview Study With Patients

Christian Schlett¹, PhD; Nicole Röttele², PhD; Piet van der Keylen³, Prof Dr; Andrea Christina Schöpf-Lazzarino¹, PhD; Miriam Klimmek¹, BA; Mirjam Körner², Prof Dr; Kathrin Schnitzius³, MSc; Sebastian Voigt-Radloff⁴, PhD; Andy Maun⁵, Prof Dr; Mario Sofroniou⁵, DMed; Erik Farin-Glattacker¹, Prof Dr

¹Section of Health Care Research and Rehabilitation Research, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

²Institute of Medical Psychology and Medical Sociology, Faculty of Medicine, University of Freiburg, Freiburg, Germany

³Institute of General Practice, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany

⁴Institute for Evidence in Medicine, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany

⁵Institute of General Practice / Family Medicine, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

Corresponding Author:

Christian Schlett, PhD

Section of Health Care Research and Rehabilitation Research

Medical Center – University of Freiburg, Faculty of Medicine

University of Freiburg

Hugstetter Str. 49

Freiburg, 79106

Germany

Phone: 49 27083732

Email: christian.schlett@uniklinik-freiburg.de

Abstract

Background: An ever-increasing number of patients seek health information via the internet. However, there is an overabundance of differing, often low-quality information available, while a lack of health literacy makes it difficult for patients to understand and assess the quality and trustworthiness of the information at hand. The web portal tala-med was thus conceived as an evidence-based, up-to-date, and trustworthy information resource for lower back pain (LBP), which could be used by primary care physicians (PCPs) and patients during and following consultations for LBP. The current evidence demonstrates that patients with LBP could benefit from web portals. However, the use of such portals by patients remains low, thus limiting their effectiveness. Therefore, it is important to explore the factors that promote or hinder the use of web portals and investigate how patients perceive their usability and utility.

Objective: In this study, we investigated the acceptance, usability, and utility of the web portal tala-med from the patient perspective.

Methods: This qualitative study was based on telephone interviews with patients who had access to the web portal tala-med from their PCP. We used a semistructured interview guide that consisted of questions about the consultation in which patients were introduced to tala-med, in addition to questions regarding patient perceptions, experiences, and utilization of tala-med. The interviews were recorded, transcribed, and analyzed through framework analysis.

Results: A total of 32 half-hour interviews were conducted with 16 female and 16 male patients with LBP. We identified 5 themes of interest: the use of tala-med by PCPs during the consultation, the use of tala-med by patients, its usability, added values derived from its use, and the resultant effects of using tala-med. PCPs used tala-med as an additional information resource for their patients and recommended the exercises. The patients appreciated these exercises and were willing to use tala-med at home. We also identified factors that promoted or hindered the use of tala-med by patients. Most patients rated tala-med positively and considered it a clear, comprehensible, trustworthy, and practical resource. In particular, the trustworthiness of tala-med was seen as an advantage over other information resources. The possibilities offered by tala-med to recap and reflect on the contents of consultations in a time-flexible and independent manner was perceived as an added value to the PCP consultation.

Conclusions: Tala-med was well accepted by patients and appeared to be well suited to being used as an add-on to PCP consultations. Patient perception also supports its usability and utility. Tala-med may therefore enrich consultations and assist patients who would otherwise be unable to find good-quality web-based health information on LBP. In addition, our findings support the future development of digital health platforms and their successful use as a supplement to PCP consultations.

International Registered Report Identifier (IRRID): RR2-10.1186/s12875-019-0925-8

(*JMIR Form Res* 2022;6(12):e38748) doi:[10.2196/38748](https://doi.org/10.2196/38748)

KEYWORDS

general practice; primary care; lower back pain; digital health intervention; web-based health information; eHealth; patient education; adherence; qualitative research; framework analysis; mobile phone

Introduction

Background

An ever-increasing number of patients access health information via the internet [1-3]. However, most of the health-related information available on the web is of low quality, while many patients are unable to adequately appraise the quality of such health information [4,5]. *Tala-med* was therefore envisaged and developed as an evidence-based, up-to-date, and easily understood web-based information resource for lower back pain (LBP) through the *Gut informierte Kommunikation zwischen Arzt und Patient* (GAP) project (well-informed communication between the general practitioner and patient) [6]. LBP in particular is a widespread health problem that causes substantial personal and financial burden [7,8]. In Germany, LBP has a 1-year prevalence of more than 60% [9,10], accounts for the most days of sick leave (6.1%) [11], and is one of the most common reasons why people visit their doctor [12].

Despite the existence of national and international clinical guidelines, approaches to treating LBP differ greatly among clinicians, institutions, and geographic regions [13,14]. However, the breadth of available web-based information often surpasses the variety of management approaches. Consequently, patients can often be confused and frustrated while searching for web-based information regarding LBP [15]. An overabundance of differing and contradictory information can make it difficult for patients to understand and assess the quality and trustworthiness of the information provided [15], often presenting a dilemma for patients searching web-based information for LBP.

Digital health interventions (DHIs) recommended by health care professionals (HCPs) may be a remedy for this dilemma, as they can provide patients with tailor-made, understandable, and high-quality information. Our web portal *tala-med* is one such DHI that could be recommended to patients by primary care physicians (PCPs). As an information resource based on national [16,17] and international clinical guidelines for LBP [18-20], *tala-med* can be classified as a DHI that can provide health content [21,22] to physicians and patients alike. *Tala-med* aims to improve the shared decision-making of PCPs, while enhancing patient-informed choices, participation, and self-management regarding LBP [6]. For interventions on shared decision-making to be most effective, Cochrane reviews have shown that these should be both physician and patient focused and include information that is indication specific [23,24].

Regarding DHIs on back pain, recent meta-analyses and systematic reviews have shown that such interventions, especially those that focus on self-management, can have clinically important effects in terms of relieving patient discomfort and improving their disability [25-27]. However, a key determinant of the effectiveness of DHIs on LBP [28] and DHIs in general [29] is adherence; that is, whether patients actually use the DHI to the intended extent [30].

Unfortunately, the extent to which patients use DHIs is often low [31-33]. In addition, promoting the use of DHIs by patients is complex, with only limited evidence available on successful strategies to do so [29]. Therefore, it is important to understand the web-based health information needs of patients [34] and other factors that may facilitate or hinder their use of a DHI. Regarding web-based information on LBP received by patients as an adjunct to their PCP consultation, Riis et al [15] found that readability, customization, design, credibility, and usability are important domains. However, these results were based on patient experiences using a variety of health-related websites and not on the use of a specific web portal provided by their PCP. Studies on patient perceptions of DHIs for LBP have also revealed that contextual factors, such as the support of HCPs, and individual factors, such as patient skill and preference, affect the acceptance of such DHIs [28,35]. A current systematic review of qualitative studies found only 4 studies that investigated the facilitators of and barriers to the use of DHIs by patients with LBP. Svendsen et al [28] thus state that “further primary research investigating the implementation of DHIs and user’s experiences is required.”

Objectives

The aim of this study was to examine patient acceptance of our web portal *tala-med* as well as its usability and utility. We were eager to see how PCPs used the portal during the consultation, how they offered it to patients, and whether it was subsequently used. We also wanted to understand which parts of the portal were considered helpful and how patients perceived *tala-med* in terms of key characteristics such as comprehensibility and trustworthiness, enabling us to identify the portal’s strengths and weaknesses. This in turn provided us with insights into how specific features of the portal or its setting may have contributed to patient perception. Finally, we examined the utility of *tala-med*, in particular, the perception of any added value brought about by its use.

Methods

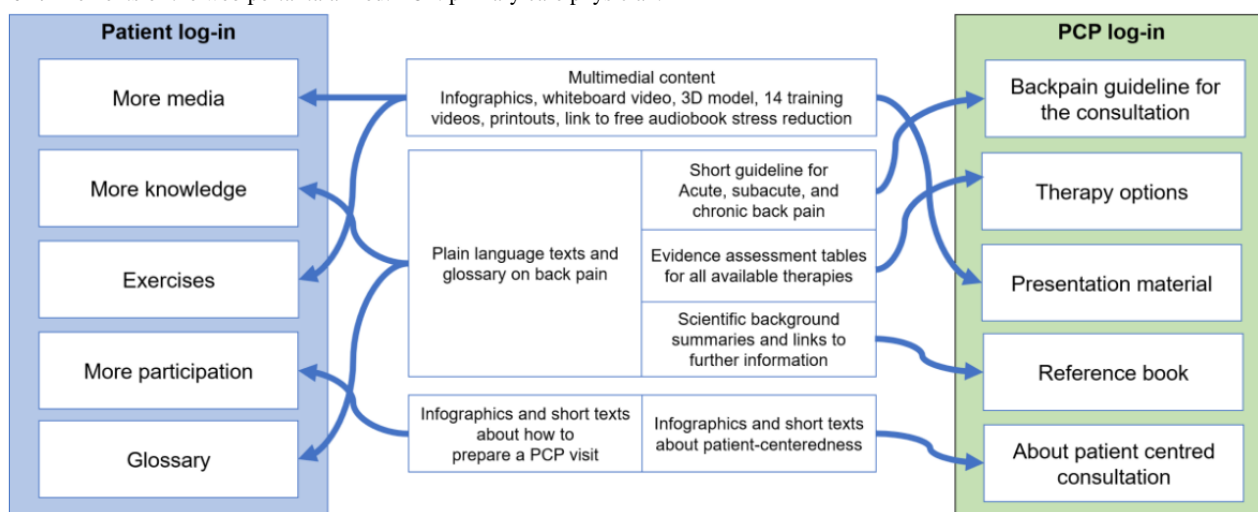
The Web Portal Tala-med

Tala-med is a German-language, evidence-based, comprehensible, and reliable internet information portal for LBP that is freely accessible and aimed at improving patient informedness and patient-doctor interaction. It was implemented within a prospective cluster-randomized controlled trial (RCT) with pre-, post- and 1 follow-up measurement [6]. *Tala-med* was designed to be used by PCPs and patients in the intervention group during and after LBP consultations. There are 2 versions that have been adapted to meet the demands and linguistic levels of PCPs and patients (Figure 1). To log in to the portal, both PCPs and patients were given an individual fictitious username consisting of a combination of 2 short animal, color, and fruit words (eg, PearOwl) and a password. This type of log-in was necessary because portal use data were also analyzed anonymously as part of the RCT [6].

The portal was developed over a 6-month period by a multidisciplinary team consisting of 1 PCP, 2 physiotherapists, 2 researchers specializing in evidence synthesis, 2 web designers, a film crew concerned with exercise video production, a specialist for whiteboard videos, and a 3D specialist. The focus was on frequently occurring issues seen in LBP, and it contained descriptive material in short and long versions, such

as an accordion style guideline, where users could click on headings and highlighted keywords to obtain further background information. The development of the portal, including information processing and design principles on which the presentation of the information was based, is described in the associated methodological guide, which is also publicly available [36]. The final version consisted of 6 infographics, with 20 illustrations, a 3D model animation of the lower back, an 11-minute explanatory whiteboard video containing 8 different topics, and 14 short exercise videos 2 to 4 minutes long, with an emphasis on teaching self-care. In addition, suggestions, infographics, and background information on well-informed shared decision-making and preventive lifestyle changes were also included. The PCPs were able to show or print material during their consultations. Study PCPs were instructed to show and explain the web portal to their patients and encourage its use. If this was not possible, for instance, if the consulting room had no computer access, PCPs were advised to promote its use, despite being unable to demonstrate the portal in real time. As a backup option, PCPs were given up to 4 hard-copy brochures, which could be handed out to patients who were unfamiliar with computer use. The restriction to 4 brochures was chosen to encourage PCPs to primarily use the web portal and only refer to a brochure when handing the portal to a patient seemed inappropriate. The brochures contained information from the web portal in printed form. Patients were able to use the patient-tailored version of the web portal at home.

Figure 1. Elements of the web portal tala-med. PCP: primary care physician.



Study Design

We used a qualitative design in which patients with LBP, who had received access to *tala-med* via their PCP, were invited by post to individual telephone interviews. For administrative reasons and owing to the delay of patient postal responses, interviews were held 1 to 2 months after the consultation, in which patients received their log-in details to the web portal. Results were reported according to the consolidated criteria for reporting qualitative research (consolidated criteria for reporting qualitative research [37]).

Recruitment

All 190 patients with LBP from the intervention group were invited to participate in an interview after the last follow-up measurement of the RCT. Although patients received a book voucher for their participation in the RCT, no incentives were offered for their participation in the subsequent interview study. Then, 35 patients (18%) accepted the invitation and returned their informed consent and contact information. Three patients were not interviewed: 2 could not be reached and 1 was no longer interested in an interview, as she had not yet used the web portal. A total of 32 patients were interviewed, and 17% of all invitees were interviewed. The patient characteristics are shown in Table 1.

Table 1. Patient characteristics (N=32).

Characteristic	Patients, n (%)
Gender	
Male	16 (50)
Female	16 (50)
Age (years)	
18-29	1 (3)
30-39	8 (25)
40-49	4 (13)
50-59	10 (31)
60-69	8 (25)
70-79	1 (3)
Level of education (highest level completed)	
Elementary school	10 (31)
Secondary school	15 (47)
High school diploma	7 (22)
Do you currently still have back pain?	
Yes	18 (56)
No, not at the moment	7 (22)
No (without any remark)	7 (22)
First experience of back pain in your life?	
<6 weeks	0 (0)
6-12 weeks	1 (3)
>12 weeks to <1 year	0 (0)
1 to <2 years	2 (6)
2 to <5 years	3 (9)
5 to <10 years	3 (9)
>10 years	23 (72)

Ethics Approval

The study was approved by the ethics committee of the Albert-Ludwigs-University Freiburg (no 559-17).

Informed Consent

All participants provided written informed consent before being interviewed. Apart from contact data, the interviewers had no prior information about the participants. At the beginning of the telephone interviews, patients were informed about data protection issues and the recording of the interviews ([Multimedia Appendix 1](#)).

Data Collection

Interviews were held between February 2019 and October 2020 by CS (31 interviews) and M Klimmek (1 interview). These were based on a semistructured interview guide developed by ACSL, CS, and NR, which contained questions about patient

experiences of the consultation and web portal ([Multimedia Appendix 1](#)). The interview guide contained skip rules to ensure that interviewees were only asked questions that they could provide answers for. For example, patients who did not use the portal were not asked questions regarding the usability of the portal and its utility. The interviews lasted an average of 30 minutes, ranging from 15 to 45 minutes in length. Interviewers did not take field notes during the interviews. The interviews were digitally recorded by connecting a digital recorder to the interviewer's landline telephone. Recordings were transcribed verbatim by a transcription service and analyzed using MAXQDA (version 20; VERBI Software GmbH). Participants did not receive the transcripts or feedback results.

Data Analysis

The transcribed interviews were analyzed using framework analysis [38-40]. [Textbox 1](#) outlines the 5 stages of the analysis as well as their implementation.

Textbox 1. Stages of the framework analysis.

- Familiarization: the purpose of this stage was to become immersed in the data to gain an insight into its range and diversity [39]. Therefore, CS read the transcripts of 10 in-depth and heterogeneous interviews.
- Identifying a thematic framework: this stage was concerned with developing a code system that covered the most important issues [39]. Codes were first developed deductively based on the main themes of the interview guide. They were then expanded inductively with codes covering the emerging issues. In this vein, CS coded 10 transcripts and created a code system. Because a single look at the data might miss important issues or overemphasize less important ones, we included a second view. Therefore, NR coded 5 transcripts independently and also created a code system; CS and NR discussed and combined their code systems into one. With this new code system, CS coded the next 14 transcripts and refined the system accordingly. To verify the refined system, 3 in-depth and heterogeneous transcripts (of the 14 last coded ones) were coded again independently by NR with the refined code system. CS and NR compared the coding of these 3 interviews. Because of high consistency, only a few changes were necessary to create a final code system.
- Indexing: this stage described the application of the code system to the entire set of data [39]. CS and M Klimmek applied the final code system to all transcripts and verified its application reciprocally.
- Charting: at this stage, coding was extracted and tabulated, with columns representing codes and rows representing patients. This allowed codes to be read horizontally for a given patient or patients to be read vertically for a given code. CS and M Klimmek condensed and summarized the coding as far as possible in the patients' own words and created a chart with codes represented as columns and patients as rows.
- Mapping and interpretation: this stage concerned the mapping out of the data and making sense of it [40]. It may include a description of the range and nature of phenomena and searching for associations between and within codes to find explanations for the research questions [39]. To lay out and make sense of the data, we used a method called one sheet of paper (OSOP) [41]. This method entailed reading the condensed extracts of each code and summarizing all the different issues of a code on OSOP. This summary of the different issues was then used as a basis for axial coding, that is, for considering which issues group into broader themes and to "develop an explanation of 'what is going on in the data' that takes account of all the issues raised" [41]. In this way, CS summarized the different issues of each code, oversaw the summary of issues, and considered how they form broader themes that could provide explanations for the research questions. Five core themes emerged from this overview: primary care physician (PCP) use of the portal during the consultation, patient use of the portal, usability, added value, and effects of the portal. Regarding these themes, CS selected the codes that provided information (Multimedia Appendix 2) and searched for associations between the codes of each theme and between the themes. CS made a first draft of results and discussed and refined it with M Klimmek. The refined results were discussed with NR, PK, ACSL, M Klimmek, M Körner, SVR, and EFG and adapted by CS.

As described in [Textbox 1](#) and in the previous section, CS, ACSL, M Klimmek, and NR were the researchers primarily involved in data collection and analysis. CS and ACSL were postdoctoral researchers in the field of health services research and rehabilitation research, who hold degrees in psychology. M Klimmek was a bachelor's student in social work who worked as a student assistant in the same field. NR is a researcher in the field of medical psychology and medical sociology who holds a degree in health education. ACSL is experienced with qualitative studies. She introduced CS and NR to the framework analysis.

Results

We identified 5 core themes using framework analysis: PCP use of the portal during the consultation (theme 1), patient use of the portal (theme 2), usability (theme 3), added value (theme 4), and effects of the portal (theme 5).

Acceptance

Two themes were related to patient acceptance of the portal: PCP use of the portal during consultation (theme 1) and patient use of the portal (theme 2).

PCP Use of the Portal During Consultation

This theme reflected on how PCPs introduced the portal to patients and the usage they encouraged. The following represents patient perceptions of PCP behavior during consultation. All 32 patients interviewed provided insight into their perception of the consultation, although 9 of these patients also expressed difficulty in recalling the consultation, as it had taken place so long ago. The degree to which PCPs introduced the portal to

their patients varied greatly: from detailed explanations with a demonstration to some explanation to no explanation at all. Many patients reported that their PCP had explained and showed them the portal, either on the screen or by using the brochure. Some also received printouts of exercises or other information from the portal of their PCP. Others reported that their PCP described aspects of the portal, such as the log-in, use, and contents, without showing it. These patients received, as a minimum, general information or a recommendation, for instance, "the portal contains exercises that might be helpful to you" (patient 3, male). Some patients were simply asked to participate in the study or were given log-in details by a physician assistant without any further information.

The PCPs who mentioned the portal mainly suggested that the patients used the exercises. Other aspects of the portal, for instance, those aimed at knowledge transfer or patient-doctor communication, were suggested to only a few patients. Even if PCPs showed their patients the portal, most patients perceived the consultation as per usual or differing only slightly from previous consultations. Only a few patients reported any noticeable changes in the consultation owing to the use of the portal, such as more in-depth and informative conversations, or more time and interest on the part of their PCP. Patient satisfaction with the consultation seemed unaffected by PCP use of the portal. Rather, it depended on the general aspects of the consultation, such as how much time the PCP devoted to them, whether the conversation was perceived as in depth, whether the PCP seemed interested in their recovery, or whether patients received the treatment they had hoped for. Most patients whose PCP used the portal endorsed the use of the portal in

future consultations. However, 1 patient expressed concerns about how the portal should be used:

Only if the PCP uses it in detail, so that it's not such an everyday project. [Patient 22, female]

I beg your pardon? Only if the PCP? [Interviewer]

Well, if the PCP really works intensively with it, so that he doesn't just say: hey, there's a program, do this and bye-bye, handing over another piece of paper, but really goes into it in more detail. [Patient 22, female]

Patient Use of the Portal

Of the 32 patients interviewed, 4 used both the web portal and the brochure, 22 used the portal alone, 2 used the brochure alone, and 4 used neither the portal nor the brochure. If a patient used the portal, the interview focused on its use, even if the patient also used the brochure.

Textbox 2 presents the facilitators of and barriers to portal usage. As in previous studies, factors relating to the initial use of the portal did not differ from those relating to the continued use of the portal [28] and were therefore reported together. Seeing the portal during the consultation raised initial patient interest and shortened the time to their first log-in.

Textbox 2. Facilitators of and barriers to portal usage (N=32).

Facilitators

- Conditions and treatments
 - Primary care physicians (PCPs) showed portal in the consultation
 - PCPs recommended exercises
 - Back pain started again
 - Comorbidity that benefits from exercise
- Organization and motivation
 - Time-flexible use of the portal
 - Emerging specific questions
 - High self-motivation of patients

Barriers

- Conditions and treatments
 - Currently absent or severe back pain
 - Other more intensive back pain therapies
 - Other health problems had priority
- Organization and motivation
 - Lack of time (due to work, household, childcare, or care for older adults)
 - Lack of motivation, patience, or concentration
 - Portal provided no new suggestions
- Technical requirements and skills
 - PC or internet problems
 - Lack of PC skills
 - Use on smartphone not possible
 - Log-in: details lost or did not work

When asked about when he first looked at the portal, a patient who had been shown the web portal during the consultation replied as follows:

Well, it was either on the same day or the day after, just out of curiosity. [Patient 5, male]

So, you logged in straight away and looked at what was available? [Interviewer]

Exactly. I read it through more carefully myself to see what exercises there are and...Yes, it was actually a good suggestion by the physician. [Patient 5, male]

In addition, patients were more willing to try exercises from the web portal if their PCP recommended them. They were also more likely to use the portal if they had a recurrence of LBP. No LBP and too much LBP were barriers to portal use: patients

saw no reason to use *tala-med* if they no longer had any LBP or equally were unable to use it if they were experiencing excessive LBP. Similarly, patients did not use the portal if they underwent more intensive treatment for LBP, such as in-patient therapy. Some patients were unable to use the portal because of other acute conditions, especially mental health conditions, such as depression. However, if they had another condition that also benefited from exercise, such as hypertension, they were more willing to use it.

The nature of the portal as a time-flexible web-based tool made it easier for patients to fit its use in their daily routines. Patients were inclined to use *tala-med* when new questions about LBP arose, as well as when they demonstrated increased motivation to improve their own health. Lack of time owing to other commitments was a frequent barrier to their use. A lack of motivation, patience, or concentration was a reason why patients used the portal less often or no longer. Patients were also abstained from further use if new suggestions within the portal could not be found. Technical requirements and skills were also important barriers to their use. If a patient's device or internet connection did not work or they were unfamiliar with their functions, they also, understandably, did not use the portal. Some patients held reservations about using the portal owing to difficulties using a computer or the internet. Lost log-in details or other log-in difficulties were also obvious barriers to patient use.

Most of the 26 patients who used the portal logged into it for the first time within the first 3 days following a consultation. The exercises were most often perceived as the most helpful part of the portal. A substantial proportion of patients were interested in web portal exercises to alleviate their pain but showed no interest in other aspects of the portal:

This theoretical background did not interest me further in this case. I just wanted to do these exercises. [Patient 5, male]

I was really only interested in the exercises, because they...well...help the most and, no, I didn't want to look up or know more. [Patient 16, female]

Nonetheless, the other 3 sections—more media, more knowledge, and more participation (Figure 1)—were also perceived by some of the patients as being the most helpful parts of the portal.

Usability

Responses regarding the usability of the portal (theme 3) relied on the answers of 26 patients. However, 10 of these patients mentioned that they had difficulty remembering the portal in detail. Textbox 3 shows how patients rated the portal and their respective reasons. In general, most patients rated the portal as positive. They mentioned that the portal was interesting and

informative, provided good information, was easy to use and easy to implement, and contained useful exercises and videos. Neutral and negative overall ratings were obtained from patients who were disappointed that the portal did not recommend their preferred treatment (injection) or lacked information about a particular type of back pain (upper back pain).

Reasons for positive design ratings were that the portal was perceived as visually well-structured and uncluttered. Patients also liked the brevity and simplicity of the exercise videos:

I also find the videos very beautiful. Not so crowded, but just a person who shows this, does that. Not so much jumping around and stuff, like some others there, when you look on the internet and everyone thinks, they have to do fancy other things, but very simple. You do that, that's how it should be done, and then you do it that way and that's it. So, I think it's good just as it is. [Patient 30, male]

Patients with neutral or negative perceptions of the portal design reported that they preferred dealing with a real person, needed more interactive elements and animations, and found that the portal was poorly designed for smartphone use. The portal was optimized for notepad-sized screens or larger screens.

Most patients perceived the portal as clear and did not have any problems interacting with it. They found it simple and well structured, well described, and thus easy to navigate. For some patients, the portal structure was unclear; 1 patient stated that he could not find everything he needed from the outset, reporting that “it would be easier for a younger person who sits in front of the computer all the time” (patient 21, male, age category: 50-59 years). The other patients had problems dealing with the portal, needing help, or finding it too convoluted and deeply structured, requiring too many clicks to find what they were searching for.

Almost all patients perceived the portal information as easy to understand, even for back pain novices, because the information was rated as simple and well described, containing only a few specific terms with no extensive texts. The portal information was unanimously perceived as trustworthy. Patients mentioned 4 main reasons for this (Textbox 4). Many patients perceived the portal as trustworthy, as it was developed or recommended by a source with a high level of expertise:

How trustworthy did you find the information on the platform? [Interviewer]

Very trustworthy. [Patient 32, male]

On what did you base that on, the trustworthiness? [Interviewer]

I put it down to the fact that my doctor recommended it to me and I actually trust her very much. [Patient 32, male]

Textbox 3. Overall and design ratings of the web portal (N=26).

Overall rating

- Positive
 - Good information
 - Well-explained exercises, easy to implement
 - Very good background and explanations
 - In sum very interesting, very informative
 - Quite good at the beginning for browsing
 - Pragmatic and practicable
 - Very easy to use, well structured
 - Simplified, everyone will get along well with it
 - Great videos
- Neutral
 - The portal did not recommend patients' preferred treatment
 - I can't judge it
- Negative
 - Too little info about upper back pain

Design

- Positive
 - Yes and no and questions and answers is good
 - Very appealing, not too cluttered
 - Videos are great
 - From the duration
 - Not so much jumping around
 - Super to see a normal person (no super athlete)
 - Visually well constructed
 - Nothing visually disturbing
- Neutral
 - No real person behind it
 - Not amazing but not bad either
 - I did not despair of it
- Negative
 - Conservative and classic, too little interactive and animated
 - Very smartphone unfriendly

Textbox 4. Clarity, comprehensibility, and trustworthiness of the web portal (N=26).

Clear

- Yes
 - I had no problems finding my way around
 - Very simple overview, for the inexperienced
 - You quickly get to where you want to be
 - It was easy to find specific information
 - Everything is well described and clear
 - Well structured
- Partly
 - It took a bit to get into it, but then it was clear
- No
 - I didn't get on with it, needed help
 - Easier for a younger person, who always sits in front of the computer
 - Too convoluted, too deeply structured
 - I have to click a lot until I find what I am looking for

Comprehensible

- Yes
 - Everything is very simple and well described
 - Few strange words or specific terms
 - No long texts
 - Many things are relatively well explained
 - Easy to understand even for back pain novices
 - Videos are very well described, they are self-explanatory
- No
 - Contained technical terms, which the patient did not comprehend

Trustworthy

- Yes
 - Trustworthy source with high expertise
 - Recommended by primary care physician (PCP)
 - Developed by experts (doctors or universities)
 - Well-founded, scientific info
 - Well-founded impression
 - Very informative and scientifically developed
 - Emphasis on info, not on fuss
 - Accurate, noncontradictory information
 - Things patient knows to be true
 - Nothing contradictory
 - Serious presentation
 - Not the impression of advertising

- Much attention to data protection
- Fictitious username and password
- Positive intention noticeable

Furthermore, underpinnings of trustworthiness were the well-founded and scientific information of the portal and its serious presentation, putting an emphasis on the clear presentation of information, while avoiding distractions or advertisements. Patients also perceived the portal as trustworthy, as it contained noncontradictory information with aspects the patients knew to be true. The use of a fictitious username with great attention to data protection also supported the impression of trustworthiness for some patients.

A total of 14 patients had suggestions for improvement. Regarding additional content, patients suggested adding the addresses of recommended specialist centers, doctors, or therapists in their region, with freely available consultation slots when a second opinion was needed at short notice, as well as an advice hotline on the contents of the portal. Including more exercises for the upper back, alternative therapies such as acupuncture, and offering support for other conditions in addition to LBP were also suggested. Design improvements were also envisaged through the use of greater customization, with separate access for patients with little experience of LBP, as well as with more experience of LBP. Regarding the exercise videos, shorter sequences for use on the go, as opposed to the current 2- to 4-minute videos on offer, would have been appreciated, as well as the ability to loop videos, adjusting the number of repetitions available for exercising in tandem with the videos. One patient would have appreciated instructions to be given throughout the program, while another patient felt the need for a diagnosis-orientated search facility, with diagnosis-specific preselection of information and exercises. Regarding accessibility, many patients felt that the portal should be made freely available to everyone, which was unfortunately not the case at the time of the study. One patient suggested that the portal be made widely available in waiting rooms and pharmacies. Some hoped that it would be made available as a smartphone app, while others were eager for a hardcopy printout to be made available to those without a computer or access to the internet. Almost all patients affirmed that they would recommend the portal to others and some had already done so.

Utility

The utility of the portal was assessed by its added value (theme 4) and its effects (theme 5).

Added Value

If patients are offered a web portal by their PCP as a supplement to the consultation, this should bring added value to the consultation or to preexisting sources of information that patients may otherwise have access to. Patients highlighted the

trustworthiness and validity of the information available through *tala-med* as bringing added value above and beyond those gained by accessing other sources of information, such as self-guided internet searches (Textbox 5).

Patients also praised the comprehensiveness of the information provided in *tala-med* and its appropriateness. In addition to the perceived value of such features, patients appreciate the time and effort saved when searching for health information. Patients also see added value in watching the exercise videos, as opposed to paper instructions or reading them on the web, which may be less easy to understand. In addition to these specific aspects of the portal, the practical relevance and the positive effects on LBP were also seen as bringing added value:

The added value for me was definitely that it comes from a clinic and my doctor recommended that I should use it. That was actually the added value for me and that I can draw the conclusion that I definitely feel better because of it. [Patient 32, male]

Two patients saw no added value in using *tala-med* as they felt it did not offer any new suggestions or corrections to existing exercises, as one would experience in a face-to-face course, the latter being preferred by one of the patients.

Compared with a PCP consultation without subsequent access to *tala-med*, patients saw added value in the possibility to repeat, deepen, and reflect on the contents of the consultation (Textbox 6).

How much (given as a percentage) does one really remember in a doctor-patient conversation? Not that much, right? And then you can just read about it [in the portal]. And that was quite good [...] You can just have another look: Ah...Now I have another question. Or [I] can take another look: Now that would have interested me, I forgot [that]. See if I can find something in there. Well, I find it good as an additional offer. Of course, it doesn't replace the personal doctor-patient consultation. [Patient 15, female]

Patients appreciated the time-flexible and independent use of the portal, which also supported their active role in taking care of their own LBP. Having access to the portal motivated patients to engage with the content and made it easier for them to implement the necessary exercises. Patients also saw the positive effects of using the portal (see the *Effects* section) as bringing added value to the PCP consultation. They stated that less would have been known about their LBP and less exercise would have been done had they not had access to *tala-med*.

Textbox 5. Added value of the web portal to other sources of information (N=26).

Trustworthiness

- Better than surfing and only coming across advertising or nonsense content
- The content and motives of the portal do not need to be verified, as the portal is trusted
- One knows that one is on the right website

Information validity

- Content is valid, more profound, well founded, and professional than what is found elsewhere on the internet
- Portal contains independent information; it contains more than just one person's experience

Comprehensive information

- Very bundled and compact; it comprises causes, treatment options, and exercises all in one
- Saves time-consuming search on the internet or elsewhere; other sources of information become unnecessary

Appropriate information

- Information is prescreened, specific to back pain (DVDs with exercises are often less specific)
- Suits what the patient is currently dealing with
- Facilitates further searches for appropriate information

Good exercise videos

- Exercise videos are better and more motivating than instructions on paper
- Exercises are well explained, which was not the case with the results of an internet search

Textbox 6. Added value of the web portal to the primary care physician (PCP) consultation (N=26).

Additional information to the PCP consultation

- To repeat:
 - To be able to read at leisure again or recall what the PCP said
 - No need to remember everything from the conversation because patients could read it again later
 - Reading up later removes uncertainty
- To deepen:
 - To deepen what the PCP said; provided targeted additional information beyond the normal consultation
 - Supplementary knowledge and new exercises
- To reflect:
 - Allows comparison with what was said by the PCP
 - Be better informed to ask my PCP questions next time

Time-flexible use

- Visit the website whenever I wanted and when I had time; always accessible
- Glance quickly when I have a pain episode; start directly with exercises, do not have to wait for an appointment

Independent use

- Practical solution for home, I could do something without needing a health professional
- Can be flexibly integrated into my daily routine, one is not dependent on someone else

Motivates engagement

- Pushes me to move more; encourages me to overcome my weaker self
- Compulsion to do something until the next consultation (to familiarize myself, to try something out)

Facilitates implementation

- Exercises are well explained and easy to follow because of the videos

Effects

Overwhelmingly positive outcomes were mentioned by patients, highlighting the effects that using the portal had on their degree of LBP, informedness, and patient participation. Many patients were able to alleviate their LBP by using the portal, specifically as a consequence of doing the exercises. Unfortunately, 1 patient with preexisting severe LBP experienced worsening LBP while performing exercises. Patients perceived the portal as an impetus to do more exercise regularly, as well as more sport or movement in general. They also felt that doing the exercises might decrease or delay their need to visit their PCP again.

Using the portal also increased the informedness of patients with LBP. It provided them with an overview of treatments, improving their understanding of LBP, as well as its causes, while offering new perspectives on their own contribution to their LBP. Increased informedness and shared information with their PCP were seen as advantageous for upcoming consultations. Patients felt that they could (1) enter consultations with better prior knowledge, (2) ask targeted questions more easily, and (3) speak to their PCP about specific topics and exercises. Patients also assumed that increased informedness might reduce their need for PCP visits:

Do you think that...if you use the portal, it influences conversations with your PCP or with other people in the healthcare system? [Interviewer]

Yes, definitely! Definitely. Because, after all, there are many people who don't know where back pain comes from. Or how to avoid it. I actually think that maybe you don't need the doctor as much. [Patient 18, female]

Using the portal also facilitated patient participation, increasing patient awareness of the importance of the interaction with their PCP and encouraging them to express their thoughts and concerns through greater participation:

I have become more open in conversations with the doctor, so that I have dared more or have understood that I actually have to say what I think, only then can we really talk about it...it became clear to me that it is also from my side, yes...that I also have thoughts about my illness, or about my pain, and that I don't just have to perceive as valid what the doctor tells me. [Patient 11, male]

For the next PCP visit, I'll write down a few questions [in advance] and I'll be more pushy about my problems or my wishes. [Patient 8, male]

Discussion

Principal Findings and Comparison With Prior Work

Patients accepted the portal well, appreciating its use both in the consultation and at home. PCPs mainly used the portal as an additional information resource for their patients and recommended the exercises as described. Although patients appreciated this use, it only partially exploited the information potential of the portal. The parts of *tala-med*, which aimed to improve the consultation and in particular shared decision-making, seem to have had little impact. With respect to these aspects, the findings suggest that *tala-med* or its implementation could be improved. Many patients perceived the exercises, presented as guided videos, as the most helpful part of the portal, this being the sole area of interest for some. This finding supports the suggestion of Wollmann et al [34] that videos and tutorials about health information would be well received by patients.

Patient use of the portal was facilitated through PCP behavior during the consultation, such as introducing the portal and recommending it to patients. This finding is consistent with previous studies showing that HCP recommendations and support for the DHI were facilitators of use [28]. The current evidence underlines the important role of HCPs in promoting patient use of DHIs. Our findings also showed that portal use was facilitated through further questions from patients and through new onset and moderate LBP, which also served as a reminder to patients to take action. Other levels of LBP, in particular no LBP or excessive LBP, hindered patient use of the portal. According to a current systematic review, DHIs for the self-management of LBP should be tailored to pain severity [28]; otherwise, patients would not use them [15,28,42]. Our finding that absent or severe LBP was a barrier to portal use may reflect the degree to which *tala-med*'s contents align with pain severity, suiting patients with mild to moderate pain but less so those with absent or severe pain. This result is also consistent with a recent study by Geraghty et al [35] on the use of a DHI in primary care, which aimed to support patients in self-managing their LBP, who found mild and severe pain as barriers to DHI use. Thus, in primary care, DHIs for the self-management of LBP seem to be used primarily by patients with sufficient but not severe pain. In line with previous evidence [28], our findings showed that the use of the portal was facilitated by the high self-motivation of patients.

Barriers to the use of the portal included other more intensive back pain therapies and other acute conditions. Similar to our findings, Geraghty et al [35] described concurrent health conditions and comorbidities as barriers, although these specific comorbidities differed. Mental health conditions, especially depression, which is often associated with LBP [43], were not mentioned in previous studies as a barrier. However, the hindering effect of depression due to reduced drive and low energy has been observed with respect to the use of face-to-face pain self-management programs [44,45]. If depression hinders

patient engagement in face-to-face programs, this hindering effect should be even stronger with DHIs such as *tala-med*, which do not involve direct contact with an HCP and thus provide less-direct guidance and encouragement for use. Furthermore, we also found that comorbidities could facilitate patient use of a DHI if this had a positive effect on both LBP and the comorbidity, for instance, in the case of hypertension and LBP. Similar to previous studies [28,29,42,46,47], we also found that technical problems with the portal or little technical skill of the patient were an important barrier to portal use. The technical requirements and required technical skills of a DHI are particularly important because they can exacerbate inequalities in access to quality health information [48,49]. For example, older people and those with lower education and lower income could be disadvantaged, as a larger proportion of these groups do not have access to or cannot use the internet [50,51]. Therefore, when developing and using a DHI, care should be taken to ensure that the information contained could be made available to patients in a nondigital format. The web portal *tala-med* considered this by offering a hard-copy brochure that covered all topics of the portal, as well as printable information graphics, fact sheets, checklists, and exercise sheets that contained summaries of single topics.

Overall, patients perceived *tala-med* as usable, rated it positively, and considered it a clear, comprehensible, trustworthy, and practical resource that they would recommend to others. Our findings are consistent with previous studies showing that approval by an HCP supported patient trust in the quality of the contents of a DHI [29,52]. In addition, they provided insights into further characteristics of the portal content or its presentation that also contributed to the portal's trustworthiness from the patients' perspective. The usability of *tala-med* could be further enhanced by offering it as a smartphone app, in addition to greater customization and the inclusion of information to help patients find or contact back pain specialists. The latter 2 suggestions corroborate previous studies in which patients with LBP felt that the DHI could be more customized to their needs and provide an opportunity to contact an HCP [15,53]. Patients perceived *tala-med* as offering added value to other sources of information. In particular, patients felt that they could trust the portal; its contents; and its provision of in-depth, comprehensive, and appropriate information. Trustworthiness, in particular, reflects an important feature of web-based health information [34] that is deficient in many websites [15]. *Tala-med* may remedy the information dilemma faced by many patients when searching for health information on the internet by adding value to self-guided web-based searches.

The practical relevance of *tala-med* and its reported positive effects underline the fact that patients were able to understand and make good use of the portal's information, while also adding value to the PCP consultation. The ability to use *tala-med* independently and flexibly is an advantage typical of DHIs [28,42]. As a consequence of these features and the incorporated exercise videos, the portal motivates and empowers patients to manage their back pain whenever there is a need or free time to do so. By providing informational, motivational, and practical support, the portal contributed to a decrease in back pain and

reduced the need for patients to visit their PCP repeatedly because of LBP. Receiving the portal from their PCP added value to patients. It enriched preceding consultations by allowing patients to repeat and reflect on the contents of the consultation, thus increasing patient informedness while also facilitating and encouraging patient participation in future consultations through increased prior knowledge. These findings are in line with a recent systematic review, which suggests that by using DHIs, “users improved understanding of LBP and enhanced communication with their HCP during subsequent consultations” [28]. Overall, patient perceptions support the utility of the portal, especially when used in combination with PCP consultations.

Clinical Implications

PCPs could make good use of well-designed web portals for LBP as a supplement to their consultations [35]. Patients welcomed this additional web-based resource if they were familiar with digital technology and the internet. To support patient acceptance of a DHI and its positive effects, it should be integrated into consultations or patient treatment plans. Even if patients can easily use the DHI on their own, PCPs play a crucial role in deciding whether their state of health, in particular pain intensity, comorbidities, and further treatments are compatible with using a given DHI, and which one may offer the most helpful content to patients in terms of prescribing exercises or further information. PCPs and patients benefit from using DHIs such as *tala-med* as a supplement to their consultations [47], as it prepares patients for upcoming consultations and increases their participation. Because the aforementioned implications and their underlying findings are not unique to patients with LBP or *tala-med*, they may be generalized to other DHIs providing health content [21,22] used during and after PCP consultations.

Strength and Limitations

With 32 half-hour interviews, our qualitative study had a comprehensive information base, including both patients who used *tala-med* and those who did not. Nonetheless, our study relied on a self-selected sample of only 17% of all invited patients. As a result, negative and rare experiences with our portal may have been missed because the patients who experienced them did not participate in our study.

Compared with previous studies, a strength of our study was that patients were able to report their actual experiences with the consultation and the web portal, rather than their anticipated preferences [15]. However, for several patients, these experiences were not particularly vivid at the time of the interview, as the interviews took place at least 4 weeks after the consultation. These patient perceptions of their PCP use of the portal during the consultation may reflect their PCP’s actual

behavior unreliably, as subtleties of the consultation, such as brief uses or remarks about the portal during the consultation, may have been forgotten. Recall difficulties may also have led patients to perceive distinguishable aspects of usability less distinctively owing to halo effects [54] and may have led to more socially desirable responses. When investigating details of the portal, such as its usability, it would have been helpful to have the portal open in front of patients, as they were interviewed.

Researchers from different departments developed the portal (AM and SVR) and conducted interviews (CS and M Klimmek). Nevertheless, patients may have mistakenly assumed that they were speaking with someone who had also developed the portal, as the interviewers worked at the same medical center as the portal developers. Thus, patients may have felt inhibited in expressing any criticisms of the portal. However, this limitation only related to patient experiences of the portal. The fact that the interviewers were employed at a university medical center independent of and in a different federal state to that of the PCPs appeared to offer an advantage, in that patients did not have to be concerned when speaking freely about their perceptions of the PCP consultation.

Future studies that quantitatively investigate how patients perceive and evaluate *tala-med* or similar informational DHIs and their effects may mitigate the aforementioned limitations. This could further expand our knowledge of the acceptance, usability, and utility of web portals for LBP.

Conclusions

Most patients accepted our web portal well. Patient perception also affirmed its usability and utility. *Tala-med* may thus mitigate the information dilemma of patients and seems well suited as a supplement to PCP consultations. The facilitators of and barriers to use in our study are consistent with previous findings and indicate that PCPs should consider pain severity, comorbidities, other therapies, and IT equipment and skills of patients to support their acceptance of the portal. The setting itself, that is, the distribution of the portal by PCPs, seems appropriate as it supports the perceived acceptance and trustworthiness of the portal by patients, as well as bringing added value to current and future consultations. Patient perceptions thus highlight the appropriateness of the portal and the setting. However, they also indicate that many PCPs do not make full use of the portal, rarely integrating it into their consultations. Beyond feedback on *tala-med* and its implementation, the insights of our study on the acceptance, usability, and utility of *tala-med* offer valuable suggestions for the development of DHIs and their successful use as supplements to PCP consultation.

Acknowledgments

The authors thank Natascha Anka for her valuable comments on this manuscript and Rebekka Allen for proofreading the translated transcripts of patients. This study was part of the *Gut informierte Kommunikation zwischen Arzt und Patient* (GAP) trial (DRKS00014279, German Clinical Trial Register), which was funded by grant 01NVF17010 from *Innovationsausschuss of the Gemeinsamer Bundesausschuss*. The funding agency has no role in any aspect of the trial, such as planning or conducting the

trial or in the analysis, interpretation, reporting, or dissemination of study results. We acknowledge support from the Open Access Publication Fund of the University of Freiburg.

Data Availability

The data sets generated and analyzed during this study are not publicly available because patients were assured in the informed consent form that their personal data would not be made public (according to the General Data Protection Regulation, transcripts of anonymized interviews are considered personal data); however, they are available from the corresponding author upon reasonable request.

Authors' Contributions

CS contributed to the development of the interview guide, conducted all but 1 interview, led the data analysis, and wrote and revised the manuscript. NR contributed to the development of the interview guide and the data analysis. PvdK was responsible for the recruitment of the patients and critically revised the manuscript. ACSL prepared the data management (eg, the data protection concept), developed the methodological design and the interview guide, and critically revised the manuscript. M Klimmek conducted an interview, contributed to the data analysis, and critically revised the manuscript. M Körner critically revised the manuscript. KS contributed to the recruitment of patients. SVR was responsible for the conception and project management of the project, in which the study was embedded. He contributed to the development of the intervention and critically revised the manuscript. AM developed the intervention, described the intervention, and critically revised the manuscript. MS critically revised the manuscript and edited it as a native speaker. EFG supervised the planning and execution of the study and critically revised the manuscript. CS, NR, PvdK, ACSL, M Klimmek, M Körner, SVR, and EFG discussed the results before preparing the first draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[DOC File, 100 KB - [formative_v6i12e38748_app1.doc](#)]

Multimedia Appendix 2

Allocation of codes to themes.

[DOC File, 90 KB - [formative_v6i12e38748_app2.doc](#)]

References

1. Kirschning S, von Kardorff E. Welche Informationen suchen internetnutzende Angehörige krebskranker Frauen und Männer? Helfen die gefundenen Informationen bei der Bewältigungsarbeit? *Med Klin (Munich)* 2007 Feb 15;102(2):136-140. [doi: [10.1007/s00063-007-1014-7](#)] [Medline: [17323020](#)]
2. Tan SS, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. *J Med Internet Res* 2017 Jan 19;19(1):e9 [FREE Full text] [doi: [10.2196/jmir.5729](#)] [Medline: [28104579](#)]
3. Schweiger W, Beck K, editors. Gesundheitskommunikation im Internet: Erscheinungsformen, Potenziale, Grenzen. In: *Handbuch Online-Kommunikation*. Wiesbaden, Germany: Springer; 2018:1-27.
4. Daraz L, Morrow AS, Ponce OJ, Beuschel B, Farah MH, Katabi A, et al. Can patients trust online health information? A meta-narrative systematic review addressing the quality of health information on the Internet. *J Gen Intern Med* 2019 Sep;34(9):1884-1891 [FREE Full text] [doi: [10.1007/s11606-019-05109-0](#)] [Medline: [31228051](#)]
5. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA* 2002 May;287(20):2691-2700. [doi: [10.1001/jama.287.20.2691](#)] [Medline: [12020305](#)]
6. Voigt-Radloff S, Schöpf AC, Boeker M, Frank L, Farin E, Kaier K, et al. Well informed physician-patient communication in consultations on back pain - study protocol of the cluster randomized GAP trial. *BMC Fam Pract* 2019 Feb 25;20(1):33 [FREE Full text] [doi: [10.1186/s12875-019-0925-8](#)] [Medline: [30803433](#)]
7. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, et al. A systematic review of the global prevalence of low back pain. *Arthritis Rheum* 2012 Jun;64(6):2028-2037 [FREE Full text] [doi: [10.1002/art.34347](#)] [Medline: [22231424](#)]
8. Hurwitz EL, Randhawa K, Yu H, Côté P, Haldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. *Eur Spine J* 2018 Sep;27(Suppl 6):796-801. [doi: [10.1007/s00586-017-5432-9](#)] [Medline: [29480409](#)]

9. Schmidt CO, Raspe H, Pflingsten M, Hasenbring M, Basler HD, Eich W, et al. Back pain in the German adult population: prevalence, severity, and sociodemographic correlates in a multiregional survey. *Spine (Phila Pa 1976)* 2007 Aug 15;32(18):2005-2011. [doi: [10.1097/BRS.0b013e318133fad8](https://doi.org/10.1097/BRS.0b013e318133fad8)] [Medline: [17700449](https://pubmed.ncbi.nlm.nih.gov/17700449/)]
10. von der Lippe E, Krause L, Porst M, Wengler A, Leddin J, Müller A, BURDEN 2020 study group. Prevalence of back and neck pain in Germany. Results from the BURDEN 2020 Burden of Disease Study. *J Health Monit* 2021 Mar;6(Suppl 3):2-14 [FREE Full text] [doi: [10.25646/7855](https://doi.org/10.25646/7855)] [Medline: [35586774](https://pubmed.ncbi.nlm.nih.gov/35586774/)]
11. Meyer M, Wiegand S, Schenkel A. Krankheitsbedingte Fehlzeiten in der deutschen Wirtschaft im Jahr 2019. In: Badura B, Ducki A, Schröder H, Klose J, Meyer M, editors. *Fehlzeiten-Report 2020: Gerechtigkeit und Gesundheit*. Berlin, Germany: Springer; 2020:365-444.
12. Grobe TG, Steinmann S, Szecsenyi J. BARMER GEK Arztreport 2016. Schriftenreihe zur Gesundheitsanalyse. Siegburg, Germany: Asgard-Verlagsservice GmbH; 2016 Feb. URL: <https://www.barmer.de/resource/blob/1026850/41528a9e5704bb8d47e25e00707af4ba/barmer-gek-arztreport-2016-data.pdf> [accessed 2022-11-14]
13. Slade SC, Kent P, Bucknall T, Molloy E, Patel S, Buchbinder R. Barriers to primary care clinician adherence to clinical guidelines for the management of low back pain: protocol of a systematic review and meta-synthesis of qualitative studies. *BMJ Open* 2015 Apr 21;5(4):e007265 [FREE Full text] [doi: [10.1136/bmjopen-2014-007265](https://doi.org/10.1136/bmjopen-2014-007265)] [Medline: [25900462](https://pubmed.ncbi.nlm.nih.gov/25900462/)]
14. Volbracht E. Gesundheitspfad Rücken. Innovative Konzepte zur Verbesserung der Versorgung von Patienten mit Rückenschmerzen. Bertelsmann Stiftung. 2007 Jun. URL: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/IN_Gesundheitspfad_Ruecken_2007.pdf [accessed 2022-11-14]
15. Riis A, Hjelmager DM, Vinther LD, Rathleff MS, Hartvigsen J, Jensen MB. Preferences for web-based information material for low back pain: qualitative interview study on people consulting a general practitioner. *JMIR Rehabil Assist Technol* 2018 Apr 02;5(1):e7 [FREE Full text] [doi: [10.2196/rehab.8841](https://doi.org/10.2196/rehab.8841)] [Medline: [29610108](https://pubmed.ncbi.nlm.nih.gov/29610108/)]
16. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Nationale VersorgungsLeitlinie Nicht-spezifischer Kreuzschmerz – Langfassung, 2nd Auflage. Berlin, Germany: Bundesärztekammer; Jul 06, 2017.
17. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). PatientenLeitlinie zur Nationale VersorgungsLeitlinie (NVL) Nicht-spezifischer Kreuzschmerz, 2. Auflage. Berlin, Germany: Ärztliches Zentrum Für Qualität In Der Medizin; 2017.
18. Low back pain and sciatica in over 16s: assessment and management. NICE guideline. National Institute for Health and Care Excellence. 2016 Nov 30. URL: <https://www.nice.org.uk/guidance/ng59/resources/low-back-pain-and-sciatica-in-over-16s-assessment-and-management-pdf-1837521693637> [accessed 2022-11-14]
19. Van Wambeke P, Desomer A, Ailliet L, Berquin A, Demoulin C, Depreitere B, et al. Low back pain and radicular pain: assessment and management. KCE Reports 287. Federaal Kenniscentrum voor de Gezondheidszorg. 2017. URL: https://kce.fgov.be/sites/default/files/2021-11/KCE_287_Low_back_pain_Report.pdf [accessed 2018-04-06]
20. Qaseem A, Wilt TJ, McLean RM, Forcica MA, Clinical Guidelines Committee of the American College of Physicians, Denberg TD, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2017 Apr 04;166(7):514-530 [FREE Full text] [doi: [10.7326/M16-2367](https://doi.org/10.7326/M16-2367)] [Medline: [28192789](https://pubmed.ncbi.nlm.nih.gov/28192789/)]
21. Gitlow S. The online community as a healthcare resource. In: Nash DB, Manfredi MP, Bozarth B, Howell S, editors. *Connecting with the New Healthcare Consumer: Defining Your Strategy*. New York, NY, USA: McGraw-Hill; 2000:113-133.
22. Eysenbach G. The impact of the Internet on cancer outcomes. *CA Cancer J Clin* 2003;53(6):356-371 [FREE Full text] [doi: [10.3322/canjclin.53.6.356](https://doi.org/10.3322/canjclin.53.6.356)] [Medline: [15224975](https://pubmed.ncbi.nlm.nih.gov/15224975/)]
23. Dwamena F, Holmes-Rovner M, Gaulden CM, Jorgenson S, Sadigh G, Sikorskii A, et al. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev* 2012 Dec 12;12:CD003267. [doi: [10.1002/14651858.CD003267.pub2](https://doi.org/10.1002/14651858.CD003267.pub2)] [Medline: [23235595](https://pubmed.ncbi.nlm.nih.gov/23235595/)]
24. Légaré F, Stacey D, Turcotte S, Cossi MJ, Kryworuchko J, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2014 Sep 15(9):CD006732. [doi: [10.1002/14651858.CD006732.pub3](https://doi.org/10.1002/14651858.CD006732.pub3)] [Medline: [25222632](https://pubmed.ncbi.nlm.nih.gov/25222632/)]
25. Du S, Liu W, Cai S, Hu Y, Dong J. The efficacy of e-health in the self-management of chronic low back pain: a meta analysis. *Int J Nurs Stud* 2020 Jun;106:103507. [doi: [10.1016/j.ijnurstu.2019.103507](https://doi.org/10.1016/j.ijnurstu.2019.103507)] [Medline: [32320936](https://pubmed.ncbi.nlm.nih.gov/32320936/)]
26. Garg S, Garg D, Turin TC, Chowdhury MF. Web-based interventions for chronic back pain: a systematic review. *J Med Internet Res* 2016 Jul 26;18(7):e139 [FREE Full text] [doi: [10.2196/jmir.4932](https://doi.org/10.2196/jmir.4932)] [Medline: [27460413](https://pubmed.ncbi.nlm.nih.gov/27460413/)]
27. Nicholl BI, Sandal LF, Stochkendahl MJ, McCallum M, Suresh N, Vasseljen O, et al. Digital support interventions for the self-management of low back pain: a systematic review. *J Med Internet Res* 2017 May 21;19(5):e179 [FREE Full text] [doi: [10.2196/jmir.7290](https://doi.org/10.2196/jmir.7290)] [Medline: [28550009](https://pubmed.ncbi.nlm.nih.gov/28550009/)]
28. Svendsen MJ, Wood KW, Kyle J, Cooper K, Rasmussen CD, Sandal LF, et al. Barriers and facilitators to patient uptake and utilisation of digital interventions for the self-management of low back pain: a systematic review of qualitative studies. *BMJ Open* 2020 Dec 12;10(12):e038800 [FREE Full text] [doi: [10.1136/bmjopen-2020-038800](https://doi.org/10.1136/bmjopen-2020-038800)] [Medline: [33310794](https://pubmed.ncbi.nlm.nih.gov/33310794/)]

29. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. *BMC Med Inform Decis Mak* 2016 Sep 15;16(1):120 [FREE Full text] [doi: [10.1186/s12911-016-0359-3](https://doi.org/10.1186/s12911-016-0359-3)] [Medline: [27630020](https://pubmed.ncbi.nlm.nih.gov/27630020/)]
30. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. *J Med Internet Res* 2017 Dec 06;19(12):e402 [FREE Full text] [doi: [10.2196/jmir.8578](https://doi.org/10.2196/jmir.8578)] [Medline: [29212630](https://pubmed.ncbi.nlm.nih.gov/29212630/)]
31. 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](https://doi.org/10.1007/s12529-016-9556-9)] [Medline: [26957109](https://pubmed.ncbi.nlm.nih.gov/26957109/)]
32. 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](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
33. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. *J Med Internet Res* 2011 Aug 05;13(3):e52 [FREE Full text] [doi: [10.2196/jmir.1772](https://doi.org/10.2196/jmir.1772)] [Medline: [21821503](https://pubmed.ncbi.nlm.nih.gov/21821503/)]
34. Wollmann K, van der Keylen P, Tomandl J, Meerpohl JJ, Sofroniou M, Maun A, et al. The information needs of Internet users and their requirements for online health information-a scoping review of qualitative and quantitative studies. *Patient Educ Couns* 2021 Aug;104(8):1904-1932. [doi: [10.1016/j.pec.2021.01.020](https://doi.org/10.1016/j.pec.2021.01.020)] [Medline: [33563502](https://pubmed.ncbi.nlm.nih.gov/33563502/)]
35. Geraghty AW, Roberts LC, Stanford R, Hill JC, Yoganantham D, Little P, et al. Exploring patients' experiences of internet-based self-management support for low back pain in primary care. *Pain Med* 2020 Sep 01;21(9):1806-1817 [FREE Full text] [doi: [10.1093/pm/pnz312](https://doi.org/10.1093/pm/pnz312)] [Medline: [31841156](https://pubmed.ncbi.nlm.nih.gov/31841156/)]
36. Voigt-Radloff S. Methoden-Leitfaden. *Cochrane Deutschland*. 2022. URL: <https://gap-medinfo.de/methoden-leitfaden/> [accessed 2022-11-14]
37. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007 Dec;19(6):349-357. [doi: [10.1093/intqhc/mzm042](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]
38. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [FREE Full text] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
39. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. *Analyzing Qualitative Data*. London, UK: Routledge; 1994:173-194.
40. Parkinson S, Eatough V, Holmes J, Stapley E, Midgley N. Framework analysis: a worked example of a study exploring young people's experiences of depression. *Qual Res Psychol* 2016;13(2):109-129. [doi: [10.1080/14780887.2015.1119228](https://doi.org/10.1080/14780887.2015.1119228)]
41. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEX (personal experiences of health and illness). *Med Educ* 2006 May;40(5):405-414. [doi: [10.1111/j.1365-2929.2006.02467.x](https://doi.org/10.1111/j.1365-2929.2006.02467.x)] [Medline: [16635119](https://pubmed.ncbi.nlm.nih.gov/16635119/)]
42. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Mirkovic J, Børøsund E, et al. Patients' needs and requirements for eHealth pain management interventions: qualitative study. *J Med Internet Res* 2019 Apr 01;21(4):e13205 [FREE Full text] [doi: [10.2196/13205](https://doi.org/10.2196/13205)] [Medline: [30877780](https://pubmed.ncbi.nlm.nih.gov/30877780/)]
43. Gore M, Sadosky A, Stacey BR, Tai KS, Leslie D. The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings. *Spine (Phila Pa 1976)* 2012 May 15;37(11):E668-E677. [doi: [10.1097/BRS.0b013e318241e5de](https://doi.org/10.1097/BRS.0b013e318241e5de)] [Medline: [22146287](https://pubmed.ncbi.nlm.nih.gov/22146287/)]
44. Bair MJ, Matthias MS, Nyland KA, Huffman MA, Stubbs DL, Kroenke K, et al. Barriers and facilitators to chronic pain self-management: a qualitative study of primary care patients with comorbid musculoskeletal pain and depression. *Pain Med* 2009 Oct;10(7):1280-1290 [FREE Full text] [doi: [10.1111/j.1526-4637.2009.00707.x](https://doi.org/10.1111/j.1526-4637.2009.00707.x)] [Medline: [19818038](https://pubmed.ncbi.nlm.nih.gov/19818038/)]
45. Damush TM, Wu J, Bair MJ, Sutherland JM, Kroenke K. Self-management practices among primary care patients with musculoskeletal pain and depression. *J Behav Med* 2008 Aug;31(4):301-307. [doi: [10.1007/s10865-008-9156-5](https://doi.org/10.1007/s10865-008-9156-5)] [Medline: [18553130](https://pubmed.ncbi.nlm.nih.gov/18553130/)]
46. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: a literature review. *Int J Med Inform* 2011 Jan;80(1):1-12 [FREE Full text] [doi: [10.1016/j.ijmedinf.2010.10.017](https://doi.org/10.1016/j.ijmedinf.2010.10.017)] [Medline: [21112244](https://pubmed.ncbi.nlm.nih.gov/21112244/)]
47. van der Keylen P, Tomandl J, Wollmann K, Möhler R, Sofroniou M, Maun A, et al. The online health information needs of family physicians: systematic review of qualitative and quantitative studies. *J Med Internet Res* 2020 Dec 30;22(12):e18816 [FREE Full text] [doi: [10.2196/18816](https://doi.org/10.2196/18816)] [Medline: [33377874](https://pubmed.ncbi.nlm.nih.gov/33377874/)]
48. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. *NPJ Digit Med* 2022 Aug 18;5(1):119 [FREE Full text] [doi: [10.1038/s41746-022-00663-0](https://doi.org/10.1038/s41746-022-00663-0)] [Medline: [35982146](https://pubmed.ncbi.nlm.nih.gov/35982146/)]
49. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Assoc* 2018 Aug 01;320(8):1080-1088 [FREE Full text] [doi: [10.1093/jama/ocv052](https://doi.org/10.1093/jama/ocv052)] [Medline: [29788380](https://pubmed.ncbi.nlm.nih.gov/29788380/)]
50. Schumacher S, Kent N. 8 charts on internet use around the world as countries grapple with COVID-19. *Pew Research Center*. 2020 Apr 2. URL: <https://www.pewresearch.org/fact-tank/2020/04/02/8-charts-on-internet-use-around-the-world-as-countries-grapple-with-covid-19/> [accessed 2022-11-14]

51. Digital Economy and Society Index (DESI) 2020. Use of internet services. European Commission. 2020. URL: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=67075 [accessed 2022-11-14]
52. Skinner H, Biscope S, Poland B. Quality of internet access: barrier behind internet use statistics. *Soc Sci Med* 2003 Sep;57(5):875-880. [doi: [10.1016/s0277-9536\(02\)00455-0](https://doi.org/10.1016/s0277-9536(02)00455-0)] [Medline: [12850112](https://pubmed.ncbi.nlm.nih.gov/12850112/)]
53. de Jong T, Heinrich J, Blatter BM, Anema JR, van der Beek AJ. The feasibility of a web-based counselling program for occupational physicians and employees on sick leave due to back or neck pain. *BMC Med Inform Decis Mak* 2009 Nov 06;9:46 [FREE Full text] [doi: [10.1186/1472-6947-9-46](https://doi.org/10.1186/1472-6947-9-46)] [Medline: [19895689](https://pubmed.ncbi.nlm.nih.gov/19895689/)]
54. Minge M, Thüring M. Hedonic and pragmatic halo effects at early stages of User Experience. *Int J Hum Comput Stud* 2018 Jan;109:13-25. [doi: [10.1016/j.ijhcs.2017.07.007](https://doi.org/10.1016/j.ijhcs.2017.07.007)]

Abbreviations

DHI: digital health intervention

GAP: Gut informierte Kommunikation zwischen Arzt und Patient

HCP: health care professional

LBP: lower back pain

OSOP: one sheet of paper

PCP: primary care physician

RCT: randomized controlled trial

Edited by A Mavragani; submitted 14.04.22; peer-reviewed by L Westerbeek, S Leitch; comments to author 23.09.22; accepted 14.11.22; published 29.12.22.

Please cite as:

Schlett C, Röttele N, van der Keylen P, Schöpf-Lazzarino AC, Klimmek M, Körner M, Schnitzius K, Voigt-Radloff S, Maun A, Sofroniou M, Farin-Glattacker E

The Acceptance, Usability, and Utility of a Web Portal for Back Pain as Recommended by Primary Care Physicians: Qualitative Interview Study With Patients

JMIR Form Res 2022;6(12):e38748

URL: <https://formative.jmir.org/2022/12/e38748>

doi: [10.2196/38748](https://doi.org/10.2196/38748)

PMID: [36580365](https://pubmed.ncbi.nlm.nih.gov/36580365/)

©Christian Schlett, Nicole Röttele, Piet van der Keylen, Andrea Christina Schöpf-Lazzarino, Miriam Klimmek, Mirjam Körner, Kathrin Schnitzius, Sebastian Voigt-Radloff, Andy Maun, Mario Sofroniou, Erik Farin-Glattacker. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 29.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Theoretical Approach and Scale Construction of Patient Privacy Protection Behavior of Doctors in Public Medical Institutions in China: Pilot Development Study

Jie Xu^{1*}, MPhil; Lu Lu^{1*}, MA; Kaichen Xing^{2*}, BA; Huwei Shi¹, MM; Ruiyao Chen¹, MEng; Yujun Yao¹, MEng; Sichen Liu¹, BSc; Zhongzhou Xiao¹, MSc; Xinwei Peng¹, MM; Shuqing Luo¹, BM; Yun Zhong¹, BBA

¹Shanghai Artificial Intelligence Laboratory, West Bank International Artificial Intelligence Center, Shanghai, China

²Université de Montpellier, Montpellier, France

*these authors contributed equally

Corresponding Author:

Jie Xu, MPhil

Shanghai Artificial Intelligence Laboratory

West Bank International Artificial Intelligence Center

701 Yunjin Road

Shanghai, 200030

China

Phone: 86 021 23537800

Email: xujie@pjlab.org.cn

Abstract

Background: Considering the high incidence of medical privacy disclosure, it is of vital importance to study doctors' privacy protection behavior and its influencing factors.

Objective: We aim to develop a scale for doctors' protection of patients' privacy in Chinese public medical institutions, following construction of a theoretical model framework through grounded theory, and subsequently to validate the scale to measure this protection behavior.

Methods: Combined with the theoretical paradigm of protection motivation theory (PMT) and semistructured interview data, the grounded theory research method, followed by the Delphi expert and group discussion methods, a theoretical framework and initial scale for doctors in Chinese public medical institutions to protect patients' privacy was formed. The adjusted scale was collected online using a WeChat electronic survey measured using a 5-point Likert scale. Exploratory and confirmatory factor analysis (EFA and CFA) and tests to analyze reliability and validity were performed on the sample data. SPSS 19.0 and Amos 26.0 statistical analysis software were used for EFA and CFA of the sample data, respectively.

Results: According to the internal logic of PMT, we developed a novel theoretical framework of a "storyline," which was a process from being unaware of patients' privacy to having privacy protection behavior, that affected doctors' cognitive intermediary and changed the development of doctors' awareness, finally affecting actual privacy protection behavior in Chinese public medical institutions. Ultimately, we created a scale to measure 18 variables in the theoretical model, comprising 63 measurement items, with a total of 208 doctors participating in the scaling survey, who were predominantly educated to the master's degree level (n=151, 72.6%). The department distribution was relatively balanced. Prior to EFA, the Kaiser-Meyer-Olkin (KMO) value was 0.702, indicating that the study was suitable for factor analysis. The minimum value of Cronbach α for each study variable was .754, which met the internal consistency requirements of the scale. The standard factor loading value of each potential measurement item in CFA had scores greater than 0.5, which signified that all the items in the scale could effectively converge to the corresponding potential variables.

Conclusions: The theoretical framework and scale to assess doctors' patient protection behavior in public medical institutions in China fills a significant gap in the literature and can be used to further the current knowledge of physicians' thought processes and adoption decisions.

(JMIR Form Res 2022;6(12):e39947) doi:[10.2196/39947](https://doi.org/10.2196/39947)

KEYWORDS

scale; Chinese public medical institutions; doctors' protection behavior of patients' privacy

Introduction

Background

Privacy disclosure [1] refers to the release, transfer, access, or disclosure of information in any way to an individual or group outside of the entity holding the information. According to a Verizon data disclosure survey, the medical industry is the only industry with an internal threat higher than an external threat, considering that a significant number of medical data leaks are associated with internal medical staff [2]. The high incidence of medical data [3-5] and patient information leakages may cause patient identity violations [6] and financial losses [7], alongside potentially more severe social effects. Currently, in China, organizations at all levels regulate the privacy protection behavior of medical staff by publishing relevant policies and setting privacy protection requirements for medical staff [8-10]. Generally, hospitals also have privacy disclosure restrictions medical practitioners must adhere to for patient privacy protection. However, even with multiple patient privacy protection requirements, the leakage of patient privacy remains frequent. Unfortunately, as those with direct contact with medical information, the negligence or improper behavior of doctors has become 1 of the primary reasons for this [11,12]. Therefore, it is vital to study doctors' privacy protection behavior in the hospital setting. Public medical institutions, which are government interventions in the medical market, are of universal significance worldwide. They provide inexpensive welfare services for the public rather than high-priced private medical services. Current studies focus on the primary influencing factors affecting doctors' motivation to comply with data protection [13], the influencing factors of electronic medical record (EMR) privacy protection by doctors [7], and the reaction mechanism of doctors toward patients' privacy protection requirements [14]. Furthermore, personal factors, including age, gender, educational background, professional title, working years, position, and understanding of laws and regulations [15,16], alongside environmental factors, such as a strict patient privacy protection system, systematic training, sufficient materials for patient privacy protection, and demonstrations by managers [17], all had an impact on doctors' privacy protection behavior toward patients. Currently, research into the privacy protection behavior of doctors, including the factors that influence behavior and intention, remains insufficient. Similarly, a theoretical basis and measurement scale of patient privacy protection behavior of doctors in public medical institutions in China are yet to be formed.

Consequently, this study aims to use the method of grounded theory to construct a theoretical model framework of the patient privacy protection behavior of doctors in public medical institutions in China. The measurement items of each variable were defined in combination with the results of coding analysis. The measurement scale was created, and the data were analyzed by exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

Theory

Noar [18] proposed that the selection of a theoretical framework requires comparisons of multiple theoretical frameworks. The theory of planned behavior [13,19], protection motivation theory (PMT) [20], and the Health Belief Model [7] are often included in the theoretical basis of privacy protection. Comparison of these 3 theories shows that for the privacy protection behavior of doctors, fear of negativity is an important motivation for behavior change. Threat appraisal (TA), coping appraisal (CA), social norms (SN), and ethical personal characteristics are supposed to influence the development of doctors' awareness of the protection of patient privacy, thus leading to modifications in the actual behavior of doctors. The criteria for designating the theoretical framework were the related behavioral theories adopted by the current studies on privacy protection, followed by domestic and international studies on the influence factors of health care workers' privacy protection behavior, in which the main influence factors should be contained in the selected dimensions. Based on the fear of negativity, the privacy protection behavior of doctors in public medical institutions in China alters the motivation of protection behavior through obtaining relevant cognitive information and thus affects the individual's protection behavior. Consequently, the framework of PMT is in line with the theoretical paradigm of this study. PMT was developed using Health Belief Model by Rogers et al [21] in 1975. The factors affecting health-related behavior include perceived severity (PSE), perceived susceptibility (PSU), self-efficacy (SE), and response cost (RC), in addition to the object's perceived intrinsic rewards (IRE) and extrinsic rewards (ERE) [22,23]. However, patients' privacy protection by doctors in public medical institutions in China is the otherness behavior of non-right holders. The effect of social norms, attitudes, personal characteristics, and other factors on behavioral intentions are not considered in PMT. The mechanism of privacy protection by public medical institution doctors in China is yet to be fully explored. Exploratory research was required to construct a theoretical model of this study.

Methods

Theoretical Construction

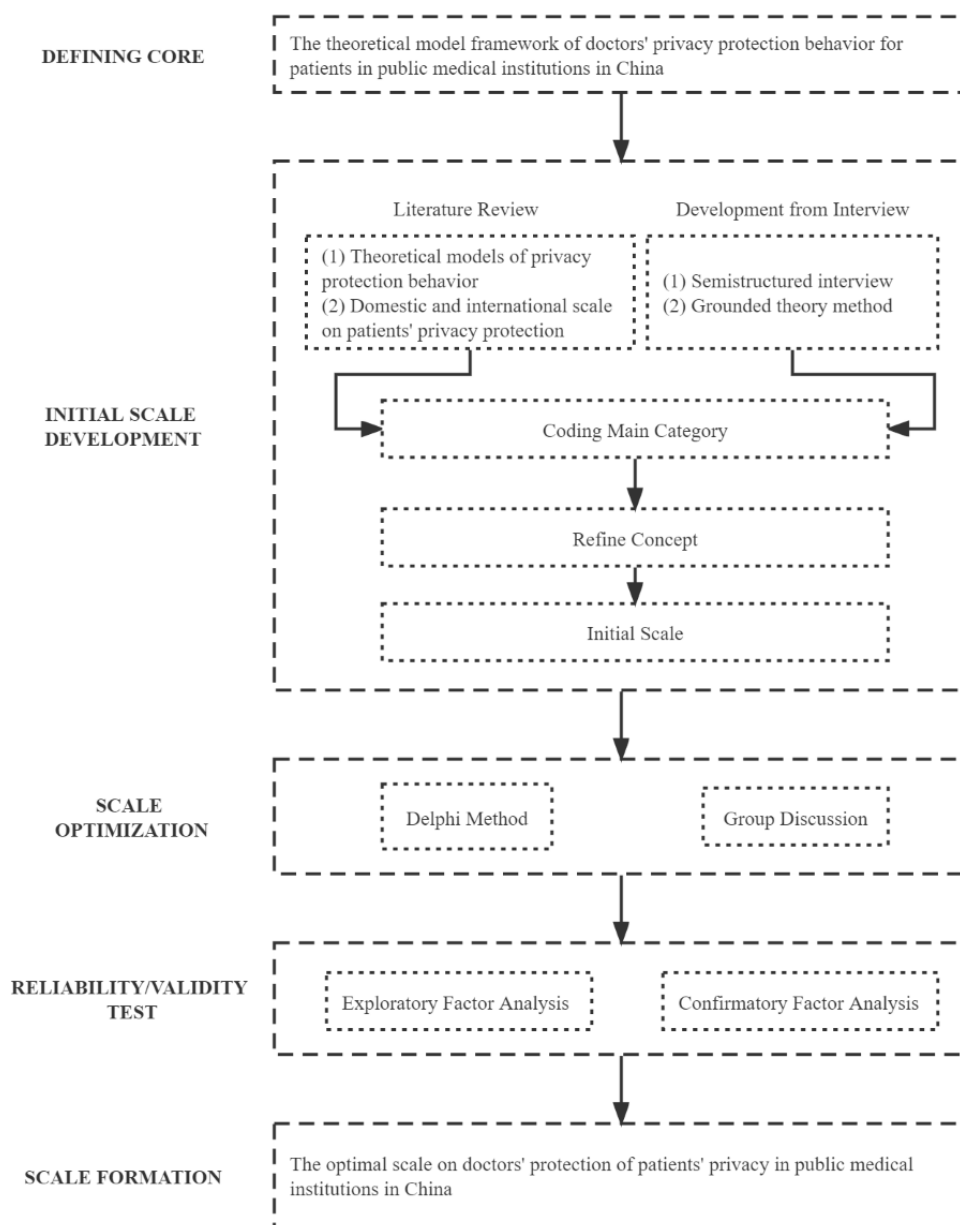
Considering the practical problem of how to promote the protection of patients' privacy in public medical institutions in China, we used the theoretical paradigm of PMT. Through theoretical sampling, we selected representative interviewees and subsequently interviewed them using a semistructured interview. The interview period spanned from January 25 to February 25, 2022, which was a total of 32 days. The original data of grounded theory coding analysis was generated from the interviews. Subsequently, program-based grounded theory [24] was used to analyze the coding based on the original interview data and refine the scope of the study, alongside discussing the logical relationship between them and building the theoretical model framework [25-27].

Scale Design and Optimization

Drawing on the methodology for scale development proposed by Churchill [28], a scale was designed based on the 3 principles of content, function, and overall uniformity. The process of scale construction is shown in Figure 1. Combined with the coding analysis results of the grounded theory method in the previous section, the key concept of interviews from the records of public medical institution doctors in China were extracted

and the measurement items of the variables in this study were processed and separated. Regarding the initial scale, we had to make some adjustments to the items of the scale following the Delphi method [29] and a group discussion, including merging and deleting items with similar meanings in the scale, merging and adjusting items with an inclusion relationship, and adjusting the wording to ensure the semantic readability was concise and that the sentences were easy to understand to prevent misunderstandings.

Figure 1. A Consolidated Standards of Reporting Trials (CONSORT) table of the process of the scale construction derived from Churchill's scale development.



Sample Preparation

The criteria for recruitment of doctors to be interviewed were as follows: (1) serving public medical institutions in China, (2) definition of “doctors” corresponding to the indicators of the *China Health Statistical Yearbook*, and (3) voluntary participation in this study. The survey adopted convenience and snowball sampling and was sent in the form of an electronic survey scale through WeChat. Data collection ran from April

1 to 25, 2022. In this study, the explicative items of threat, coping, support, and ethical appraisal were measured in the form of a 5-point Likert scale [30]: 1=completely disagree, 2=disagree, 3=uncertain, 4=agree, and 5=fully agree. In addition, IRE, ERE, and RC were inversely assigned. The explained variables, such as consciousness formation, body privacy, information privacy, and related privacy, were measured and assigned “yes” or “no” using binary variables, corresponding to 1 or 0, respectively.

Data Analysis

First, we conducted EFA, including a reliability evaluation using Cronbach α , which was the optimum method to evaluate the reliability of internal consistency. It is generally accepted that a Cronbach α score above .8 indicates excellent internal consistency, .6-.8 implies good consistency, and below .6 suggests poor internal consistency [31,32]. To ensure that the items involved in the measurement comprehensively and accurately measured the corresponding variables, we used EFA to assess the content validity of the scale. Although EFA is not an accurate method to test theoretical assumptions, it allowed us to draw conclusions regarding the construct validity of the proposed scale [33,34]. Prior to EFA, the Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity test was conducted on the scale to determine whether the scale is suitable for factor analysis. According to the judgment standard of KMO values, when the KMO value is above 0.6, the scale can be subject to factor analysis, while if it is above 0.8, the scale is suitable for factor analysis [35]. The variance maximization rotation for principal component analysis of the measurement scale and the Kaiser normalization maximum variance method for rotation were used in the EFA of this study. The rotation was confirmed to have converged following 7 iterations. In this study, SPSS 19.0 statistical analysis software was used for EFA of the sample data.

CFA was helpful to verify whether the subordinate relationship between the items in the scale and the extracted factors were correct or whether there were any wrong attributions to dimension problems [36,37]. Concerning the CFA result, if the standardized factor load of the item is greater than 0.5, it is accepted that the item can converge to its corresponding latent variable. The maximum likelihood was used in the model estimation. The χ^2 (*df*) value, the root-mean-square error of approximation (RMSEA), the standardized root-mean-square residual (SRMR), the Tucker-Lewis index (TLI), and the comparative fit index (CFI) were calculated to evaluate the model fit [38,39]. A model with good fit is achieved if χ^2 (*df*) is lower than 3 [40]. An RMSEA value below 0.05 indicates that the model is good [40]. SRMR values below 0.1 suggest that the model is acceptable [41]. The CFI and TLI values should be greater than 0.95 [42]. In addition, we evaluated the reliability of the scale by calculating the comprehensive reliability score [43,44] and analyzed both the convergence and discrimination effectiveness by comparing the average variance extracted (AVE) and the square correlation value [45,46]. If the variance of potential structure interpretation is greater than the variance

according to measurement error (if the AVE value is higher than 0.5), the convergence effectiveness is clear [46]. If the AVE value is greater than the square correlation value, the discrimination effectiveness is obvious [46]. In conclusion, aggregate validity and discriminant validity are powerful indicators of structural validity [45,47]. Throughout this research, Amos 26.0 statistical analysis software was used for CFA of the samples.

Ethical Considerations

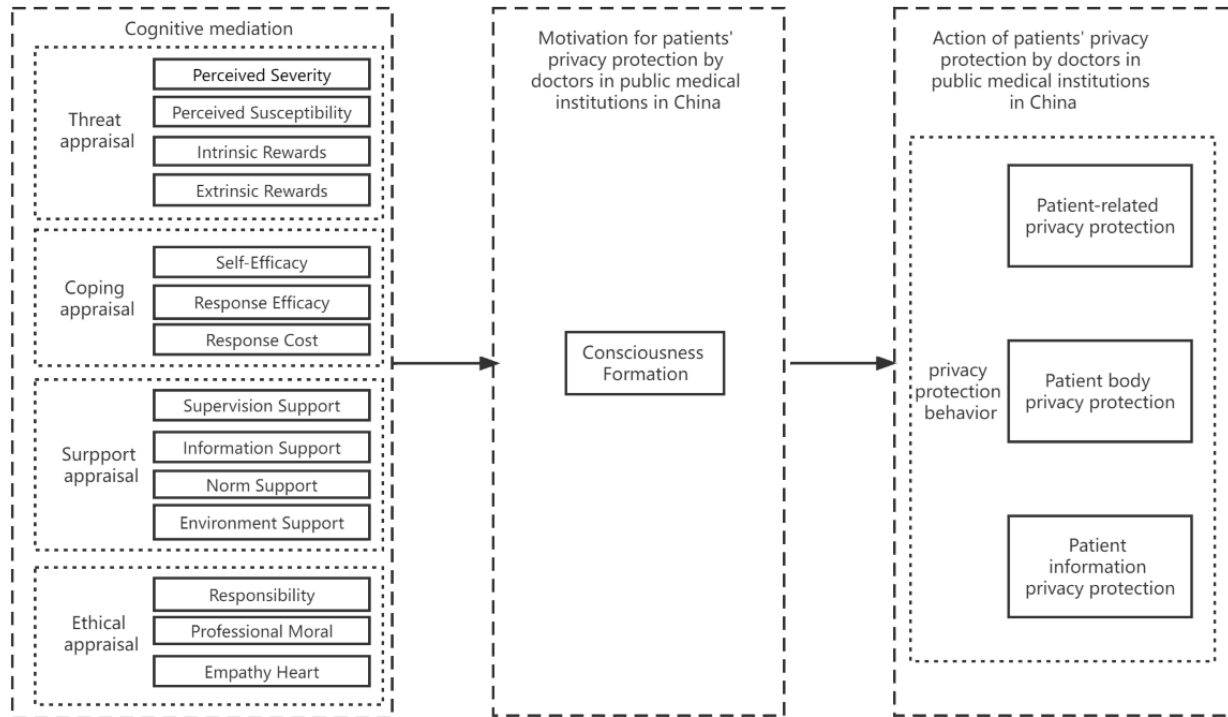
Survey recipients were informed that participation was anonymous and voluntary, that all responses would be kept confidential, and that the collected data would be used for academic research only. The survey was approved by the Institutional Review Board at Huadong Sanatorium (approval no. (2022)13 of the Ethic Committee), and all participants provided written informed consent.

Results

Theoretical Model

In total, 26 public medical institution doctors in China were selected, 10 (38.5%) of whom had personal, in-depth interviews, while the remaining 16 (61.5%) had 2 online focus group interviews according to their time arrangement. Finally, we obtained 12 interview records of over 50,000 words. Based on coding analysis and the theoretical saturation test, the results are detailed in [Multimedia Appendix 1](#). According to the internal logic of PMT, the cognitive intermediary of doctors in public medical institutions in China regarding patient privacy protection would affect the doctors' behavior by altering the development of their awareness of the protection of patient privacy. The protection of patients' privacy under the awareness of public medical institutions would lead to modifications in the actual behavior of doctors. Consequently, in this study, we considered doctors in public medical institutions in China to move through a process, from being unaware of patients' privacy to having privacy protection behavior. We regarded that this process affected doctors' cognitive intermediary and changed the development of their awareness of patients' privacy protection in public medical institutions in China in order to affect actual privacy protection behavior. According to the logical relationship of this "storyline," we developed a novel theoretical framework, which is the theoretical model framework of the mechanism of doctors' protection of patients' privacy in public medical institutions in China, as illustrated in [Figure 2](#).

Figure 2. The theoretical model framework of doctors' privacy protection behavior for patients in public medical institutions in China.



Optimization of the Initial Scale

The results of the initial scale are displayed in [Multimedia Appendix 2](#). A total of 15 items were corrected. Following the aforementioned correction and adjustment of measurement items, an initial scale to measure 18 direct measurement

variables in the theoretical model of doctors' behavior mechanism of protecting patients' privacy in public medical institutions in China was formed. This also included 63 measurement items, which were coded. [Tables 1-5](#) illustrates the specific codes and corresponding measurement items.

Table 1. Item setting of the initial scale for TA^a.

Variable and code	Item
PSE^b	
PSE1	I think it is very serious and dangerous that the disclosure of patient privacy information will incur punishment by laws and regulations.
PSE2	I think it is very serious and dangerous that the disclosure of patient privacy information will protect patients' rights and deepen the contradiction between doctors and patients.
PSE3	I think it is very serious and dangerous that the disclosure of patient privacy information will incur punishment according to the hospital standard system.
PSU^c	
PSU1	I think that laws and regulations pay increasing attention to the protection of patients' privacy and have the tendency to make mandatory punishment measures for privacy disclosure.
PSU2	I think that patients' awareness of protecting rights is progressively becoming stronger, and the protection of personal privacy is paid increasing attention. The leakage of patient privacy will further deepen the contradiction between doctors and patients.
PSU3	I think hospitals pay increasing attention to the privacy protection of patients, and the standards and systems will be more and more rigorous, and privacy disclosure incidents will be punishable.
IRE^d	
IRE1	I think that the disclosure of patients' privacy can be exchanged for certain financial returns.
IRE2	I think it is inevitable that patient privacy will be leaked in the process of scientific research output.
IRE3	I think meeting celebrities or attending new events at work will 'get out' on personal social platforms.
ERE^e	
ERE1	I've heard about the exchange of property through patient privacy information.
ERE2	I hear that the easier it is for individuals or institutions to get patient data, the greater the output of scientific research.
ERE3	I have heard that doctors have exposed some medical information or personal information about celebrities and related people on social platforms.

^aTA: threat appraisal.

^bPSE: perceived severity.

^cPSU: perceived susceptibility.

^dIRE: intrinsic rewards.

^eERE: extrinsic rewards.

Table 2. Item setting of the initial scale for CA^a.

Variable and code	Item
SE^b	
SE1	I think it's easy for me to protect the privacy of patients.
SE2	I think it's convenient for me to protect the privacy of patients.
SE3	I have the ability to protect the privacy of patients from being disclosed.
RE^c	
RE1	I think the doctors' protection measures to ensure the privacy of patients can effectively prevent the leakage of patients' privacy.
RE2	I think the privacy protection measures of doctors can keep patients' privacy in a safe environment.
RE3	I think the privacy protection measures of doctors for patients can better protect the privacy of patients.
RC^d	
RC1	I think that paying attention to the protection of patients' privacy will affect the output of my overall scientific research results.
RC2	I think that paying attention to the privacy protection of patients will affect the development and efficiency of my clinical work and teaching.
RC3	I think that paying attention to the privacy protection of patients will increase my work pressure.

^aCA: coping appraisal.

^bSE: self-efficacy.

^cRE: response efficacy.

^dRC: response cost.

Table 3. Item setting of the initial scale for SA^a.

Variable and code	Item
SS^b	
SS1	I think the protection of patients' privacy needs the full-time supervision and management of a hospital department.
SS2	I think it is necessary for the hospital to regularly organize training and assessment according to the laws and regulations related to patient privacy protection.
SS3	I think it is necessary for the hospital to regularly organize training and assessment for the hospital system related to patient privacy protection and other contents related to patient privacy protection.
IS^c	
IS1	I think it is necessary to use information technology, artificial intelligence, and other technologies to improve the information construction level of hospitals for patient privacy protection.
IS2	I think it is necessary to carry out reasonable authority management on the information system to protect the patient's private information.
IS3	I think it is necessary to impose reasonable data transmission restrictions on the information system to protect patients' private information.
NS^d	
NS1	I think it is necessary to build a patient privacy protection system and carry it out effectively to ensure the rationalization process of patient privacy protection in doctors' work.
NS2	I think it is necessary to combine the patient privacy protection system with the doctor's daily work, so that the doctor's behavior of protecting the patient's privacy becomes a daily aspect of the work.
NS3	I think it is necessary to formulate a reasonable scientific research application system and conduct scientific research efficiently on the basis of legal and compliant patient privacy protection.
ES^e	
ES1	I think improving the medical environment (such as independent consulting room, sound insulation treatment of consulting room) can better protect the privacy of patients.
ES2	I think it is necessary to maintain the order of medical treatment (for example, prevent irrelevant patients from gathering in the consulting room), which can better protect the privacy of patients.
ES3	I think facilities that provide patient privacy protection (such as curtains and privacy processing of bedside card information) can better protect patient privacy.

^aSA: support appraisal.

^bSS: supervision support.

^cIS: information support.

^dNS: norm support.

^eES: environment support.

Table 4. Item setting of the initial scale for EA^a.

Variable and code	Item
RS^b	
RS1	I think it is the duty of doctors to protect patients' privacy.
RS2	I believe that doctors should protect patients' privacy.
RS3	I think my sense of responsibility urges me to protect patients' privacy in my daily work.
PM^c	
PM1	I think doctors' protection of patients' privacy is a requirement of their own professional ethics.
PM2	I think my sense of professional ethics urges me to protect patients' privacy in my daily work.
PM3	From education to work, the protection of patients' private information is a professional ethic repeatedly emphasized by doctors.
EH^d	
EH1	I think doctors should consider the harm of privacy information disclosure from the perspective of patients, to become more aware of protecting the privacy of patients.
EH2	I have had a personal information disclosure experience as a patient, so I am more aware of protecting the privacy of patients.
EH3	I think that I can 'push myself to others' to protect my patients' privacy in my daily work.

^aEA: ethical appraisal.

^bRS: responsibility.

^cPM: professional moral.

^dEH: empathy heart.

Table 5. Item setting of the initial scale for CF^a, BP^b, IP^c, and RP^d.

Variable and code	Item
CF	
CF1	I think I have developed a sense of privacy protection in my clinical work.
CF2	I think I have formed a sense of privacy protection in my teaching.
CF3	I think I have formed a sense of privacy protection in my own research work.
BP	
BP1	Protect the patient's privacy during surgery or examination, such as curtain pulling and preventing a third party from breaking in.
BP2	Effectively block the privacy of patients during live operations.
BP3	Medical observation or teaching requires the consent of the patient.
BP4	No illegal touch or peek at the patient's privacy.
IP	
IP1	In the situations of outpatient, ward check, case discussion, medical education and observation, the patient shall obtain the consent of the patient himself and take confidentiality measures. The privacy information of the patient shall not be publicized or publicly discussed orally, including the personal information and disease information with identifiable characteristics, such as avoiding calling the full name of the patient loudly, avoiding 'listening' or 'breaking in' by people other than patients without the consent of the patient.
IP2	In the face of the condition inquiry, strictly confirm and ask the status of the patient's condition personnel, confirm as me or with my consent.
IP3	For patients with special conditions (for example infectious diseases involving privacy), it is necessary to talk to the patients individually.
IP4	Deliberately disclose and disseminate the privacy of patients without using their duties, such as taking the bedside card test sheets of celebrities to the internet.
IP5	Protect medical documents such as inspection and medical records without random placing, damage, loss, and prevent theft and being wrongly picked up.
IP6	Under the unnecessary diagnosis and treatment process, without the consent of the patient, the medical documents shall not be checked, copied, or borrowed during the hospitalization of the patient.
IP7	Use personal information system account number as required, and login to view patient information without borrowing non-authorized people.
IP8	Not disclose the privacy information of the patient for any benefit reasons to obtain business, advertise or defraud.
IP9	When leaving the office seat, protect the pages with patient privacy information and lock the screen of the computer.
IP10	Scientific research, including the mining of electronic medical record information, whether it is the steps of data acquisition, viewing, processing or analysis, is strictly done to de privacy.
IP11	In the form of talks or written (case discussion, writing medical treatises, scientific research papers), for example, when communicating and learning on medical social network platform to share typical cases, do well in privacy treatment.
RP	
RP1	Do not disclose information about family members and other personal relationships of any patient.
RP2	Do not disclose family members and other personal relationship information of any patient on social platforms.
RP3	Do not verbally promote or publicly discuss family members and other personal relationship information of any patient.

^aCF: consciousness formation.

^bBP: body privacy.

^cIP: information privacy.

^dRP: related privacy.

Descriptive Analysis Results

The survey we issued was scanned a total of 278 times, and 208 valid questionnaires were identified following recovery; thus, the effective recovery rate was 74.8%. The gender ratio of men to women was relatively balanced, with 46.6% (n=97) of the participants being men. The majority of the respondents were

26-35 years old, accounting for 63.9% (n=133) of the total sample. The respondents were predominantly educated to the master's degree level, accounting for 72.6% (n=151), which was consistent with the general education background of doctors. The department distribution was relatively balanced. Regarding urban distribution, Shanghai accounted for the highest proportion (n=131, 63.0%) due to convenience sampling. The

detailed characteristics of the pretest samples are displayed in [Table 6](#).

Table 6. Statistical results of sample characteristics.

Sample characteristics and measurement items	Sample size (N=208), n (%)
Gender	
Male	97 (46.6)
Female	111 (53.4)
Age (years)	
18-25	9 (4.3)
26-35	133 (63.9)
36-50	61 (29.3)
>50	5 (2.4)
Educational background	
Undergraduate	41 (19.7)
Master	151 (72.6)
Doctor	16 (7.7)
Department	
Otorhinolaryngology	5 (2.4)
Infectious diseases	5 (2.4)
Pulmonology	16 (7.7)
Severe medicine	9 (4.3)
Clinical laboratory	5 (2.4)
Endocrinology	5 (2.4)
Anesthesiology	24 (11.5)
Pediatrics	23 (11.1)
Internal medicine	23 (11.1)
Burns and plastic surgery	5 (2.4)
Internal medicine: cardiovascular	16 (7.7)
Surgery	9 (4.3)
Ophthalmology	9 (4.3)
Medical service	9 (4.3)
Imaging	18 (8.7)
Oncology	9 (4.3)
Dental	18 (8.7)
City	
Beijing	20 (9.6)
Changzhou, Jiangsu Province	5 (2.4)
Huai'an, Jiangsu Province	5 (2.4)
Nanjing, Jiangsu Province	5 (2.4)
Nantong, Jiangsu Province	5 (2.4)
Wuxi, Jiangsu Province	32 (15.4)
Zhenjiang, Jiangsu Province	5 (2.4)
Shanghai	131 (63.0)

Exploratory and Verifiable Analysis

The Cronbach α of the whole scale was determined to be .768 by calculating the consistency coefficient of the scale, which was between 0.6 and 0.8, indicating that the scale possessed good internal consistency. According to the results illustrated in Table 7, the KMO value was greater than 0.6, indicating that the study is suitable for factor analysis, theoretically.

Following the factor analysis operation, 18 common factors were screened. The cumulative interpretation total variance of factor analysis was 71.49%, implying that this research had good explanatory ability. The contribution of single-factor variance was less than 40%, demonstrating that this scale could exclude any possible homologous deviation [48]. In general, this scale conformed to the preassumed theoretical structure and possessed good content validity. From the factor loading of each dimension, it was clear that the measurement items were independent of common factors, and the load was above 0.7, far greater than the standard of 0.4. Conversely, the absolute value of the load of the 18-factor measurement items on other factors was below 0.4. This indicates that the items of variables in this study could both effectively converge on their own common factors and effectively be different from other common factors. The minimum value of the Cronbach α for each study variable was .754; thus, the coefficients for each study variable

were between 0.6 and 0.8, which met the internal consistency requirements of the scale. Consequently, the scale in this study could be considered to have good convergence validity and internal consistency [49]. The results are stated in Multimedia Appendix 3.

According to the fitting results of the model data in Table 8, most of the fitting indexes met the requirements of a critical value, indicating that the confirmatory factor model was well fit. According to the results of CFA in Table 9, the standard factor loading value of each potential measurement item in this study had scores greater than 0.5. The critical ratio (CR) values were greater than 7, and all had statistical significance levels within the range of $P < .001$. All the composite reliability (CPR) values were greater than 0.7, and the value of AVE scores were in excess of 0.5. These indexes were in accordance with the standards, which signified that all the items in the scale can effectively converge to the corresponding potential variables. The data displayed in Table 10 show that the value of the square root of the AVE of each variable was greater than 0.7. The score was significantly greater than the correlation value of the row and column in which it was located, namely the value below the diagonal. This demonstrated that there were significant differences among variables and confirmed that the scale in this study has good discriminative validity.

Table 7. Results of KMO^a and Bartlett sphericity test.

Variable	Value
KMO metric for sufficient sampling	0.702
Bartlett sphericity test	
Approximate χ^2 (df)	6130.640 (1953)
Significance	0.000

^aKMO: Kaiser-Meyer-Olkin.

Table 8. Judgment of fitting indexes of the CFA^a model.

Model fitting index	Critical value	Model fitting index value	Model fit judgment
χ^2/df	<3	1.138	Yes
RMSEA ^b	<0.08	0.026	Yes
SRMR ^c	<0.1	0.0474	Yes
CFI ^d	>0.95	0.951	Yes
TLI ^e	>0.95	0.945	Yes

^aCFA: confirmatory factor analysis.

^bRMSEA: root-mean-square error of approximation.

^cSRMR: standardized root-mean-square residual.

^dCFI: comparative fit index.

^eTLI: Tucker-Lewis index.

Table 9. Results of validation factor analysis.

Variable and item	Standard factor load	SE	CR ^a	P value
PSE^b (AVE^c=0.301, CPR^d=0.819)				
PSE1	0.760	N/A ^e	N/A	N/A
PSE2	0.794	0.095	9.997	<.001
PSE3	0.771	0.095	9.869	<.001
PSU^f (AVE=0.580, CPR=0.805)				
PSU1	0.720	N/A	N/A	N/A
PSU2	0.752	0.112	9.199	<.001
PSU3	0.810	0.119	9.454	<.001
IRE^g (AVE=0.567, CPR=0.797)				
IRE1	0.778	N/A	N/A	N/A
IRE2	0.708	0.101	8.951	<.001
IRE3	0.770	0.130	9.341	<.001
ERE^h (AVE=0.541, CPR=0.780)				
ERE1	0.706	N/A	N/A	N/A
ERE2	0.747	0.149	8.513	<.001
ERE3	0.753	0.133	8.535	<.001
SEⁱ (AVE=0.603, CPR=0.819)				
SE1	0.697	N/A	N/A	N/A
SE2	0.841	0.122	9.699	<.001
SE3	0.784	0.111	9.515	<.001
RE^j (AVE=0.569, CPR=0.798)				
RE1	0.799	N/A	N/A	N/A
RE2	0.792	0.110	9.441	<.001
RE3	0.665	0.100	8.641	<.001
RC^k (AVE=0.567, CPR=0.796)				
RC1	0.702	N/A	N/A	N/A
RC2	0.731	0.137	8.753	<.001
RC3	0.820	0.145	9.048	<.001
SS^l (AVE=0.532, CPR=0.773)				
SS1	0.719	N/A	N/A	N/A
SS2	0.769	0.132	8.670	<.001
SS3	0.699	0.125	8.290	<.001
IS^m (AVE=0.568, CPR=0.797)				
IS1	0.719	N/A	N/A	N/A
IS2	0.783	0.118	9.266	<.001
IS3	0.757	0.108	9.124	<.001
NSⁿ (AVE=0.586, CPR=0.809)				
NS1	0.770	N/A	N/A	N/A
NS2	0.768	0.104	9.694	<.001
NS3	0.758	0.100	9.630	<.001

Variable and item	Standard factor load	SE	CR ^a	P value
ES^o (AVE=0.567, CPR=0.797)				
ES1	0.768	N/A	N/A	N/A
ES2	0.737	0.098	9.329	<.001
ES3	0.753	0.095	9.448	<.001
RS^p (AVE=0.547, CPR=0.784)				
RS1	0.719	N/A	N/A	N/A
RS2	0.770	0.108	8.685	<.001
RS3	0.729	0.120	8.524	<.001
PM^q (AVE=0.545, CPR=0.782)				
PM1	0.695	N/A	N/A	N/A
PM2	0.776	0.131	8.427	<.001
PM3	0.742	0.122	8.350	<.001
EH^r (AVE=0.539, CPR=0.778)				
EH1	0.722	N/A	N/A	N/A
EH2	0.748	0.094	8.471	<.001
EH3	0.732	0.097	8.407	<.001
CF^s (AVE=0.513, CPR=0.757)				
CF1	0.801	N/A	N/A	N/A
CF2	0.610	0.10	7.074	<.001
CF3	0.724	0.122	7.480	<.001
BP^t (AVE=0.552, CPR=0.831)				
BP1	0.701	N/A	N/A	N/A
BP2	0.736	0.111	9.072	<.001
BP3	0.792	0.119	9.539	<.001
BP4	0.740	0.116	9.110	<.001
IP^u (AVE=0.532, CPR=0.926)				
IP1	0.743	N/A	N/A	N/A
IP2	0.718	0.094	10.391	<.001
IP3	0.717	0.094	10.381	<.001
IP4	0.723	0.094	10.468	<.001
IP5	0.744	0.090	10.796	<.001
IP6	0.806	0.095	11.802	<.001
IP7	0.733	0.094	10.632	<.001
IP8	0.67	0.091	9.703	<.001
IP9	0.765	0.091	11.136	<.001
IP10	0.704	0.093	10.171	<.001
IP11	0.689	0.089	9.942	<.001
RP^v (AVE=0.574, CPR=0.802)				
RP1	0.736	N/A	N/A	N/A
RP2	0.781	0.116	9.043	<.001

Variable and item	Standard factor load	SE	CR ^a	<i>P</i> value
RP3	0.756	0.110	8.984	<.001

^aCR: critical ratio.

^bPSE: perceived severity.

^cAVE: average variance extracted.

^dCPR: composite reliability.

^eN/A: not applicable.

^fPSU: perceived susceptibility.

^gIRE: intrinsic rewards.

^hERE: extrinsic rewards.

ⁱSE: self-efficacy.

^jRE: response efficacy.

^kRC: response cost.

^lSS: supervision support.

^mIS: information support.

ⁿNS: norm support.

^oES: environment support.

^pRS: responsibility.

^qPM: professional moral.

^rEH: empathy heart.

^sCF: consciousness formation.

^tBP: body privacy.

^uIP: information privacy.

^vRP: related privacy.

Table 10. Results of validity test.

Variables	RP ^a	IP ^b	BP ^c	CF ^d	EH ^e	PM ^f	RS ^g	ES ^h	NS ⁱ	IS ^j	SS ^k	RC ^l	RE ^m	SE ⁿ	ERE ^o	IRE ^p	PSU ^q	PSE ^r
RP	<i>0.758^s</i>	N/A ^t	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
IP	0.046	<i>0.730^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
BP	-0.145	0.015	<i>0.743^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CF	-0.110	-0.069	0.045	<i>0.716^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EH	0.010	-0.006	0.105	-0.001	<i>0.734^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PM	-0.021	0.079	-0.029	-0.041	0.336	<i>0.738^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RS	0.058	-0.038	0.034	-0.034	0.436	0.357	<i>0.740^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ES	0.066	0.044	-0.026	-0.027	-0.062	-0.082	-0.056	<i>0.753^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
NS	-0.036	-0.010	-0.047	-0.053	-0.107	0.028	-0.116	0.425	<i>0.765^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
IS	-0.104	-0.012	0.047	-0.154	0.055	0.002	-0.013	0.489	0.408	<i>0.753^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SS	0.008	-0.019	-0.009	-0.014	-0.075	0.038	-0.160	0.482	0.396	0.451	<i>0.730^s</i>	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RC	-0.073	-0.010	0.062	0.025	-0.021	0.035	0.060	-0.054	0.060	-0.013	0.008	<i>0.753^s</i>	N/A	N/A	N/A	N/A	N/A	N/A
RE	-0.056	0.053	0.035	0.011	-0.041	0.027	0.009	-0.017	0.108	0.058	0.104	0.355	<i>0.755^s</i>	N/A	N/A	N/A	N/A	N/A
SE	-0.046	0.005	-0.012	-0.043	0.013	0.078	-0.014	-0.051	0.001	0.038	0.112	0.458	0.350	<i>0.776^s</i>	N/A	N/A	N/A	N/A
ERE	0.043	0.012	0.106	-0.068	0.039	0.053	0.076	0.014	-0.020	0.028	-0.001	-0.022	0.001	-0.080	<i>0.736^s</i>	N/A	N/A	N/A
IRE	0.049	0.068	0.103	0.071	-0.017	0.083	-0.021	0.060	0.035	0.007	0.076	-0.029	0.010	-0.124	0.392	<i>0.753^s</i>	N/A	N/A
PSU	0.118	0.017	0.066	-0.002	-0.081	0.056	0.002	-0.003	-0.028	0.023	-0.065	0.030	0.079	-0.041	0.438	0.369	<i>0.762^s</i>	N/A
PSE	0.108	0.050	0.060	-0.032	-0.071	0.047	0.008	-0.008	0.001	-0.014	0.028	0.030	0.020	0.036	0.395	0.335	0.372	<i>0.775^s</i>

^aRP: related privacy.

^bIP: information privacy.

^cBP: body privacy.

^dCF: consciousness formation.

^eEH: empathy heart.

^fPM: professional moral.

^gRS: responsibility.

^hES: environment support.

ⁱNS: norm support.

^jIS: information support.

^kSS: supervision support.

^lRC: response cost.

^mRE: response efficacy.

ⁿSE: self-efficacy.

^oERE: extrinsic rewards.

^pIRE: intrinsic rewards.

^qPSU: perceived susceptibility.

^rPSE: perceived severity.

^sItalicized values are significant.

^tN/A: not applicable.

Scale Determination

We ultimately designed a scale with high reliability and validity and a correct subordination structure between factors. The scale

determined 63 measurement items around 18 direct measurement variables in the theoretical model of this study to create a formal scale for this study. The specific contents are illustrated in [Table 11](#).

Table 11. Measurement items of the formal scale of this study.

Variable and code	Measurement item
PSE^a	
PSE1	I think it is very serious and dangerous that the disclosure of patient privacy information will incur punishment by laws and regulations.
PSE2	I think it is very serious and dangerous that the disclosure of patient privacy information will protect patients' rights and deepen the contradiction between doctors and patients.
PSE3	I think it is very serious and dangerous that the disclosure of patient privacy information will incur punishment according to the hospital standard system.
PSU^b	
PSU1	I think that laws and regulations pay increasing attention to the protection of patients' privacy and have the tendency to make mandatory punishment measures for privacy disclosure.
PSU2	I think that patients' awareness of protecting rights is progressively becoming stronger, and the protection of personal privacy is paid increasing attention. The leakage of patient privacy will further deepen the contradiction between doctors and patients.
PSU3	I think hospitals pay increasing attention to the privacy protection of patients, and the standards and systems will be more and more rigorous, and privacy disclosure incidents will be punishable.
IRE^c	
IRE1	I think that the disclosure of patients' privacy can be exchanged for certain financial returns.
IRE2	I think it is inevitable that patient privacy will be leaked in the process of scientific research output.
IRE3	I think meeting celebrities or attending new events at work will 'get out' on personal social platforms.
ERE^d	
ERE1	I've heard about the exchange of property through patient privacy information.
ERE2	I hear that the easier it is for individuals or institutions to get patient data, the greater the output of scientific research.
ERE3	I have heard that doctors have exposed some medical information or personal information about celebrities and related people on social platforms.
SE^e	
SE1	I think it's easy for me to protect the privacy of patients.
SE2	I think it's convenient for me to protect the privacy of patients.
SE3	I have the ability to protect the privacy of patients from being disclosed.
RE^f	
RE1	I think the doctors' protection measures to ensure the privacy of patients can effectively prevent the leakage of patients' privacy.
RE2	I think the privacy protection measures of doctors can keep patients' privacy in a safe environment.
RE3	I think the privacy protection measures of doctors for patients can better protect the privacy of patients.
RC^g	
RC1	I think that paying attention to the protection of patients' privacy will affect the output of my overall scientific research results.
RC2	I think that paying attention to the privacy protection of patients will affect the development and efficiency of my clinical work and teaching.
RC3	I think that paying attention to the privacy protection of patients will increase my work pressure.
SS^h	
SS1	I think the protection of patients' privacy needs the full-time supervision and management of a hospital department.
SS2	I think it is necessary for the hospital to regularly organize training and assessment according to the laws and regulations related to patient privacy protection.
SS3	I think it is necessary for the hospital to regularly organize training and assessment for the hospital system related to patient privacy protection and other contents related to patient privacy protection.
ISⁱ	

Variable and code	Measurement item
IS1	I think it is necessary to use information technology, artificial intelligence, and other technologies to improve the information construction level of hospitals for patient privacy protection.
IS2	I think it is necessary to carry out reasonable authority management on the information system to protect the patient's private information.
IS3	I think it is necessary to impose reasonable data transmission restrictions on the information system to protect patients' private information.
NS^j	
NS1	I think it is necessary to build a patient privacy protection system and carry it out effectively to ensure the rationalization process of patient privacy protection in doctors' work.
NS2	I think it is necessary to combine the patient privacy protection system with the doctor's daily work, so that the doctor's behavior of protecting the patient's privacy becomes a daily aspect of the work.
NS3	I think it is necessary to formulate a reasonable scientific research application system and conduct scientific research efficiently on the basis of legal and compliant patient privacy protection.
ES^k	
ES1	I think improving the medical environment (such as independent consulting room, sound insulation treatment of consulting room) can better protect the privacy of patients.
ES2	I think it is necessary to maintain the order of medical treatment (for example, prevent irrelevant patients from gathering in the consulting room), which can better protect the privacy of patients.
ES3	I think facilities that provide patient privacy protection (such as curtains and privacy processing of bedside card information) can better protect patient privacy.
RS^l	
RS1	I think it is the duty of doctors to protect patients' privacy.
RS2	I believe that doctors should protect patients' privacy.
RS3	I think my sense of responsibility urges me to protect patients' privacy in my daily work.
PM^m	
PM1	I think doctors' protection of patients' privacy is a requirement of their own professional ethics.
PM2	I think my sense of professional ethics urges me to protect patients' privacy in my daily work.
PM3	From education to work, the protection of patients' private information is a professional ethic repeatedly emphasized by doctors.
EHⁿ	
EH1	I think doctors should consider the harm of privacy information disclosure from the perspective of patients, to become more aware of protecting the privacy of patients.
EH2	I have had a personal information disclosure experience as a patient, so I am more aware of protecting the privacy of patients.
EH3	I think that I can 'push myself to others' to protect my patients' privacy in my daily work.
CF^o	
CF1	I think I have developed a sense of privacy protection in my clinical work.
CF2	I think I have formed a sense of privacy protection in my teaching.
CF3	I think I have formed a sense of privacy protection in my own research work.
BP^p	
BP1	Protect the patient's privacy during surgery or examination, such as curtain pulling and preventing a third party from breaking in.
BP2	Effectively block the privacy of patients during live operations.
BP3	Medical observation or teaching requires the consent of the patient.
BP4	No illegal touch or peek at the patient's privacy.
IP^q	

Variable and code	Measurement item
IP1	In the situations of outpatient, ward check, case discussion, medical education and observation, the patient shall obtain the consent of the patient himself and take confidentiality measures. The privacy information of the patient shall not be publicized or publicly discussed orally, including the personal information and disease information with identifiable characteristics, such as avoiding calling the full name of the patient loudly, avoiding 'listening' or 'breaking in' by people other than patients without the consent of the patient.
IP2	In the face of the condition inquiry, strictly confirm and ask the status of the patient's condition personnel, confirm as me or with my consent.
IP3	For patients with special conditions (for example infectious diseases involving privacy), it is necessary to talk to the patients individually.
IP4	Deliberately disclose and disseminate the privacy of patients without using their duties, such as taking the bedside card test sheets of celebrities to the internet.
IP5	Protect medical documents such as inspection and medical records without random placing, damage, loss, and prevent theft and being wrongly picked up.
IP6	Under the unnecessary diagnosis and treatment process, without the consent of the patient, the medical documents shall not be checked, copied, or borrowed during the hospitalization of the patient.
IP7	Use personal information system account number as required, and login to view patient information without borrowing non-authorised people.
IP8	Not disclose the privacy information of the patient for any benefit reasons to obtain business, advertise or defraud.
IP9	When leaving the office seat, protect the pages with patient privacy information and lock the screen of the computer.
IP10	Scientific research, including the mining of electronic medical record information, whether it is the steps of data acquisition, viewing, processing or analysis, is strictly done to de privacy.
IP11	In the form of talks or written (case discussion, writing medical treatises, scientific research papers), for example, when communicating and learning on medical social network platform to share typical cases, do well in privacy treatment.
RP^f	
RP1	Do not disclose information about family members and other personal relationships of any patient.
RP2	Do not disclose family members and other personal relationship information of any patient on social platforms.
RP3	Do not verbally promote or publicly discuss family members and other personal relationship information of any patient.

^aPSE: pperceived severity.

^bPSU: perceived susceptibility.

^cIRE: intrinsic rewards.

^dERE: extrinsic rewards.

^eSE: self-efficacy.

^fRE: response efficacy.

^gRC: response cost.

^hSS: supervision support.

ⁱIS: information support.

^jNS: norm support.

^kES: environment support.

^lRS: responsibility.

^mPM: professional moral.

ⁿEH: empathy heart.

^oCF: consciousness formation.

^pBP: body privacy.

^qIP: information privacy.

^rRP: related privacy.

Discussion

Principal Findings

The results of both EFA and CFA revealed the development of awareness and the behavior of public medical institutions in China regarding patients' privacy protection behavior, which

could be measured by the scale proposed in this research. The scale consisted of 18 diverse dimensions, which were theoretically meaningful.

We suggest that the threat assessment process involves 4 dimensions: PSE, PSU, IRE, and ERE. Rogers [21] defines PSE as the judgment of the degree of harmfulness caused by a threat

event. PSU refers to the probability of feeling as though a threat event will occur. IRE and ERE both represent the “benefits” introduced by considering hazard factors from the researcher’s own feelings and from the outside world, respectively [22,23]. In this study, PSE included the punishments faced by doctors who disclosed patients’ private information in line with laws and regulations, in addition to the punishment related to hospital norms and the threat of patients’ rights protection. PSU includes the possibility of introducing laws and regulations to punish doctors for disclosing patients’ private information, the possibility of hospitals penalizing doctors for disclosing patients’ private information, and the possibility of patients safeguarding their own rights. IRE includes obtaining financial returns and scientific research achievements, in addition to satisfying vanity by divulging patients’ private information. ERE refers to information doctors hear from the outside world. CA is the assessment of doctors’ ability to protect patients’ privacy, including SE, RE, and RC [22,23]. Throughout this study, SE refers to doctors’ cognition regarding whether they are capable of protecting patients’ privacy. RE refers to the judgment of whether the privacy protection measures undertaken by doctors can protect patient information to an effective degree. RC refers to the cost that doctors are required to pay when taking protective actions, spanning the impact on scientific research and clinical work, that is, doctors’ protection of patients’ privacy will hinder their scientific research work or clinical work.

PMT is used to explain the way in which individuals seek self-protection from a harmful or stressful life [50]. Doctors do not protect patients’ privacy of their own accord. Based on the results of grounded theory, 2 cognitive evaluation processes applicable to this study are summarized: support evaluation and ethical evaluation. Various doctors in the interview revealed that, “...for clinical work, doctors are too busy to notice so much, if possible, hospitals should improve information construction and better protect patient information. In fact, we can make a proposition on how to protect patient privacy in the data flow process...” Throughout this study, support evaluation was predominantly in the context of hospitals. All aspects of hospital support are greatly significant for doctors to successfully carry out patient privacy protection, which can be summarized as SS, IS, NS, and ES. SS means that the hospital must designate a department to perform full-time supervision of patients’ privacy protection by doctors, as well as providing training and assessments on privacy protection laws and regulations. IS refers to the improvement in the overall information construction of the hospital and ensuring the privacy and safety of patients are protected. NS includes the establishment and development of a patient privacy protection system, determining the privacy protection process and a reasonable scientific research application system, and carrying out scientific research efficiently on the basis of legal and compliant patient privacy protection. ES refers to improving the medical environment, maintaining the medical order, and providing facilities to prevent patient privacy leakage, for example, curtains, and privacy processing of bedside card information. EA means to analyze and extract relevant factors that promote doctors’ privacy protection behavior according to medical ethics, including RS, PM, and EH. Previous studies have demonstrated that to protect medical privacy, all medical

staff who protect others have the responsibility to protect medical privacy. Responsibility ethics is a form of moral thinking that includes other thinking. The lack of moral quality in data application subjects is 1 of the primary subjective factors of anomie of data privacy ethics [51,52]. The original interview stated that “...the main responsibility of doctors is to treat and save patients...” In this study, RS determined that protecting patients’ privacy is both the doctor’s job and social responsibility. The medical ethics of doctors is mentioned in the norms and implementation measures for medical ethics of medical personnel, according to the Ministry of Health [8] issued on December 15, 1988. It discusses the ideological qualities that medical personnel should possess and the sum of the relationship between medical personnel, patients, and society. The implication of medical ethics is to attempt to do everything within your means to be good for patients [53]. “Keeping medical secrets for patients and not divulging patients’ privacy and secrets” is 1 of the provisions mentioned in the code of medical ethics. Similarly, the international code of medical ethics stipulates that “due to the trust of patients, a doctor must absolutely keep patients’ privacy” [54]. The original interview stated that “...for my own consideration of professional ethics, I try my best to protect patients’ privacy in clinical and scientific research work...” In this study, the protection of the privacy of patients is a requirement of the doctors’ PM. It is repeatedly emphasized that patients’ private information must be protected by doctors from the educational stage to the work stage. Empathy was first introduced to the field of psychology by the humanistic psychologist Rogers. This refers to an individual understanding of the experience and emotional state of others [55]. Empathetic doctors will convey their understanding to patients [56] and even map the patient’s experience, transforming the patient’s point of view to become a self-centered point of view [57]. According to the analysis of the interview data, a conclusion of this study on empathy was formed, including that doctors can consider the harm of disclosing private information from the perspective of patients and the experience of privacy information disclosure when doctors are patients. Referring to the original text “...imagine that your information will be leaked, commercial and public. Isn’t it terrible...do multi-centre research, and I especially emphasize whether the information is desensitized...lack of empathy to protect the patient’s privacy. It’s different after thinking about yourself...”

The motivation for protection in PMT has typical characteristics of motivation, which can cause, maintain, and guide activities. Generally, the results of cognitive assessment cause individuals to display greater intention of protective behavior [58]. The concept of protection motivation, in combination with various research scenarios, has evolved into individual intentions of protection behavior, for example, the intention of hot-spring tourists to revisit [59], the intention regarding self-care of elderly patients with chronic diseases [60], and the intention of self-management of diabetic retinopathy [61]. The original interview data discussed that “...the doctors’ work is too busy, and the awareness of patients’ privacy protection is not strong...” Throughout this study, the meaning of doctors’ motivation for patients’ privacy protection in public medical institutions in

China is whether doctors develop an awareness of the importance of patient privacy protection.

According to the results of the coding analysis, the privacy protection behavior of doctors includes protecting the private space of patients (BP), protecting the patients' data (IP), and protecting family members and other personal relations (RP). The protection of BP refers to the protection of the patient's body privacy by the doctor. Meanwhile, the protection of patient IP refers to the doctor's protection of the patient's information privacy. Finally, the protection of patient RP refers to the doctor's protection of the patient's associated privacy, including not disclosing or promoting the family members and other personal relationship information about any patient.

Our interviews and scales were created with input from the expected target population, that is, Chinese medical institutions. EFA revealed a unified structure of factor loads without cross loads, allowing a clear interpretation of all potential configurations. CFA confirmed the structure of the scale, meaning the subordinate relationship between each item and the extracted factors was correct. The primary fitting indexes met the requirements of the critical value, and the fitting of CFA was acceptable. It is known that model fit assessment is challenging, considering a small sample size [42], and thus, we agree that further goodness-of-fit assessments with a larger sample size will be necessary in future studies.

Limitations

This study was subject to various limitations. First, the sample size used was small. Although a general minimum sample size is yet to be defined [34], the authors recommended a sample size at least 5-10 times that of the items [30]. Another limitation was the poor representativeness of samples. In the process of selecting interviewees, the feasibility of obtaining interview samples was considered. Due to the full nature of doctors' schedules, it was challenging to coordinate a time and place, especially when conducting focus group interviews. Thus, the overall sampling was concentrated in East China. The scale survey was also predominantly based on the principle of

convenience sampling, and the regions from which samples were taken mainly included East and North China. Considering the understanding of doctors' privacy protection behavior by interviewees and respondents, hospitals in different regions have different norms and requirements for doctors' privacy protection behavior, which would lead to a certain selection bias. In addition, doctors' privacy protection for patients was not merely a problem of doctors' behavior but was also due to hospital management. Furthermore, when doctors were interviewed, they might have reservations regarding their true ideas.

Implications for Future Research

First, we must sample public medical institutions across all 7 regions of China to further verify the reliability and validity of the scale and of doctors' privacy protection behavior. Subsequently, under the influence of various cognitive mediating factors, doctors' awareness of patient privacy protection will be affected, resulting in actual privacy protection behavior. The variables will be measured using the scale proposed in this research, and the key factors to promote doctors' privacy protection behavior will be identified, with the aim of improving it from the perspective of management. In addition to the theoretical model factors, personal factors, including age, gender, educational background, professional title, religious belief, and working years, as well as the environmental policy factors of the region and department of hospitals, also affect doctors' privacy protection behavior. Nevertheless, to gain additional insights into these relationships and the relative importance of different factors, further research including the scale is needed.

Conclusion

The theoretical framework and the scale of doctors' protective behavior of patients' privacy in public medical institutions fill a crucial gap in the literature and can be used to further the current knowledge of physicians' thought processes and decisions regarding patients' privacy protection.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The coding results of qualitative research.

[\[DOCX File, 75 KB - formative_v6i12e39947_app1.docx\]](#)

Multimedia Appendix 2

The overall cognitive intermediary measurement scale.

[\[DOCX File, 46 KB - formative_v6i12e39947_app2.docx\]](#)

Multimedia Appendix 3

Results of EFA. EFA: exploratory factor analysis.

[\[DOCX File, 39 KB - formative_v6i12e39947_app3.docx\]](#)

References

1. Kloss L, Brodник M, Rinehart-Thompson L. Access and disclosure of personal health information: a challenging privacy landscape in 2016-2018. *Yearb Med Inform* 2018 Aug 29;27(1):60-66 [FREE Full text] [doi: [10.1055/s-0038-1667071](https://doi.org/10.1055/s-0038-1667071)] [Medline: [30157506](https://pubmed.ncbi.nlm.nih.gov/30157506/)]
2. Verizon. 2021 Data Breach Investigations Report. URL: <https://www.verizon.com/business/resources/reports/2021/2021-data-breach-investigations-report.pdf> [accessed 2022-02-21]
3. Ware SG, Young RM. Intentionality and conflict in the best laid plans interactive narrative virtual environment. *IEEE Trans Comput Intell AI Games* 2016 Dec;8(4):402-411. [doi: [10.1109/tciaig.2015.2489159](https://doi.org/10.1109/tciaig.2015.2489159)]
4. Ma SY, Wenbin C. Investigation on the current situation and demand of patients' medical information privacy protection. *Chin J Hospital Manag* 2018;34(1):4.
5. Yu PB, Wu J. Investigation on privacy status of emergency patients and analysis of influencing factors. *Chin J Modern Nurs* 2014;9:3.
6. Hou X, Qian F, Xiaoqing J, Lei Q, Zhaorong S, Yuxiu L. Ethical issues in the construction of health care big data. *J PLA Hospital Manag* 2020;27(06):552-554. [doi: [10.16770/J.cnki.1008-9985.2020.06.016](https://doi.org/10.16770/J.cnki.1008-9985.2020.06.016)]
7. Sher M, Talley PC, Cheng T, Kuo K. How can hospitals better protect the privacy of electronic medical records? Perspectives from staff members of health information management departments. *Health Inf Manag* 2017 May 04;46(2):87-95. [doi: [10.1177/1833358316671264](https://doi.org/10.1177/1833358316671264)] [Medline: [27702792](https://pubmed.ncbi.nlm.nih.gov/27702792/)]
8. Ministry of Health. Norms and implementation measures of medical ethics of medical personnel. *Chin J Hospital Manag* 1989;9(3):8-8.
9. Li RL. *Medical Ethics* (Vol 33). Changsha, China: Central South University Press; 2003.
10. China PsGotsRo. Law of the People's Republic of China on Licensed Doctors (2005). URL: http://www.gov.cn/banshi/2005-08/01/content_18970.htm [accessed 2022-02-22]
11. Sichun M, Li L, Lin S, Meiping J. Survey and analysis on the current status of privacy preservation for urological patients. *J Tradit Chin Med Manag* 2015;23(16):2.
12. Zheng L, Jianjun J, Lu S, Yue S, Ping D. Survey on awareness and protection of patient privacy among clinical trainees. *J Hospital Admin of PLA* 2020;27(5):5.
13. Foth M. Factors influencing the intention to comply with data protection regulations in hospitals: based on gender differences in behaviour and deterrence. *EurJ Inf Syst* 2017 Dec 19;25(2):91-109. [doi: [10.1057/ejis.2015.9](https://doi.org/10.1057/ejis.2015.9)]
14. Ting P. *Study on the Behavioral Mechanism of Medical Staff's Protection of Patients' Privacy*. Dalian, China: Dalian University of Technology; 2020.
15. Chen LF, Li L, Yang X. Analysis on the current situation and influencing factors of nurses' privacy protection behavior towards patients. *Nurs Manag China* 2018;18(1):6.
16. Dehua L, Rongmei W, Liu M, Yuyang W. Investigation on nurses' cognition of privacy protection of electronic medical record. *China Health Admin* 2017;5:3.
17. Jingya H, Yawen M, Chunyan S, Li S, Xia F. Study on the current situation and influencing factors of hemodialysis nurses' protection of patients' privacy. *J Nurs* 2019;34(11):4.
18. Noar SM. A health educator's guide to theories of health behavior. *Int Q Community Health Educ* 2005;24(1):75-92. [doi: [10.2190/DALP-3F95-GCT3-M922](https://doi.org/10.2190/DALP-3F95-GCT3-M922)] [Medline: [17690053](https://pubmed.ncbi.nlm.nih.gov/17690053/)]
19. Ho SS, Lwin MO, Yee AZ, Lee EW. Understanding factors associated with Singaporean adolescents' intention to adopt privacy protection behavior using an extended theory of planned behavior. *Cyberpsychol Behav Soc Netw* 2017 Sep;20(9):572-579. [doi: [10.1089/cyber.2017.0061](https://doi.org/10.1089/cyber.2017.0061)] [Medline: [28876967](https://pubmed.ncbi.nlm.nih.gov/28876967/)]
20. Zhang X, Liu S, Chen X, Wang L, Gao B, Zhu Q. Health information privacy concerns, antecedents, and information disclosure intention in online health communities. *Inf Manag* 2018 Jun;55(4):482-493. [doi: [10.1016/j.im.2017.11.003](https://doi.org/10.1016/j.im.2017.11.003)]
21. Xiangyun Q, Wei H, Guiling G, Haixia L. Research progress of protection motivation theory and its application. *Chin J Modern Nurs* 2012;18(4):3.
22. Rogers RW. A protection motivation theory of fear appeals and attitude change. *J Psychol* 1975 Sep 02;91(1):93-114. [doi: [10.1080/00223980.1975.9915803](https://doi.org/10.1080/00223980.1975.9915803)] [Medline: [28136248](https://pubmed.ncbi.nlm.nih.gov/28136248/)]
23. Rogers R, Prentice-Dunn S. Protection motivation theory. In: Gochman DS, editor. *Handbook of Health Behavior Research 1: Personal and Social Determinants*. New York, NY: Plenum Press; 1997:113-132.
24. Kenny M, Fourie R. Contrasting classic, Straussian, and constructivist grounded theory: methodological and philosophical conflicts. *TQR* 2015 Aug 17;20(8):1270-1289. [doi: [10.46743/2160-3715/2015.2251](https://doi.org/10.46743/2160-3715/2015.2251)]
25. Eich D. grounded theory of high-quality programs in higher education: perspectives from student leadership development programs in higher education. *J Leadersh Organ Stud* 2008 Jun 30;15(2):176-187. [doi: [10.1177/1548051808324099](https://doi.org/10.1177/1548051808324099)]
26. Anyuan G. *Research on the Influencing Factors and Mechanism of Psychological Contract Breach Based on Grounded Theory*. Wuhan, China: Wuhan University; 2015.
27. Yanbo Y, Yixue L. Conceptual model and operation mechanism of tourism market order: an exploratory study based on grounded theory. *Tourism Tribune* 2019;34(5):14.
28. Churchill GA. A paradigm for developing better measures of marketing constructs. *J Mark Res* 2018 Nov 28;16(1):64-73. [doi: [10.1177/002224377901600110](https://doi.org/10.1177/002224377901600110)]

29. Côrte-Real N, Ruivo P, Oliveira T, Popovič A. Unlocking the drivers of big data analytics value in firms. *J Bus Res* 2019 Apr;97:160-173. [doi: [10.1016/j.jbusres.2018.12.072](https://doi.org/10.1016/j.jbusres.2018.12.072)]
30. Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;22:140.
31. Gliem J, Gliem R. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. 2003 Presented at: Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education; 2003; Columbus, Ohio.
32. Ma WJ, Bo P. The reliability and validity of the questionnaire and how to analyze it with SAS software. *China Health Stat* 2000;17(6):364-365.
33. O'Leary-Kelly SW, Vokurka RJ. The empirical assessment of construct validity. *J Oper Manag* 1998 Nov 25;16(4):387-405. [doi: [10.1016/s0272-6963\(98\)00020-5](https://doi.org/10.1016/s0272-6963(98)00020-5)]
34. Williams B, Onsman A, Brown T. Exploratory factor analysis: a five-step guide for novices. *Australas J Paramed* 2010 Aug 02;8(3):1-13. [doi: [10.33151/ajp.8.3.93](https://doi.org/10.33151/ajp.8.3.93)]
35. Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behav Res* 1990 Apr;25(2):173-180. [doi: [10.1207/s15327906mbr2502_4](https://doi.org/10.1207/s15327906mbr2502_4)]
36. Qi C, Wannian L, Qun M. Structural equation model and its application. *China Health Stat* 2004;21(2):70-74.
37. Bo Q, Haiqiang G, Jiping R, Gao S. Structural equation model and its application. *China Health Stat* 2005;22(6):405-407.
38. Jackson DL, Gillaspay JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Methods* 2009 Mar;14(1):6-23. [doi: [10.1037/a0014694](https://doi.org/10.1037/a0014694)] [Medline: [19271845](https://pubmed.ncbi.nlm.nih.gov/19271845/)]
39. Schmitt TA. Current methodological considerations in exploratory and confirmatory factor analysis. *J Psychoeduc Assess* 2011 Aug 10;29(4):304-321. [doi: [10.1177/0734282911406653](https://doi.org/10.1177/0734282911406653)]
40. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 1999 Sep;4(3):272-299. [doi: [10.1037/1082-989x.4.3.272](https://doi.org/10.1037/1082-989x.4.3.272)]
41. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. *Counsel Psychol* 2016 Jun 30;34(6):806-838. [doi: [10.1177/0011000006288127](https://doi.org/10.1177/0011000006288127)]
42. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999 Jan;6(1):1-55. [doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)]
43. Geldhof GJ, Preacher KJ, Zyphur MJ. Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychol Methods* 2014 Mar;19(1):72-91. [doi: [10.1037/a0032138](https://doi.org/10.1037/a0032138)] [Medline: [23646988](https://pubmed.ncbi.nlm.nih.gov/23646988/)]
44. Werts C, Linn R, Jöreskog K. Intraclass reliability estimates: testing structural assumptions. *Educ Psychol Meas* 2016 Jul 02;34(1):25-33. [doi: [10.1177/001316447403400104](https://doi.org/10.1177/001316447403400104)]
45. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh* 2007 Jun;39(2):155-164. [doi: [10.1111/j.1547-5069.2007.00161.x](https://doi.org/10.1111/j.1547-5069.2007.00161.x)] [Medline: [17535316](https://pubmed.ncbi.nlm.nih.gov/17535316/)]
46. Bagozzi RP. Evaluating structural equation models with unobservable variables and measurement error: a comment. *J Mark Res* 1981 Aug;18(3):375. [doi: [10.2307/3150979](https://doi.org/10.2307/3150979)]
47. Thompson B, Daniel LG. Factor analytic evidence for the construct validity of scores: a historical overview and some guidelines. *Educ Psychol Meas* 2016 Jul 02;56(2):197-208. [doi: [10.1177/0013164496056002001](https://doi.org/10.1177/0013164496056002001)]
48. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* 2003 Oct;88(5):879-903. [doi: [10.1037/0021-9010.88.5.879](https://doi.org/10.1037/0021-9010.88.5.879)] [Medline: [14516251](https://pubmed.ncbi.nlm.nih.gov/14516251/)]
49. Ming Long W. *Statistical Analysis of Survey Practices: SPSS Operation and Application (Chinese Edition)*. China: Chongqing University Press; 2010.
50. Maddux JE, Rogers RW. Protection motivation and self-efficacy: a revised theory of fear appeals and attitude change. *J Exp Soc Psychol* 1983 Sep;19(5):469-479. [doi: [10.1016/0022-1031\(83\)90023-9](https://doi.org/10.1016/0022-1031(83)90023-9)]
51. Tang KL. Research on big data privacy ethics. *Stud Ethics* 2016;6:5.
52. Chenglong X. Ethical analysis of privacy in the era of big data. *Popular Literature Art* 2021;4:215-216.
53. Nanshan Z. Brief discussion on the connotation of medical ethics. *Chin Med Ethics* 2006;19(3):3.
54. Kun Y. Brief discussion on medical confidentiality. *Chin Med Ethics* 1999;4:3. [doi: [10.1136/jme.10.1.3](https://doi.org/10.1136/jme.10.1.3)] [Medline: [6708061](https://pubmed.ncbi.nlm.nih.gov/6708061/)]
55. Shamay-Tsoory SG. The neural bases for empathy. *Neuroscientist* 2011 Feb 11;17(1):18-24. [doi: [10.1177/1073858410379268](https://doi.org/10.1177/1073858410379268)] [Medline: [21071616](https://pubmed.ncbi.nlm.nih.gov/21071616/)]
56. Nightingale SD, Yarnold PR, Greenberg MS. Sympathy, empathy, and physician resource utilization. *J Gen Intern Med* 1991 Sep;6(5):420-423. [doi: [10.1007/bf02598163](https://doi.org/10.1007/bf02598163)]
57. Jensen N. The empathic physician. *Arch Intern Med* 1994;154(1):106a. [doi: [10.1001/archinte.154.1.106a](https://doi.org/10.1001/archinte.154.1.106a)]
58. Boss SR, Galletta DF, Lowry PB, Moody GD, Polak P. What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly* 2015 Apr 4;39(4):837-864. [doi: [10.25300/misq/2015/39.4.5](https://doi.org/10.25300/misq/2015/39.4.5)]
59. Wang WL, Huang L. A study on the evaluation of the theoretical model of hot spring tourists' willingness to revisit: from the perspective of conservation motivation theory. *Tourism Tribune* 2021;36(06):133-142. [doi: [10.19765/j.cnki.1002-5006.2021.06.014](https://doi.org/10.19765/j.cnki.1002-5006.2021.06.014)]

60. Aihong X. Effect of nursing intervention based on protective motivation theory in elderly patients with chronic diseases. *J Nurs* 2021;28(16):5.
61. Sujing L, Qin C. Study on self-management of patients with diabetes retinopathy by standard health education based on protective motivation theory. *China Standardiz* 2021;10:2.

Abbreviations

AVE: average variance extracted
BP: body privacy
CA: coping appraisal
CF: consciousness formation
CFA: confirmatory factor analysis
CFI: comparative fit index
CPR: composite reliability
CR: critical ratio
EA: ethical appraisal
EFA: exploratory factor analysis
EH: empathy heart
ERE: extrinsic rewards
ES: environment support
IP: information privacy
IRE: intrinsic rewards
IS: information support
KMO: Kaiser-Meyer-Olkin
NS: norm support
PM: professional moral
PMT: protection motivation theory
PSE: perceived severity
PSU: perceived susceptibility
RC: response cost
RE: response efficacy
RMSEA: root-mean-square error of approximation
RP: related privacy
RS: responsibility
SA: support appraisal
SE: self-efficacy
SRMR: standardized root-mean-square residual
SS: supervision support
TA: threat appraisal
TLI: Tucker-Lewis index

Edited by A Mavragani; submitted 29.05.22; peer-reviewed by M Roguljić, SF Fung; comments to author 01.11.22; revised version received 09.11.22; accepted 21.11.22; published 14.12.22.

Please cite as:

Xu J, Lu L, Xing K, Shi H, Chen R, Yao Y, Liu S, Xiao Z, Peng X, Luo S, Zhong Y
Theoretical Approach and Scale Construction of Patient Privacy Protection Behavior of Doctors in Public Medical Institutions in China: Pilot Development Study
JMIR Form Res 2022;6(12):e39947
URL: <https://formative.jmir.org/2022/12/e39947>
doi: [10.2196/39947](https://doi.org/10.2196/39947)
PMID: [36515996](https://pubmed.ncbi.nlm.nih.gov/36515996/)

©Jie Xu, Lu Lu, Kaichen Xing, Huwei Shi, Ruiyao Chen, Yujun Yao, Sichen Liu, Zhongzhou Xiao, Xinwei Peng, Shuqing Luo, Yun Zhong. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 14.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR*

Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Job Disengagement Among Physical Education Teachers: Insights From a Cross-sectional Web-Based Survey With Path Modeling Analysis

Nasr Chalghaf^{1*}, PhD; Wen Chen^{2*}, MSc; Amayra Tannoubi³, MSc; Noomen Guelmami³, PhD; Luca Puce⁴, PhD; Nouredine Ben Said⁵, PhD; Maher Ben Khalifa⁶, PhD; Fairouz Azaiez¹, PhD; Nicola Luigi Bragazzi⁷, MPH, MD, PhD

¹Higher Institute of Sport and Physical Education of Gafsa, University of Gafsa, Gafsa, Tunisia

²Department of Child Psychology, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, National Children's Regional Medical Center, Hangzhou, Zhejiang, China

³Higher Institute of Sport and Physical Education of Kef, University of Jendouba, Jendouba, Tunisia

⁴Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOEMI), University of Genoa, Genoa, Italy

⁵Department of Biomechanics and Motor Behavior, College of Sport Sciences and Physical Activity, King Saud University, Riyadh, Saudi Arabia

⁶Department of Human Sciences, Higher Institute of Sport and Physical Education of Sfax, University of Sfax, Sfax, Tunisia

⁷Laboratory for Industrial and Applied Mathematics (LIAM), Department of Mathematics and Statistics, York University, Toronto, ON, Canada

*these authors contributed equally

Corresponding Author:

Nicola Luigi Bragazzi, MPH, MD, PhD

Laboratory for Industrial and Applied Mathematics (LIAM)

Department of Mathematics and Statistics

York University

4700 Keele St

Toronto, ON, M3J 1P3

Canada

Phone: 1 416 736 2100 ext 66093

Email: bragazzi@yorku.ca

Abstract

Background: Physical education teachers often experience stress and job disengagement.

Objective: This study's aims were as follows: (1) to adapt in the Arabic language and test the reliability and the validity of the work–family conflict (WFC) and family–work conflict (FWC) scales, (2) to develop and assess the psychometric properties of work disengagement among physical education teachers, and (3) to evaluate an explanatory model by presenting the mediating role of perceived stress as a major influencing factor in work disengagement and job satisfaction.

Methods: A total of 303 primary and secondary school physical education teachers, comprising 165 (54.5%) men and 138 (45.5%) women participated voluntarily in our study. The measuring instruments are the Work Disengagement Scale, the Perceived Stress Scale, the WFC scale, the FWC scale, and the 9-item Teacher of Physical Education Job Satisfaction Inventory.

Results: The Arabic language versions of the WFC and FWC scales had reasonably adequate psychometric properties, which were justified by confirmatory factor analyses and by the measurement of reliability, convergent, and discriminant validity through the measurement model using SmartPLS software. Similarly, the structural model established with SmartPLS confirmed strong links of the concepts of FWC, WFC, the job satisfaction questionnaire, and perceived stress with work disengagement among teachers of physical education.

Conclusions: There is a growing interest in helping teachers cope with the daily pressures of work and family. A positive organizational context is a context with clear values regarding work priorities, which constitutes the basis of a feeling of shared responsibility and professional support. Good conditions can act as protective factors reducing work stress and positively influencing personal well-being, work attitudes, work commitment, and professional efficiency. Additional research on teachers is needed to examine the relationship between perceived work stress and the role of families, along with the extent to which this association can have a significant impact on teachers' commitment to work.

KEYWORDS

Work Disengagement Scale; work; job; job satisfaction; family–work conflict; perceived stress; physical education; PLS-SEM; SmartPLS; teacher; engagement; Arab; stress; primary school; secondary school; development; measurement; scale; tool; fitness; teacher; educator; school; satisfaction; digital tool; mental health; family; cross-sectional; survey; modelling; psychology

Introduction

The problem of perceived stress has a great impact on the psychological health and well-being of employees, and it is always a very important factor to study [1].

Perceived stressors have been associated with a wide range of both mental and physical health issues, including depression and anxiety disorders, suicide, workplace accidents and injuries, and cardiovascular risk. Researchers generally agree that stress is a serious problem in many countries and has a very negative impact in the professional world. Contemporary organizations have opted for individual and organizational stress management [2]. However, it is frequently reported that the majority of stress management interventions are ineffective [3] and lead to disengagement from work. Other psychosocial factors, such as work–family conflict (WFC) and family–work conflict (FWC), can also increase stress. WFC can occur when demands from one role at home can affect one’s ability to meet the demands associated with another role at the workplace. The reverse is known as FWC, with issues at work clashing with family responsibilities and duties. FWC refers to “a form of inter-role conflict in which the general demands, the time spent, and the tensions created by the family interfere with the execution of work-related responsibilities” [4].

The features of a successful teacher include personal characteristics, topic knowledge, and highly qualified teaching skills, all of which can have an impact on pupils [5]. Looking into how students perceive their physical education teachers’ skill sets may help improve academic quality and boost participation in both extracurricular sports and physical education programs. This strategy for increasing physical activity has been developed as a result of the discovery that students’ perceptions of the school setting were in fact associated with their well-being, academic success, and positive attitudes toward school-based physical exercises. The physical education teacher, as physical activity specialist, should give a positive primary representation of this environment and should act as a role model for the pupils in terms of preserving physical health and leading an active life. Teachers’ influences are much more nuanced than merely imparting knowledge and transferring abilities. Recent social and economic developments have prompted researchers to look for new strategies for enhancing teachers’ professionalism, given that, in recent years, the condition of physical education teachers has changed because of the introduction of new programs and stringent quality checks: perceived stress and professional disengagement linked to the profession among many teachers has been increasing and has become dramatically remarkable [1].

It is, therefore, very important to focus on the professional disengagement linked to the professional stress among physical

education teachers and to highlight 4 factors that are likely to increase it: relations with superiors, the professional environment, the family circle, and job satisfaction. In this regard, relationships and interactions with colleagues and students sometimes create a critical organizational environment and are, therefore, often potential sources of stress. Good interpersonal relationships help achieve personal goals among individuals and organizational goals of whole teams, while poor interpersonal relationships cause stress and affect teachers’ performance and well-being.

In Arab countries including Tunisia, we lack valid and reliable tools to measure concepts such as WFC, FWC, and work engagement. Therefore, to test relational models, it is necessary to adapt or validate measurement scales in Arabic.

The objective of this study was to develop a measurement scale of job disengagement among physical education teachers to verify the psychometric properties of WFC and FWC and to present an explanatory model describing the mediating role of perceived stress and job satisfaction, along with the relationships between family and work as an indirect effect.

Methods

Participants

We used a snowball sampling procedure to collect cross-sectional data for physical education teachers in Tunisia. Overall, 303 physical education teachers with a mean age of 36.46 (SD 7.92) years participated in the study. The sample comprised 165 (54.5%) males and 138 (45.5%) females. Furthermore, 162 (53.5%) of them were physical education teachers at primary school and 141 (46.5%) of them taught at secondary school. All teachers have more than 15 years of experience in their professional careers.

Ethics Approval

The study protocol received ethical clearance from the United Nations Educational, Scientific and Cultural Organization chair “Health Anthropology Biosphere and Healing Systems”; University of Genoa, Genoa, Italy; the Higher Institute of Sport and Physical Education of Sfax, Sfax, Tunisia; the Faculty of Letters and Human Sciences of Sfax, Sfax; and the Higher Institute of Sport and Physical Education of Kef, Kef, Tunisia. The study was approved by the ethical committee of the University of Sfax (020/2021).

All study participants provided written informed consent. Teachers were extensively informed about the purposes and procedure of the study and were advised that the results would be made available to them upon completion of the study only in aggregate form, with no possibility to trace back to individual teachers’ scores, thus ensuring anonymity and preserving the privacy of each participant.

This study was carried out following the ethical principles of the 1964 Helsinki declaration and its subsequent amendments.

Instruments

Arabic Version of the Perceived Stress Scale

The 10-item Perceived Stress Scale [6] measures global perceived stress experienced across the past 30 days, on a 5-point scale (0="never," 1="almost never," 2="once in a while," 3="often," and 4="very often"). The validity and reliability of the Arabic version of the Perceived Stress Scale were assessed with acceptable results. The overall Cronbach α coefficient was .80 for the Arabic version of the Perceived Stress Scale. The test-retest reliability had an intraclass correlation coefficient of .90 [7].

WFC and FWC

We used 5 items of 2 Arabic versions of the WFC and FWC scales [4]. The psychometric properties of the 2 original instruments were satisfactory and tested on 3 different samples. The reference model shows an adequate fit and factor invariance across the goodness-of-fit index (GFI) and the comparative fit and Tucker-Lewis indices, which are in the range of .90 and above. Responses were rated on a 5-point Likert scale ranging from 1="strongly disagree" to 5="strongly agree."

A discussion group was organized with experts of the Arabic language, sports, and physical education, as well as the humanities and applied sciences to translate the items of the 2 tools. Then, several corrections made it possible to reformulate the questions that seemed unclear. This has improved the tool. Finally, a pilot study was carried out to test the preliminary properties and the readability of the questionnaires—this test confirmed the validity of this work.

Job Dissatisfaction

Job dissatisfaction was assessed with an inverse score based on a 5-point Likert scale—the Arabic version of the 9-item Teacher of Physical Education Job Satisfaction Inventory (TPEJSI-9) [8].

Internal consistency α coefficients of the TPEJSI were all $>.80$: for satisfaction with colleagues, $\alpha=.87$; for satisfaction with parents, $\alpha=.87$; and for satisfaction with students, $\alpha=.86$. The tool also provides good exploratory factor analysis factor loadings, acceptable confirmatory factor analysis (CFA) fit indices, and excellent convergent validity.

The Work Disengagement Scale

The Work Disengagement Scale (WDES) was developed in the Arabic language from the Utrecht Work Engagement Scale validated by Schaufeli et al [9]. This scale contains 9 items that present disengagement from work on a 6-point Likert scale.

Procedure

The first step of the validation process was the translation of the original English versions of the WFC and FWC scales to classical Arabic by a committee.

After informed consent was obtained, the selected participants answered a paper version structured questionnaire that included all scales (Multimedia Appendix 1). The entire procedure of questionnaire administration lasted over 2 months. A proper time period (approximately 60 minutes) was ensured for each participant to answer the questionnaire thoroughly.

Statistical Analysis

Analyses were performed in SPSS Statistics software (version 22.0; IBM Corp) and SmartPLS (version 3.2.9; SmartPLS GmbH).

Before commencing any statistical analysis, data were visually inspected for potential outliers. The normality of data distribution was checked using the Pearson-D'Agostino omnibus test. Means and SDs for ordinal data were computed for the entire sample.

Questionnaires' scores were also checked for skewness and kurtosis, computing the Mardia multivariate skewness and kurtosis statistics [10].

CFA was used to verify the psychometric properties of the WFC and FWC scales and the WDES. Several fit indices were evaluated to determine model fit.

When the model was tested using CFA, as suggested and recommended by many scholars, a wide range of fit indices was calculated and reported, including the following: (1) discrepancy indices (including the chi-square and the Steiger and Lind [11] root mean square error of approximation), (2) tests comparing the target model with the null model (including the Bentler and Bonett [12] normed fit index; the Bentler and Bonett [12] nonnormed fit index, known also as the Tucker-Lewis index; the Bentler comparative fit index; and the James-Mulaik-Brett parsimony GFI [13]), and (3) information theory goodness-of-fit measures (the Joreskog GFI and the Joreskog adjusted GFI).

For the estimation of the model, we used the SmartPLS software (version 3.2.9) to analyze the path model. The PLS method uses least squares regression techniques to estimate the models. The objective of PLS modeling is to maximize the explained variance of the dependent latent variable, whereas that of covariance-based methods is to reproduce the theoretical covariance matrix.

The measurement model defines the relationships between the latent and observable variables. It contains indications of the operationalization of the theoretical concepts of a study. The need to evaluate the measurement model is an essential need since the relationships between the latent variables and the theoretical concept or among the different concepts may be inaccurate. For the different measurement scales, the reflective model was adopted following bibliographic recommendations. In a reflective measurement scale, the direction of causality is from the latent variable to the indicators (represented in blue and yellow, respectively, in Figures 1 and 2). If the indicators are highly correlated and interchangeable, they are reflective and their reliability and validity must be thoroughly examined [14-16].

Figure 1. SmartPLS model. FWC: family–work conflict; JD: job dissatisfaction; PS: perceived stress; WD: work disengagement; WDES: Work Disengagement Scale; WFC: work–family conflict.

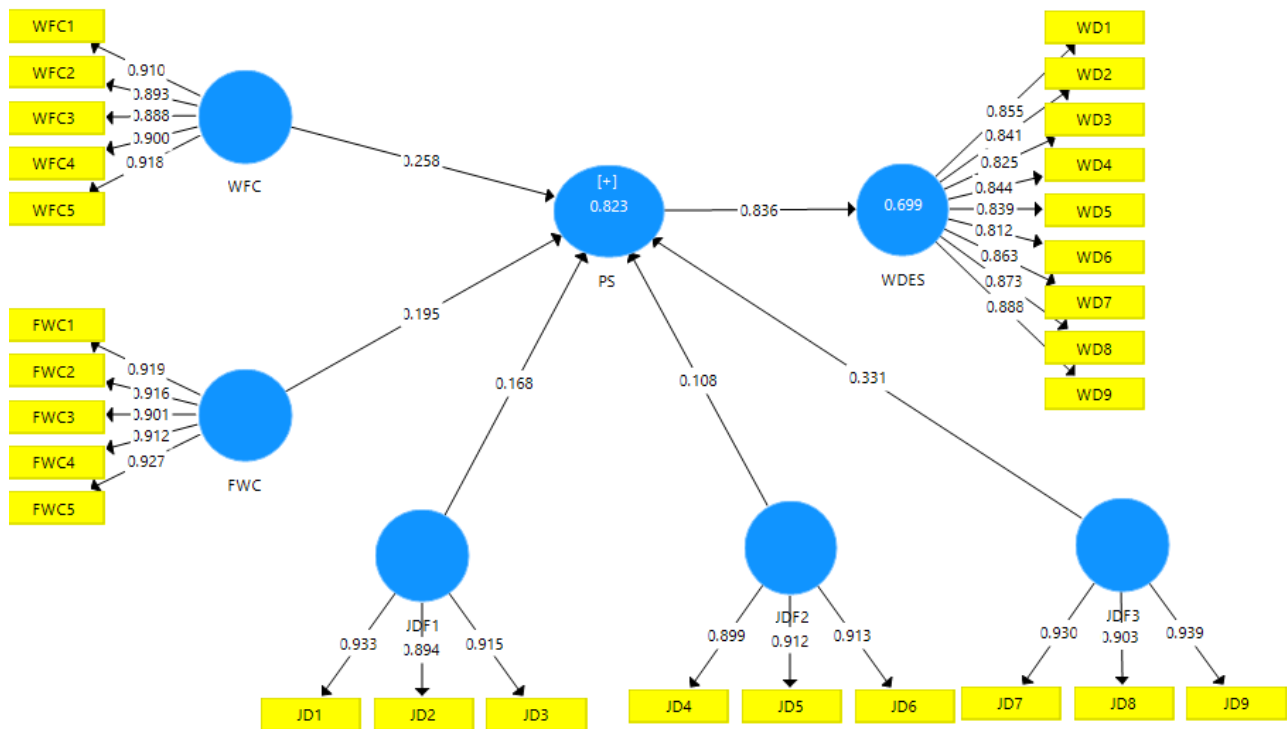
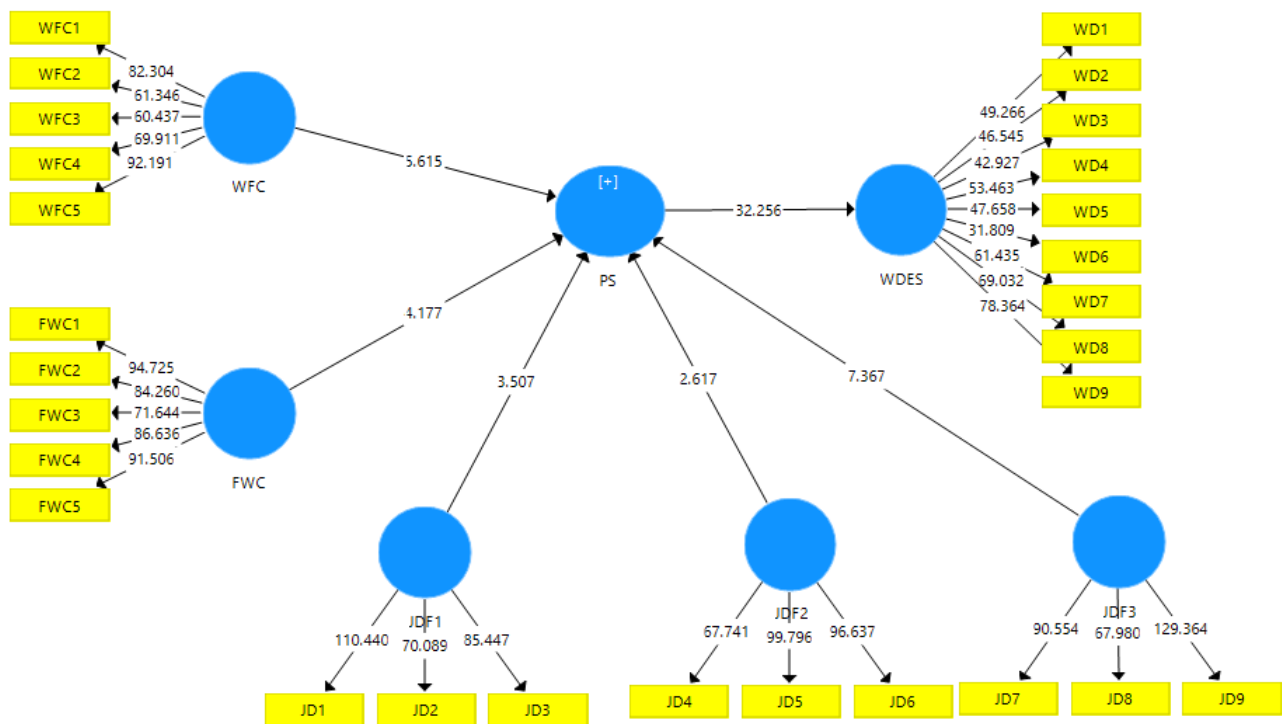


Figure 2. SmartPLS bootstrapping model. FWC: family–work conflict; JD: job dissatisfaction; PS: perceived stress; WD: work disengagement; WDES: Work Disengagement Scale; WFC: work–family conflict.



Results

Overview

We started with the factorial examination of the 2 adapted

scales—WFC and FWC scales—as well as the constructed WDES (Table 1).

The Cronbach α coefficients were remarkable for the 3 scales: Cronbach α values were .943, .952, and .952 for the WFC scale, FWC scale, and WDES.

Table 1. Confirmatory factor analysis fit index for the work–family conflict scale, family–work conflict scale, and Work Disengagement Scale.

Scales	Chi-square	df	Chi-square / df	Normed fit index	Goodness-of-fit index	Adjusted goodness-of-fit index	Tucker-Lewis index	Comparative fit index	Root mean residual	Root mean square error
Work–family conflict scale	3.47	5	0.70	.99	.99	.98	1	1	.02	0
Family–work conflict scale	7.41	5	1.48	.995	.99	.97	.997	.998	.020	.040
Work Disengagement Scale	59.62	27	2.21	.97	.96	.93	.98	.98	.03	.063

Evaluation of the Measurement Model

The findings of the SmartPLS algorithm are pictorially shown in Figure 1.

Hair et al [17] argue that to verify the measurement model, it is necessary to verify its reliability using composite reliability (CR) and its validity by measuring convergent validity and

discriminant validity. According to Fornell and Larcker [18], the CR must be $\geq .6$ to have a reliable construct. Here, we obtained values of .963, .939, .934, .946, .923, .959, and .956 for the variables “FWC,” “JDF1,” “JDF2,” “JDF3,” “PS,” “WDES,” and “WFC” respectively (Table 2). These values are well above the recommended threshold. This indicates that our measurement model is reliable.

Table 2. Reliability and average variance extracted (AVE).

Variables	Cronbach α	PA	Composite reliability	AVE
Family–work conflict	.952	.952	.963	.838
Job dissatisfaction (F1)	.902	.904	.939	.836
Job dissatisfaction (F2)	.893	.894	.934	.824
Job dissatisfaction (F3)	.914	.917	.946	.854
Perceived stress	.907	.907	.923	.545
Work disengagement	.952	.952	.959	.721
Work–family conflict	.943	.943	.956	.814

Convergent validity is measured using the loading factor, which must be $> .6$ [19] and the average variance extracted (AVE), which must be $> .5$ [18]. Our results show that all measures—Cronbach α , ρ_A , and CR—of the variables are $> .7$ and the AVE is $> .5$ (Table 2). These values respect the recommendations of the authors, which indicates that the measures of each variable of the converging model allow us to properly represent the variable in question.

The discriminant validity of a construct can be assessed by comparing the square root of the values of the AVE with correlations of latent variables [18]. The square roots of the

AVE coefficients are presented in the correlation matrix along the diagonal. The square root of the AVE of each construct must be greater than its strongest correlation with any other construct to demonstrate discriminant validity [15].

Table 3 shows that the square root of the AVE of each variable is greater than its correlation coefficients with other variables. These results indicate that the measurement model has good discriminant validity. Based on the results of CR, convergent validity, and discriminant validity, we can thus deduce that our measurement model is reliable and valid. Hence, after evaluating the measurement model, we will now evaluate our structural model.

Table 3. Reliability and average variance extracted.

Variable	PS ^a	FWC ^b	JD ^c F1	JDF2	JDF3	WDES ^d	WFC ^e
PS	.738	— ^f	—	—	—	—	—
FWC	.794	.915	—	—	—	—	—
JDF1	.752	.675	.914	—	—	—	—
JDF2	.684	.585	.644	.908	—	—	—
JDF3	.799	.642	.679	.614	.924	—	—
WDES	.836	.691	.644	.577	.670	.849	—
WFC	.792	.815	.614	.586	.631	.677	.902

^aPS: perceived stress.

^bFWC: family–work conflict.

^cJD: job dissatisfaction.

^dWDES: Work Disengagement Scale.

^eWFC: work–family conflict.

^fNot applicable.

Assessment of the Structural Model

Evaluation of the structural model was carried out by calculating the path coefficients to evaluate the hypotheses, the coefficient of determination (R^2), the effect size (f^2), and the predictive relevance (Q^2) [17,20].

R^2 represents the explanation's power of the dependent variable by the independent variables. Falk and Miller [21] propose .10 as the minimum value of R^2 for it to be accepted. The results of our model show that the R^2 of the variables perceived stress (PS) and WDES are equal to .823 and .699, respectively. This indicates that the influence of job dissatisfaction (JD), WFC, and FWC represents 82% of the variance of PS. Similarly, the influence of all these variables explains 70% of the variance in work disengagement.

The f^2 represents the influence of each independent variable on the dependent variable. According to Cohen [22], (1) $f^2 > .35$ implies that the independent variable has a large effect on the dependent variable, (2) $.15 < f^2 < .35$ implies that the independent variable has a medium effect on the dependent variable, and (3) $.02 < f^2 < .15$ implies that the independent variable has a small effect on the dependent variable. Our results show that PS has an f^2 of 2.32, whereas the f^2 for WFC, FWC, JDF1, JDF2, and JDF3 are .12, .07, .06, .03, and .27, respectively, and this has a small effect.

The predictive relevance Q^2 represents the predictive capacity of the model in measuring endogenous variables. Using the blindfolding method, we found that the Q^2 is equal to .44 and .50 for the variables PS and WDES, respectively. All these values are positive, which indicates that the model has good predictive capacity [17,23], which suggests that the Q^2 is greater than 0. Hypothesis testing is shown in Table 4.

Table 4. Hypothesis testing.

	Path coefficient	SD	t-value	P value	Hypothesis
FWC ^a →PS ^b	0.195	0.047	4.177	<.001	Confirmed
FWC→WDES ^c	0.163	0.039	4.161	<.001	Confirmed
JD ^d F1→PS	0.167	0.048	3.507	<.001	Confirmed
JDF1→WDES	0.140	0.041	3.455	.001	Confirmed
JDF2→PS	0.107	0.041	2.617	.009	Confirmed
JDF2→WDES	0.089	0.035	2.581	.01	Confirmed
JDF3→PS	0.330	0.045	7.367	<.001	Confirmed
JDF3→WDES	0.276	0.039	7.151	<.001	Confirmed
PS→WDES	0.837	0.026	32.256	<.001	Confirmed
WFC ^e →PS	0.260	0.046	5.615	<.001	Confirmed
WFC→WDES	0.217	0.038	5.641	<.001	Confirmed

^aFWC: family–work conflict.

^bPS: perceived stress.

^cWDES: Work Disengagement Scale.

^dJD: job dissatisfaction.

^eWFC: work–family conflict.

Discussion

Principal Findings

Our study aimed to (1) adapt in Arabic language and examine the reliability and the validity of the WFC and FWC scales, (2) design and analyze the psychometric properties of the work disengagement among physical education instructors, and (3) explore an explanatory model by demonstrating the mediation function of perceived stress as a significant influencing factor in work disengagement and job satisfaction.

The Arabic versions of the WFC and FWC scales had reasonably adequate psychometric properties, which were justified by CFA and the measure of reliability, convergent, and discriminant validity. Similarly, the developed tool WDES showed good reliability and adequate fit indices on CFA. These results of the WDES have been supported by measurement model review established in SmartPLS.

The structural model established with SmartPLS software confirmed strong links between PS and FWC, WFC, and job satisfaction among physical education teachers. These results are in line with those of Caesens [24], which revealed relationships among perceived organizational support, job satisfaction, and PS.

In line with our results, the study by Watson et al [25] conducted with 53 beginning teachers, revealed through regression analysis that holistic well-being and PS contributed significantly to the variance in job satisfaction. Furthermore, Orgambidez-Ramos et al [26] revealed by hierarchical models that job satisfaction was significantly predicted by stress and work engagement.

Besides, perceived work stress has been found to mitigate the effect of high job demands on WFC [27]. Furthermore,

significant gender differences were found in PS levels, FWCs, and commitment to work.

In another study by Gandhi et al [28] performed with 150 nurses (both male and female), the results showed that PS and perceived job satisfaction were negatively correlated. These results confirmed those of Guppy and Gutteridge [29], who found that job satisfaction was negatively linked to stress.

Among physical education teachers, Koustelios and Tsigilis [30] examined the multivariate relationship between job satisfaction and burnout as an extremely stressful situation experienced by Greek physical education teachers in school. Canonical correlation analysis revealed a negative multivariate relationship between the two constructs ($R_c=.61$).

To explain these findings, Laugaa et al [31] claimed that teaching working conditions have been deteriorating. Various researchers in Quebec, Canada, have also identified several sources of stress among teachers: workload (source most often noted) [32], lack of time [33], resources [32], and recognition or respect [33]. However, these studies have not examined relationships with families.

Furthermore, a certain degree of stress can sometimes have positive effects such as learning [34], hope, joy, passion, or satisfaction [33]; however, prolonged stress among teachers can have many negative consequences [35]: early retirement [32], effects on family life and relationships [33], and effects on satisfaction (the most stressed individuals are also the least satisfied ones [35]). Similarly, Chu [36] studied the influence of PS on WFC and mental health. The results show that PS is an effective predictor of WFC and mental health.

Limitations

This research has a number of limitations. First, the sample size did not allow a confirmatory model to be generated using the structural equation modeling approach instead of the partial least square predictive technique. Second, other factors that are related to work engagement such as grit and personality traits of study participants were not examined. Third, moderating effects such as expertise and gender were not examined. It is recommended that these moderating effects should be assessed in further work.

Conclusions

The results of the reliability and CFA suggest that the WFC scale, FWC scale, and WDES are valid and reliable and can measure all 3 concepts in a Tunisian context. These results were supported by the SmartPLS model. In addition, the structural

model's results support stress as a major influencing factor for work disengagement in physical education teachers. This model also shows that this PS stems from bidirectional FWC.

Therefore, there is growing interest in helping teachers cope with the daily pressures at work and family. A positive organizational context is a context with clear values regarding work priorities, which constitutes the basis of a feeling of shared responsibility and professional support.

Good conditions can act as protective factors reducing work stress and positively influencing personal well-being, work attitudes, work commitment, and professional efficiency. Additional research is needed to examine the relationship between perceived work stress and the role of teachers' families and the extent to which this association can have a significant impact on teachers' commitment to work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Scales in Arabic.

[[DOCX File , 15 KB - formative_v6i12e29130_app1.docx](#)]

References

1. Montgomery C, Rupp AA. A meta-analysis for exploring the diverse causes and effects of stress in teachers. *Can J Educ* 2005;28(3):458. [doi: [10.2307/4126479](#)]
2. Cooper CL, Cartwright S. Healthy mind; healthy organization— a proactive approach to occupational stress. *Human Relations* 1994;47(4):455-471. [doi: [10.1177/001872679404700405](#)]
3. Reynolds S, Briner RB. Stress management at work: with whom, for whom and to what ends? *Br J Guid Couns* 1994;22(1):75-89. [doi: [10.1080/03069889408253667](#)]
4. Netemeyer RG, Boles JS, McMurrian R. Development and validation of work–family conflict and family–work conflict scales. *J Appl Psychol* 1996 Aug;81(4):400-410. [doi: [10.1037/0021-9010.81.4.400](#)]
5. Bidner S. Reform in music teacher preparation. *J Music Teach Educ* 2001;10(2):3-4. [doi: [10.1177/10570837010100020101](#)]
6. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983 Dec;24(4):385. [doi: [10.2307/2136404](#)]
7. Almadi T, Cathers I, Hamdan Mansour AM, Chow CM. An Arabic version of the perceived stress scale: translation and validation study. *Int J Nurs Stud* 2012 Jan;49(1):84-89. [doi: [10.1016/j.ijnurstu.2011.07.012](#)] [Medline: [21851941](#)]
8. Chalghaf N, Guelmami N, Re TS, Maldonado Briegas JJ, Garbarino S, Azaiez F, et al. Trans-cultural adaptation and validation of the "Teacher Job Satisfaction Scale" in Arabic language among sports and physical education teachers ("Teacher of Physical Education Job Satisfaction Inventory"-TPEJSI): insights for sports, educational, and occupational psychology. *Front Psychol* 2019 Oct 22;10:2234 [FREE Full text] [doi: [10.3389/fpsyg.2019.02234](#)] [Medline: [31695634](#)]
9. Schaufeli WB, Bakker AB, Salanova M. The measurement of work engagement with a short questionnaire. *Educ Psychol Meas* 2006;66(4):701-716. [doi: [10.1177/0013164405282471](#)]
10. Mardia KV. Measures of multivariate skewness and kurtosis with applications. *Biometrika* 1970;57(3):519-530. [doi: [10.1093/biomet/57.3.519](#)]
11. Steiger JH. A note on multiple sample extensions of the RMSEA fit index. *Struct Equ Model* 1998;5(4):411-419. [doi: [10.1080/10705519809540115](#)]
12. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 1980 Nov;88(3):588-606. [doi: [10.1037/0033-2909.88.3.588](#)]
13. James LR, Mulaik SA, Brett JM. *Causal analysis: assumptions, models, and data*. Beverly Hills, CA: Sage Publications; 1982.
14. Haenlein M, Kaplan AM. A beginner's guide to partial least squares analysis. *Understanding Statistics* 2004 Nov;3(4):283-297. [doi: [10.1207/s15328031us0304_4](#)]
15. Hair JF, Ringle CM, Sarstedt M. Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. *Long Range Planning* 2013 Feb;46(1-2):1-12. [doi: [10.1016/j.lrp.2013.01.001](#)]

16. Petter S, Straub D, Rai A. Specifying formative constructs in information systems research. *MIS Quarterly* 2007;31(4):623. [doi: [10.2307/25148814](https://doi.org/10.2307/25148814)]
17. Hair J, Hollingsworth CL, Randolph AB, Chong AYL. An updated and expanded assessment of PLS-SEM in information systems research. *IMDS* 2017 Apr 10;117(3):442-458. [doi: [10.1108/imds-04-2016-0130](https://doi.org/10.1108/imds-04-2016-0130)]
18. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 2018 Nov 28;18(1):39-50. [doi: [10.1177/002224378101800104](https://doi.org/10.1177/002224378101800104)]
19. Hair JF, Ringle CM, Sarstedt M. Partial least squares: the better approach to structural equation modeling? *Long Range Planning* 2012 Oct;45(5-6):312-319. [doi: [10.1016/j.lrp.2012.09.011](https://doi.org/10.1016/j.lrp.2012.09.011)]
20. Bielby WT, Hauser RM. Structural equation models. *Annu Rev Sociol* 1977 Aug;3(1):137-161. [doi: [10.1146/annurev.so.03.080177.001033](https://doi.org/10.1146/annurev.so.03.080177.001033)]
21. Falk F, Miller N. *A primer for soft modelling*. Akron, OH: University of Akron Press; 1992.
22. Cohen J. *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
23. Henseler J. On the convergence of the partial least squares path modeling algorithm. *Comput Stat* 2009 Aug 1;25(1):107-120. [doi: [10.1007/s00180-009-0164-x](https://doi.org/10.1007/s00180-009-0164-x)]
24. Caesens G, Marique G, Stinglhamber F. The relationship between perceived organizational support and affective commitment. *J Pers Psychol* 2014 Jan;13(4):167-173. [doi: [10.1027/1866-5888/a000112](https://doi.org/10.1027/1866-5888/a000112)]
25. Watson J, Harper S, Ratliff L, Singleton S. Holistic wellness and perceived stress: predicting job satisfaction among beginning teachers. *Research In The Schools* 2010;17(1):29-37.
26. Orgambidez-Ramos A, Borrego-Alés Y, Mendoza-Sierra I. Role stress and work engagement as antecedents of job satisfaction in Spanish workers. *JTEM* 2014 Apr 07;10(1). [doi: [10.3926/jiem.992](https://doi.org/10.3926/jiem.992)]
27. Mache S, Bernburg M, Groneberg DA, Klapp BF, Danzer G. Work family conflict in its relations to perceived working situation and work engagement. *Work* 2016 Apr 15;53(4):859-869. [doi: [10.3233/wor-162257](https://doi.org/10.3233/wor-162257)]
28. Gandhi S, Sangeetha G, Ahmed N, Chaturvedi S. Somatic symptoms, perceived stress and perceived job satisfaction among nurses working in an Indian psychiatric hospital. *Asian J Psychiatr* 2014 Dec;12:77-81. [doi: [10.1016/j.ajp.2014.06.015](https://doi.org/10.1016/j.ajp.2014.06.015)] [Medline: [25440565](https://pubmed.ncbi.nlm.nih.gov/25440565/)]
29. Guppy A, Gutteridge T. Job satisfaction and occupational stress in UK general hospital nursing staff. *Work Stress* 1991 Oct;5(4):315-323. [doi: [10.1080/02678379108257029](https://doi.org/10.1080/02678379108257029)]
30. Koustelios A, Tsigilis N. The relationship between burnout and job satisfaction among physical education teachers: a multivariate approach. *Eur Phys Educ Rev* 2005;11(2):189-203. [doi: [10.1177/1356336x05052896](https://doi.org/10.1177/1356336x05052896)]
31. Laugaa D, Rasche N, Bruchon-Schweitzer M. Stress and burnout among French elementary school teachers: A transactional approach. *Eur Rev Soc Psychol* 2008 Dec;58(4):241-251. [doi: [10.1016/j.erap.2008.09.007](https://doi.org/10.1016/j.erap.2008.09.007)]
32. Howard S, Johnson B. Resilient teachers: resisting stress and burnout. *Soc Psychol Educ* 2004;7(4):399-420. [doi: [10.1007/s11218-004-0975-0](https://doi.org/10.1007/s11218-004-0975-0)]
33. Naylor C. *Teacher Workload and Stress: An International Perspective on Human Costs and Systemic Failure*. BCTF Research Report. British Columbia Teachers' Federation. 2001. URL: <http://files.eric.ed.gov/fulltext/ED464028.pdf> [accessed 2022-09-14]
34. Karasek R, Theorell T. *Healthy work: stress, productivity, and the reconstruction of working life*. New York, NY: Basic Books; 1990.
35. King A, Peart M. The satisfaction and stress of being a teacher. *Worklife Report* 1992;8(6):12-14.
36. Chu L. The influence of perceived stress on work-family conflict and mental health: the moderating effect of person-environment fit. *J Nurs Manag* 2014 Jul 04;22(5):613-620. [doi: [10.1111/jonm.12014](https://doi.org/10.1111/jonm.12014)] [Medline: [25041801](https://pubmed.ncbi.nlm.nih.gov/25041801/)]

Abbreviations

- AVE:** average variance extracted
- CFA:** confirmatory factor analysis
- CFI:** comparative fit index
- CR:** composite reliability
- FWC:** family-work conflict
- GFI:** goodness-of-fit index
- JD:** job dissatisfaction
- PS:** perceived stress
- TPEJSI-9:** 9-item Teacher of Physical Education Job Satisfaction Inventory (
- WDES:** Work Disengagement Scale
- WFC:** work-family conflict

Edited by A Mavragani; submitted 27.03.21; peer-reviewed by R Meng, S Choudhary; comments to author 18.04.22; revised version received 22.05.22; accepted 29.05.22; published 01.12.22.

Please cite as:

Chalghaf N, Chen W, Tannoubi A, Guelmami N, Puce L, Ben Said N, Ben Khalifa M, Azaiez F, Bragazzi NL

Job Disengagement Among Physical Education Teachers: Insights From a Cross-sectional Web-Based Survey With Path Modeling Analysis

JMIR Form Res 2022;6(12):e29130

URL: <https://formative.jmir.org/2022/12/e29130>

doi: [10.2196/29130](https://doi.org/10.2196/29130)

PMID: [36084318](https://pubmed.ncbi.nlm.nih.gov/36084318/)

©Nasr Chalghaf, Wen Chen, Amayra Tannoubi, Noomen Guelmami, Luca Puce, Noureddine Ben Said, Maher Ben Khalifa, Fairouz Azaiez, Nicola Luigi Bragazzi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 01.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Perceptions and Aspirations Toward Peer Mentoring in Social Media–Based Electronic Cigarette Cessation Interventions for Adolescents and Young Adults: Focus Group Study

Joanne Chen Lyu¹, PhD; Aliyyat Afolabi¹, BSc; Justin S White^{1,2}, PhD; Pamela M Ling¹, MPH, MD

¹Center for Tobacco Control Research and Education, University of California, San Francisco, San Francisco, CA, United States

²Philip R Lee Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:

Joanne Chen Lyu, PhD

Center for Tobacco Control Research and Education

University of California, San Francisco

530 Parnassus Ave

San Francisco, CA, 94143-1390

United States

Phone: 1 415 502 4181

Email: chenjoanne.lyu@ucsf.edu

Abstract

Background: Social media offer a promising channel to deliver e-cigarette cessation interventions to adolescents and young adults (AYAs); however, interventions delivered on social media face challenges of low participant retention and decreased engagement over time. Peer mentoring has the potential to ameliorate these challenges.

Objective: The aim of this study was to understand, from both the mentee and potential mentor perspective, the needs, expectations, and concerns of AYAs regarding peer mentoring to inform the development of social media–based peer mentoring interventions for e-cigarette cessation among AYAs.

Methods: Seven focus groups, including four mentee groups and three potential mentor groups, were conducted with 26 AYAs who had prior experience with e-cigarette use and attempts to quit in the context of a social media–based e-cigarette cessation intervention. Discussion focused on preferred characteristics of peer mentors, expectations about peer mentoring, mentoring mode, mentor training, incentives for peer mentors, preferred social media platforms for intervention delivery, supervision, and concerns. Focus group transcripts were coded and analyzed using a thematic analysis approach.

Results: Overall, participants were receptive to peer mentoring in social media–based cessation interventions and believed they could be helpful in assisting e-cigarette cessation. Participants identified the most important characteristics of peer mentors to be of similar age and to be abstinent from e-cigarette use. Participants expected peer mentors would share personal experiences, provide emotional support, and send check-ins and reminders. Peer mentors supporting a group of mentees in combination with one-on-one mentoring as needed was the preferred mentoring mode. A group of 10 mentees with a mentor:mentee ratio of 1:3-5 was deemed acceptable for most participants. Participants expressed that mentor training should include emotional intelligence, communication skills, and the scientific evidence about e-cigarettes. Although monetary incentives were not the main motivating factor for being a peer mentor, they were viewed as a good way to compensate mentors' time. Instagram was considered an appropriate social media platform to deliver a peer-mentored intervention due to its functionality. Participants did not express many privacy concerns about social media–based peer mentoring, but mentioned that boundaries and community agreements should be set to keep relationships professional.

Conclusions: This study reflects the needs and preferences of young people for a peer mentoring intervention to complement a social media program to support e-cigarette cessation. The next step will be to establish the feasibility, acceptability, and preliminary efficacy of such a peer mentoring program.

(*JMIR Form Res* 2022;6(12):e42538) doi:[10.2196/42538](https://doi.org/10.2196/42538)

KEYWORDS

peer mentoring; electronic nicotine delivery systems; cessation; social media; adolescents and young adults

Introduction

Since 2014, electronic nicotine delivery systems (or “e-cigarettes”) have been the most commonly used tobacco product among US adolescents and young adults (AYAs), whose interest in these products continues to grow [1]. From 2017 to 2018, current e-cigarette use increased by 46.2% (5.2% to 7.6%) among young adults [2], 77.8% (from 11.7% to 20.8%) among high school students, and 48.5% (from 3.3% to 4.9%) among middle school students [3]. Despite a decline in e-cigarette use among young people during the COVID-19 pandemic [4], approximately 25% of adolescents (aged 15-17 years) and young adults (aged 18-24 years) reported current e-cigarette use in May 2020 [5]. Any tobacco use, including e-cigarettes, poses risks to youth and young adults. Compared with those of older adults, the brains of AYAs are more vulnerable to the harmful health effects of nicotine, because development and maturation of the prefrontal cortex occur primarily during adolescence and are fully accomplished at the age of 25 years [6]. Potential risks include nicotine addiction, reduced impulse control, mood disorders, and poor attention and thinking skills [1,7]. A report from the Truth Longitudinal Cohort, a national probability-based survey, found that among current e-cigarette users aged 15-36 years, 54.2% reported a general intention to quit and 33.3% reported a past-year quit attempt [8], suggesting that there would be interest in cessation interventions for young people.

Social media are widely used by young Americans, with 84% of adults aged 18 to 29 years and 95% of adolescents ever using one or more social media platforms [9,10]. Young people’s reliance on social media to pass time, connect with others, and learn new things [11] highlights these platforms as promising channels to deliver interventions to this group. Although few social media-based interventions addressing AYA e-cigarette use have been conducted to date [12], smoking cessation interventions delivered on social media have demonstrated feasibility, acceptability, and early efficacy [13,14]. A systematic review found that five of seven social media interventions for smoking cessation increased quit attempts and abstinence, and reduced relapse [13]. Despite the great potential for social media as a platform for intervention delivery, tobacco treatment and other behavioral interventions delivered on social media have shown low participant retention and declining engagement over time [15,16]. There is a need for strategies that can overcome these obstacles to improve intervention outcomes.

A promising finding from prior research is that having a human support component could increase participant engagement in online interventions. For instance, one study found that when participants received encouragement and perceived social support on Facebook, they engaged more deeply in cessation interventions [17]. In addition, peer mentors are a particularly promising source of support, because people feel more comfortable with those who are like them [18]. Integrating social support from peers into social media interventions has the potential to improve participant engagement and the efficacy of social media-based interventions. Although widely adopted in health behavior change interventions such as weight management and addiction recovery [19-22], and implemented in some smoking cessation programs [23,24], peer mentoring

has not been studied in social media-based tobacco interventions. Therefore, the aim of this study was to understand (1) the needs and expectations for peer mentoring of AYAs who had prior experience with e-cigarette use and trying to quit (ie, the mentee perspective), and (2) the preferences and ideas of those who had interest in becoming a peer mentor to help others quit using e-cigarettes (ie, the mentor perspective). Such evidence could inform the development of peer mentoring programs for AYA e-cigarette cessation interventions delivered on social media platforms.

Methods

Focus Group Guide Development

The focus group discussion guide was adapted from the iQuit program (led by coauthor JSW), a peer mentoring program designed for a text message-based smoking cessation intervention for adult smokers [24]. Through an iterative process, the content was adapted to address young people, e-cigarettes, and interventions delivered on social media. An expert on qualitative research was consulted and provided feedback on the draft discussion guide. The final version of the guide contained seven topics: (1) characteristics of peer mentors, (2) expectations about peer mentoring, (3) mentoring mode, (4) training for peer mentors, (5) incentives for peer mentors, (6) preferred social media platforms for intervention delivery, and (7) supervision and concerns. Since the aim of the study was to understand both the mentee perspective and mentor perspective, we organized two types of focus groups: one for mentees and one for potential mentors. The wording of questions was slightly modified for each group.

Participant Recruitment

We recruited participants with the support of our media partner who has been running our team’s Quit the Hit (QTH) program in California, South Carolina, and Minnesota. The QTH program is an Instagram-based e-cigarette cessation intervention for AYAs. Participants receive up to 3 posts per weekday for 5 weeks in Instagram groups. The post content incorporated motivational interviewing, cognitive behavioral coping skills, and the transtheoretical model of behavior change [25-27], and utilized images, videos, and text designed to reflect the vaping experience of young people and elicit participation. Recruitment messages were sent via emails, text messages, and/or Instagram direct messages to QTH current participants and graduates who had indicated in previous surveys a willingness to participate in future studies. Given this, all participants had prior experience with e-cigarette use and trying to quit in the Instagram support groups. Any individuals from the QTH program were eligible to participate in the mentee group, regardless of whether or not they were currently using e-cigarettes. Those eligible for the mentor group had quit using e-cigarettes for at least 1 month and had at least some interest in being a peer mentor. We recruited 26 participants in total and conducted 3 mentee groups and 4 mentor groups.

Focus Group Procedures

Seven focus groups were conducted via web conference in February-June 2022. In addition to the discussion, participants

completed a short questionnaire on their demographic characteristics. Two to three trained moderators facilitated each group. Within each group, one moderator was responsible for asking the questions from the discussion guide. Order and the specific language of the questions varied slightly from the guide to ensure a natural and smooth flow to the conversation. The other moderator(s) was responsible for asking follow-up questions and encouraging participants to elaborate on specific responses. Moderators encouraged less active participants to participate in the conversation. Focus groups lasted 75 to 90 minutes. The number of participants per group ranged from 2 to 9. Discussions were video-recorded and professionally transcribed.

Data Analysis

We used a thematic analysis approach [28] to analyze participant responses to questions in Dedoose, a widely used qualitative data management software. Two authors (JCL and AA) developed the initial coding guide and independently coded one randomly selected transcript from the mentor group and one from the mentee group, and then resolved any coding discrepancies through discussion. A final version of the coding guide was developed from these discussions. One coder (AA) coded the remaining transcripts in accordance with the final coding guide. Major themes for each topic were categorized through research team discussion. All authors in the research team contributed to the selection of quoted responses from participants to ensure the quotes were representative of their corresponding themes. Analytic memos were written to

characterize emerging themes that did not fall into any of the seven topics and were synthesized in the results.

Ethical Considerations

This study was approved by the WCG Institutional Review Board (IRB), an independent IRB that partners with more than 3300 institutions ranging from small research sites to large academic medical centers and universities (IRB tracking number 20204627). Online informed consent was obtained before participants joined in the focus group. Participants were asked not to disclose any information regarding the focus group to others. The first author deidentified all of the transcripts before uploading them to the data management software Dedoose for analysis. Each participant received a US \$60 gift card for participation in the focus group.

Results

Participant Characteristics

Of the 26 participants, 16 participated in the mentee groups and 10 in the mentor groups. Participants had a mean age of 19.4 years and 50% were women; over half self-identified as lesbian, gay, bisexual, transgender, queer, or other sexual orientation (LGBTQ+) and as non-Hispanic white; 39% reported that they lived comfortably. A large majority of participants (92%) believed that peer mentoring would help them to quit e-cigarettes. Among the participants, 35% had not quit e-cigarettes when interviewed. Details of the participant characteristics are listed in [Table 1](#).

Table 1. Focus group participant characteristics (N=26).

Characteristics	Value
Demographics	
Age (years), mean (SD)	19.4 (2.8)
Sex at birth, n (%)	
Male	13 (50)
Female	13 (50)
Sexual orientation, n (%)	
Heterosexual	12 (46)
LGBTQ+ ^a	14 (54)
Race/ethnicity, n (%)	
NH ^b white	15 (58)
NH Black	3 (12)
NH Asian	2 (8)
Hispanic	2 (8)
Other/multirace	4 (15)
Financial situation, n (%)	
Live comfortably	10 (39)
Meet needs with a little left	6 (23)
Just meet basic expenses	7 (27)
Don't meet basic expenses	2 (8)
Attitude toward peer mentoring: peer mentoring will help with quitting, n (%)	
Disagree a lot	0 (0)
Disagree a little	0 (0)
Not sure	2 (8)
Agree a little	8 (31)
Agree a lot	16 (62)
Quit status, n (%)	
Not yet	9 (35)
Yes, <1 month	1 (4)
Yes, >1 month but <3 months	7 (27)
Yes, >3 months	9 (35)

^aLGBTQ+: lesbian, gay, bisexual, transgender, queer, or other sexual orientation.

^bNH: non-Hispanic.

Characteristics of Peer Mentors

Most participants expressed that similar age and e-cigarette abstinence were the two most important characteristics for a mentor to have. They believed that mentors of the same age would be easier to talk to and connect with, and would better relate to participants' e-cigarette experiences. They especially disliked having significantly younger mentors because they lacked "the same life experience" to understand the participants' struggles. Most participants preferred older mentors but indicated that they should be no more than 10 years older (and less than 5 years older preferred). Despite this preference, no participants objected to mentoring younger mentees (although

a gap of less than 5 years was preferred). Moreover, they believed it would be "very beneficial for them" because they could provide wisdom and advice to younger mentees.

Most participants emphasized the importance of having mentors who had already achieved e-cigarette abstinence. Otherwise, participants would view the mentoring as less credible and less legitimate. In discussing mentors who had not yet quit using e-cigarettes, one participant stated:

I'd feel like "oh you might as well just be someone else in the group." Why should I take your advice?

Because clearly your advice hasn't worked. Because you're still vaping. [male, age 18 years]

Although a few participants expressed it would be “nice” to have a mentor of the same gender, most participants viewed concordance of characteristics such as gender, race, and sexual orientation as less important as long as the group was an inclusive and friendly space.

Some LGBTQ+ participants also mentioned that having an LGBTQ+-friendly group would be welcoming for members of this community. Agreeing with this comment, one participant from the LGBTQ+ group stated: “I'd probably want to be mentoring an LGBTQ group, because I can easily relate to them...It's just easier to interact with your own community” [male, age 21 years]

Expectations About Peer Mentoring

Sharing personal experiences, emotional support, and check-ins/reminders were three items that participants expected from a peer mentor. Most participants mentioned that the mentor's prior personal experience with e-cigarette use and trying to quit would differentiate peer mentors from a professional counselor; thus, peer mentoring can be a model to motivate others. Sharing personal experiences with e-cigarettes, including challenges with e-cigarette cessation and quit methods, would give those who were trying to quit a sense of credibility that the mentors had undertaken the same journey as them, were knowledgeable on the topic of e-cigarette cessation, and could provide useful guidance.

Most participants considered emotional support as essential to peer mentoring. For example, “I think that [emotional support]'s probably the biggest factor that I've found helped me. Just being in a group where everybody's really open and the mentors facilitating that.” [female, age 20 years]

They expected mentors not only to be online but also to be available emotionally, and hoped mentors would be people who they could talk to when having a bad day and who would not shame them for failure to quit. The key words participants used to describe an ideal peer mentor were youthful, compassionate, understanding, relatable, emotionally intelligent, open-minded, optimistic, passionate, supportive, kind, gentle, peaceful, friendly, LGBTQ-friendly, respectful, trustworthy, accountable, experienced, knowledgeable, perspective, persistent, and consistent.

Many participants expected both to be able to check in with peer mentors and to receive reminders from them. They believed that would allow both the mentors and themselves to see their progress and keep on track with their goal of cessation. Many also stated that the check-ins could initiate conversation in group chats and allow the group to share more personal issues with each other, open up about their struggles, and receive helpful tips and advice from other mentees. However, participants disagreed on the appropriate frequency of check-ins. Some suggested 2-3 times a day; others felt daily check-ins were “overwhelming” because check-in messages would constantly remind them of their struggle and how they need to quit. Such participants preferred receiving messages 2-3 times a week.

Mentoring Mode

Most participants preferred that peer mentors would post questions and give advice in the group (ie, team mentoring) rather than one-on-one direct messages seen only by each individual participant. Participants reported that this type of team mentoring “[would be] helpful [for] seeing how other people are dealing with things and to all support each other through it.” It was also a nice reminder that they were not going through this process alone. At the same time, many participants also wanted the opportunity to have one-on-one interactions with peer mentors because it would make the process more personalized and would not overwhelm the group.

The majority of participants expressed a dislike of large groups, such as a 15-people group, and many mentioned that in small groups it might be easier to probe for stressors that contribute to their continued e-cigarette use. However, they also worried that small groups would easily fall silent if everyone was not actively posting. Although participants varied in their opinions of the ideal group size, most agreed that 10 mentees per group would achieve a good tradeoff. Three to five mentees per mentor was the modal preferred ratio from the perspectives of both mentors and mentees. One participant stated: “I feel like five for me would be the best numbers, like not too little, not too much, just the right amount where I feel like I would be able to keep in touch with all of them” [female, age 16 years].

As to the mentee-mentor assignment, many participants agreed that mentees should not be assigned to a mentor who they already knew in order to keep the cessation process confidential and professional. A few participants preferred same-gender mentors. However, most of the participants were fine with random assignment.

I don't see an issue with it being random, because I feel like we can all find some level to identify with each other since we're the same age and we're going through similar experiences. So, I feel like the small details wouldn't matter as much [male, age 17 years]

Training for Peer Mentors

Participants mentioned three types of training needed for a mentor to be competent: emotional intelligence training, communication skills training, and e-cigarette science training. Participants further emphasized emotional intelligence as essential to having a successful program and determining the appropriate way to help mentees.

So, if somebody needs tough love, you should be able to give them tough love. If somebody needs emotional support, to just coddle them in a way, you should be able to understand that and do that for them [male, age 18 years]

Some participants suggested that sensitivity training be included as part of the emotional intelligence training to help foster a safe space of inclusivity in the group chat. One participant stated:

There should be some sensitivity training. I think it would be pretty off-putting and probably harmful if your mentor said something that was, for example,

discriminatory or insensitive based on the way you identify in any type of way [female, age 23 years]

While talking about emotional intelligence, many participants mentioned that mentors should be trained in communication skills, including active listening and communicating in a nonjudgmental way. Participants expressed different preferences for e-cigarette science training. Most participants in the mentor group liked the idea of having science training in which mentors learned basic knowledge related to nicotine addiction, harmful effects of e-cigarette use, and relevant scientific findings to share the evidence with others. In contrast, many participants in the mentee group did not express much enthusiasm for scientific training. Instead, some participants reported they would feel “bad,” “stressful,” or “overwhelmed” if their peer mentors talked about topics they would expect from a medical professional.

It's nice to not get things from a medical point of view because then it's, for me, that's really stressful and I start to panic...that tends to be more of, “Why are you doing this to your body?” And then, you feel bad even though you already feel bad because you don't want to be doing it [female, age 23 years]

Incentives for Peer Mentors

Generally, the potential mentors expressed that they were self-motivated to help others quit using e-cigarettes because of the struggle they had experienced in attempting to quit and the “pay it forward” mindset they had cultivated during their journey toward achieving abstinence.

I would say if you know the struggle of addiction, whether it's vaping or something else, that's kind of my motivation. It's just, addiction's never a good thing. So, I would like to help people not be addicted to anything [male, age 17 years]

Although many participants mentioned money would not be the main motivating factor for becoming a peer mentor, potential mentors and mentees agreed that receiving money would be a good way to compensate mentoring time. Most participants agreed that being well compensated would allow mentors to take their role more seriously. Some participants in the potential mentor role noted that although there might be scheduled hours in the day to respond, a mentor would need to be flexible if a mentee needed them after these hours; thus, the role of the peer mentor could be emotionally taxing and important. When it came to the incentive amount, there was a heated discussion in one mentor group. Most participants tended to estimate the appropriate amount factoring in the time commitment, hourly wage, and the difference of this online mentoring from a job requiring physical presence. Taking a 5-week program in which peer mentoring constituted half an hour per day, despite varying answers ranging from US \$100 to \$350, most participants believed US \$200-250 for the whole program to be reasonable.

A letter of recommendation and leadership certificate were considered attractive incentives as well. A few participants would like to become a mentor to build leadership, communication, and counseling skills that might be helpful to have on a resume and in their future career.

Preferred Social Media Platforms for Intervention Delivery

Participants were asked to openly comment on social media platforms they have used that might be appropriate for delivery of a peer-mentored vaping cessation intervention. They mentioned many social media platforms such as Snapchat, TikTok, Facebook, and Twitter, and expressed a preference for Instagram over other platforms. Many participants mentioned Instagram as a professional platform where they can communicate with mentors and mentees. In contrast, many participants said they would feel uncomfortable adding their mentors or mentees on Snapchat because they share personal things there for close friends to see. TikTok was viewed as too casual:

I use TikTok and my Tik group is full of random, funny things. I don't think I could take the group chat seriously if it was on my TikTok. Instagram just seems like the most professional and efficient out of all of it [female, age 16 years]

Some participants mentioned they barely used Facebook and Twitter, and felt “nobody my age really uses them.” They observed that most young people already have Instagram on their phones, which would make it easy and convenient to receive program messages. Participants also stated that they like many of Instagram's tools such as group messaging, direct messaging, audio recordings, and video and photo sharing that can make Instagram content funny and engaging. Instagram Live was also recommended by some participants to host live streams during which mentors could share their journey to cessation and respond to questions synchronously.

Supervision and Concerns

Participants generally expressed trust and comfort with offering the peer mentoring on social media and did not express much concern, provided that mentors and mentees behaved respectfully such as no sexual comments and no attacks on race, religious belief, or physical features. Most participants were open to using their social media account rather than a business account, either in the mentor or mentee role, because it would be able to provide depth to the person. They were also comfortable with being directly messaged as long as the volume of messages was not overwhelming. Maintaining contact after the program ended would not be a problem for many who believed it beneficial to have a support system outside their own family and friends.

While showing openness, most participants mentioned the need to set boundaries to keep the mentoring “professional.” They wanted the direct messaging to focus on e-cigarette cessation rather than personal details unless they related to e-cigarette use. Many participants stated that there should be a designated time of day when mentors and mentees should interact. Although the appropriate time varied by person, most participants favored the 10 AM-10 PM window. All participants opposed any romantic relationships between a mentee and mentor:

Don't date someone in the group, especially as a mentor. Don't make any inappropriate advances. I don't think topics like that should be talked about in

the group...It would just be harmful for the group as a whole [female, age 24 years]

Discussion

Principal Findings

Peer mentoring has the potential to improve decreased engagement over time that is a critical barrier to successful social media interventions. This study is among the first to explore peer mentoring in social media-based interventions among AYAs. We conducted focus groups with AYAs with prior experience of e-cigarette use and attempts to quit in the context of a social media-based e-cigarette cessation intervention. Through this, we aim to understand, from both the mentee and potential mentor perspective, the needs, expectations, and concerns of AYAs regarding peer mentoring. We identified several key features of peer mentoring that can inform the development of social media-based peer mentoring interventions for e-cigarette cessation among AYAs.

In general, participants indicated high receptivity to the peer mentoring concept and believed that it could be helpful for facilitating e-cigarette abstinence. The opinions expressed by mentors and mentees were similar on this point. As to the mentor characteristics, although participants were generally open to other sociodemographic characteristics, they almost unanimously expressed that the mentors should be of similar age and a little bit older, and believed the same life experience of similar age would help them to more easily relate to each other. A large majority of participants also mentioned the importance of mentors' e-cigarette abstinence. Younger age and having not achieved e-cigarette cessation would decrease mentor credibility. Therefore, the findings of this study indicate that future development of a peer mentoring program for social media interventions with AYAs should take age and abstinence status into consideration. A few LGBTQ+ participants expressed interest in mentoring and being mentored by members of the same community. Sexual and gender minority young adults may require more tobacco cessation support compared to their peers [29], and prefer tailored interventions to address the particular stressors associated with tobacco use among the LGBTQ+ group, such as internalized stigma, prejudice, and discrimination [30,31]. This may indicate that peer mentoring may be especially needed for this group and deserves more exploration.

The core of peer support is to provide social and/or emotional support that combines expertise from lived experience to assist the support recipient to make a change [32,33]. This is consistent with our participants' expectations about peer mentoring in e-cigarette cessation interventions. They valued the peer mentors' ability to share personal experiences and believed that the mentors' prior experiences with e-cigarette and e-cigarette abstinence, and corresponding experiential knowledge, would facilitate others to quit using e-cigarettes. Having similar struggles with quitting as the mentees' differentiates peer mentors from professional counselors and makes the treatment-seeking AYAs feel a stronger bond. Participants also expected to receive emotional support from peer mentors, and some considered emotional support a key to successful peer

mentoring. These expectations about peer mentoring were also reflected in the training that participants believed a mentor should have. For them, scientific knowledge about e-cigarettes was not the most important part of mentor training; instead, peer mentor training should place emphasis on emotional intelligence and communication skills. One reason for this finding may be because our participants came from the QTH program where evidence-based quitting strategies were provided by a professional counselor; thus, the participants did not need additional scientific information from peer mentors unless they had personal expertise. Another possible reason is that many participants believed they had adequate exposure to e-cigarette science, and too much scientific information from peer mentors would increase the mentee's stress. In other words, the perceived needed training for mentors reflected a need for emotional support rather than informational support. Developing an effective peer training curriculum will require consideration of how mentors can provide emotional support together with insights from their personal experiences to facilitate e-cigarette cessation.

While participants showed sincerity, openness, and desire for connectedness to the peer-mentored social media intervention, they also clearly expressed the need for boundaries and moderation in communication. Most participants did not express concern about privacy disclosure, and they preferred using a personal account rather than a business account for the program. They also preferred having team mentoring so that everyone can benefit from the group discussion and sharing from others. However, while proposing the group size and ratio of mentees to mentors not to be too big to hinder rapport-building or overloading the mentors, they also disliked groups that are too small (<3 mentees per mentor), which would increase their pressure to be active contributors. Similarly, while one-one-one mentoring and direct messaging were acceptable for most participants, they wanted the communication to be professional and focused on e-cigarette use rather than unrelated personal topics. Under the umbrella of being professional, they called for establishing guidelines for contact frequency and scheduling allowable times of day for contact. Although there were large individual differences, the consensus was that overly frequent contact could backfire, such as leading participants to turn off notifications or even leave the group. The balance between being connected and being professional may be a critical challenge for peer mentoring, especially for the mentoring embedded in social media-based interventions in which participants do not have the interpersonal relationship pressure that is found using in-person interventions and thus can more easily drop out [34].

This study suggested that Instagram is a promising and well-accepted platform for delivering peer-mentored e-cigarette cessation support to AYAs. Instagram has already been an integral part of the daily lives of many young people; therefore, an Instagram-based intervention and the add-on peer mentoring would be highly accessible and easily delivered in a way that is familiar, comfortable, and convenient to AYAs. In addition, many Instagram functions can support peer mentoring. For instance, the group messaging function of Instagram makes team mentoring feasible so that intervention participants can

benefit from the group chat and sharing; direct messaging allows one-on-one mentoring to discuss relatively personal issues without involving other participants in the group. Therefore, Instagram can easily accommodate a need for a combination of team mentoring and one-on-one mentoring. However, it should be noted that all participants of the focus groups had already signed up for Instagram groups, and therefore this sample likely has a disproportionate affinity for Instagram. Therefore, our participants' preference for Instagram may not be generalized to the conclusion that Instagram is the best platform in general.

Potential mentors generally felt that mentoring on social media would not be burdensome as they are frequently online. While monetary compensation was felt to be a fair compensation for time, many participants mentioned that due to their own experiences receiving help from others, they were quite self-motivated to provide mentoring to help others quit. Social media interventions are not confined to geographic restriction and may also assuage privacy concerns that could inhibit participation in face-to-face groups [29]. Motivating those who have successfully quit using e-cigarettes to become a peer mentor is a critical factor to scale up peer-mentored social media interventions in order to reach a large number of e-cigarette users.

Limitations

There were several limitations of this study. First, conducting the focus groups remotely by web conference does not provide

the same communicative cues (eg, eye contact) as an in-person session. Second, although research staff invited a similar number of AYAs to each focus group and followed a standardized protocol to send reminders, the number of participants in each focus group varied. Varied group size may influence the group dynamics and interactive discussion among participants. Third, most of the focus group participants showed interest in being a peer mentor. While this indicates the high promise of recruiting AYA mentors to help their peers quit using e-cigarettes, we may have limited insight into how to motivate unwilling quitters to become a peer mentor.

Conclusions

AYAs who had completed a social media support group intervention were very receptive to incorporating peer mentoring into e-cigarette cessation interventions on social media. The participants reached a consensus on many key features of peer mentoring, providing important information for the development of a peer mentoring program to complement social media-based e-cigarette cessation interventions. Developing such a peer-mentored program and testing its feasibility and preliminary efficacy is a logical next step toward leveraging peer mentoring to improve social media-based e-cigarette cessation intervention outcomes.

Acknowledgments

This study was funded by California Tobacco Related Diseases Research Program (T31IR1910) and the UCSF Helen Diller Comprehensive Cancer Center SF CAN program. JCL was supported by the National Cancer Institute (F32CA265059).

Data Availability

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. National Center for Chronic Disease Prevention and Health Promotion (U.S.). Office on Smoking and Health. E-Cigarette use among youth and young adults: a report of the Surgeon General. US Department of Health and Human Services. 2016. URL: https://e-cigarettes.surgeongeneral.gov/documents/2016_sgr_full_report_non-508.pdf [accessed 2022-11-19]
2. Dai H, Leventhal AM. Prevalence of e-Cigarette use among adults in the United States, 2014-2018. JAMA 2019 Nov 12;322(18):1824-1827 [FREE Full text] [doi: [10.1001/jama.2019.15331](https://doi.org/10.1001/jama.2019.15331)] [Medline: [31524940](https://pubmed.ncbi.nlm.nih.gov/31524940/)]
3. Gentzke AS, Creamer M, Cullen KA, Ambrose BK, Willis G, Jamal A, et al. Vital signs: tobacco product use among middle and high school students - United States, 2011-2018. MMWR Morb Mortal Wkly Rep 2019 Feb 15;68(6):157-164. [doi: [10.15585/mmwr.mm6806e1](https://doi.org/10.15585/mmwr.mm6806e1)] [Medline: [30763302](https://pubmed.ncbi.nlm.nih.gov/30763302/)]
4. Gaiha SM, Lempert LK, Halpern-Felsher B. Underage youth and young adult e-Cigarette use and access before and during the coronavirus disease 2019 pandemic. JAMA Netw Open 2020 Dec 01;3(12):e2027572 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.27572](https://doi.org/10.1001/jamanetworkopen.2020.27572)] [Medline: [33270127](https://pubmed.ncbi.nlm.nih.gov/33270127/)]
5. Kreslake JM, Simard BJ, O'Connor KM, Patel M, Vallone DM, Hair EC. E-Cigarette use among youths and young adults during the COVID-19 pandemic: United States, 2020. Am J Public Health 2021 Jun;111(6):1132-1140. [doi: [10.2105/AJPH.2021.306210](https://doi.org/10.2105/AJPH.2021.306210)] [Medline: [33856888](https://pubmed.ncbi.nlm.nih.gov/33856888/)]
6. Arain M, Haque M, Johal L, Mathur P, Nel W, Rais A, et al. Maturation of the adolescent brain. Neuropsychiatr Dis Treat 2013;9:449-461. [doi: [10.2147/NDT.S39776](https://doi.org/10.2147/NDT.S39776)] [Medline: [23579318](https://pubmed.ncbi.nlm.nih.gov/23579318/)]

7. Talking to youth and young adults to prevent e-cigarette use. Centers for Disease Control and Prevention (CDC). 2021 Sep 20. URL: <https://www.cdc.gov/tobacco/features/back-to-school/index.html> [accessed 2022-06-17]
8. Cuccia A, Patel M, Amato M, Stephens D, Yoon S, Vallone D. Quitting e-cigarettes: quit attempts and quit intentions among youth and young adults. *Prev Med Rep* 2021 Mar;21:101287 [FREE Full text] [doi: [10.1016/j.pmedr.2020.101287](https://doi.org/10.1016/j.pmedr.2020.101287)] [Medline: [33489721](https://pubmed.ncbi.nlm.nih.gov/33489721/)]
9. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/> [accessed 2022-06-17]
10. Vogel E, Gelles-Watnick R, Mssarat N. Teens, social media and technology. Pew Research Center. 2022 Sep 10. URL: <https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/> [accessed 2022-09-05]
11. Schaeffer K. Most U.S. teens who use cellphones do it to pass time, connect with others, learn new things. Pew Research Center. 2019 Aug 23. URL: <https://www.pewresearch.org/fact-tank/2019/08/23/most-u-s-teens-who-use-cellphones-do-it-to-pass-time-connect-with-others-learn-new-things/> [accessed 2022-06-16]
12. Quit the hit: a ready-made, Instagram-based cessation program to help young people quit vaping. Rescue Agency. URL: <https://rescueagency.com/ready-made/quit-the-hit> [accessed 2022-11-21]
13. Naslund JA, Kim SJ, Aschbrenner KA, McCulloch LJ, Brunette MF, Dallery J, et al. Systematic review of social media interventions for smoking cessation. *Addict Behav* 2017 Oct;73:81-93 [FREE Full text] [doi: [10.1016/j.addbeh.2017.05.002](https://doi.org/10.1016/j.addbeh.2017.05.002)] [Medline: [28499259](https://pubmed.ncbi.nlm.nih.gov/28499259/)]
14. Thurl J, Tormohlen KN, Meacham MC. Social media for tobacco smoking cessation intervention: a review of the literature. *Curr Addict Rep* 2019 Jun;6(2):126-138 [FREE Full text] [Medline: [31485388](https://pubmed.ncbi.nlm.nih.gov/31485388/)]
15. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelandotte C. Are health behavior change interventions that use online social networks effective? A systematic review. *J Med Internet Res* 2014 Feb 14;16(2):e40 [FREE Full text] [doi: [10.2196/jmir.2952](https://doi.org/10.2196/jmir.2952)] [Medline: [24550083](https://pubmed.ncbi.nlm.nih.gov/24550083/)]
16. Thurl 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](https://doi.org/10.2196/jmir.4575)] [Medline: [26561529](https://pubmed.ncbi.nlm.nih.gov/26561529/)]
17. Kim SJ, Marsch LA, Brunette MF, Dallery J. Harnessing Facebook for smoking reduction and cessation interventions: Facebook user engagement and social support predict smoking reduction. *J Med Internet Res* 2017 May 23;19(5):e168 [FREE Full text] [doi: [10.2196/jmir.6681](https://doi.org/10.2196/jmir.6681)] [Medline: [28536096](https://pubmed.ncbi.nlm.nih.gov/28536096/)]
18. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res* 2011 Mar 10;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
19. Hibbert C, Trottier E, Boville M, Hahn M, Hernandez A, Siddiqui A. The effect of peer support on knowledge and self-efficacy in weight management: a prospective clinical trial in a mental health setting. *Community Ment Health J* 2021 Jul 25;57(5):979-984. [doi: [10.1007/s10597-020-00703-7](https://doi.org/10.1007/s10597-020-00703-7)] [Medline: [32840689](https://pubmed.ncbi.nlm.nih.gov/32840689/)]
20. Gagne CA, Finch WL, Myrick KJ, Davis LM. Peer workers in the behavioral and integrated health workforce: opportunities and future directions. *Am J Prev Med* 2018 Jun;54(6 Suppl 3):S258-S266 [FREE Full text] [doi: [10.1016/j.amepre.2018.03.010](https://doi.org/10.1016/j.amepre.2018.03.010)] [Medline: [29779550](https://pubmed.ncbi.nlm.nih.gov/29779550/)]
21. Long JA, Jahnle EC, Richardson DM, Loewenstein G, Volpp KG. Peer mentoring and financial incentives to improve glucose control in African American veterans: a randomized trial. *Ann Intern Med* 2012 Mar 20;156(6):416-424 [FREE Full text] [doi: [10.7326/0003-4819-156-6-201203200-00004](https://doi.org/10.7326/0003-4819-156-6-201203200-00004)] [Medline: [22431674](https://pubmed.ncbi.nlm.nih.gov/22431674/)]
22. Tracy K, Burton M, Miescher A, Galanter M, Babuscio T, Frankforter T, et al. Mentorship for Alcohol Problems (MAP): a peer to peer modular intervention for outpatients. *Alcohol Alcohol* 2012;47(1):42-47 [FREE Full text] [doi: [10.1093/alcalc/agr136](https://doi.org/10.1093/alcalc/agr136)] [Medline: [22045903](https://pubmed.ncbi.nlm.nih.gov/22045903/)]
23. Ford P, Clifford A, Gussy K, Gartner C. A systematic review of peer-support programs for smoking cessation in disadvantaged groups. *Int J Environ Res Public Health* 2013 Oct 28;10(11):5507-5522 [FREE Full text] [doi: [10.3390/ijerph10115507](https://doi.org/10.3390/ijerph10115507)] [Medline: [24169412](https://pubmed.ncbi.nlm.nih.gov/24169412/)]
24. White JS, Toussaert S, Thurl J, Bontemps-Jones J, Abrams L, Westmaas JL. Peer mentoring and automated text messages for smoking cessation: a randomized pilot trial. *Nicotine Tob Res* 2020 Mar 16;22(3):371-380. [doi: [10.1093/ntr/ntz047](https://doi.org/10.1093/ntr/ntz047)] [Medline: [30892616](https://pubmed.ncbi.nlm.nih.gov/30892616/)]
25. Hall SM, Humfleet GL, Muñoz RF, Reus VI, Prochaska JJ, Robbins JA. Using extended cognitive behavioral treatment and medication to treat dependent smokers. *Am J Public Health* 2011 Dec;101(12):2349-2356. [doi: [10.2105/ajph.2010.300084](https://doi.org/10.2105/ajph.2010.300084)]
26. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol* 1983 Jun;51(3):390-395. [doi: [10.1037//0022-006x.51.3.390](https://doi.org/10.1037//0022-006x.51.3.390)] [Medline: [6863699](https://pubmed.ncbi.nlm.nih.gov/6863699/)]
27. DiClemente CC, Prochaska JO, Fairhurst SK, Velicer WF, Velasquez MM, Rossi JS. The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. *J Consult Clin Psychol* 1991 Apr;59(2):295-304. [doi: [10.1037//0022-006x.59.2.295](https://doi.org/10.1037//0022-006x.59.2.295)] [Medline: [2030191](https://pubmed.ncbi.nlm.nih.gov/2030191/)]
28. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]

29. Vogel EA, Thrul J, Humfleet GL, Delucchi KL, Ramo DE. Smoking cessation intervention trial outcomes for sexual and gender minority young adults. *Health Psychol* 2019 Jan;38(1):12-20 [FREE Full text] [doi: [10.1037/hea0000698](https://doi.org/10.1037/hea0000698)] [Medline: [30489104](https://pubmed.ncbi.nlm.nih.gov/30489104/)]
30. Blossnich JR, Horn K. Associations of discrimination and violence with smoking among emerging adults: differences by gender and sexual orientation. *Nicotine Tob Res* 2011 Dec 12;13(12):1284-1295 [FREE Full text] [doi: [10.1093/ntr/ntr183](https://doi.org/10.1093/ntr/ntr183)] [Medline: [21994344](https://pubmed.ncbi.nlm.nih.gov/21994344/)]
31. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychol Bull* 2003 Sep;129(5):674-697 [FREE Full text] [doi: [10.1037/0033-2909.129.5.674](https://doi.org/10.1037/0033-2909.129.5.674)] [Medline: [12956539](https://pubmed.ncbi.nlm.nih.gov/12956539/)]
32. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. *Psychiatr Rehabil J* 2004;27(4):392-401. [doi: [10.2975/27.2004.392.401](https://doi.org/10.2975/27.2004.392.401)] [Medline: [15222150](https://pubmed.ncbi.nlm.nih.gov/15222150/)]
33. Salzer MS, Shear SL. Identifying consumer-provider benefits in evaluations of consumer-delivered services. *Psychiatr Rehabil J* 2002;25(3):281-288. [doi: [10.1037/h0095014](https://doi.org/10.1037/h0095014)] [Medline: [11860001](https://pubmed.ncbi.nlm.nih.gov/11860001/)]
34. Thrul J, Tormohlen KN, Meacham MC. Social media for tobacco smoking cessation intervention: a review of the literature. *Curr Addict Rep* 2019 Jun;6(2):126-138 [FREE Full text] [Medline: [31485388](https://pubmed.ncbi.nlm.nih.gov/31485388/)]

Abbreviations

AYA: adolescent and young adult

IRB: Institutional Review Board

LGBTQ+: lesbian, gay, bisexual, transgender, queer, or other sexual orientation

QTH: Quit the Hit

Edited by A Mavragani; submitted 10.09.22; peer-reviewed by M Jacobs; comments to author 25.10.22; revised version received 15.11.22; accepted 16.11.22; published 01.12.22.

Please cite as:

Lyu JC, Afolabi A, White JS, Ling PM

Perceptions and Aspirations Toward Peer Mentoring in Social Media-Based Electronic Cigarette Cessation Interventions for Adolescents and Young Adults: Focus Group Study

JMIR Form Res 2022;6(12):e42538

URL: <https://formative.jmir.org/2022/12/e42538>

doi: [10.2196/42538](https://doi.org/10.2196/42538)

PMID: [36454628](https://pubmed.ncbi.nlm.nih.gov/36454628/)

©Joanne Chen Lyu, Aliyyat Afolabi, Justin S White, Pamela M Ling. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 01.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Persian Version of the Mobile Application Rating Scale (MARS-Fa): Translation and Validation Study

Saeed Barzegari¹, PhD; Ali Sharifi Kia², MSc; Marco Bardus³, BA, MA, PhD; Stoyan R Stoyanov⁴, MSc; Marjan GhaziSaeedi⁵, PhD; Mouna Rafizadeh⁵, MSc

¹Department of Paramedicine, Faculty of Paramedical Sciences, Mazandaran University of Medical Sciences, Sari, Iran

²Department of Health Information Management, School of Health Management and Information Science, Iran University of Medical Sciences, Tehran, Iran

³Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, United Kingdom

⁴Creative Industries Faculty, School of Design, Queensland University of Technology, Brisbane, Australia

⁵Department of Health Information Management, School of Allied Medical Sciences, Tehran University of Medical Science, Tehran, Iran

Corresponding Author:

Marco Bardus, BA, MA, PhD

Institute of Applied Health Research

College of Medical and Dental Sciences

University of Birmingham

Edgbaston campus

Edgbaston, B15 2TT

United Kingdom

Phone: 44 0121 414 3344

Email: m.bardus@bham.ac.uk

Abstract

Background: Approximately 110 million Farsi speakers worldwide have access to a growing mobile app market. Despite restrictions and international sanctions, Iran's internal mobile health app market is growing, especially for Android-based apps. However, there is a need for guidelines for developing health apps that meet international quality standards. There are also no tools in Farsi that assess health app quality. Developers and researchers who operate in Farsi could benefit from such quality assessment tools to improve their outputs.

Objective: This study aims to translate and culturally adapt the Mobile Application Rating Scale in Farsi (MARS-Fa). This study also evaluates the validity and reliability of the newly developed MARS-Fa tool.

Methods: We used a well-established method to translate and back translate the MARS-Fa tool with a group of Iranian and international experts in Health Information Technology and Psychology. The final translated version of the tool was tested on a sample of 92 apps addressing smartphone addiction. Two trained reviewers completed an independent assessment of each app in Farsi and English. We reported reliability and construct validity estimates for the objective scales (engagement, functionality, aesthetics, and information quality). Reliability was based on the evaluation of intraclass correlation coefficients, Cronbach α and Spearman-Brown split-half reliability indicators (for internal consistency), as well as Pearson correlations for test-retest reliability. Construct validity included convergent and discriminant validity (through item-total correlations within the objective scales) and concurrent validity using Pearson correlations between the objective and subjective scores.

Results: After completing the translation and cultural adaptation, the MARS-Fa tool was used to assess the selected apps for smartphone addiction. The MARS-Fa total scale showed good interrater reliability (intraclass correlation coefficient=0.83, 95% CI 0.74-0.89) and good internal consistency (Cronbach α =.84); Spearman-Brown split-half reliability for both raters was 0.79 to 0.93. The instrument showed excellent test-retest reliability ($r=0.94$). The correlations among the MARS-Fa subdomains and the total score were all significant and above $r=0.40$, suggesting good convergent and discriminant validity. The MARS-Fa was positively and significantly correlated with subjective quality ($r=0.90$, $P<.001$), and so were the objective subdomains of engagement ($r=0.85$, $P<.001$), information quality ($r=0.80$, $P<.001$), aesthetics ($r=0.79$, $P<.001$), and functionality ($r=0.57$, $P<.001$), indicating concurrent validity.

Conclusions: The MARS-Fa is a reliable and valid instrument to assess mobile health apps. This instrument could be adopted by Farsi-speaking researchers and developers who want to evaluate the quality of mobile apps. While we tested the tool with a

sample of apps addressing smartphone addiction, the MARS-Fa could assess other domains or issues since the Mobile App Rating Scale has been used to rate apps in different contexts and languages.

(*JMIR Form Res* 2022;6(12):e42225) doi:[10.2196/42225](https://doi.org/10.2196/42225)

KEYWORDS

mobile application rating scale; Farsi; mobile apps; validation; smartphone addiction; Persian; Iran; development; mobile health; mHealth; scale; validate; reliability; measurement tool; assessment tool

Introduction

Background

In the last 2 decades, advancements in mobile phone technology have allowed users the possibility to access health information from anywhere through nearly ubiquitous internet connectivity; at the same time, health care and public health organizations can diffuse health messages and provide diverse, continuous, indiscriminate support through mobile phones [1]. In July 2022, smartphones accounted for 4 in 5 mobile handsets available worldwide, with a global user base reaching 5.34 billion (an increase of 93 million since 2021), representing a penetration rate of 67% [2].

Undoubtedly, the public health and research communities consider mobile phones as preferred delivery modes in interventions addressing various health issues such as physical inactivity, substance misuse, and mental health [3-5]. Even a systematic review of mobile health (mHealth) interventions conducted in Iran showed that mobile phones (particularly SMS text messages) were increasingly used to deliver health interventions [6]. Some works argue that mobile apps include system design features that would prompt behavior change [7], with positive effects reported on physical function, pain intensity [8], physical activity [3], and mental health [9]. However, very little evidence exists on the sustained impacts of mobile apps on behaviors and health outcomes [3,4].

Nevertheless, the global mobile health app market does not seem to stop; it was valued at US \$38.2 billion in 2021 and is expected to grow by nearly 12% between 2022 and 2030 [10]. According to Statista, the Iranian digital health market also follows a similar growth trend [11]. Some recent studies have highlighted the proliferation of mHealth in low- and middle-income countries such as Iran [12]. There are no official statistics about the number of smartphone users in Iran. Still, there are about 40 million active social media users, which could indicate technological adoption across the population [13]. With an estimated 150-220 million native speakers [14,15], mobile app development in the Persian language (or Farsi) seems particularly promising for local developers' profitability. Industry-driven mHealth apps may offer a variety of advanced functions and capabilities. However, without the involvement of scientific expertise and evaluation, they may risk delivering unhelpful or potentially hazardous interventions [16-18].

Developers can easily leverage the limited application of guidelines in unregulated app markets that rely on open platforms such as Android in Iran, whose population has limited or no access to the global app markets on Google Play and Apple App Store. Two recently published reviews of the Iranian health

app market identified about 3300 [19] and 3500 apps in the Android marketplace, which is the largest [20]. Two other reviews of COVID-19 apps in Iran searched for apps in different stores for iOS, such as CafeBazar, ParsHub, Charkhooneh, SibBazar, Sibche, SibApp, and SibIrani [21,22]. However, these stores are considered unsafe and unreliable by most Iranian citizens.

The proliferation of mobile health apps globally and in Iran raises concerns about their quality, accuracy, reliability, and efficacy [23]. According to some recent systematic reviews, various mHealth evaluation tools and rating scales have been developed to address this need [17,24]. These assessment tools vary from adapted website assessment tools to the use of consumers' reviews or rating [25]. App store ratings are subjective and, by nature, a poor indicator of quality, medical usefulness, safety, or effectiveness. Quality reviews by trusted third parties can serve as landmarks in assessing the security, validity, and quality of mHealth apps [26]. According to a review of health app evaluation tools by BinDhim et al [25], the most frequently used were the Royal College of Physicians' Health Informatics Unit Checklist [27], the Organization for the Review of Care and Health Applications-24 Question Assessment (ORCHA-24) [28], and the Mobile Application Rating Scale (MARS) [29]. The Royal College of Physicians' Health Informatics Unit Checklist only looks at the developer, the functionality, and whether the app has been evaluated effectively in related interventions [27]. The Organization for the Review of Care and Health Applications-24 Question Assessment focuses on data governance, clinical impact and assurance, and user experience and engagement as quality aspects [28], but it fails to provide a comprehensive, multidimensional evaluation of app quality. Conversely, the MARS assesses app quality on a broader and more diverse range of criteria or domains, such as engagement, functionality, aesthetics, and information quality. According to Azad-Khanegah and colleagues [24], the MARS provides a multidimensional, reliable, and flexible app-quality rating scale for researchers, developers, and health care professionals [29]. The MARS has been used to evaluate apps in user-based heuristic evaluations [30] and expert-driven content analyses of apps [31-33]. The MARS has been validated across multiple studies [32] and translated into Italian [34], and more recently into German [35], Spanish [36], Arabic [37], Japanese [38], Korean [39], French [40], and Turkish [41]. However, this instrument has no translation or cultural adaptation for the Farsi language.

Objectives

This study aimed to (1) translate and culturally adapt the MARS in the Farsi language (MARS-Fa) and (2) validate the tool by examining its psychometric properties.

Methods

Study Design

This study followed a 2-step process, starting with the translation and cultural adaptation of the MARS in English to Farsi, as done in the validation studies mentioned above [34-41]. The second step involved a statistical evaluation of the MARS-Fa's reliability and validity.

Original Instrument: The MARS

The MARS [29] consists of 29 items divided into the following 4 objective subscales: *engagement* (items 1-5), *functionality* (items 6-9), *aesthetics* (items 10-12), and *information* (items 13-19); it also comprises a subjective subscale, which is *app subjective quality* (items 20-23). The MARS also includes items intended to measure the *perceived impact of the app* for the intended end users. The perceived impact scale includes 6 additional items that evaluate the app's potential to affect users' knowledge, awareness, and intentions to perform the target behaviors. However, it is intended for the end users and is generally not used to assess app quality or to compare apps. All items are rated on 5-point scales, usually ranging from 1 ("poor") to 5 ("excellent"), except for the *perceived impact* items, which are based on 5-point Likert-type scales, where 1 is "strongly disagree," and 5 is "strongly agree." According to the guidelines from the original MARS study [29], an average score is calculated for each subscale. A total app quality score represents the average of the 4 objective subscales. The original MARS study reported high internal consistency (Cronbach $\alpha=0.90$) and reliability, with intraclass correlation coefficients (ICCs) averaging 0.79 [29].

Translation and Adaptation Process

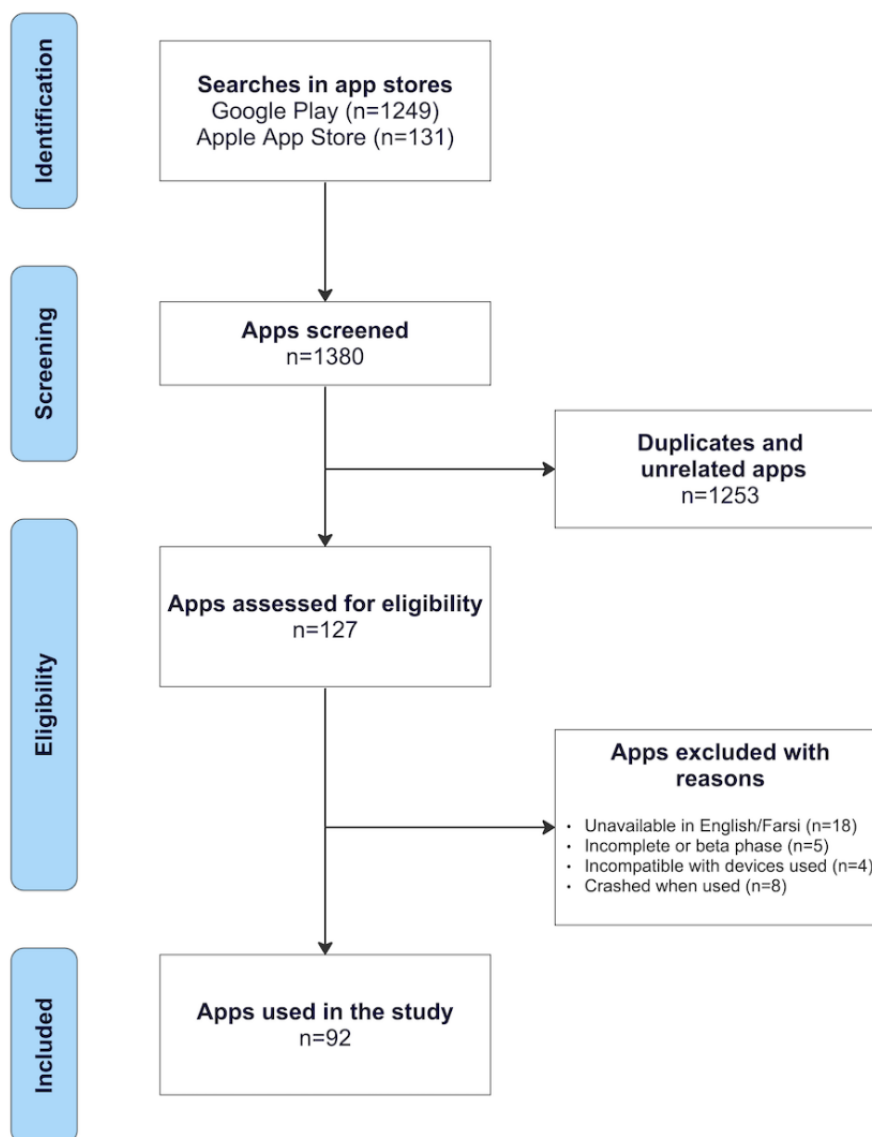
Following the so-called "universalist approach" [42] applied in other MARS validation studies [37], the translation and

adaptation process consisted of the following steps. First, a translation was conducted, including essential item and conceptual equivalences, which were evaluated and validated by a panel of 8 experts, including 4 PhD students in health information management and health information technology, and 4 researchers with PhDs in psychology and nursing. In the next step, 2 English translators familiar with IT concepts independently translated the MARS tool into Farsi. A semantic evaluation was also performed to check the ambiguity and simplicity of the Farsi translation among the potential target population. Finally, to ensure that the Farsi version was perceived as the original English scale, it was translated back to English by a bilingual translator and compared with the original version. The back-translated version was finally checked and validated by the developer of the original MARS [29], and a few amendments were made.

Sample Selection for Scale Validation

To assess the reliability and validity of the MARS-Fa, we selected a sample of health apps targeting smartphone addiction and available on the Android and iOS stores. While the MARS is intended to address health apps of any domain, our study team included researchers with solid expertise in health information technology and smartphone addiction.

A systematic process was followed to select the smartphone addiction apps for evaluation. All steps are presented in the diagram in Figure 1. Two Health Information Technology experts independently searched the Google Play and Apple App stores on May 22 and June 1, 2019. The keywords included "Smartphone Addiction," "Phone Addiction," "Mobile Phone Addiction," "Cellphone Addiction," and "Nomophobia." To be included in the sample, apps had to (1) be available in either English or Farsi languages, (2) address smartphone addiction, and (3) be free of charge. Exclusion criteria were as follows: (1) apps being underdevelopment or not released yet; (2) apps that were unavailable or that could not be downloaded due to device incompatibility; (3) apps failing to launch after 3 attempts or apps crashing. The app selection process is summarized in Figure 1.

Figure 1. Flow diagram of the app selection process.

Validation Process

Two raters had a session to study the MARS tool and discuss their perception regarding its concepts. As a result, both raters came to a shared understanding of how to use the MARS for the app target group. Both raters downloaded each selected app on both iOS and Android-based smartphones. They completed an independent assessment of each app in both Farsi and English.

Initially, the 2 raters independently evaluated 10 apps for about 10 minutes each. The similarity between the reviewers' judgments was assessed by comparing ICCs, as done in the original MARS study [29]. This step was introduced to establish a minimum interrater reliability level and allow the raters to identify and discuss differences and address inconsistencies before assessing the remaining apps. After 2 weeks, 10 apps were randomly selected and evaluated for the second time by the same 2 raters to evaluate their test-retest reliability.

In the next step, out of the selected 92 apps, 45 (49%) were randomly chosen for the validation exercise. This number was deemed sufficient to reach an empirical assurance of 90% and

an assurance probability of .15, as done in the study that brought about the development of the Italian version of the MARS [34].

Ethical Considerations

This study involved secondary analyses of research data without including human participants; as such, no ethical approval was needed.

Statistical Analyses

Descriptive statistics were calculated for all items, subscales, and the total MARS scale, including means, standard deviations, and asymmetry coefficients. Subsequently, the reliability and validity measures of the MARS were evaluated separately for both raters. Interrater differences were assessed for subscale scores.

ICCs, using a mixed 2-factor model, were used to evaluate the interrater reliability [43]. This method has been deemed appropriate, as it accounts for the proximity of scores rather than an absolute agreement between raters. ICC values less than 0.50, between 0.50 and 0.75, between 0.75 and 0.90, and greater than 0.90 are respectively considered poor, moderate, good, and

excellent interrater reliability [43]. Cronbach α was used to assess internal consistency and was interpreted as excellent (≥ 0.90), good (0.80-0.89), acceptable (0.70-0.79), questionable (0.60-0.69), poor (0.50-0.59), and unacceptable (< 0.50), as reported in [30]. Split-half reliability was used to evaluate the internal consistency of the average of the 2 raters using the Spearman-Brown prophecy formula, as used in the Italian MARS validation study [34]. Pearson correlations were used to assess the test-retest reliability [43].

To determine construct and concurrent validity, we replicated the approach of Yamamoto et al [38], who validated the Japanese MARS. Construct validity was based on evaluating item-subscale correlations [38] for the objective scales only, considering the intrinsic subjectivity of the “subjective quality” scale. Convergent validity was deemed satisfactory when an item achieved a correlation above $r=0.20$ with the respective subscale, a threshold used in the Italian [34] and Japanese [38] validation studies. Discriminant validity was deemed satisfactory if more than 80% of the correlation coefficients were higher than those with other subscales [38]. To establish concurrent validity, we examined the correlations between the MARS-Fa objective scales and the subjective quality, given that there are

no gold-standard app quality indicators other than the MARS itself [38]. Other studies have compared the MARS objective and subjective scores to the average app store ratings for each app [37,38]; however, these were deemed inappropriate as the Farsi version of the app pages include few reviews and ratings that might be biased and manipulated, hence being unreliable indicators of app quality.

Results

Translation and Adaptation Process

In the forward and backward translation and face validation phases, we used IT and health experts, who identified common words, phrases, and sentences in both disciplines. There were some corrections made after the backward-translated version was reviewed and edited by 2 authors (one of them was the corresponding author of the original scale). The final version of the translation was deemed clear and understandable for both groups and not in conflict with the original version. Table 1 shows the words that were corrected in the process. The final version of the MARS-Fa tool is available in Multimedia Appendix 1.

Table 1. Corrections on the back-translated Mobile Application Rating Scale (MARS, in English) and the Farsi version (MARS-Fa).

First Farsi translation	Retranslated word	Correct word	Corrected Farsi word
	Entertainment	Engagement	
	Constant use	Frequent use	
	Incompatible	Incoherent	
	Overwhelming	Was explained as “too much for the user to know where to start”	
	Obvious	Intuitive	

App Selection Process

Initial searches in the app stores yielded 1380 apps from both Android and iOS stores. After removing duplicates and irrelevant apps ($n=1253$), 127 apps were screened for inclusion. Of these 127 apps, 18 (14.2%) were excluded because they were not available in either Farsi or English, 5 (3.9%) were excluded because they were incomplete (beta versions), 4 (3.1%) because they were incompatible with the devices used to test the apps, and 8 (6.3%) could not load (crashed when launching them), leaving a final set of 92 (72.4%) apps for the validation study (Figure 1). An ID was assigned to each app. In the next step, 45 apps (Multimedia Appendix 2) were randomly and with equal proportions selected from the two app stores (Google Play: $n=30$, 67%; App Store: $n=15$, 33%) for preliminary testing. The apps included in this study lacked any peer-reviewed publications of formal efficacy trials. Hence, item 19 of the information domain, “Evidence base,” which aims to assess the app’s reported efficacy based on randomized controlled trials, was excluded from the calculations as none of the apps were formally trialed.

Reliability and Validity Analyses

Table 2 presents the descriptive statistics for each subscale and the total MARS-Fa score separately for each rater. As the

responses followed a nonnormal distribution, nonparametric tests were used to check the differences between raters. The paired t test and the Wilcoxon test (2-tailed) showed no significant differences between the raters’ mean scores.

Table 3 reports the results of the reliability analyses. Interrater reliability was good for “engagement” ($ICC=0.85$), “information quality” ($ICC=0.76$), and “aesthetics” ($ICC=0.75$), and moderate for “functionality” ($ICC=0.60$). ICC was also good for the MARS-Fa total score (0.83) and “subjective quality” (0.78). The Spearman-Brown split-half reliability estimates ranged between 0.79 and 0.93, confirming good interrater reliability.

Cronbach α coefficients for each of the MARS-Fa subdomains, total, score, and subjective quality (Table 3) ranged from .51 to .89 for the first rater and .56 to .84 for the second rater. The average alpha coefficient was .84 for the total MARS-Fa and subjective quality. Spearman-Brown split-half reliability indicators were very good and excellent, ranging from 0.79 for functionality and impact and 0.93 for the MARS-Fa total score.

The MARS-Fa total score and subscales had excellent and good test-retest reliability, with correlations above 0.90, indicating no significant change over time ($P>.05$) for all objective subscales, total score, and subjective quality score. Overall, the average test-retest correlation between the 2 raters was high ($r=0.94$).

Table 2. Descriptive statistics and interrater comparisons.

Scale	Minimum-maximum		Skewness		Shapiro-Wilk (<i>P</i> value)		Mean (SD)		<i>P</i> value ^a	Cohen <i>d</i>
	R1 ^b	R2	R1	R2	R1	R2	R1	R2		
Engagement	1.60-4.60	1.60-4.60	-0.40	-0.43	0.94 (.02)	0.97 (.40)	3.40 (0.87)	3.31 (0.69)	.11	0.24
Functionality	2.00-4.75	2.25-4.50	-0.50	-0.96	0.96 (.14)	0.91 (<.001)	3.68 (0.57)	3.73 (0.54)	.47	0.11
Aesthetics	2.33-5.00	2.33-5.00	-0.74	-0.46	0.92 (<.001)	0.93 (.01)	3.90 (0.64)	3.79 (0.74)	.16	0.21
Information	1.75-4.50	1.75-4.17	-0.54	-0.28	0.96 (.14)	0.96 (.08)	3.28 (0.57)	3.31 (0.54)	.67	0.06
MARS-Fa ^c total score	2.17-4.66	2.48-4.50	-0.40	-0.29	0.97 (.30)	0.97 (.36)	3.57 (0.56)	3.54 (0.52)	.53	0.001
Subjective quality	1.00-4.75	1.20-4.75	-0.52	-0.19	0.91 (<.001)	0.94 (.03)	3.32 (1.19)	3.24 (1.00)	.34	0.14

^a*P* value of the Wilcoxon *W* or *t* test.

^bR: reviewer.

^cMARS-Fa: Mobile Application Rating Scale in Farsi.

Table 3. Interrater reliability, internal consistency, and test-retest reliability results.

Scale	Cronbach α		ICC ^a (95% CI)	Spearman-Brown split-half reliability	Test-retest reliability (Pearson <i>r</i>)	
	R1 ^b	R2			R1	R2
Engagement	.89	.83	0.85 (0.72-0.90)	0.92	0.94	0.96
Functionality	.51	.56	0.60 (0.39-0.74)	0.79	0.93	0.96
Aesthetics	.71	.83	0.75 (0.62-0.84)	0.85	0.95	0.91
Information	.77	.65	0.76 (0.63-0.84)	0.86	0.92	0.89
MARS-Fa ^c total score	.84	.84	0.83 (0.74-0.89)	0.93	0.92	0.95
Subjective quality	.84	.84	0.78 (0.67-0.86)	0.82	0.94	1.00

^aICC: intraclass correlation coefficient.

^bR: reviewer.

^cMARS-Fa: Mobile Application Rating Scale in Farsi.

Construct Validity

The item-total correlations are shown in Table 4, all of which were above 0.40 in the objective subscales except for functionality item 7, "Ease of use" ($r=0.27$). Success rate was deemed satisfactory for convergent validity. Overall, success rate was also deemed satisfactory for divergent validity, with all items being above the threshold in all subdomains except functionality (item 7), and information quality (item 13, "Accuracy of app description"), which had the lowest correlation with the total among the other items of the domain.

Pearson correlations between the MARS-Fa total score, the respective objective subdomains, and the subjective quality score are shown in Table 5. The MARS-Fa was positively and significantly correlated with subjective quality ($r=0.90$, $P<.001$), and so were the objective subdomains of engagement ($r=0.85$, $P<.001$), information quality ($r=0.80$, $P<.001$), aesthetics ($r=0.79$, $P<.001$), and functionality ($r=0.57$, $P<.001$). The relationships between the MARS-Fa and the objective domains are not reported because the MARS-Fa is their composite score. The relationships among the objective domains were also significant ($P<.001$).

Table 4. Construct validity indicators.

MARS ^a objective subscale items	Corrected item-total correlations (Pearson <i>r</i>)	Success rate	
		Convergent validity	Divergent validity
Engagement		5/5	5/5
Entertainment	0.72		
Interest	0.70		
Customization	0.64		
Interactivity	0.74		
Target group	0.83		
Functionality		3/4	3/4
Performance	0.65		
Ease of use	0.27		
Navigation	0.50		
Gestural design	0.38		
Aesthetics		3/3	3/3
Layout	0.70		
Graphics	0.79		
Visual appeal	0.81		
Information		6/6	5/6
Accuracy of app description	0.43		
Goals	0.67		
Quality of information	0.57		
Quantity of information	0.68		
Visual information	0.51		
Credibility	0.47		
Evidence base	N/A ^b	N/A	N/A

^aMARS: Mobile Application Rating Scale.

^bN/A: not applicable.

Table 5. Correlations between the Mobile Application Rating Scale in Farsi (MARS-Fa) objective scores and subjective quality.

Relationship	Pearson <i>r</i>	Lower 95% CI	Upper 95% CI
MARS-Fa total score-subjective quality	0.90	0.86	0.93
Subjective quality-engagement	0.85	0.77	0.90
Subjective quality-functionality	0.57	0.41	0.69
Subjective quality-aesthetics	0.79	0.73	0.85
Subjective quality-information quality	0.80	0.73	0.86
Engagement-functionality	0.38	0.20	0.53
Engagement-aesthetics	0.76	0.70	0.84
Engagement-information quality	0.75	0.66	0.82
Functionality-aesthetics	0.54	0.41	0.67
Functionality-information quality	0.52	0.35	0.68
Aesthetics-information quality	0.69	0.58	0.77

Discussion

This is the first study that developed a translation and cultural adaptation of the MARS scale into the Farsi language. This study also included validation of the MARS-Fa tool with a sample of apps targeting smartphone addiction. The results show that the MARS-Fa is a reliable and valid tool that can be used to assess app quality. Health care professionals, researchers, authorities, organizations, and app developers can use this tool when developing new or evaluating existing apps in Farsi.

Translation and Cultural Adaptation

Translating IT terminology tends to be challenging, especially in unrelated contexts such as health care. The employment of experts in the translation process facilitates the adaptation of scales, as discussed in the Italian MARS validation study [34]. To translate the tool, we employed 2 experts in IT concepts; the concepts were subsequently translated and then back translated into English by a third bilingual translator, similar to the Arabic MARS validation study [37]. The original scale developer and another English expert were asked to check each version. In the process, we excluded item 19, "Evidence base," as it was not applicable because no apps were used in randomized controlled studies, as done in previous studies [29,34,44]. Nevertheless, given the complex terminology of the scale, it is recommended to develop a dedicated training module for Farsi-speaking app reviewers, such as the one developed for the original MARS [29] and the German version of the tool [35]. The training module will likely improve the interrater reliability, test-retest reliability, and possibly the validity of the MARS-Fa, but this needs to be formally tested in future studies, possibly with a different set of apps.

Reliability

The MARS-Fa showed a good degree of interrater reliability, with ICCs ranging from 0.60 to 0.85, with results that are aligned with the original study (ICCs=0.79) [29] and other similar validation studies, such as the Italian (0.96) [34], Spanish (0.96) [36], German (0.83) [35], Arabic (0.84) [37], French (0.89) [40], Japanese (0.70) [38], and Turkish (0.94) [41] studies. Functionality was the domain with the lowest ICC value, as in the original MARS study (0.50) [29] and the Japanese study (0.40) [38]. This might be due to the nature of mHealth apps for mental health used in both studies, similar to the ones targeting smartphone addiction in this paper. It can also be due to differences in how raters interpreted the items. Training raters before using the instrument will likely reduce the likelihood of misinterpretations, as in the German study [35].

The MARS-Fa displayed a good internal consistency, with Cronbach α coefficients of both raters deemed "good" for the MARS total score. The Spearman-Brown split-half reliability indicated good internal consistency among the raters, as reported in the Italian MARS validation study [34]. Altogether, the internal consistency estimates of the MARS-Fa are aligned with similar MARS validation studies [34,36,37,40,41]. The functionality domain had a relatively lower level of internal consistency, as reported in the original MARS study (0.80) [29], and in other MARS-validation studies such as the Italian (0.82 between 2 raters) [34], Arabic (0.72) [37], French (0.79) [40],

and Turkish (0.78 between 2 raters) [41] studies. A relatively low level of internal consistency estimates for the information quality subscale was also reported in other validation studies, such as the Italian (0.72 between 2 raters) [34], German (0.72) [35], and French (0.61) [40]. These differences might be due to the diverse nature of the tested apps, as functionality, navigation features, ease of use, and information included in each app can vary significantly between apps, depending on the type of health issue addressed and within apps, because content and format can vary across platforms and devices, as reported in Bardus et al [31].

Additionally, the MARS-Fa showed excellent test-retest reliability, as testified by significant and high Pearson correlations over time; all subscales and the total score were more than or equal to 0.90, according to methodology literature [45], indicating an excellent test-retest reliability.

Validity

Overall, the MARS-Fa shows good construct validity, as all items seemed to correlate well within each objective subdomain. Similar to the Japanese validation study [38], one item of the functionality domain appeared to have the lowest correlation with the other items, "ease of use," which might indicate a wide variability in the usability of the apps analyzed. As for concurrent validity, the MARS-Fa total score (objective quality) was significantly correlated with subjective quality. However, this might be interpreted with caution as the subjective quality might be influenced by the reviewers' completing the objective quality evaluation in the same instance, as discussed in the original MARS study [29] and reported in the Japanese validation study [38]. In the absence of other benchmarks, the correlation between the MARS-Fa total score and its subjective counterpart indicates that the two measures are somehow aligned.

Strengths and Limitations

This is the first study reporting on the translation and cultural adaptation of the MARS scale into Farsi and its subsequent validation. A major strength of this study is the systematic process followed in translating and validating the MARS-Fa tool, using well-established and sound methodologies. The translation and cultural adaptation process involved IT and health sciences experts, who checked the content for and provided face validity. Furthermore, the construct and scale validation process followed a robust approach. The study involved 2 raters who independently assessed a systematically selected sample of apps. Through this process, the MARS-Fa tool can be reliably used to evaluate the quality of health apps in the Farsi language.

Limitations of this study include the fact that we tested the tool with a selected sample of apps for smartphone addiction. While the tool is intended to assess health apps in any domain, there might be some variability in the type of apps analyzed. Hence, we suggest that future studies test the MARS-Fa using other health apps. One of the limitations is that the MARS-Fa total score and objective subscales were validated against the subjective quality in the absence of an equivalent app quality evaluation scale. Future studies could compare the MARS-Fa

to other app quality evaluation tools identified in the literature [24] to ascertain concurrent validity.

Conclusions

The Farsi version of the MARS tool (MARS-Fa) is a reliable and valid instrument to assess mobile health app quality, as demonstrated by a sample of apps targeting smartphone

addiction. Health experts, researchers, and app developers can use the MARS-Fa, to evaluate their apps or to assess groups of apps of the same kind. It can be easily accessed ([Multimedia Appendix 1](#)) free of charge. We hope that the MARS-Fa could be used as a criterion for evaluating apps before these are prescribed to patients.

Acknowledgments

The authors thank Zahra Mahmoudvand, Shahab Abhari, Meysam Rahmani, Hadi Kazemi, and Mohammad Hosseini for participating in the expert panel and helping with the translation of the MARS-Fa tool. This research was partially supported by grant number 14738 by Mazandaran University of Medical Sciences, awarded to the first author.

Data Availability

The list of the tested mobile apps with the respective ratings is available in [Multimedia Appendix 2](#). The Mobile App Rating Scale in Farsi (MARS-Fa) is available in [Multimedia Appendix 1](#).

The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SB conceived and designed the study with intellectual input from MR, MGS, and ASK. MB and SRS provided oversight to the study implementation. SRS reviewed the back-translated version of the tool. SB and ASK undertook initial data analyses, and MR, MGS, MB, and SRS contributed to data interpretation. SB and MR drafted the manuscript, which all authors then edited. MB revised the manuscript and finalized all edits. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of mobile apps tested.

[\[DOCX File , 76 KB - formative_v6i12e42225_app1.docx \]](#)

Multimedia Appendix 2

The Mobile App Rating Scale in Farsi (MARS-Fa).

[\[DOCX File , 39 KB - formative_v6i12e42225_app2.docx \]](#)

References

1. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. *JMIR Mhealth Uhealth* 2015 Feb 26;3(1):e20 [[FREE Full text](#)] [doi: [10.2196/mhealth.3335](https://doi.org/10.2196/mhealth.3335)] [Medline: [25803705](https://pubmed.ncbi.nlm.nih.gov/25803705/)]
2. Kemp S. Digital 2022: April Global Statshot Report. Data Reportal. URL: <https://datareportal.com/reports/digital-2022-april-global-statshot> [accessed 2022-06-06]
3. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR Mhealth Uhealth* 2020 Mar 18;8(3):e17046 [[FREE Full text](#)] [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
4. Serrano-Ripoll MJ, Zamanillo-Campos R, Fiol-DeRoque MA, Castro A, Ricci-Cabello I. Impact of Smartphone App-Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials. *JMIR Mhealth Uhealth* 2022 Jan 27;10(1):e29621 [[FREE Full text](#)] [doi: [10.2196/29621](https://doi.org/10.2196/29621)] [Medline: [35084346](https://pubmed.ncbi.nlm.nih.gov/35084346/)]
5. Bahadoor R, Alexandre J, Fournet L, Gellé T, Serre F, Auriacombe M. Inventory and Analysis of Controlled Trials of Mobile Phone Applications Targeting Substance Use Disorders: A Systematic Review. *Front Psychiatry* 2021;12:622394 [[FREE Full text](#)] [doi: [10.3389/fpsy.2021.622394](https://doi.org/10.3389/fpsy.2021.622394)] [Medline: [33692708](https://pubmed.ncbi.nlm.nih.gov/33692708/)]
6. Ershad SR, Sadoughi F, Jamshidi OR, Bahaadinbeigy K. Role of Mobile Technology in Iran Healthcare System: A review study. *jhbmi* 2018;4(4):313-326.
7. Oyeboode O, Ndulue C, Alhasani M, Orji R. Persuasive Mobile Apps for Health and Wellness: A Comparative Systematic Review. In: *Persuas Technol Des Future Change*. Cham, Switzerland: Springer; 2020:163-181.

8. Elbers S, Wittink H, Pool JJM, Smeets RJEM. The effectiveness of generic self-management interventions for patients with chronic musculoskeletal pain on physical function, self-efficacy, pain intensity and physical activity: A systematic review and meta-analysis. *Eur J Pain* 2018 Oct;22(9):1577-1596 [FREE Full text] [doi: [10.1002/ejp.1253](https://doi.org/10.1002/ejp.1253)] [Medline: [29845678](https://pubmed.ncbi.nlm.nih.gov/29845678/)]
9. Poon SK. Pacifica: stressed or worried? An app to help yourself (Mobile App User Guide). *Br J Sports Med* 2016 Jan 18;50(3):191-192. [doi: [10.1136/bjsports-2015-095747](https://doi.org/10.1136/bjsports-2015-095747)]
10. Grand View Research. 2022. URL: <https://www.grandviewresearch.com/industry-analysis/mhealth-app-market> [accessed 2022-07-29]
11. Digital Health - Iran. Statista. URL: <https://www.statista.com/outlook/dmo/digital-health/iran> [accessed 2022-07-29]
12. Katibeh M, Mousavi B, Kalantarion M, Sabbaghi H, Abdolahi E, Nikkhah H, et al. Using mHealth to improve eye care in remote areas of Iran. *Community Eye Health* 2019;32(107):65-66. [Medline: [32123481](https://pubmed.ncbi.nlm.nih.gov/32123481/)]
13. Kheradmand A, Amirlatifi ES, Sohrabi M, Mazaheri Meybodi A. Validation of the Persian Smartphone Addiction Scale Among Tehran University Students, Iran. *Int J High Risk Behav Addict* 2019 Mar 18:e81176. [doi: [10.5812/ijhrba.81176](https://doi.org/10.5812/ijhrba.81176)]
14. Windfuhr G, editor. *The Iranian Languages*. Routledge. London, UK: Routledge; 2009.
15. Mahmoodi-Bakhtiari J. *Iranian Languages*. In: *The World's Major Languages*. London, UK: Routledge; 2018.
16. Ashall-Payne L, Leigh S. Health Apps, the Good, the Bad, and the Ugly: The Importance of information. *Digital Health London*. URL: <https://digitalhealth.london/health-apps-good-bad-ugly-importance-information> [accessed 2022-07-29]
17. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. *J Med Internet Res* 2014 Feb 19;16(2):e36 [FREE Full text] [doi: [10.2196/jmir.2925](https://doi.org/10.2196/jmir.2925)] [Medline: [24554442](https://pubmed.ncbi.nlm.nih.gov/24554442/)]
18. Barton AJ. The regulation of mobile health applications. *BMC Med* 2012 May 08;10:46 [FREE Full text] [doi: [10.1186/1741-7015-10-46](https://doi.org/10.1186/1741-7015-10-46)] [Medline: [22569114](https://pubmed.ncbi.nlm.nih.gov/22569114/)]
19. Naderi H, Etminani K. An evaluation of released mobile health apps in popular Iranian app stores. *IJMEI* 2019;11(4):320. [doi: [10.1504/ijmei.2019.104961](https://doi.org/10.1504/ijmei.2019.104961)]
20. Fadaizadeh L, Sanaat M, Yousefi E, Alizadeh N. Mobile health: A comparative study of medicalhealth applications in Iran. *Biomed Biotechnol Res J* 2022;249-254. [doi: [10.4103/bbrj.bbrj_31_22](https://doi.org/10.4103/bbrj.bbrj_31_22)]
21. Nouri R, Salari R, Kalhori SRN, Ayyoubzadeh SM, Gholamzadeh M. Persian mobile health applications for COVID-19: A use case-based study. *J Educ Health Promot* 2022;11:100 [FREE Full text] [doi: [10.4103/jehp.jehp_759_21](https://doi.org/10.4103/jehp.jehp_759_21)] [Medline: [35573610](https://pubmed.ncbi.nlm.nih.gov/35573610/)]
22. Erfannia L, Amraei M, Arji G, Yazdani A, Sabzehgar M, Yaghoobi L. Reviewing and Content Analysis of Persian Language Mobile Health Apps for COVID-19 Management. *Stud Health Technol Inform* 2022 Jan 14;289:106-109. [doi: [10.3233/SHTI210870](https://doi.org/10.3233/SHTI210870)] [Medline: [35062103](https://pubmed.ncbi.nlm.nih.gov/35062103/)]
23. 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](https://doi.org/10.1001/jamapsychiatry.2017.0262)] [Medline: [28384700](https://pubmed.ncbi.nlm.nih.gov/28384700/)]
24. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disabil Rehabil Assist Technol* 2021 Oct;16(7):712-721. [doi: [10.1080/17483107.2019.1701103](https://doi.org/10.1080/17483107.2019.1701103)] [Medline: [31910687](https://pubmed.ncbi.nlm.nih.gov/31910687/)]
25. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. *Telemed J E Health* 2015 Feb;21(2):97-104. [doi: [10.1089/tmj.2014.0088](https://doi.org/10.1089/tmj.2014.0088)] [Medline: [25469795](https://pubmed.ncbi.nlm.nih.gov/25469795/)]
26. Powell AC, Landman AB, Bates DW. Certification of mobile apps for health care--reply. *JAMA* 2014 Sep 17;312(11):1156-1157. [doi: [10.1001/jama.2014.9007](https://doi.org/10.1001/jama.2014.9007)] [Medline: [25226488](https://pubmed.ncbi.nlm.nih.gov/25226488/)]
27. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. *Clin Med (Lond)* 2015 Dec;15(6):519-521 [FREE Full text] [doi: [10.7861/clinmedicine.15-6-519](https://doi.org/10.7861/clinmedicine.15-6-519)] [Medline: [26621937](https://pubmed.ncbi.nlm.nih.gov/26621937/)]
28. Leigh S, Ouyang J, Mimnagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. *Evid Based Ment Health* 2017 Nov;20(4):e20. [doi: [10.1136/eb-2017-102751](https://doi.org/10.1136/eb-2017-102751)] [Medline: [28947676](https://pubmed.ncbi.nlm.nih.gov/28947676/)]
29. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
30. Bardus M, Ali A, Demachkieh F, Hamadeh G. Assessing the Quality of Mobile Phone Apps for Weight Management: User-Centered Study With Employees From a Lebanese University. *JMIR Mhealth Uhealth* 2019 Jan 23;7(1):e9836 [FREE Full text] [doi: [10.2196/mhealth.9836](https://doi.org/10.2196/mhealth.9836)] [Medline: [30672742](https://pubmed.ncbi.nlm.nih.gov/30672742/)]
31. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *Int J Behav Nutr Phys Act* 2016 Mar 10;13:35 [FREE Full text] [doi: [10.1186/s12966-016-0359-9](https://doi.org/10.1186/s12966-016-0359-9)] [Medline: [26964880](https://pubmed.ncbi.nlm.nih.gov/26964880/)]
32. Terhorst Y, Philipp P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). *PLoS One* 2020;15(11):e0241480 [FREE Full text] [doi: [10.1371/journal.pone.0241480](https://doi.org/10.1371/journal.pone.0241480)] [Medline: [33137123](https://pubmed.ncbi.nlm.nih.gov/33137123/)]

33. Knitza J, Tascilar K, Messner E, Meyer M, Vossen D, Pulla A, et al. German Mobile Apps in Rheumatology: Review and Analysis Using the Mobile Application Rating Scale (MARS). *JMIR Mhealth Uhealth* 2019 Aug 05;7(8):e14991 [[FREE Full text](#)] [doi: [10.2196/14991](https://doi.org/10.2196/14991)] [Medline: [31381501](https://pubmed.ncbi.nlm.nih.gov/31381501/)]
34. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Med Inform Decis Mak* 2016 Jul 07;16:83 [[FREE Full text](#)] [doi: [10.1186/s12911-016-0323-2](https://doi.org/10.1186/s12911-016-0323-2)] [Medline: [27387434](https://pubmed.ncbi.nlm.nih.gov/27387434/)]
35. Messner E, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German Version of the Mobile App Rating Scale (MARS-G): Development and Validation Study. *JMIR Mhealth Uhealth* 2020 Mar 27;8(3):e14479 [[FREE Full text](#)] [doi: [10.2196/14479](https://doi.org/10.2196/14479)] [Medline: [32217504](https://pubmed.ncbi.nlm.nih.gov/32217504/)]
36. Martin Payo R, Fernandez Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov S, Llana Suárez E. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. *Int J Med Inform* 2019 Sep;129:95-99. [doi: [10.1016/j.ijmedinf.2019.06.005](https://doi.org/10.1016/j.ijmedinf.2019.06.005)] [Medline: [31445295](https://pubmed.ncbi.nlm.nih.gov/31445295/)]
37. Bardus M, Awada N, Ghandour LA, Fares E, Gherbal T, Al-Zanati T, et al. The Arabic Version of the Mobile App Rating Scale: Development and Validation Study. *JMIR Mhealth Uhealth* 2020 Mar 03;8(3):e16956 [[FREE Full text](#)] [doi: [10.2196/16956](https://doi.org/10.2196/16956)] [Medline: [32130183](https://pubmed.ncbi.nlm.nih.gov/32130183/)]
38. Yamamoto K, Ito M, Sakata M, Koizumi S, Hashisako M, Sato M, et al. Japanese Version of the Mobile App Rating Scale (MARS): Development and Validation. *JMIR Mhealth Uhealth* 2022 Apr 14;10(4):e33725 [[FREE Full text](#)] [doi: [10.2196/33725](https://doi.org/10.2196/33725)] [Medline: [35197241](https://pubmed.ncbi.nlm.nih.gov/35197241/)]
39. Hee Ko KK, Kim SK, Lee Y, Lee JY, Stoyanov SR. Validation of a Korean version of mobile app rating scale (MARS) for apps targeting disease management. *Health Informatics J* 2022 Apr 11;28:14604582221091975 [[FREE Full text](#)] [doi: [10.1177/14604582221091975](https://doi.org/10.1177/14604582221091975)] [Medline: [35404685](https://pubmed.ncbi.nlm.nih.gov/35404685/)]
40. Saliassi I, Martinon P, Darlington E, Smentek C, Tardivo D, Bourgeois D, et al. Promoting Health via mHealth Applications Using a French Version of the Mobile App Rating Scale: Adaptation and Validation Study. *JMIR Mhealth Uhealth* 2021 Aug 31;9(8):e30480 [[FREE Full text](#)] [doi: [10.2196/30480](https://doi.org/10.2196/30480)] [Medline: [34463623](https://pubmed.ncbi.nlm.nih.gov/34463623/)]
41. Mendi O, Kiyimac Sari M, Stoyanov S, Mendi B. Development and validation of the Turkish version of the Mobile App Rating Scale - MARS-TR. *Int J Med Inform* 2022 Oct;166:104843. [doi: [10.1016/j.ijmedinf.2022.104843](https://doi.org/10.1016/j.ijmedinf.2022.104843)] [Medline: [35964383](https://pubmed.ncbi.nlm.nih.gov/35964383/)]
42. Herdman M, Fox-Rushby J, Badia X. A Model of Equivalence in the Cultural Adaptation of HRQoL Instruments: The Universalist Approach. *Qual Life Res* 1998 May 1;7(4):323-335. [doi: [10.1023/A:1024985930536](https://doi.org/10.1023/A:1024985930536)]
43. 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](https://doi.org/10.1016/j.jcm.2016.02.012)] [Medline: [27330520](https://pubmed.ncbi.nlm.nih.gov/27330520/)]
44. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. *JMIR Mhealth Uhealth* 2015 Aug 19;3(3):e82 [[FREE Full text](#)] [doi: [10.2196/mhealth.4328](https://doi.org/10.2196/mhealth.4328)] [Medline: [26290327](https://pubmed.ncbi.nlm.nih.gov/26290327/)]
45. Clark L, Watson D. Constructing validity: Basic issues in objective scale development. *Psychological Assessment* 1995 Sep;7(3):309-319. [doi: [10.1037/1040-3590.7.3.309](https://doi.org/10.1037/1040-3590.7.3.309)]

Abbreviations

- ICC:** intraclass correlation coefficient
MARS: Mobile Application Rating Scale
MARS-Fa: Mobile Application Rating Scale in Farsi
mHealth: mobile health

Edited by A Mavragani; submitted 27.08.22; peer-reviewed by O Mendi, C Fakhri El Khoury, M Kapsetaki; comments to author 13.10.22; revised version received 26.11.22; accepted 28.11.22; published 05.12.22.

Please cite as:

Barzegari S, Sharifi Kia A, Bardus M, Stoyanov SR, GhaziSaeedi M, Rafizadeh M
The Persian Version of the Mobile Application Rating Scale (MARS-Fa): Translation and Validation Study
JMIR Form Res 2022;6(12):e42225
URL: <https://formative.jmir.org/2022/12/e42225>
doi: [10.2196/42225](https://doi.org/10.2196/42225)
PMID: [36469402](https://pubmed.ncbi.nlm.nih.gov/36469402/)

©Saeed Barzegari, Ali Sharifi Kia, Marco Bardus, Stoyan R Stoyanov, Marjan GhaziSaeedi, Mouna Rafizadeh. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 05.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted

use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Social Determinants of Digital Health Adoption: Pilot Cross-sectional Survey

Sharvil Piyush Patel¹, BS; Elizabeth Sun¹, BA, BS; Alec Reinhardt², BS; Sanjaly Geevarghese¹, BA; Simon He¹, BA; Julie A Gazmararian³, MPH, PhD

¹Omnimed Inc, Columbus, GA, United States

²Emory University, Atlanta, GA, United States

³Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:

Sharvil Piyush Patel, BS

Omnimed Inc

5363 Veterans Parkway

Suite C

Columbus, GA, 31904

United States

Phone: 1 706 905 2971

Email: spatel@omnimedinc.org

Abstract

Background: Interest in and funding for digital health interventions have rapidly grown in recent years. Despite the increasing familiarity with mobile health from regulatory bodies, providers, and patients, overarching research on digital health adoption has been primarily limited to morbidity-specific and non-US samples. Consequently, there is a limited understanding of what personal factors hold statistically significant relationships with digital health uptake. Moreover, this limits digital health communities' knowledge of equity along digital health use patterns.

Objective: This study aims to identify the social determinants of digital health tool adoption in Georgia.

Methods: Web-based survey respondents in Georgia 18 years or older were recruited from mTurk to answer primarily closed-ended questions within the following domains: participant demographics and health consumption background, telehealth, digital health education, prescription management tools, digital mental health services, and doctor finder tools. Participants spent around 15 to 20 minutes on a survey to provide demographic and personal health care consumption data. This data was analyzed with multivariate linear and logistic regressions to identify which of these determinants, if any, held statistically significant relationships with the total number of digital health tool categories adopted and which of these determinants had absolute relationships with specific categories.

Results: A total of 362 respondents completed the survey. Private insurance, residence in an urban area, having a primary care provider, fewer urgent emergency room (ER) visits, more ER visits leading to inpatient stays, and chronic condition presence were significantly associated with the number of digital health tool categories adopted. The separate logistic regressions exhibited substantial variability, with 3.5 statistically significant predictors per model, on average. Age, federal poverty level, number of primary care provider visits in the past 12 months, number of nonurgent ER visits in the past 12 months, number of urgent ER visits in the past 12 months, number of ER visits leading to inpatient stays in the past 12 months, race, gender, ethnicity, insurance, education, residential area, access to the internet, difficulty accessing health care, usual source of care, status of primary care provider, and status of chronic condition all had at least one statistically significant relationship with the use of a specific digital health category.

Conclusions: The results demonstrate that persons who are socioeconomically disadvantaged may not adopt digital health tools at disproportionately higher rates. Instead, digital health tools may be adopted along social determinants of health, providing strong evidence for the digital health divide. The variability of digital health adoption necessitates investing in and building a common framework to increase mobile health access. With a common framework and a paradigm shift in the design, evaluation, and implementation strategies around digital health, disparities can be further mitigated and addressed. This likely will begin with a coordinated effort to determine barriers to adopting digital health solutions.

KEYWORDS

digital health; health accessibility; utilization; mobile health; mHealth; telemedicine

Introduction

In the midst of the Digital Revolution and growing focus on health care access, the US health care system has begun to openly welcome digital health with an increased expansion of mobile apps, wearable electronic devices (eg, smartwatches and fitness trackers), artificial intelligence, and telemedicine [1]. Digital health and mobile health (mHealth) encompass the clinical application of information and communications technologies to improve health and wellness management [2]. In the context of the COVID-19 pandemic, interest and funding for digital health ventures reached US \$24 billion in 2020, an increase from over US \$5 billion in 2015 [3]. Not only is digital health increasingly used in clinical settings, but it is also prevalent in consumer settings in which individuals can access digital health products on everyday gadgets such as smartphones. A 2015 cross-sectional survey found that over half of mobile phone users had downloaded a health-related app [4]. Allowing individuals to manage their health on everyday devices has transformed the way that consumers interact with their health. Simultaneously, it has created a “divided digital revolution” due to the fact that many mHealth interventions may not be socioculturally developed and consequently exclude diverse groups [5].

With increasing interest from the Food and Drug Administration, Centers for Medicare and Medicaid Services, and other regulatory agencies and payers to bring mHealth into their care strategies and oversight, the potential for digital health solutions has evolved alongside consumer wellness offerings. Digital health is now inclusive of diagnostic, disease management, and clinical decision support tools [6]. New technologies aid health care providers by reducing the repetitiveness of their work and supporting their clinical decisions, workflow, and productivity [7]. In addition to digital health technology’s ability to standardize and improve the clinical experience, digital health’s value may be even greater as a mechanism to improve accessibility in a cost-effective manner [8].

Despite the potential benefits, digital health continues to lack a proper validation system for digital health interventions. This is a particularly pressing matter given the current landscape in which many technologies have been developed and as the implementation of new technology is shifting into a growing focus [9]. Many venture-funded digital health start-ups lack clinical robustness, once again indicating the need for increased emphasis on evidence-based approaches to product development [10]. Some researchers have called for validation domains spanning technical, clinical, and system validation [1]. While some literature exists on possible solutions that would increase the validation and utility of digital health tools, there exists a gap in digital health literature regarding uptake and use by the general population [1,11]. Moreover, much of the published literature supporting the increased implementation of digital health technology in improving patient outcomes is specific to

certain patient populations and morbidities or was conducted outside the United States [1,12-16]. Such literature has provided valuable insight into the uptake of digital health tools. For example, one review focusing on older adults described 14 themes that affect the uptake of digital health, with the most common being technical literacy, lack of desire, and cost [17]. Another review focused on digital health interventions for adults with overweight and obesity highlighted attrition as a barrier to digital solutions and emphasized the need for digital health tools to be engaging for users [12]. The current specificity of research around the mHealth landscape not only provides useful information regarding digital health adoption but also indicates a need for a broader view of how individuals perceive and adopt digital health in the US health care system.

Fundamentally, empirical research in digital health and mHealth has a large gap in individual experience and adoption of broad digital health categories outside of condition-specific subgroups in the United States. Understanding the larger demographic trends and socioeconomic patterns in digital health uptake presents an important opportunity to numerically characterize the effect of social determinants within the burgeoning field of digital health. Ideally, the digital health innovation community can use original research about social determinants of digital health to think about and respond to the structural pillars of the digital health divide. In doing so, innovators can design mHealth interventions with an eye for accessibility. To achieve this goal, this study represents an initial, pilot investigation of the “social determinants of digital health” throughout the state of Georgia. The analysis statistically describes the use patterns of common digital health tools among web-based survey respondents in Georgia. By doing so, this investigation aims to understand how socioeconomic factors and care-seeking behaviors affect digital health uptake, and compares the adoption distribution across 8 common categories of digital health tools.

Methods

Study Design

This was a cross-sectional study in the area of digital health technology using survey data collected from June to November 2021. Participants responded to a web-based survey administered through REDCap that consisted of 172 possible questions (available upon request) to explore the use of and attitudes toward 8 common digital health tools. Common demographic questions and health care consumption questions were used as predictors of the social determinants of digital health use and were based on questions used in the National Health Interview Survey project [18]. Conditional and branching logic based on prior responses were used to pare down the questionnaire to deliver the right questions for the right participant.

Study Population

Individuals were recruited from Amazon's mTurk platform [19]. While patients were initially going to be recruited in a clinic setting, this became a safety hazard because of the COVID-19 pandemic and required the study team to use digital recruitment methods. In the context of this issue, mTurk provided access to a large heterogeneous population of willing research participants. The inclusion criteria required participants with existing mTurk worker accounts to be aged ≥ 18 years and be residents of the state of Georgia in the United States. Participants were compensated US \$3 through mTurk's internal platform. The data set began with 1022 records from which 630 duplicates, 29 incomplete responses, and 1 invalid survey were removed. Thus, the final analytic sample included 362 participants' responses.

Data Source

Survey questions were designed by the study team to investigate the use rates of different digital health tools and what factors may shape an individual's use of digital health tools. The survey instrument contained primarily closed-ended question types with some opportunities to provide open-ended responses if the participant was willing. The questionnaire encompassed 172 possible questions in the English language that entailed the following domains: participant demographics and health consumption background, telehealth, digital health education, prescription management tools, digital mental health services, and doctor finder tools (available on request). The survey took approximately 15 to 20 minutes for participants to complete. Questions were presented to each participant in the same order. Demographic data and questions about health care use were also implemented to understand how social factors may impact digital health tool use.

Study data were collected and managed using REDCap electronic data capture tools hosted at Emory University [20,21]. REDCap is a secure web-based software platform designed to support data capture for research studies, providing an intuitive interface for validated data capture, audit trails for tracking data manipulation and export procedures, automated export procedures for seamless data downloads to common statistical packages, and procedures for data integration and interoperability with external sources.

Data Measures

The analysis included 17 independent variables. There were 2 continuous independent variables: age (as reported by the respondent) and federal poverty level (calculated using household income and household size). There were 15 self-reported categorical variables: race, gender, ethnicity, insurance status, educational attainment, living area, access to the internet, access to health care, primary care-seeking behavior in non-life-threatening events (as defined by the type of clinic

visited during said events), having a routine primary care provider (PCP), amount of PCP visits in the last 12 months, amount of emergency room (ER) visits for nonurgent events in the last 12 months, amount of ER visits for urgent events in the last 12 months, amount of ER visits leading to inpatient stays in the last 12 months, and presence of chronic conditions (actual survey items available on request).

There was 1 continuous dependent variable (the number of digital health tool categories used by a patient) and 8 yes or no questions for self-reported digital health tool use in the following categories: telehealth, digital health education, prescription management tools, doctor finders, social services referral tools, digital mental health tools, digital insurance navigators, and patient portals.

Statistical Analysis

First, descriptive statistics were calculated to determine the frequency of using the 8 digital health categories. Second, an adjusted linear regression was completed between the independent variables and the number of digital health tool categories adopted by a respondent. Third, separate logistic regressions were used to identify predictors of use for each category of tools. Statistical significance was determined by *P* values $< .05$. All statistical analyses were completed using R version 4.0.5 (R Foundation for Statistical Computing) with the packages tidyverse, MASS, and mice [22-25]. Multiple imputation based on mean-matching was used to impute missing data within the sample. Imputation required at least 50% or more of the existing data, and all of the included independent and dependent variables required some imputation.

Ethical Considerations

This study was approved by Emory University's institutional review board (reference STUDY00001999). Participants were recruited from Amazon's mTurk platform and were compensated US \$3 for their participation through mTurk's internal system. All participation was voluntary, and no participant was subject to any harm. Informed consent was obtained from each participant regarding what their involvement in the study entailed and how their responses would be handled. The privacy of research participants was maintained throughout the study, and all responses were deidentified.

Results

Over half ($n=189$, 52.2%) of the 362 respondents reported having private insurance, and more than half ($n=160$, 60.2%) of them had earned at least a bachelor's degree (Table 1). Almost all ($n=302$, 98.6%) respondents reported always or almost always having reliable access to the internet. About one-third ($n=118$, 34.7%) of the respondents reported having at least one chronic condition. On average, respondents used more than 3 digital health tools.

Table 1. Descriptive statistics of the respondents presented as percentages, means, or medians (n=362).

Variables	Value
Age (years)	
Mean (SD)	37 (0.6)
Median (range)	34 (18-73)
Federal poverty line ratio, mean (SD)	3 (0.28)
Number of PCP^a visits in the past 12 months	
Mean (SD)	4 (0.14)
Median (range)	3 (1-11)
Number of ER^b visits for nonurgent issues in the past 12 months	
Mean (SD)	2 (0.1)
Median (range)	1 (1-10)
Number of ER visits for urgent issues in the past 12 months	
Mean (SD)	2 (0.11)
Median (range)	1 (1-11)
Number of ER visits leading to inpatient stay	
Mean (SD)	4 (0.05)
Median (range)	1 (1-10)
Difficulty accessing health care (Likert scale: 1 is very hard, 5 is very easy)	
Mean (SD)	4 (0.05)
Median (range)	4 (1-5)
Seek care in non-life-threatening situations, n (%)	
Primary care physician	191 (52.8)
Urgent care	111 (30.7)
Emergency room	30 (8.3)
Other	30 (8.3)
Has PCP, n (%)	
Yes	275 (76.0)
No	87 (24.0)
Presence of chronic condition, n (%)	
Yes	118 (34.7)
No	222 (61.3)
No response	22 (6.0)
Race, n (%)	
Non-White	104 (28.7)
White	253 (65.3)
No response	5 (1.4)
Gender, n (%)	
Male	218 (60.9)
Female	140 (39.1)
No response	4 (1.1)
Ethnicity, n (%)	
Hispanic	32 (8.8)

Variables	Value
Not Hispanic	322 (89.0)
No response	8 (2.2)
Insurance, n (%)	
Public insurance	95 (26.2)
Private insurance	189 (52.2)
A mix of both public and private insurance	23 (6.4)
I have insurance, but I'm not sure what type	8 (2.2)
Uninsured	47 (13.0)
Education, n (%)	
High school or less	27 (7.5)
Some college, no degree	85 (23.5)
Associate's degree	32 (8.8)
Bachelor's degree	160 (44.2)
Master's degree	49 (13.5)
Professional/doctoral degree	9 (2.5)
Living area, n (%)	
Urban	104 (28.7)
Suburban	187 (51.7)
Rural	71 (19.6)
Internet access, n (%)	
Always have access	302 (83.4)
Does not always have access	55 (15.2)
No response	5 (1.4)
Number of digital health tool categories adopted	
Mean (SD)	4 (0.11)
Median (range)	4 (0-8)
Telehealth use, n (%)	
No	140 (38.7)
Yes	222 (61.3)
Health education tool use, n (%)	
No	124 (34.3)
Yes	238 (65.7)
Prescription management tool use, n (%)	
No	232 (64.1)
Yes	130 (35.9)
Doctor finder tool use, n (%)	
No	185 (51.1)
Yes	177 (48.9)
Social service referral tool use, n (%)	
No	289 (79.8)
Yes	73 (20.2)
Mental health tool use, n (%)	
No	239 (66.0)

Variables	Value
Yes	123 (34.0)
Insurance management tool use, n (%)	
No	239 (66.0)
Yes	123 (34.0)
Patient portal use, n (%)	
No	123 (34.0)
Yes	239 (66.0)

^aPCP: primary care provider.

^bER: emergency room.

Results from the linear regression analysis indicate that for total digital tool adoption count, the model was statistically significant ($P < .001$; adjusted $R^2 = 0.28$; Table 2). Six variables were statistically significant at the $\alpha < .05$ level. On average, respondents with private insurance used 0.55 more digital health tools than respondents with public insurance ($P = .03$). Respondents living in rural or suburban areas used 0.75 fewer digital health tools than respondents in urban areas ($P = .003$ and $.02$, respectively). Patients without a PCP used 1.25 fewer digital health tools than respondents with a PCP ($P = .005$), while respondents with no chronic condition used 0.73 fewer tools than respondents with chronic conditions ($P = .001$). A 1-unit increase in the number of urgent ER visits was associated with a 0.4 decrease in digital health tools used ($P = .04$), but a 1-unit

increase in the number of ER visits leading to inpatient stays resulted in the use of 0.64 more digital health tools ($P = .002$), on average. No multicollinearity was observed.

None of the predictors held a statistically significant relationship with the outcomes across all 8 logistic regression models (Multimedia Appendix 1). In fact, the total number of significant variables across each logistic regression model varied greatly, with an average of 3.5 statistically significant predictors in a model. For example, the logistic regression model for the doctor finder only had a significant relationship with 1 variable (residential area), whereas telehealth use had 6 significant predictors (number of PCP visits in the past 12 months, self-reported race, insurance status, residential area, PCP status, and chronic condition status).

Table 2. Estimates for the linear regression model on the total number of digital health tool categories adopted (n=362).

Variables	β (95% CI)	P value
Intercept	4.08 (3.06 to 5.1)	<.001
Age	-0.06 (-0.24 to 0.12)	.49
Federal poverty level	-0.01 (-0.18 to 0.17)	.96
Number of PCP ^a visits in past 12 months	0.05 (-0.21 to 0.31)	.71
Number of nonurgent ER ^b visits in past 12 months	0.31 (0 to 0.63)	.05
Number of urgent ER visits in past 12 months	-0.31 (-0.67 to 0.04)	.08
Number of ER visits leading to inpatient stays in past 12 months	0.58 (0.2 to 0.95)	<.001
Difficulty	-0.02 (-0.21 to 0.18)	.87
Seek care in non-life-threatening situations		
Primary care physician (reference)	N/A ^c	N/A
Urgent care	0.42 (-0.01 to 0.85)	.06
Emergency room	0.28 (-0.43 to 0.98)	.44
Other	0.51 (-0.19 to 1.21)	.15
Has PCP		
Has PCP (reference)	N/A	N/A
No PCP	-0.61 (-1.73 to 0.05)	.07
Presence of chronic condition		
Has chronic condition (reference)	N/A	N/A
No chronic condition	-0.72 (-1.1 to -0.33)	<.001
Race		
Non-White (reference)	N/A	N/A
White	0.20 (-0.2 to 0.59)	.34
Gender		
Female (reference)	N/A	N/A
Male	-0.38 (-0.76 to 0)	.05
Ethnicity		
Hispanic (reference)	N/A	N/A
Non-Hispanic	0.00 (-0.59 to 0.58)	.99
Insurance		
Public (reference)	N/A	N/A
Private	0.37 (-0.09 to 0.83)	.11
Mix	0.24 (-0.53 to 1.01)	.54
Unsure	-0.18 (-1.41 to 1.06)	.78
Uninsured	-0.52 (-1.21 to 0.18)	.14
Education		
Professional degree (reference)	N/A	N/A
Master's	-0.29 (-1.57 to 0.98)	.65
Bachelor's	-0.20 (-1.01 to 0.61)	.63
Associate's	-0.18 (-0.88 to 0.51)	.60
Some college	-0.23 (-1.09 to 0.63)	.60
Less than high school or high school graduate/GED ^d	-0.65 (-1.37 to 0.08)	.08

Variables	β (95% CI)	P value
Living area		
Urban (reference)	N/A	N/A
Suburban	-0.61 (-1.06 to -0.16)	.01
Rural	-0.62 (-1.17 to -0.06)	.03
Internet		
Always has access (reference)	N/A	N/A
Does not always have access	-0.25 (-0.76 to 0.25)	.32

^aPCP: primary care provider.

^bER: emergency room.

^cN/A: not applicable.

^dGED: General Educational Development.

Discussion

Principal Findings

Given the paucity of literature surrounding the use of common digital health tools by the general population in the United States, this paper is a novel investigation on how 8 common forms of digital health management tools are adopted and used in a diverse population in Georgia. Surprisingly, there are large variations in how many and which factors predict digital health use uptake. The large variability indicates that there is likely diversity in the populations using various types of digital health tools. Moreover, this analysis' emphasis on the variability in the adoption of digital health potentially highlights a gap between mHealth's ideal and actual implementations; this is especially true when relationships that would logically be meaningful did not appear in the analysis. Ideally, mHealth would be a tool that allows medically indigent populations to improve their access to health care and to community-based resources that empower their care navigation efforts. Specifically, these findings highlight that digital health innovations are not always distributed equitably or to the people who would ostensibly need it the most. For example, the use of social services referral tools did not vary with education level, federal poverty level ratio, or lack of insurance. Ideally, socioeconomically disadvantaged populations—proxied for by education level, federal poverty level, and insurance status—would have significantly higher rates of adoption for social services referral tools since these tools are meant to expand this population's access to social services. Similarly, digital prescription management tools had no statistically significant relationship with insurance status, which levies a similar concern since patients who are uninsured should be more likely to use digital tools to lower out-of-pocket prescription costs when holding the chronic condition status constant. Fundamentally, this initial investigation spotlights the need for incorporating more real-world evidence into digital health distribution, use, and outcomes beyond the controlled trial environment.

Interventions can and should be targeted to reach patients that traditionally cannot access health care options; however, these results indicate that, in practice, the patients with the greatest need may not find these technologies accessible. If this is

evidence that the most socially and medically indigent patients cannot access new innovations, inequality can either become a substantial barrier to improving health outcomes at scale or even exacerbate health disparities across a digital divide. With the incorporation of on-the-ground use data, innovators and researchers can assess whether inequalities in uptake are a result of the inequalities in patient access to prerequisite infrastructure (eg, smartphones, Wi-Fi, or computers) or if some current investigations of digital health interventions may simply lack external validity, thus requiring a paradigm shift in the design, evaluation, and implementation strategies around digital health. These observations align with the fundamental shift in health intervention evaluation methodologies. More specifically, there is an increasing emphasis on generating and using real-world evidence and data to measure digital health efficacy [26].

These results functionally support the extant literature's depiction of digital health inequality across various digital health domains and diverse populations. For example, Tappen et al [27] described a "deep digital health divide" in older populations that is associated with age, education, income, and ethnicity. Likewise, Saeed and Masters [28] noted that in psychiatric conditions, telehealth use is negatively correlated with lower socioeconomic status. Moreover, Brown et al [16] found that significant disparities exist in access to telemedical care among cardiovascular patients that are low-income, older adults, or Black or Hispanic [16]. Perhaps unsurprisingly, historic drivers of inequities in health care continue to exist in a similar manner in the digital health space. The impact of social determinants of health goes beyond access to medical care and pure health outcomes. There is an apparent impact on how the social determinants can affect access to digital health technologies, which may be a mediating variable for health outcomes and access to care. Fortunately, there is a movement toward prioritizing contextually tailored mHealth interventions that can mitigate inequalities in access to digital health [5].

Strengths and Limitations

The self-reported data collected from this investigation has at least three strengths. First, the large sample size provided minimizes concerns about response bias and improves statistical strength. Moreover, the use of mTurk to recruit respondents allows for a much larger pool than the original target population,

focusing on patient-defined populations. As a result, this study included a more diverse panel of respondents. Second, the survey examined data across a broad array of digital health tools rather than focusing on one specific category of tools. As a result, this novel data set provides a broad view of digital health uptake across multiple subsets of patients. This also provides better comparisons across the various classes of commercialized technology. Third, the variety of independent predictors allows for better adjustment in the linear and logistic regression models to control for confounding variables' bias.

Despite these strengths, there are at least three limitations. First, respondents were recruited and surveyed using a web-based platform (mTurk), which inherently introduces the effects of self-selection. For instance, this study analyzes how digital (ie, internet-connected and internet-enabled) health tools are used by asking questions to respondents who therefore are adept with technology and have substantial access to digital platforms before the survey. Second, the platform does not provide the opportunity to randomly sample respondents, limiting generalizability. Third, the cross-sectional nature of the data and analysis precludes causal induction. This analysis cannot link respondent behavior over time with the changes in their independent predictors' statuses. In the future, semicausal and causal methods should be used to establish cause-and-effect relationships.

Future Directions

Further work to identify and reduce barriers to entry for digital health tools is vital to expanding its impact and promoting equity with the advent of new technologies. Understanding how digital health disseminates will be crucial to intervention design, implementation, and evaluation. The existing variability in the adoption of mHealth underscores the lack of a common framework to increase mHealth use across diverse patient populations, which ultimately limits an intervention's potential

for success. Innovative strategies can help improve access moving forward; however, this will require digital health innovators and regulators to regularly collect and review data regarding mHealth adoption. This can help evaluate the social and clinical returns on funding for digital health interventions, especially those aimed at improving health care access.

The large variability in adoption across digital health tools indicates that a one-size-fits-all approach to deploying these tools will likely mitigate their potential impact. Consequently, additional research is needed to better understand these patterns at larger scales and across more diverse populations, and the full range of factors that contribute to intervention uptake. Similarly, the digital health innovation community should rebuild the evaluation framework for evaluating tool distribution and adoption throughout the intervention's lifecycle to ensure that patients are able to equitably access these services. More specifically, special consideration should be given to how new tools can contribute to the digital health divide, so when new tools are deployed, innovators should look at the characteristics of users to determine if adoption is equitably distributed.

Conclusions

This investigation establishes an initial portrait of how variable the use of digital health tools is across various patient demographics. Moreover, these results indicate that populations who could benefit the most from using certain tools (eg, patients who are socioeconomically disadvantaged and who would benefit from a social services referral tool) are not using these technologies. Although these tools are already publicly available, populations that could realize substantial benefits from this technology experience larger barriers to entry and sustained use (eg, information, internet, and cost obstacles). Ultimately, while these tools can be valuable, user uptake is the most important prerequisite to clinical and social utility.

Acknowledgments

This study was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award UL1TR002378.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Estimates for the logistic regression models of use of tools in digital health categories (n=362).

[[PDF File \(Adobe PDF File\), 66 KB - formative_v6i12e39647_app1.pdf](#)]

References

1. Whitelaw S, Pellegrini D, Mamas M, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. *Eur Heart J Digit Health* 2021 Mar;2(1):62-74 [[FREE Full text](#)] [doi: [10.1093/ehjdh/ztab005](https://doi.org/10.1093/ehjdh/ztab005)] [Medline: [34048508](https://pubmed.ncbi.nlm.nih.gov/34048508/)]
2. Ronquillo Y, Meyers A, Korvek SJ. *Digital Health*. Treasure Island, FL: StatPearls Publishing; Jan 2022.
3. Aitken M. Digital health trends 2021: innovation, evidence, regulation, and adoption. IQVIA. 2021 Jul. URL: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/digital-health-trends-2021/iqvia-institute-digital-health-trends-2021.pdf?_af=1652132450364 [accessed 2022-11-23]

4. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov 04;3(4):e101 [[FREE Full text](#)] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
5. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the future: achieving health equity through health informatics and digital health. *JMIR Mhealth Uhealth* 2020 Jan 14;8(1):e14512 [[FREE Full text](#)] [doi: [10.2196/14512](https://doi.org/10.2196/14512)] [Medline: [31934874](https://pubmed.ncbi.nlm.nih.gov/31934874/)]
6. Sharma A, Harrington RA, McClellan MB, Turakhia MP, Eapen ZJ, Steinhubl S, et al. Using digital health technology to better generate evidence and deliver evidence-based care. *J Am Coll Cardiol* 2018 Jun 12;71(23):2680-2690 [[FREE Full text](#)] [doi: [10.1016/j.jacc.2018.03.523](https://doi.org/10.1016/j.jacc.2018.03.523)] [Medline: [29880129](https://pubmed.ncbi.nlm.nih.gov/29880129/)]
7. Mesko B, Györfy Z. The rise of the empowered physician in the digital health era: viewpoint. *J Med Internet Res* 2019 Mar 26;21(3):e12490 [[FREE Full text](#)] [doi: [10.2196/12490](https://doi.org/10.2196/12490)] [Medline: [30912758](https://pubmed.ncbi.nlm.nih.gov/30912758/)]
8. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851 [[FREE Full text](#)] [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
9. Buis L. Implementation: the next giant hurdle to clinical transformation with digital health. *J Med Internet Res* 2019 Nov 20;21(11):e16259 [[FREE Full text](#)] [doi: [10.2196/16259](https://doi.org/10.2196/16259)] [Medline: [31746763](https://pubmed.ncbi.nlm.nih.gov/31746763/)]
10. Day S, Shah V, Kaganoff S, Powelson S, Mathews SC. Assessing the clinical robustness of digital health startups: cross-sectional observational analysis. *J Med Internet Res* 2022 Jun 20;24(6):e37677 [[FREE Full text](#)] [doi: [10.2196/37677](https://doi.org/10.2196/37677)] [Medline: [35723914](https://pubmed.ncbi.nlm.nih.gov/35723914/)]
11. Shaw RJ, Steinberg DM, Bonnet J, Modarai F, George A, Cunningham T, et al. Mobile health devices: will patients actually use them? *J Am Med Assoc* 2016 May;23(3):462-466 [[FREE Full text](#)] [doi: [10.1093/jamia/ocv186](https://doi.org/10.1093/jamia/ocv186)] [Medline: [26911820](https://pubmed.ncbi.nlm.nih.gov/26911820/)]
12. Beileigoli AM, Andrade AQ, Caçado AG, Paulo MN, Diniz MDFH, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. *J Med Internet Res* 2019 Jan 08;21(1):e298 [[FREE Full text](#)] [doi: [10.2196/jmir.9609](https://doi.org/10.2196/jmir.9609)] [Medline: [30622090](https://pubmed.ncbi.nlm.nih.gov/30622090/)]
13. Aapro M, Bossi P, Dasari A, Fallowfield L, Gascón P, Geller M, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. *Support Care Cancer* 2020 Oct;28(10):4589-4612 [[FREE Full text](#)] [doi: [10.1007/s00520-020-05539-1](https://doi.org/10.1007/s00520-020-05539-1)] [Medline: [32533435](https://pubmed.ncbi.nlm.nih.gov/32533435/)]
14. Slevin P, Kessie T, Cullen J, Butler MW, Donnelly SC, Caulfield B. Exploring the potential benefits of digital health technology for the management of COPD: a qualitative study of patient perceptions. *ERJ Open Res* 2019 Apr;5(2):00239-2018 [[FREE Full text](#)] [doi: [10.1183/23120541.00239-2018](https://doi.org/10.1183/23120541.00239-2018)] [Medline: [31111039](https://pubmed.ncbi.nlm.nih.gov/31111039/)]
15. Foley K, Freeman T, Ward P, Lawler A, Osborne R, Fisher M. Exploring access to, use of and benefits from population-oriented digital health services in Australia. *Health Promot Int* 2021 Aug 30;36(4):1105-1115. [doi: [10.1093/heapro/daaa145](https://doi.org/10.1093/heapro/daaa145)] [Medline: [33367568](https://pubmed.ncbi.nlm.nih.gov/33367568/)]
16. Brown K, Mathenge N, Crousillat D. Social determinants of telemedicine utilization in ambulatory cardiovascular patients during the COVID-19 pandemic. *Eur Heart J* 2021;2(2):244-253. [doi: [10.1093/ehjdh/ztab039](https://doi.org/10.1093/ehjdh/ztab039)]
17. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. *JMIR Med Inform* 2020 Aug 12;8(8):e20359 [[FREE Full text](#)] [doi: [10.2196/20359](https://doi.org/10.2196/20359)] [Medline: [32784177](https://pubmed.ncbi.nlm.nih.gov/32784177/)]
18. National Center for Health Statistics. National health interview survey. Centers for Disease Control and Prevention. 2022 Aug 24. URL: <https://www.cdc.gov/nchs/nhis/index.htm> [accessed 2022-11-23]
19. Amazon Mechanical Turk. URL: <https://www.mturk.com/> [accessed 2022-04-28]
20. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
21. 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](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
22. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Software* 2011;45(3):1-67. [doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03)]
23. Venables WN, Ripley BD. *Modern Applied Statistics with S*. New York, NY: Springer; 2002.
24. R Core Team. The R Project for Statistical Computing. 2021. URL: <https://www.R-project.org/> [accessed 2022-11-23]
25. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *J Open Source Software* 2019 Nov;4(43):1686. [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
26. Khosla S, Tepie MF, Nagy MJ, Kafatos G, Seewald M, Marchese S, et al. The alignment of real-world evidence and digital health: realising the opportunity. *Ther Innov Regul Sci* 2021 Jul;55(4):889-898 [[FREE Full text](#)] [doi: [10.1007/s43441-021-00288-7](https://doi.org/10.1007/s43441-021-00288-7)] [Medline: [33914297](https://pubmed.ncbi.nlm.nih.gov/33914297/)]
27. Tappen RM, Cooley ME, Luckmann R, Panday S. Digital health information disparities in older adults: a mixed methods study. *J Racial Ethn Health Disparities* 2022 Feb;9(1):82-92 [[FREE Full text](#)] [doi: [10.1007/s40615-020-00931-3](https://doi.org/10.1007/s40615-020-00931-3)] [Medline: [33415705](https://pubmed.ncbi.nlm.nih.gov/33415705/)]

28. Saeed SA, Masters RM. Disparities in health care and the digital divide. *Curr Psychiatry Rep* 2021 Jul 23;23(9):61 [FREE Full text] [doi: [10.1007/s11920-021-01274-4](https://doi.org/10.1007/s11920-021-01274-4)] [Medline: [34297202](https://pubmed.ncbi.nlm.nih.gov/34297202/)]

Abbreviations

ER: emergency room

mHealth: mobile health

PCP: primary care provider

Edited by A Mavragani; submitted 08.06.22; peer-reviewed by N Shah, H Ngusie, YC Liu; comments to author 17.08.22; revised version received 29.08.22; accepted 12.09.22; published 06.12.22.

Please cite as:

*Patel SP, Sun E, Reinhardt A, Geevarghese S, He S, Gazmararian JA
Social Determinants of Digital Health Adoption: Pilot Cross-sectional Survey
JMIR Form Res 2022;6(12):e39647*

URL: <https://formative.jmir.org/2022/12/e39647>

doi: [10.2196/39647](https://doi.org/10.2196/39647)

PMID: [36472905](https://pubmed.ncbi.nlm.nih.gov/36472905/)

©Sharvil Piyush Patel, Elizabeth Sun, Alec Reinhardt, Sanjaly Geevarghese, Simon He, Julie A Gazmararian. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Data Quality and Study Compliance Among College Students Across 2 Recruitment Sources: Two Study Investigation

Abby L Braitman¹, PhD; Megan Strowger¹, MS; Jennifer L Shipley¹, MS, MPH; Jordan Ortman¹, BS; Rachel I MacIntyre^{1,2}, PhD; Elizabeth A Bauer³, BS

¹Department of Psychology, Old Dominion University, Norfolk, VA, United States

²Department of Psychology, Millersville University, Millersville, PA, United States

³Department of Psychological and Brain Sciences, Texas A&M University, College Station, TX, United States

Corresponding Author:

Abby L Braitman, PhD
Department of Psychology
Old Dominion University
250 Mills Godwin Bldg
Norfolk, VA, 23529
United States
Phone: 1 757 683 3708
Email: abraitma@odu.edu

Abstract

Background: Models of satisficing suggest that study participants may not fully process survey items and provide accurate responses when survey burden is higher and when participant motivation is lower. Participants who do not fully process survey instructions can reduce a study's power and hinder generalizability. Common concerns among researchers using self-report measures are data quality and participant compliance. Similarly, attrition can hurt the power and generalizability of a study.

Objective: Given that college students comprise most samples in psychological studies, especially examinations of student issues and psychological health, it is critical to understand how college student recruitment sources impact data quality (operationalized as attention check items with directive instructions and correct answers) and retention (operationalized as the completion of follow-up surveys over time). This examination aimed to examine the following: whether data quality varies across recruitment sources, whether study retention varies across recruitment sources, the impact of data quality on study variable associations, the impact of data quality on measures of internal consistency, and whether the demographic qualities of participants significantly vary across those who failed attention checks versus those who did not.

Methods: This examination was a follow-up analysis of 2 previously published studies to explore data quality and study compliance. Study 1 was a cross-sectional, web-based survey examining college stressors and psychological health (282/407, 69.3% female; 230/407, 56.5% White, 113/407, 27.8% Black; mean age 22.65, SD 6.73 years). Study 2 was a longitudinal college drinking intervention trial with an in-person baseline session and 2 web-based follow-up surveys (378/528, 71.6% female; 213/528, 40.3% White, 277/528, 52.5% Black; mean age 19.85, SD 1.65 years). Attention checks were included in both studies to assess data quality. Participants for both studies were recruited from a psychology participation pool (a pull-in method; for course credit) and the general student body (a push-out method; for monetary payment or raffle entry).

Results: A greater proportion of participants recruited through the psychology pool failed attention checks in both studies, suggesting poorer data quality. The psychology pool was also associated with lower retention rates over time. After screening out those who failed attention checks, some correlations among the study variables were stronger, some were weaker, and some were fairly similar, potentially suggesting bias introduced by including these participants. Differences among the indicators of internal consistency for the study measures were negligible. Finally, attention check failure was not significantly associated with most demographic characteristics but varied across some racial identities. This suggests that filtering out data from participants who failed attention checks may not limit sample diversity.

Conclusions: Investigators conducting college student research should carefully consider recruitment and include attention checks or other means of detecting poor quality data. Recommendations for researchers are discussed.

(*JMIR Form Res* 2022;6(12):e39488) doi:[10.2196/39488](https://doi.org/10.2196/39488)

KEYWORDS

data quality; attention checks; recruitment; retention; college students; mobile phone

Introduction

Background

The validity of the findings of any study hinges on the integrity of the data collected. Common concerns among psychology researchers using self-report measures are data quality and participant compliance. Participants may not fully read or process the self-report measure instructions or items, adding noise to data rather than reflecting the constructs being assessed, or they may not complete the study protocol, reducing the number of assessments used for analyses. Both reduce the study's power [1] and may hinder the generalizability of its findings [2,3]. Given that clinical trials of psychological treatments are chronically underpowered [4], reduced power due to poor data quality can exacerbate the lack of trust in the study findings. To prevent these negative effects on statistical power and external validity, researchers may aim to recruit compliant participants, better incentivize compliance, and detect and remove noncompliant participants from data sets. Given that college students comprise most samples in psychological studies [5-7], it is necessary to understand the associations between recruitment sources targeting college students and participant compliance. For this paper, we defined participant compliance as providing high quality data (ie, putting in reasonable effort and fully reading each item before responding) and completing follow-up assessments (ie, retention; only applicable for longitudinal studies).

One approach to identifying participant noncompliance that affects data quality is to use attention check items with instructions to select a particular answer (eg, "Select 'slightly agree' for this item") or that have factual answers (eg, "Which number is largest?"); these are also called instructional manipulation checks [1], bogus items [8], infrequency scales [9], or random-responding indicators [10]. These items can identify participants who are satisficing (ie, putting in minimal effort and potentially not fully reading or comprehending each item), which is sometimes called careless responding. Removing these participants may increase statistical power such that correlations are stronger among relevant study variables and experimental effects across conditions are larger [1,10-12] or may otherwise reduce "noise" among study variable associations [13]. Identifying poor quality data where responses may not reflect the true study construct via attention checks and removing these satisficing cases may reduce random error and increase statistical power, but it also may result in removing a meaningful group and introducing bias, as certain demographics (eg, gender, age, race, education, and intrinsic motivation) can be associated with satisficing [11,14]. It is possible that some recruitment sources may yield participants who are not only less inclined to satisfice but are also more demographically diverse, allowing for the removal of satisficing participants without limiting the diversity and generalizability of the sample.

In a model explaining how individuals formulate and respond to survey questions, Tourangeau et al [15] proposed that poor

data quality comes from failing to engage in at least one of the four stages of the cognitive processing: (1) understanding the meaning of the item, (2) finding information in memory relevant to the item, (3) summarizing the information found, and (4) using that information to choose a response given the options. Satisficing could result from failing to engage in any one of these stages of cognitive processing. However, the attention checks used in this examination as an indicator of data quality are designed to detect the most egregious forms of satisficing (stage 1: not fully reading and understanding the item) rather than the later forms of cognitive processing (eg, taking the time to process exactly how anxious they remember being or whether this is best reflected as an endorsement of 4 vs 5 for a given item). The study by Krosnick [16] gave an overview of a variety of response strategies that respondents may choose when engaging in satisficing to conserve their mental energy and suggested that some individuals may choose to engage in satisficing to conserve their mental energy when faced with great task difficulty (eg, many items in a survey) and that the likelihood of satisficing increases as the burden of the survey increases, and participants become less motivated to perform well as they become more fatigued. Moreover, it suggested that respondents may first be less diligent about the later stages of cognitive processing (eg, taking the time to decide between 4 and 5) before fully omitting stages (eg, not fully reading the item). A systematic review of 141 studies that included various indicators of satisficing revealed that 74% of the studies found that task difficulty was significantly associated with satisficing, and 68% of the studies found that respondent motivation was significantly associated with satisficing [17], suggesting that satisficing is a result of the qualities of both the survey task (highly burdensome) and the participant (low motivation).

Another issue of participant compliance particularly important for longitudinal research such as examinations of psychological health is study retention rates. Many studies require multiple assessments, such as observing natural developmental trajectories or following changes in behaviors, symptoms, or attitudes after the administration of an intervention. Although interventions offer additional benefits to participants compared with nonintervention research (eg, potential improvements in mental or physical health), these benefits are often obtained immediately and do not extend to incentivizing retention for follow-up surveys. Moreover, follow-up assessments are typically administered remotely and after substantial time has passed (eg, weeks or months). Therefore, retention rates typically drop with each additional follow-up. For example, challenges with retention have been noted in a meta-analysis of cohort studies on mental and behavioral health [18], a meta-analysis of dissonance-based interventions for health behavior change [19], and a meta-analysis of digital interventions for the treatment and prevention of eating disorders [20]. This is particularly challenging for studies involving college students. An integrative analysis of 24 studies of brief interventions for college drinking found retention rates as low as 46% for the 6-month follow-ups and 51% for the 9- to

12-month follow-ups [21]. The authors of these studies have noted how challenging it can be to retain college students, particularly when administering interventions such as those for college drinking [22,23] or web-based programs for students with depression, anxiety, or stress [24]. Identifying recruitment methods that are associated with better data quality and higher compliance (eg, passing attention checks and completing study protocols including follow-up assessments) may reduce the costs associated with longitudinal research, increase the benefits and practicality of the study designs, and strengthen the trust in the study findings.

Recruitment Sources

Undergraduate college students serve as study samples in most psychology studies, a consistent trend over time. An investigation of 6 top journals in multiple fields of psychology over 20 years (1975, 1985, and 1995) included 1559 articles with human participants [7]. Researchers found that most studies (68%) exclusively used undergraduate college student samples and that this finding was consistent over time (69.8% in 1975, 66.7% in 1985, and 68.2% in 1995). An investigation into 1 specific premier journal, the *Journal of Personality and Social Psychology*, revealed that 67% of American study samples were specifically undergraduate students enrolled in psychology courses, and this rose to 80% of samples for non-American studies [5]; however, this number has dropped to 42% overall (39% of American-based studies and 54% of non-American studies) in more recent years [6]. This makes it imperative to examine study compliance among college students and, in particular, whether the use of psychology student pools has an impact.

Student participant pools have received both praise and criticism from the psychology community [25,26]. They provide researchers with a low-cost and efficient recruitment source, which may be particularly important for student researchers who do not have funding [26]. Although there are concerns that student participant pools mainly comprise female, White, and young psychology majors [26,27] and that this can result in samples that do not generalize beyond Western, educated, industrialized, rich, and democratic societies [28], student participation pools are becoming more demographically diverse, mirroring the increasing diversity among those attending college [26]. Moreover, some research questions focus specifically on student populations (eg, studies that focus on unique college stressors and their links to mental health or interventions targeting college drinking), creating a need for student participant sources.

Student participation pools may potentially reflect the true population of interest in addition to being convenient; however, these pools can also be associated with lower enrollment and study compliance. Sharpe and Poets [26] found that as many as 56.7% of students in 2 large introductory courses chose not to participate in research or earn any research credits. Motivational issues and time commitment are the 2 primary factors linked to student nonparticipation in research pools [29,30]. It is possible that factors contributing to low motivation to participate may also impact study compliance among those who choose to participate.

Antoun et al [31] suggested that the type of recruitment method may explain the differences in data quality, differentiating between pull-in and push-out recruitment. They defined pull-in recruitment methods as those that post studies to participant pools already opting into research to some degree, such as Amazon Mechanical Turk (MTurk) workers or individuals looking for paid research opportunities on Craigslist (similar to student participant pools at academic institutions), whereas push-out recruitment uses methods in which advertisements are posted in venues not already focused on research, such as advertisements on websites not dedicated to the purpose of recruiting study participants (eg, Facebook advertisements), flyers, and email blasts. In a study comparing pull-in versus push-out approaches to recruit iPhone users for a cross-sectional web-based survey, Antoun et al [31] found that pull-in methods (using Craigslist and MTurk) were more efficient in recruiting participants, in that the rate of enrollment was faster and the cost per participant who enrolled in the study was lower, than push-out methods (using paid advertisements on Google and Facebook). Although no attention checks were included in the data collection, the authors concluded that the participants recruited through pull-in methods provided better data (ie, fewer “don’t know” responses and fewer skipped or incomplete responses), possibly indicating less satisficing. Multiple studies have extended these findings by recruiting samples across both pull-in and push-out approaches and including attention checks as indicators of data quality. One such study recruited participants from MTurk (pull-in), Facebook (push-out), and Qualtrics panels (pull-in) and included 1 attention check [32]. They found that the rate of passing the attention check question was highest among the participants recruited through MTurk (93%) compared with those recruited through Facebook advertisements (66%) and Qualtrics panels (40%). The participants recruited through MTurk endorsed “don’t know” responses only 0.4% of the time compared with those recruited through Facebook (4%) and Qualtrics panels (5%). These findings suggested that the push-out and pull-in distinction may be less important than the sources, given that both the highest (MTurk) and lowest (Qualtrics panels) rates of attention check failure were associated with pull-in sources. A similar study found that MTurk samples were more likely to pass the attention check (97.5%) than panel respondents via Dynata (91.6%), which were both pull-in sources [33]. These findings suggested that further research into the pull-in versus push-out distinction is necessary; differences in participant compliance have not been explored using the pull-in and push-out recruitment methods more commonly used on college campuses (ie, pull-in: a psychology student participation pool; push-out: email announcements to the general undergraduate student body). Moreover, no study to date has explored study compliance across these 2 sources among college students.

This Examination: A Two Study Investigation

This examination explored study compliance (ie, data quality and retention) by recruitment source across 2 studies of college students in the United States with varying design protocols. These were follow-up analyses of published studies with different primary research goals. Study 1 focused on unique college stressors and links to mental health [34] and involved

remotely distributing a web-based survey (fully remote and cross-sectional design). Study 2 examined an intervention targeting college drinking (ClinicalTrials.gov NCT03440463) [35] and involved an in-person baseline procedure with a computerized survey and remote web-based follow-up surveys 1 month and 3 months later (in-person component and longitudinal design). For both studies, participants were recruited from (1) a psychology student participation pool, receiving research credit in psychology courses as compensation and (2) the general student body via emailed announcements, receiving either a raffle entry (study 1) or monetary compensation (study 2).

This examination aimed to examine the following: (1) whether data quality varied across recruitment sources, (2) whether study retention varied across recruitment sources, (3) the impact of data quality on study variable associations, (4) the impact of data quality on measures of internal consistency, and (5) whether the demographic qualities of participants significantly varied across those who failed attention checks versus those who did not. Data quality was examined with attention checks that were used in both study 1 and study 2, and retention was operationalized as follow-up completion rates in study 2 only. Given the limited research on study compliance by recruitment methods with college samples, the analyses for aims 1, 2, and 5 were exploratory in nature. For aims 3 and 4, consistent with previous findings that satisficing can add noise to the assessment and reduce the strength of effects [1,11], we hypothesized that both internal consistency indicators and study variable associations would be stronger after eliminating those who failed attention checks. In particular, satisficing participants tend to endorse midpoints across multiple measures [10], potentially reducing the strength of association among variables, and the inclusion of satisficing participants can mask the strong effects that are revealed after their removal [1], supporting our hypothesis for aim 3. In addition to tendencies to endorse scale

midpoints, satisficing participants also fail to notice scale reversals (ie, reverse-scored items) [11], potentially reducing indicators of internal consistency, supporting our hypothesis for aim 4.

Study 1

Methods

Study 1 was a cross-sectional examination of worry as a mediator between psychosocial stressors and anxiety, stress, and depression [34]. The main outcomes of interest to the original study included worry, stress, depression, and anxiety.

Participants

Undergraduate students (282/407, 69.3% female; 230/407, 56.5% White; mean age 22.65, SD 6.73 years) from a large, public, minority-serving university in the mid-Atlantic region of the United States were recruited via university-wide student announcements (a push-out approach; n=257) as well as through the psychology student research pool (a pull-in approach; n=150) to complete a web-based survey. They were relatively evenly distributed across the year in school. Refer to [Table 1](#) for relevant demographic information for the full sample as well as categorized based on recruitment source. Both recruitment advertisements mentioned that the study was a web-based survey and the type of information assessed (eg, anxiety, worry, and related cognitions). Both indicated an estimate of how long the survey would take and information about compensation. Only the university-wide student announcement included a sentence about how the data would be used and that their data would remain confidential, as that detail was already clear for participants from the psychology pool. Different links were provided for recruitment through the psychology pool versus university-wide announcements. The 2 data sets were coded to reflect how participants accessed the survey and then merged.

Table 1. Descriptive information of the study 1 sample categorized by recruitment source^a.

Variable	General student announcements (n=257)	Psychology pool (n=150)	Total (N=407)	<i>P</i> value
Gender, n (%)				.32
Female	168 (75.7)	114 (76.5)	282 (69.3)	
Male	47 (21.2)	34 (22.8)	81 (19.9)	
Transgender	5 (2.3)	0 (0)	5 (1.2)	
Other	2 (0.9)	1 (0.7)	3 (0.9)	
Ethnicity, n (%)				.24
Hispanic or Latinx	16 (7.2)	16 (10.7)	32 (8.6)	
Not Hispanic or Latinx	205 (92.8)	133 (89.3)	338 (91.4)	
Race^b, n (%)				
Asian	32 (12.5)	15 (10)	47 (11.5)	.46
Black or African American	54 (21)	59 (39.3)	113 (27.8)	<i><.001^c</i>
Native American	7 (2.7)	3 (2)	10 (2.5)	.75
Other	10 (3.9)	7 (4.7)	17 (4.2)	.71
White	146 (56.8)	84 (56)	230 (56.5)	.87
Year in school, n (%)				.001
Freshman	53 (24)	65 (43.6)	118 (31.9)	
Sophomore	44 (19.9)	21 (14.1)	65 (17.6)	
Junior	56 (25.3)	27 (18.1)	83 (22.4)	
Senior	68 (30.8)	36 (24.2)	104 (28.1)	
Employment, n (%)				.02
Employed	134 (60.6)	68 (45.6)	202 (54.6)	
Not employed	82 (37.1)	71 (47.7)	153 (41.4)	
Other	5 (2.3)	10 (6.7)	15 (3.7)	
Age (years), mean (SD)	22.71 (6.19)	22.58 (7.48)	22.65 (6.73)	.86

^aCategories with <5 participants per cell were not included in the chi-square examinations.

^bThe participants could select >1 response option for race, so tallies may sum up to more than the total sample size.

^cSignificant *P* values are indicated in italics.

Procedure

A study advertisement was included in the student announcements emailed to all the students at the host institution. Interested students could click on a link to complete the web-based survey. A similar advertisement was included in the web-based portal for the psychology research pool, which was linked to the same survey. The psychology participation pool included students enrolled in psychology courses. In exchange for their participation in the studies posted, they were provided research credit that they can apply to a course in which they were enrolled. Instructors may build these credits into the grading criteria for the course or offer the students extra credit. The students might sign up for any study for which they were eligible. Volunteering as a study participant was not required to earn these research credits; students might alternatively complete scientific article critiques. Student announcements were emailed to every student enrolled in the university each day. They included announcements for academic workshops,

research studies, social activities, and employment opportunities available to the students. Data collection for study 1 took place from July to September 2017.

Ethical Considerations

We complied with American Psychological Association's ethical standards in the treatment of our sample. The Human Subjects Review Committee of the institution determined the study to be exempt from ongoing oversight (reference number 1103992-1). All participants provided informed consent before completing the survey. The participants recruited via student announcements had the choice of being compensated with either a raffle entry (one of 4 US \$50 Amazon gift cards or one of 12 US \$25 Amazon gift cards) or research credit (if applicable). Participants recruited via the psychology student research pool were compensated with research credit. Because information for compensation purposes was collected in a separate questionnaire not linked to survey responses, the study data were anonymous.

Materials

Worry

Worry was assessed using the Penn State Worry Questionnaire [36], which is a 16-item measure that assesses the severity of participants' worries (eg, my worries overwhelm me). Response options ranged from 1=*not at all typical of me* to 5=*very typical of me*.

Stress, Depression, and Anxiety

Stress, depression, and anxiety were assessed with the 21-item Depression Anxiety Stress Scale [37]. Each construct was assessed with 7 items, including stress (eg, I found myself getting upset rather easily), depression (eg, I felt that life was meaningless), and anxiety (eg, I felt I was close to panic). Responses ranged from 0=*did not apply to me at all over the last week* to 3=*applied to me very much or most of the time over the past week*.

Attention Checks

In total, 8 attention check questions were added to the surveys to assess data quality or, more specifically, to detect satisficing, where inattentive participants were not fully reading survey items or instructions. Of these, 4 were separate questions (eg, Select the highest number), and 4 were integrated into questionnaires (eg, Select "5-7 days" for this answer). The number of incorrect responses was summed and then recoded into a series of variables that represented whether the participants answered any of the attention checks incorrectly ($n=55$), ≥ 2 incorrectly ($n=16$), or ≥ 3 incorrectly ($n=9$; 0=*no* and 1=*yes* for all variables). A variable was not created for answering ≥ 4 incorrectly, as this represented only 1 participant.

Analysis Approach

The demographic characteristics of the sample were compared across recruitment sources using chi-square tests for categorical variables (eg, year in school and employment) and 2-tailed t tests for continuous variables (ie, age). To test study aim 1 (data quality across recruitment sources), the proportion of participants failing attention checks (coded for failing ≥ 1 , ≥ 2 , or ≥ 3 as *yes vs no*) was compared against the recruitment source (general student body vs psychology pool) using a series of chi-square tests of independence (or Fisher exact when the expected value for any cell was < 5). This examination was repeated as a series of logistic regressions to control for any demographic characteristics that significantly varied across recruitment sources. We considered the survey completion time as another marker of data quality. However, time spent on the survey could be impacted by multiple factors, such as satisficing (potentially resulting in faster completion times than other participants) or distraction (potentially resulting in slower completion times than other participants). However, survey completion time could also be impacted by external factors not related to the quality of the data such as poor internet connection or taking a break and coming back, which would result in slower completion times, but responses may still be of high quality. It could also be impacted by skipping some items or not completing the full survey, which would result in faster completion times, but the completed responses may still be of high quality. Moreover, if researchers use fast survey completion

times to throw out cases, throwing out those who only complete part of the survey, they could introduce systematic bias by using a complete-case analysis, which is labeled as one of the worst methods for addressing missing data by the American Psychological Association Task Force on Statistical Inference [38]. As such, we chose to focus exclusively on failed attention checks as a marker of poor data quality. For the same reason, we chose to focus on failing attention checks (ie, completing the item but getting it wrong) as opposed to answering the item correctly, as this approach allowed the participants to drop out of the survey and not complete all attention check items while still potentially providing good quality data for the items answered.

Aim 2 (whether study retention varied across recruitment sources) was not examined for study 1 because it was not longitudinal. To test study aim 3 (the impact of data quality on study variable associations), a series of bivariate correlations were conducted among the variables of interest to the original study (ie, worry, stress, depression, and anxiety). These were conducted once for the full sample and then again for only those who did not fail any attention checks, those who failed ≤ 1 attention checks, those who failed ≤ 2 attention checks, and those who failed ≤ 3 attention checks (ie, retaining those who were not *engaging* in satisficing using various cutoffs). Finally, they were conducted again for those who failed at least one attention check (ie, among those who were *engaging* in satisficing). Correlations were not conducted among those who failed ≥ 2 (or ≥ 3) attention checks because of the small number of participants meeting these criteria (ie, ≤ 16). The largest discrepancies between the correlations for those who failed any attention checks and those who did not were examined via Fisher z for independent samples. The comparisons were only conducted for those who failed any attention checks versus none, as they represented a split of the full sample (ie, none of the participants were in both groups). This allowed us to detect whether significant noise was introduced to the sample by those engaging in satisficing, potentially reducing the strength of associations or increasing SEs via random error.

To examine aim 4 (the impact of data quality on measures of internal consistency), Cronbach α and McDonald omega were calculated for the key study measures using the full sample and then again only for those who did or did not fail varying numbers of attention checks. Both indicators were provided because McDonald omega has more realistic and attainable assumptions and thus may be a more accurate indicator of internal consistency in many circumstances, but Cronbach α is more widely used and understood [39]. Finally, a series of 2-tailed t tests and chi-square tests were conducted to test study aim 5, which is to explore whether the demographic qualities of participants significantly varied across those who failed any attention checks versus those who did not. All analyses were conducted using SPSS statistical software (version 26; IBM Corp; (including using a macro by Hayes and Coutts [39] for McDonald omega). Sample size for the original examination [34] was determined via a power analysis using G*Power [40], specifying a 2-tailed test, an α of .05, and a power of 0.80. Power analysis was not repeated for this study because it was a secondary analysis.

Results

Overview

As shown in Table 1, significantly more Black or African American students were recruited through the psychology pool (59/150, 39.3%) than through the general student body (54/257, 21%; $P < .001$). In addition, the sample recruited via the psychology pool had significantly more first-year students (almost a majority; 65/150, 43.6%), whereas the sample recruited via the general student body was generally more balanced across years in school ($P = .001$). Finally, significantly more participants recruited via the general student body were employed (134/257, 60.6%) than those from the psychology pool (68/150, 45.6%; $P = .02$). The sample did not significantly vary across recruitment methods for gender, age, ethnicity, or other racial identities.

Aim 1: Data Quality by Recruitment Source

Across the total sample, 86.5% (352/407) of the participants did not fail any attention checks, 9.6% (39/407) failed 1 check, 1.7% (7/407) failed 2 checks, 2% (8/407) failed 3 checks, and 0.2% (1/407) failed 5 checks. No one failed >5 (out of 8) checks. The recruitment method type was associated with data quality such that more psychology pool participants (29/150, 19.3%) failed any attention checks than the general student body participants (26/257, 10.1%; $\chi^2_1 = 6.9$, $P = .009$). Similarly, more psychology pool participants (10/150, 6.7%) failed ≥ 2 attention checks than the general student body participants (6/257, 2.3%; $\chi^2_1 = 4.7$, $P = .03$). Although the trend was in the same direction for failing ≥ 3 attention checks (6/150, 4% vs 3/257, 1.2% for psychology pool participants compared with the general student body participants), this did not reach statistical significance based on Fisher exact test ($P = .08$).

These comparisons were repeated as a series of logistic regressions, controlling for the demographics that were significantly different across recruitment sources (year in school, employment, and endorsing Black or African American for race). Year in school and employment did not significantly predict attention check failure and were dropped as predictors. The model controlling for the endorsement of Black or African American identity was consistent with the chi-square analysis,

finding that the recruitment source was significantly associated with any attention check failure, with the participants recruited via the psychology pool significantly more likely to fail the attention checks ($B = 0.63$; $P = .04$; $\exp[B] = 1.87$, 95% CI 1.04-3.37). Controlling for race, the recruitment source was not significantly associated with failing ≥ 2 attention checks ($B = 0.83$; $P = .13$; $\exp[B] = 2.28$, 95% CI 0.79-6.60), or ≥ 3 attention checks ($B = 0.76$; $P = .30$; $\exp[B] = 2.13$, 95% CI 0.51-8.99).

Aim 3: Impact of Data Quality on Variable Associations

Correlations among the key variables for the original study (ie, worry, stress, depression, and anxiety) were conducted for the full sample, those who did not fail any attention checks, those who had failed at least one attention check, those who failed <2 , and those who failed <3 attention checks (Table 2). Before the analysis, the variables were examined for extreme values (ie, outliers) and normality. All variables were found to be normally distributed, and no extreme outliers were identified. Overall, when comparing the full sample with those who did not fail any attention checks, there was no clear pattern of differences; some correlations became smaller, whereas others were larger. Similarly, there were mixed findings when comparing the strength of correlations between the participants who did not fail any attention checks and those who did. As expected, correlations for those who failed <2 or <3 attention checks were midrange between those who did not fail any and those who failed at least one.

The changes in correlations between those who did not fail any attention checks and those who did were compared with Fisher z independent sample comparisons to examine the magnitude of difference. Contrary to what was hypothesized, the association between depression and anxiety was significantly stronger among the participants who failed at least one attention check than among those who did not fail any (z score -3.11 ; $P = .001$), as was the association between stress and anxiety (z score -3.66 ; $P < .001$). The next largest differences (between stress and worry: $z = 1.60$; $P = .06$ and between anxiety and worry: $z = 1.54$; $P = .06$) were in the expected direction but were not significant. The differences between all other correlations were smaller in magnitude and were not significantly different across groups by attention check failure.

Table 2. Correlations among key study 1 variables categorized based on attention check failure^a.

Measure	1	2	3	4
Full sample (N =407)				
1. Worry	— ^b	—	—	—
2. Stress	.62 ^a	—	—	—
3. Depression	.49 ^a	.73 ^a	—	—
4. Anxiety	.57 ^a	.79 ^a	.69 ^a	—
Did not fail any attention checks (n=352)				
1. Worry	—	—	—	—
2. Stress	.64 ^a	—	—	—
3. Depression	.51 ^a	.72 ^a	—	—
4. Anxiety	.60 ^a	.77 ^a	.66 ^a	—
Failed >1 attention checks (n=54)				
1. Worry	—	—	—	—
2. Stress	.47	—	—	—
3. Depression	.43	.81	—	—
4. Anxiety	.43	.92	.85	—
Failed <2 attention checks (n=391)				
1. Worry	—	—	—	—
2. Stress	.62	—	—	—
3. Depression	.50	.72	—	—
4. Anxiety	.59	.78	.68	—
Failed <3 attention checks (n=394)				
1. Worry	—	—	—	—
2. Stress	.63	—	—	—
3. Depression	.51	.73	—	—
4. Anxiety	.58	.79	.68	—

^aAll correlations were significant at $P < .001$.

^bNot applicable.

Aim 4: Impact of Data Quality on Internal Consistency

As shown in [Table 3](#), differences in Cronbach α and McDonald omega were negligible across samples restricted in size by attention check failure.

Table 3. Internal consistency measures among key study 1 variables by attention check failure.

Variables	No attention failures (n=352)		No more than 1 failure (n=391)		No more than 2 failures (n=398)		No more than 3 failures (n=406)		Full sample (N=407)	
	α	Ω	α	Ω	α	Ω	α	Ω	α	Ω
Worry	.935	0.941	.933	0.940	.933	0.939	.932	0.938	.931	0.938
Stress	.865	.0868	.868	0.870	.871	0.873	.872	0.874	.872	0.874
Anxiety	.853	0.856	.859	0.861	.861	0.863	.863	0.865	.862	0.865
Depression	.905	0.907	.912	0.914	.913	0.914	.912	0.913	.912	0.913

Aim 5: Demographics by Attention Check Failure

A series of chi-square analyses revealed that attention check failure was not significantly associated with gender ($\chi^2_3=1.7$, $P=.63$), ethnicity ($\chi^2_1=0.0$, $P=.86$), year in school ($\chi^2_3=0.9$, $P=.81$), employment status ($\chi^2_4=4.2$, $P=.38$), or age ($t_{365}=0.68$; $P=.50$). Although it was not associated with the endorsement of some racial identities (ie, identifying as Asian: $\chi^2_1=0.09$, $P=.77$ or Native American: $\chi^2_1=0.4$, $P=.54$), it was significantly associated with identifying as Black or African American ($\chi^2_1=7.9$, $P=.005$) and as White ($\chi^2_1=10.5$, $P=.001$). More participants who identified as Black or African American failed at least one attention check (21.2%) than those who did not identify as Black (10.5%), whereas fewer participants who identified as White failed the attention checks (8.7%) than those who did not identify as White (19.8%). Given the different demographic breakdown by recruitment source, we also examined attention check failure and race within recruitment source. Identifying as Black was still significantly associated with attention check failure within the psychology pool ($\chi^2_1=5.3$, $P=.03$). A similar trend was observed for the announcement pool ($\chi^2_1=1.0$, $P=.31$), but it was not significant (likely because of the smaller sample size).

Comparing failing ≥ 2 attention checks with failing < 2 checks revealed the same general pattern of findings. There were no significant associations between attention check failure and most demographic variables. However, more participants who identified as Black or African American failed ≥ 2 attention checks (8.8%) than those who did not identify as Black (2%; Fisher exact $P=.003$), whereas fewer participants who identified as White failed ≥ 2 attention checks (1.7%) than those who did not identify as White (6.8%; $\chi^2_1=6.7$, $P=.009$).

Comparing failing ≥ 3 attention checks with failing < 3 checks also revealed the same general pattern of findings. There were no significant associations between attention check failure and most demographic variables. However, more participants who identified as Black or African American failed ≥ 2 attention

checks (7.1%) than participants who did not identify as Black (0.3%; Fisher exact $P<.001$), whereas fewer participants who identified as White failed ≥ 2 attention checks (0%) than those who did not identify as White (5.1%; Fisher exact $P<.001$).

Study 2

Methods

Study 2 was a longitudinal (ClinicalTrials.gov NCT03440463) randomized control trial designed to examine the effects on drinking outcomes of personalized normative feedback booster emails sent after completing a web-based alcohol intervention administered in person [35]. The main outcomes of interest to the original study included alcohol consumption, alcohol-related problems, and descriptive normative perceptions (ie, how much one thinks relevant others drink).

Participants

Participants (378/528, 71.6% female; 281/528, 53.2% Black or African American; 215/528, 40.9% White; mean age 19.85 years, SD 1.65 years) were recruited from the same university as the one from where study 1 participants were recruited (a large, public, minority-serving institution in the mid-Atlantic region of the United States) through 2 recruitment sources: via student announcement emails (a push-out approach; $n=127$) and a psychology research pool (a pull-in approach; $n=401$). Eligible participants were current students aged between 18 and 24 years who had consumed at least one alcoholic beverage in the past 2 weeks. Refer to Table 4 for relevant demographic information for the full sample as well as categorized by recruitment source. Both recruitment advertisements mentioned that the study required in-person attendance for the first session and that it investigated the effects of a computerized intervention on student health behaviors, such as drinking, over an extended period. Both indicated eligibility criteria, an estimate of how long the first session would take, and information about compensation. Only the university-wide student announcement mentioned that the data would remain confidential, as that detail was already clear for participants from the psychology pool.

Table 4. Study 2 sample descriptive information categorized by recruitment source^a.

Variable	General student announcements (n=127)	Psychology pool (n=401)	Total (N=528)	<i>P</i> value
Gender, n (%)				.85
Female	91 (71.7)	287 (71.6)	378 (71.6)	
Male	36 (28.3)	113 (28.2)	149 (28.2)	
Transgender	0 (0)	0 (0)	0 (0)	
Other	0 (0)	1 (0.2)	1 (0.2)	
Ethnicity, n (%)				.75
Hispanic or Latinx	15 (11.8)	43 (10.8)	58 (11)	
Not Hispanic or Latinx	112 (88.2)	355 (89.2)	467 (89)	
Race^b, n (%)				
Asian	12 (9.4)	39 (9.9)	51 (9.8)	.88
Black or African American	64 (50.4)	213 (54.1)	277 (53.2)	.47
Native American	8 (6.3)	11 (2.8)	19 (3.6)	.07
Other	12 (9.4)	24 (6.1)	36 (6.9)	.19
White	49 (38.6)	164 (41.6)	213 (40.9)	.54
Year in school, n (%)				<.001 ^c
Freshman	21 (16.5)	156 (38.9)	177 (33.5)	
Sophomore	27 (21.3)	123 (30.7)	150 (28.4)	
Junior	32 (25.2)	64 (16)	96 (18.2)	
Senior	46 (36.2)	56 (14)	102 (19.3)	
Graduate	0 (0)	1 (0.2)	1 (0.2)	
Other	1 (0.8)	1 (0.2)	2 (0.4)	
Age (years), mean (SD)	20.51 (1.66)	19.65 (1.60)	19.85 (1.65)	<.001

^aCategories with <5 participants per cell were not included in the chi-square examinations.

^bParticipants could select >1 response option for race, so tallies may sum up to more than the sample size.

^cSignificant *P* values indicated in italics.

Procedure

A study advertisement was included in the emailed student announcements sent to all the students at the host institution. Interested students could click on a link to complete a screener survey. Eligible individuals were directed to a web-based scheduler to select an upcoming appointment. A similar advertisement was included in the web-based portal for the psychology research pool. The structure of the psychology participation pool was identical to that of the first study; students enrolled in psychology courses could earn research credit in exchange for their participation in the studies posted or by writing scientific article critiques. The portal allowed for the advertisement to be viewed only by students who met the restricted age criterion. The psychology pool participants did not need to complete the screener survey (the alcohol criterion was prominently displayed in the study description) and could immediately access a web-based scheduler to select an upcoming appointment. All the participants were informed that participating in the study involved attending a 90-minute time slot at the research laboratory where they would be instructed

to complete a web-based survey (information of the nature of the constructs were provided) and an alcohol intervention. The participants attended their baseline session in an on-campus research laboratory. After providing informed consent, they completed the baseline survey before completing the web-based alcohol intervention. Participants were randomized into 3 conditions that varied based on the feedback they were provided 2 weeks later via email. All the participants received follow-up surveys via emailed invitations 1 month and 3 months after baseline; those who opted in received reminders via text messages as well. These follow-up surveys were shorter than the initial baseline survey and were completed on the web, so the participants did not have to return to the research laboratory. Baseline data were collected from April 2017 to December 2017.

Ethical Considerations

We complied with American Psychological Association's ethical standards in the treatment of our sample. The Old Dominion University Institutional Review Board approved the study (reference number 690348-2). The participants provided

informed consent before beginning the survey during the baseline session. The participants who were recruited through the psychology research pool could choose to earn research credit or monetary compensation (US \$20) for completing the baseline survey. Participants who were recruited through the general student body received monetary compensation (US \$20) for completing the baseline survey. All the participants received monetary compensation for completing the follow-up surveys (US \$10 each) and a bonus (US \$10) for completing both follow-up surveys. To ensure confidentiality, after data collection and cleaning were complete, all data were deidentified.

Materials

Drinking Outcomes

The Daily Drinking Questionnaire [41] was used to assess alcohol consumption during each day of a typical week in the past 30 days. The participants were asked to enter the total number of standard drinks consumed on each day of the week as well as the number of hours that passed while they were drinking on those days. Typical drinks per drinking day were calculated by dividing the typical quantity of drinks consumed per week by the total number of drinking days per week. Typical estimated blood alcohol concentration (eBAC) was calculated by averaging the eBAC levels for each drinking day. These levels were calculated based on the number of drinks consumed, hours passed while drinking, and body composition based on sex and weight [42].

Descriptive Norms

The Daily Drinking Questionnaire [41] was modified so that participants reported how many standard drinks they believe their close friends consume on each day of a typical week. Descriptive norms reflecting perceived drinks per drinking day for their close friends were calculated by dividing the total number of drinks in a typical week by the number of drinking days.

Alcohol-Related Problems

The Young Adult Alcohol Consequences Questionnaire [43] was used to measure the total number of consequences a participant reported for the past 30 days. A total of 48 items assessed consequences across 8 domains (eg, impaired control, academic or occupational consequences, and social or interpersonal consequences). The participants reported whether they experienced the consequence (*yes*) or not (*no*); the number of reported consequences were summed.

Attention Checks

In total, 4 attention check questions were added to the surveys to assess data quality or, more specifically, to detect satisficing, where inattentive participants were not fully reading survey items or instructions. Of these questions, 2 were separate questions (eg, Which is the highest number?) and 2 were integrated into questionnaires (eg, Select “Neutral” for this question). The number of incorrect responses was summed and then recoded into a variable that represented whether the participants answered any of the attention checks incorrectly ($n=64$) or ≥ 2 incorrectly ($n=16$; 0=*no* and 1=*yes* for all

variables). A variable was not created for answering ≥ 3 questions incorrectly, as this represented only 3 participants.

Analysis Approach

As with study 1, the demographic characteristics of the sample were compared across recruitment sources using chi-square tests for categorical variables (eg, year in school and employment) and 2-tailed *t* tests for continuous variables (ie, age). To test study aim 1 (data quality across recruitment sources), the proportion of participants failing attention checks (*yes* vs *no*) was compared against recruitment source (psychology pool vs general student body) using 3 chi-square tests of independence (1 for each wave of data collection). This examination was repeated as a series of logistic regressions to control for any demographic characteristics that significantly varied across recruitment sources.

To test study aim 2 (study compliance or retention across recruitment sources), the proportion of participants who completed each follow-up survey (*yes* vs *no*) was compared against the recruitment source using 2 chi-square tests of independence (1 for each follow-up survey). These comparisons were also repeated as a pair of logistic regressions to control for any demographic characteristics that significantly varied across recruitment sources. To test study aim 3 (the impact of data quality on study variable associations), a series of bivariate correlations were conducted among the variables of interest to the original study (ie, typical alcohol consumption, typical eBAC, alcohol-related problems, and descriptive norms). These were conducted once for the full sample, then again for only those who did not fail any attention checks, a third time only for those who failed at least one attention check, and, finally, a fourth time for those who failed < 2 attention checks. Correlations were not conducted among those who failed ≥ 2 attention checks because of the small number of participants meeting this criterion (ie, ≤ 16). The largest discrepancies between the correlations for those who failed attention checks and those who did not were examined using Fisher *z* for independent samples.

To examine aim 4 (the impact of data quality on internal consistency indicators), Cronbach α and McDonald omega were calculated for the only traditional measure (alcohol-related problems) using the full sample and then again only for those who did or did not fail varying numbers of attention checks. Finally, to test study aim 5, a series of 2-tailed *t* tests and chi-square tests were conducted to explore whether the demographic qualities of participants significantly varied across those who failed any attention checks versus those who did not. All analyses were conducted using SPSS (version 26; IBM) (including using a macro by Hayes and Coutts [39] for McDonald omega). The sample size for the original examination [35] was determined via a power analysis using Monte Carlo simulation methods, specifying a 2-tailed test, an α of .05, and a power of 0.80. Power analysis was not repeated for this study because it was a secondary analysis.

Results

Overview

As shown in Table 4, the sample recruited via the psychology pool had significantly more freshmen (156/401, 38.9%) and sophomore (123/401, 30.7%) students than upper classmen, whereas the sample recruited via the general student body was generally more balanced across year in school ($P < .001$). The psychology pool participants were also slightly younger (mean age 19.65, SD 1.60 years) than the participants recruited via the general student body (mean age 20.51 years, SD 1.66 years; $P < .001$). The sample did not significantly vary across recruitment methods for gender, ethnicity, or race.

Aim 1: Data Quality by Recruitment Source

Across the total sample for study 2, at baseline, 87.9% (464/528) of the participants did not fail any attention checks, 9.1% (48/528) failed 1 check, 2.5% (13/528) failed 2 checks, and 0.6% (3/528) failed 3 checks. No one failed 4 attention checks. Recruitment type was associated with data quality for the baseline protocol ($\chi^2_1 = 4.0$, $P = .046$), with more psychology pool participants (55/401, 13.7%) failing any attention checks than those from the general student body (9/127, 7.1%). Failing ≥ 2 attention checks was not significantly different between the psychology pool participants (15/401, 3.7%) and those from the general student body (1/127, 0.8%; Fisher exact $P = .14$).

At the 1-month follow-up, 80.3% (285/355) of the participants did not fail any attention checks, 13.2% (47/355) failed 1 check, 6.2% (22/355) failed 2 checks, and 0.3% (1/355) failed 3 checks. No one failed 4 attention checks. Similar to the baseline, recruitment type was significantly associated with data quality for the 1-month follow-up ($\chi^2_1 = 4.6$, $P = .03$), with more psychology pool participants (55/241, 22.8%) failing attention checks than those from the general student body (15/114, 13.2%). In addition, failing ≥ 2 attention checks was significantly more prevalent among the psychology pool participants (20/241, 8.3%) than among those from the general student body (3/114, 2.6%; $\chi^2_1 = 4.1$, $P = .04$).

At the 3-month follow-up, 81.7% (250/306) of the participants did not fail any attention checks, 13.1% (40/306) failed 1 check, 4.6% (14/306) failed 2 checks, and 0.7% (2/306) failed 3 checks. No one failed 4 attention checks. Significant differences in data quality were also observed for the 3-month follow-up ($\chi^2_1 = 4.2$, $P = .04$), with more psychology pool participants (43/199, 21.6%) failing attention checks than the participants from the general student body (13/107, 12.1%). However, failing ≥ 2 attention checks was not significantly different between the psychology pool participants (13/199, 6.5%) and those from the general student body (3/107, 2.8%; Fisher exact $P = .19$).

We also examined attention check failures over time. Among the individuals who completed the 1-month follow-up survey of the 33 individuals who failed ≥ 1 attention checks at baseline, 16 (48%) also failed ≥ 1 attention checks in the follow-up survey. By contrast, of the 322 individuals who did not fail an attention check at baseline, 54 (17%) failed ≥ 1 attention checks in the follow-up survey. This suggested that the attention check failure

at baseline was associated with the attention check failure in the follow-up survey ($\chi^2_1 = 19.01$, $P < .001$). Similarly, among the individuals who completed the 3-month follow-up survey of the 31 individuals who failed ≥ 1 attention checks at baseline, 13 (42%) also failed ≥ 1 attention checks in the follow-up survey. By contrast, of the 275 individuals who did not fail an attention check at baseline, 43 (16%) failed ≥ 1 attention checks in the follow-up survey. This suggested that the attention check failure at baseline was again associated with the failure in the follow-up survey ($\chi^2_1 = 12.9$, $P < .001$).

Controlling for year in school and age, logistic regressions with recruitment type predicting any attention check failure were not significant for baseline ($B = 0.65$; $P = .10$; $\exp[B] = 1.91$, 95% CI 0.89-4.08) or the 3-month follow-up ($B = 0.69$; $P = .06$; $\exp[B] = 1.99$, 95% CI 0.98-4.06), but were significant for the 1-month follow-up ($B = 0.77$; $P = .02$; $\exp[B] = 2.16$, 95% CI 1.12-4.16). Controlling for class year and age, logistic regressions with recruitment type predicting failing ≥ 2 attention checks were not significant for baseline ($B = 1.70$; $P = .11$; $\exp[B] = 5.46$, 95% CI 0.69-43.12); the 1-month follow-up ($B = 1.23$; $P = .06$; $\exp[B] = 3.42$, 95% CI 0.96-12.27); or the 3-month follow-up ($B = 1.27$; $P = .06$; $\exp[B] = 3.56$, 95% CI 0.94-13.57).

Aim 2: Study Retention by Recruitment Source

Recruitment type was associated with retention at the 1-month follow-up ($\chi^2_1 = 38.5$, $P < .001$), where more participants recruited from the general student body (114/127, 89.8%) completed the 1-month follow-up than the psychology pool participants (241/401, 60.1%); similarly, more participants from the general student body completed the 3-month follow-up assessment (107/127, 84.3%) than the psychology pool participants (199/401, 49.6%; $\chi^2_1 = 47.5$, $P < .001$). These comparisons were to be repeated as logistic regressions, controlling for demographics that were significantly different across recruitment sources (year in school and age). However, year in school and age did not significantly predict attention check failure at any time point, nor did they predict retention for either follow-up survey. Thus, the original chi-square comparisons served as the final models.

Aim 3: Impact of Data Quality on Variable Associations

Correlations among key variables for study 2 (ie, typical alcohol consumption, typical eBAC, alcohol-related problems, and descriptive norms) were conducted for the full sample, those who did not fail any attention checks, those who had failed at least one attention check, and those who failed < 2 attention checks (Table 5). Before the analysis, the variables were examined for extreme values (ie, outliers) and normality. In total, 2 outliers were winsorized (or reduced to less extreme values while maintaining rank) for consumption (ie, drinks per drinking day), 5 outliers were winsorized for eBAC, 3 values were winsorized for alcohol-related problems, and 2 cases were winsorized for drinking norms (ie, perceived drinks per drinking day for close friends). Normality was confirmed for all variables. Overall, the changes in correlations were small, with several relationships increasing in strength from the full sample to the participants who did not fail an attention check. When

comparing the participants who did not fail any attention checks with those who failed ≥ 1 , the patterns of differences were larger, but there were a few relationships in a direction that was not anticipated (ie, stronger correlations among those who had failed). As expected, correlations for those who failed < 2 attention checks were midrange between those who did not fail any checks and those who failed at least one check.

The largest change in correlations among those who did not fail any attention checks versus those who did were compared with Fisher z independent sample comparison (eBAC with alcohol-related problems), finding that the correlation was significantly stronger among those who did not fail any attention checks than among those who did ($z=1.67$; $P=.048$). The differences between all other correlations were smaller in magnitude and were not significantly different across groups by attention check failure.

Table 5. Correlations among the key study 2 variables categorized by attention check failure^a.

Measure	1	2	3	4
Full sample (N=528)				
1. Drinks per drinking day	— ^b	—	—	—
2. Typical eBAC ^c	.82	—	—	—
3. Alcohol-related problems	.41	.42	—	—
4. Descriptive norms ^d	.66	.56	.33	—
Did not fail any attention checks (n=464)				
1. Drinks per drinking day	—	—	—	—
2. Typical eBAC	.82	—	—	—
3. Alcohol-related problems	.42	.44	—	—
4. Descriptive norms	.66	.55	.34	—
Failed >1 attention checks (n=64)				
1. Drinks per drinking day	—	—	—	—
2. Typical eBAC	.82	—	—	—
3. Alcohol-related problems	.35	.23	—	—
4. Descriptive norms	.67	.62	.26	—
Failed <2 attention checks (n=512)				
1. Drinks per drinking day	—	—	—	—
2. Typical eBAC	.82	—	—	—
3. Alcohol-related problems	.42	.42	—	—
4. Descriptive norms	.66	.56	.33	—

^aAll correlations were significant at $P<.001$ except for those italicized.

^bNot available.

^ceBAC: estimated blood alcohol concentration.

^dDescriptive norms refer to perceived consumption (drinks per drinking day) for close friends.

Aim 4: Impact of Data Quality on Internal Consistency

The differences in Cronbach α for alcohol-related problems were negligible across the full sample ($\alpha=.918$), omitting those who failed ≥ 2 attention checks ($\alpha=.917$) or those who failed any attention checks ($\alpha=.917$). Differences in McDonald omega was also negligible across the full sample ($\alpha=.921$), omitting those who failed ≥ 2 attention checks ($\alpha=.920$) or those who failed any attention checks ($\alpha=.920$).

Aim 5: Demographics by Attention Check Failure

A series of chi-square analyses revealed that failing any attention checks versus none of the attention checks was not significantly associated with gender ($\chi^2_1=0.3$, $P=.57$), ethnicity ($\chi^2_1=0.7$,

$P=.40$), year in school ($\chi^2_5=4.3$, $P=.51$), or age ($t_{526}=0.38$; $P=.71$). Although it was not associated with the endorsement of some racial identities (ie, identifying as Asian: $\chi^2_1=0.3$, $P=.60$ or Native American: $\chi^2_1=0.3$, $P=.62$), it was significantly associated with identifying as Black or African American ($\chi^2_1=6.6$, $P=.01$), and a trend was present for identifying as White ($\chi^2_1=3.4$, $P=.07$), although it failed to reach significance. A similar pattern was observed in study 1, where more participants who identified as Black or African American failed at least one attention check (43/277, 15.5%) than those who did not identify as Black (20/244, 8.2%), whereas fewer participants who identified as White failed attention checks (19/213, 8.9%)

than those who did not identify as White (44/308, 14.3%). When comparing failing ≥ 2 attention checks with failing < 2 attention checks, there were no significant associations between attention check failure and the demographic variables.

Discussion

Overview

An examination of whether data quality varies across recruitment sources (aim 1) revealed that a greater proportion of college student participants recruited through the psychology pool failed attention checks than that of those recruited through general emailed announcements, suggesting poorer data quality through satisficing. An examination of whether retention varies across recruitment sources (aim 2) revealed that the psychology pool was also associated with worse compliance via lower retention rates for the web-based follow-up surveys at 1 month and 3 months after baseline (study 2 only). For the examination of the impact of data quality on study variable associations (aim 3), there was no clear pattern of differences when comparing the strength of correlations between participants who did not fail any attention checks and those who did. The direction of the significant effect was consistent with our hypothesis for study 2 (ie, a stronger correlation was found among those who did not fail any attention checks) but was contrary to what was hypothesized for the 2 significant findings of study 1 (ie, stronger correlations were found among those who failed at least one attention check). Study 1 had 2 additional findings that were consistent with our hypothesis but did not reach significance. As for the impact of data quality on measures of internal consistency (aim 4), the impact of omitting those who failed attention checks was negligible on measures of internal consistency. Finally, when examining whether the demographic qualities of participants significantly varied across those who failed attention checks versus those who did not (aim 5), attention check failure was significantly greater among those who identified as Black or African American (both studies) and significantly lower among those who identified as White (study 1 only). It was not significantly associated with other racial identities, ethnicity, gender, age, year in school, or employment status.

Studies 1 and 2 were consistent in their findings that attention check failure rates were lower among the students recruited via general emailed announcements than among the psychology pool participants, suggesting better data quality (aim 1). This was true for both a remote, web-based, cross-sectional survey focused on college stressors and mental health (study 1) and an in-person longitudinal design examining an intervention for college drinking (study 2). However, the difference in rates was greater for the completely remote web-based study protocol (study 1: 19.3% vs 10.1% for failing any attention checks) than for the in-person baseline protocol (study 2: 13.5% vs 7.1%) for failing any attention checks, and rates were also generally higher for the web-based protocol. This finding became nonsignificant for study 2, as the sample was split into a smaller, more unbalanced proportion for examinations of failing ≥ 2 attention checks. Ward and Pond [44] found that having a researcher present via virtual meeting reduced careless responses

by 2.13%, so having a researcher present for the in-person protocol at baseline may have reduced satisficing among participants who would otherwise have satisficed on the web (ie, changing the behavior of those enrolled in the study). In addition, the on-site protocol required that students sign up for a specific time slot and show up at a particular location on campus, requiring greater commitment. This may reflect greater motivation to participate (ie, impacting those who enrolled in the study), which has been linked to reduced satisficing among college students in prior work [14] and is consistent with the suppositions by Krosnick [16]. Thus, in-person protocols with specific sessions and researchers present may result in higher data quality through both who enrolls (self-selection of only those with greater motivation to participate) and through the protocol impacting the behavior of those enrolled (increasing motivation).

Push-in Versus Pull-out Recruitment Approaches

Antoun et al [31] noted that pull-in recruitment sources were more efficient (ie, faster rate of enrollment and lower cost) than push-out recruitment sources. This was true for study 2, a longitudinal study with an in-person baseline session. Enrollment was much higher using the psychology pool ($n=401$) than the general student body contacted via emailed announcements ($n=127$). Similarly, the psychology pool cost is lower (using research credits as compensation rather than monetary payments). However, the findings were contradictory for study 1, which yielded lower enrollment using the psychology pool ($n=127$) than the emailed announcements to the general student body ($n=257$). Both recruitment methods were relatively low cost, with participants from the general student body compensated only with entry into one of a handful of raffles for relatively low-cost gift cards.

The finding that satisficing was greater in the pull-in recruitment source (the psychology pool) than the push-out recruitment source (emailed general announcements) was contrary to the findings of Antoun et al [31]. In total, 3 studies comparing pull-in versus push-out recruitment focused on recruitment not specific to the college population, where participant presence in the pool or panel was completely through self-selection (eg, MTurk and Qualtrics or Dynata panels) [31-33]. These individuals joined the panel specifically to participate in research and earn money. By contrast, the pull-in source for this study included students enrolled in psychology courses who could participate in research studies for course credit (either as extra credit or as part of the requirements for the class). Although they could participate in research to earn a reward, and this is the sole purpose of the panel, their existence in the pool or panel was determined through course enrollment. This could suggest that their presence in the panel was less voluntary. However, equivalent credit could be earned through article critiques rather than study participation, making study participation fully voluntary. This could suggest that which recruitment method is best depends on whether the source is college specific or general sources. One study we are aware of has compared satisficing across a general pull-in source (MTurk for US \$0.50) versus a college-specific pull-in source (a psychology pool for course credit), where they operationalized satisficing using nondifferentiation (ie, selecting the same response option for

all items within a scale) [45]. They found the MTurk sample engaged in more satisficing than the college psychology pool. Given that both are considered pull-in methods, it may be that compensation structure (money vs course credit) was driving this difference, with the participants completing the survey for course credit providing better-quality data. Conversely, in our study comparing 2 college sources, we found that financial compensation was associated with lower rates of satisficing, whether these payments were larger and guaranteed (US \$20 for baseline in study 1) or based on chance (raffles for study 2). In particular, the pull-in approach used in this study (a psychology student participation pool at a single institution) is widely used, and results to other psychology pools are highly generalizable. However, many pull-in approaches (eg, Amazon MTurk) contain participant panels of individuals from across the country and often the globe. This makes the findings of this study less generalizable to pull-in approaches more broadly. With no robust findings across studies regarding push-in versus pull-out methods or financial compensation versus course credit, it appears that there is no guaranteed method to minimize satisficing, making its detection critically important. Attention is a prerequisite for receiving the treatment in most survey experiments, and attention checks effectively reveal who receives the treatment and who does not, such as when Berinsky et al [11] found large condition effects among those who passed the attention screener and no condition effects among those who failed. Detecting and eliminating satisficing is critical for researchers conducting studies examining treatments for psychological health.

Longitudinal Research

The same recruitment source (announcements emailed to the general student body) provided both greater study retention (aim 2) and higher data quality (aim 1); thus, longitudinal researchers can choose a recruitment method that optimizes study compliance for both minimizing satisficing and promoting retention. It is worth noting that retention may have been better for the students who participated through student announcements because their compensation for follow-up surveys was consistent with their compensation for baseline (financial), unlike the psychology pool (course credit). Attrition for longitudinal psychological treatment studies is particularly critical, as meta-analyses have shown dropout rates of 24% to 35% for smartphone-delivered mental health interventions [46], 26% for cognitive behavioral therapy [47], 21% for eating disorder e-treatments [20], and 25% for individual college drinking interventions [48], among others. Researchers striving to minimize satisficing in their clinical trials must still try to optimize retention, and choosing an appropriate recruitment method may help with both concerns.

Satisficing Impact on Study Findings

Contrary to our expectations, correlations did not show consistent strengthening of study variable associations or effects across the 2 studies (aim 3). Select correlations did change significantly in both studies, but the effects went in both directions (sometimes stronger and sometimes weaker). The strengthening of some correlations is consistent with multiple prior studies finding stronger effects after screening out

participants who satisficed [1,11,12]. However, falsely inflated values in the full sample were similar to those reported by Huang et al [13], still pointing to a bias introduced by including these participants. Moreover, Credé [10] noted that whether random responding inflates or deflates the true value of correlations may be influenced by whether the measures examined naturally peak around the lower end of the response option continuum (such as with suicide ideation, psychopathy, and depression) versus around the higher end of the continuum (such as with self-esteem and altruistic behavior) as well as by whether the correlation among those not satisficing is positive or negative, suggesting that both inflated and deflated correlations can be expected with satisficing. Moreover, how participants are carelessly responding may influence the direction of bias. King et al [49] found that when data are not skewed, uniformly responding (ie, each response option has an equal chance of being selected) falsely deflates estimates, whereas long-string responding (ie, selecting the same response option for many items in a row) falsely inflates estimates. Thus, our findings demonstrating correlations changing in both directions support the notion that screening out participants who are satisficing does impact study findings, potentially reducing bias.

Also contrary to our expectations, measures of internal consistency were not stronger after dropping participants who failed attention checks (aim 4). The differences were negligible across the 2 studies. If satisficing by participants adds noise to the data set, researchers might expect it to add measurement error as well. Oppenheimer et al [1] found that internal consistency was reduced among those failing an instructional manipulation for a measure containing reverse-scored items, but these findings were not replicated in this study. However, only 1 measure in this study (the Penn State Worry Questionnaire) [36] contained reverse-scored items, which might be more sensitive to satisficing.

Satisficing Detection Decisions (Number of Failures; Dropping vs Feedback)

In this study, examinations were repeated for multiple cutoffs for satisficing (ie, failing any attention checks vs failing a larger number of attention checks such as 2 or 3). A zero-tolerance approach for identifying satisficing, excluding participants who had ≥ 1 incorrect responses to attention checks, is consistent with what is most commonly reported by researchers [9]. Although a recent examination revealed that the zero-tolerance approach can result in excluding more participants, in particular those who do not demonstrate satisficing on other indices, it is the most common way of screening participants for data quality [9]. This study did not reveal major differences in the pattern of findings across zero tolerance versus basing the cutoff on a larger number, suggesting that there may be some flexibility in which approach a researcher might choose.

One concern raised by prior researchers is whether screening out satisficing participants could introduce a different source of bias, namely reducing the demographic diversity of the sample [11,14]. This study found that attention check failure was not significantly associated with ethnicity, gender, age, year in school, or employment status, indicating that bias is not introduced for these dimensions. However, the participants who

identified as Black or African American were more likely to fail at least one attention check, suggesting that screening out participants could introduce a concern relating to reducing the diversity of the sample. Researchers should be thoughtful regarding recruitment strategies to access larger numbers of participants who could be lost to this screening process so that the final sample still has a substantial representation of this group, as in this study.

In addition to screening out those who fail attention checks, another possible approach that would allow researchers to retain everyone in the sample is to use live feedback to inform participants that the researchers have noticed that they are not paying attention and ask them to read the items carefully. Prompting respondents who completed survey items very quickly to note that this was likely too fast to respond accurately and asking whether they want to reconsider their answers led to reduced satisficing and more accurate responses [50]. Similarly, providing feedback when someone fails an attention check can increase measurement quality [51]. King et al [49] noted that almost no published research in the addiction literature reports screened their data for satisficing. It may be that for research that heavily invests resources in obtaining each data point (such as with longitudinal research and clinical trials, common in the addiction field), throwing out cases results in heavy resource loss, and researchers may be more motivated to try to detect and eliminate satisficing as it occurs. Berinsky et al [11] used different strategies to improve attention during data collection, including warning participants that their data would be monitored before beginning the survey, pairing this warning with a message thanking the participants for their time and careful attention, and providing live feedback (ie, “There was a problem with your response. Please try again”). All 3 approaches resulted in higher rates of passing the attention check items. However, these approaches did not result in reduced noise or bias in associations among study variables or larger treatment effects. It may be that the framing of these messages matter. A systematic review of studies examining prompts in health promotion or health behavior interventions found that messages were more effective if they were tailored with a personal touch [52]. Similarly, a review of retention in panel studies emphasized that explaining the importance of the project and the contributions of the participants is key to engaging participants and promoting good study retention [53]. The same approaches can be used to promote good data quality. Accentuating the purpose and importance of the study, how the participants are helping, and that their responses are of great value to the researchers may have an effect on not just passing attention checks but also actually increasing attention and minimizing noise. Pairing this warm introduction to the fact that the responses will be monitored with a similarly framed live feedback message when attention checks are failed (eg, “Your answer for this question is not correct. Your contributions to our research are extremely valuable. Please be sure to read questions thoroughly and answer carefully”) may have more of an impact on data quality.

Recommendations

To promote data quality and minimize bias, we have several recommendations for researchers. (1) Use attention checks to

detect satisficing. Failure of attention checks was prevalent in both studies across both recruitment sources, suggesting that satisficing is occurring among college students regardless of the study design or recruitment method. Moreover, findings changed after screening out those who failed at least one attention check, suggesting that ignoring this phenomenon could introduce bias into study conclusions. Attention checks can help researchers identify who is providing higher quality data. (2) Carefully consider the recruitment source. Although using psychology pools can cost less and be more efficient (as in study 2, a longitudinal study with an in-person baseline session) and potentially be more convenient, recruiting using broader methods to reach students may result in a better-quality sample (ie, lower satisficing and greater retention). Moreover, the broader recruitment source was more efficient in study 1 (remote and cross-sectional), suggesting that researchers may want to consider their study design in making this determination. When possible, researchers might use multiple recruitment sources to diversify their samples. (3) Weigh the benefits of screening out the participants who fail attention checks (demonstrated to reduce bias in study findings) versus including live feedback (very limited research on this approach). Related to this, (4) consider whether screening out participants could reduce demographic diversity. It could be problematic to increase internal validity to the detriment of external validity. If researchers intend to use attention checks for screening purposes, then they might oversample from populations more likely to be screened out (if possible). Alternatively, letting participants know that their responses will be monitored for data quality and providing live feedback could minimize attention check failure. For treatment studies or other longitudinal studies where tossing cases is problematic, live feedback may be a better option. Finally, (5) researchers interested in minimizing satisficing rather than detecting and removing the data from these participants might consider holding time-specific sessions with a researcher present (in person or on the web).

Limitations

This investigation was a 2-study examination using different study designs (cross-sectional web-based survey vs in-person baseline for a longitudinal randomized controlled trial) and different domains of inquiry (mental health vs drinking behaviors) to maximize the external validity and relevance for other psychological health researchers. However, several limitations should be noted. First, attention checks were the only indicators of data quality used. More robust approaches have included additional indicators, such as psychometric antonyms or synonyms (ie, within-person correlations of similar items), LongStrings (ie, length of response patterns with the same value), Mahalanobis distance values (ie, multivariate outliers for similar items), and self-report items of attention and effort [8,34]. The main advantages of this approach are ease of use, nonrequirement of specialized data management skills, and speed of the screening process. However, researchers may consider using an error-balancing approach that takes multiple indices into account, particularly if working with smaller data sets of specialized populations that are harder to access, where keeping more cases is much more critical.

Another limitation of this examination was that the recruitment sources were limited to single-site data collection methods at 1 institution in the United States. Amazon MTurk is another approach that researchers can use to access the student population more broadly, potentially increasing the demographic diversity of the study samples while maintaining high data quality, including lower rates of satisficing [54,55]. Other web-based approaches, not limited to a single site, could include advertisements on Facebook, Craigslist, etc.

This may increase the demographic and geographic diversity of the sample [54], although the confirmation of student status may be harder. In addition, these pull-in methods may result in a sample with lower income that engages in greater risky behaviors [56] if such qualities are relevant to the research questions being examined. Moreover, although the samples used in this study had a strong representation of Black or African American and White racial identities, other identities were not as well represented. In particular, aim 5 had low sample sizes for some examinations. Although study 2 used a protocol that allowed us to identify all the participants and prohibited repeat sign-ups, study 1's fully web-based protocol did not. The psychology pool participants likely also saw the survey via university-wide announcements, although the recruitment materials requested that students complete the survey only once. Unfortunately, the nature of the system used for the psychology pool uses only anonymous identifiers to issue research credits in the system, so we could not verify for ourselves that the psychology pool participants did not also complete the survey via university announcements.

Finally, although we focused on recruitment sources to label the differences between these 2 groups, compensation was also different. Students in the psychology pool were compensated with research credits that could be applied to their course grades. Students in the emailed announcement group were compensated monetarily (with a raffle entry in study 1 and direct payments in study 2). We believe that this is consistent with most studies using these recruitment sources and feel that compensation is part of these approaches. What is notable is that the pattern of reduced satisficing in the emailed announcement group was true even when the compensation was weak (raffle entry) rather than strong (direct monetary payments), suggesting that the strength of compensation is not driving the effect.

Conclusions

This investigation examined participant compliance (ie, data quality and retention) by recruitment source across 2 studies of college students with varying design protocols (study 1: a fully remote, cross-sectional design examining college stressors and psychological health; study 2: a longitudinal design with an in-person baseline session that examined an intervention targeting college drinking). For both studies, the participants were recruited from (1) a psychology student participation pool, receiving research credit in psychology courses as compensation, and (2) the general student body via emailed announcements, receiving either a raffle entry (study 1) or monetary compensation (study 2). The examination revealed that a greater proportion of college student participants recruited through the psychology pool failed attention checks than that of those recruited through general emailed announcements, suggesting poorer data quality through satisficing in both studies. Moreover, the psychology pool was also associated with worse compliance via lower retention rates in the web-based follow-up surveys at 1 month and 3 months after baseline (study 2 only). After screening out those who failed at least one attention check, some correlations among the study variables were strengthened (potentially due to reducing noise), some were weakened, and some were fairly similar; this mixed pattern potentially points to a bias introduced by including these participants. Finally, attention check failure was not significantly associated with most demographic characteristics (ethnicity, gender, age, year in school, employment status, and select racial identities) but was greater among those who identified as Black or African American (both studies) and significantly lower among those who identified as White (study 1 only). Investigators focused on student research should carefully consider recruitment in their study design and include attention checks or other means of detecting poor quality data. Satisficing was detected across both sources, although it was worse in the psychology pool than in the general student body. Researchers should carefully consider how the study design could promote engagement (eg, live sessions with a researcher), weigh screening participants versus providing live feedback, and consider oversampling demographics that are more likely to be screened out, if possible.

Acknowledgments

Study 2 was supported by the National Institute on Alcohol Abuse and Alcoholism award K01 AA023849 (principal investigator: ALB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Alcohol Abuse and Alcoholism.

Data Availability

Data will be available from the principal investigator upon reasonable request.

Conflicts of Interest

None declared.

References

1. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: detecting satisficing to increase statistical power. *J Experimental Social Psychol* 2009 Jul;45(4):867-872. [doi: [10.1016/j.jesp.2009.03.009](https://doi.org/10.1016/j.jesp.2009.03.009)]
2. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013 May;14(5):365-376. [doi: [10.1038/nrn3475](https://doi.org/10.1038/nrn3475)] [Medline: [23571845](https://pubmed.ncbi.nlm.nih.gov/23571845/)]
3. Palaniappan K, Kum IY. Underlying causes behind research study participants' careless and biased responses in the field of sciences. *Curr Psychol* 2017 Nov 15;38(6):1737-1747. [doi: [10.1007/s12144-017-9733-2](https://doi.org/10.1007/s12144-017-9733-2)]
4. Sakaluk J, Williams A, Kilshaw R, Rhyner K. Evaluating the evidential value of empirically supported psychological treatments (ESTs): a meta-scientific review. *J Abnorm Psychol* 2019 Aug;128(6):500-509. [doi: [10.1037/abn0000421](https://doi.org/10.1037/abn0000421)] [Medline: [31368729](https://pubmed.ncbi.nlm.nih.gov/31368729/)]
5. Arnett JJ. The neglected 95%: why American psychology needs to become less American. *Am Psychol* 2008;63(7):602-614. [doi: [10.1037/0003-066x.63.7.602](https://doi.org/10.1037/0003-066x.63.7.602)]
6. Thalmayer AG, Toscanelli C, Arnett JJ. The neglected 95% revisited: is American psychology becoming less American? *Am Psychol* 2021 Jan;76(1):116-129. [doi: [10.1037/amp0000622](https://doi.org/10.1037/amp0000622)] [Medline: [32271027](https://pubmed.ncbi.nlm.nih.gov/32271027/)]
7. Gallander Wintre M, North C, Sugar LA. Psychologists' response to criticisms about research based on undergraduate participants: a developmental perspective. *Canadian Psychology / Psychologie canadienne* 2001 Aug;42(3):216-225. [doi: [10.1037/h0086893](https://doi.org/10.1037/h0086893)]
8. Meade AW, Craig SB. Identifying careless responses in survey data. *Psychol Methods* 2012 Sep;17(3):437-455. [doi: [10.1037/a0028085](https://doi.org/10.1037/a0028085)] [Medline: [22506584](https://pubmed.ncbi.nlm.nih.gov/22506584/)]
9. Kim DS, McCabe CJ, Yamasaki BL, Louie KA, King KM. Detecting random responders with infrequency scales using an error-balancing threshold. *Behav Res Methods* 2018 Oct 21;50(5):1960-1970. [doi: [10.3758/s13428-017-0964-9](https://doi.org/10.3758/s13428-017-0964-9)] [Medline: [28936811](https://pubmed.ncbi.nlm.nih.gov/28936811/)]
10. Credé M. Random responding as a threat to the validity of effect size estimates in correlational research. *Educational Psychol Measurement* 2010 Apr 28;70(4):596-612. [doi: [10.1177/0013164410366686](https://doi.org/10.1177/0013164410366686)]
11. Berinsky AJ, Margolis MF, Sances MW. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am J Political Sci* 2013 Nov 06;58(3):739-753. [doi: [10.1111/ajps.12081](https://doi.org/10.1111/ajps.12081)]
12. Maniaci MR, Rogge RD. Caring about carelessness: participant inattention and its effects on research. *J Res Personality* 2014 Feb;48:61-83. [doi: [10.1016/j.jrp.2013.09.008](https://doi.org/10.1016/j.jrp.2013.09.008)]
13. Huang JL, Liu M, Bowling NA. Insufficient effort responding: examining an insidious confound in survey data. *J Appl Psychol* 2015 May;100(3):828-845. [doi: [10.1037/a0038510](https://doi.org/10.1037/a0038510)] [Medline: [25495093](https://pubmed.ncbi.nlm.nih.gov/25495093/)]
14. Anduiza E, Galais C. Answering without reading: IMCs and strong satisficing in online surveys. *Int J Public Opin Res* 2016 May 19;edw007. [doi: [10.1093/ijpor/edw007](https://doi.org/10.1093/ijpor/edw007)]
15. Tourangeau R, Rips L, Rasinski K. *The Psychology of Survey Response*. Cambridge, United Kingdom: Cambridge University Press; 2000.
16. Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl Cognit Psychol* 1991 May;5(3):213-236. [doi: [10.1002/acp.2350050305](https://doi.org/10.1002/acp.2350050305)]
17. Roberts C, Gilbert E, Allum N, Eisner L. Research synthesis : satisficing in surveys: a systematic review of the literature. *Public Opin Q* 2019;83(3):598-626. [doi: [10.1093/poq/nfz035](https://doi.org/10.1093/poq/nfz035)]
18. Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-Tyszkiewicz M, SEED Lifecourse Sciences Theme. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. *BMC Med Res Methodol* 2018 Nov 26;18(1):151 [FREE Full text] [doi: [10.1186/s12874-018-0586-7](https://doi.org/10.1186/s12874-018-0586-7)] [Medline: [30477443](https://pubmed.ncbi.nlm.nih.gov/30477443/)]
19. Freijy T, Kothe EJ. Dissonance-based interventions for health behaviour change: a systematic review. *Br J Health Psychol* 2013 May 11;18(2):310-337. [doi: [10.1111/bjhp.12035](https://doi.org/10.1111/bjhp.12035)] [Medline: [23480429](https://pubmed.ncbi.nlm.nih.gov/23480429/)]
20. Linardon J, Shatte A, Messer M, Firth J, Fuller-Tyszkiewicz M. E-mental health interventions for the treatment and prevention of eating disorders: an updated systematic review and meta-analysis. *J Consult Clin Psychol* 2020 Nov;88(11):994-1007. [doi: [10.1037/ccp0000575](https://doi.org/10.1037/ccp0000575)] [Medline: [32852971](https://pubmed.ncbi.nlm.nih.gov/32852971/)]
21. Mun E, de la Torre J, Atkins DC, White HR, Ray AE, Kim S, Project INTEGRATE Team. Project INTEGRATE: an integrative study of brief alcohol interventions for college students. *Psychol Addict Behav* 2015 Mar;29(1):34-48 [FREE Full text] [doi: [10.1037/adb0000047](https://doi.org/10.1037/adb0000047)] [Medline: [25546144](https://pubmed.ncbi.nlm.nih.gov/25546144/)]
22. Braitman AL, Henson JM. Personalized boosters for a computerized intervention targeting college drinking: the influence of protective behavioral strategies. *J Am Coll Health* 2016 Oct;64(7):509-519 [FREE Full text] [doi: [10.1080/07448481.2016.1185725](https://doi.org/10.1080/07448481.2016.1185725)] [Medline: [27148633](https://pubmed.ncbi.nlm.nih.gov/27148633/)]
23. Braitman AL, Lau-Barraco C. Personalized boosters after a computerized intervention targeting college drinking: a randomized controlled trial. *Alcohol Clin Exp Res* 2018 Sep;42(9):1735-1747 [FREE Full text] [doi: [10.1111/acer.13815](https://doi.org/10.1111/acer.13815)] [Medline: [29995326](https://pubmed.ncbi.nlm.nih.gov/29995326/)]
24. Day V, McGrath PJ, Wojtowicz M. Internet-based guided self-help for university students with anxiety, depression and stress: a randomized controlled clinical trial. *Behav Res Ther* 2013 Jul;51(7):344-351. [doi: [10.1016/j.brat.2013.03.003](https://doi.org/10.1016/j.brat.2013.03.003)] [Medline: [23639300](https://pubmed.ncbi.nlm.nih.gov/23639300/)]
25. Sears DO. College sophomores in the laboratory: influences of a narrow data base on social psychology's view of human nature. *J Personality Social Psychol* 1986 Sep;51(3):515-530. [doi: [10.1037/0022-3514.51.3.515](https://doi.org/10.1037/0022-3514.51.3.515)]

26. Sharpe D, Poets S. Canadian psychology department participant pools: closing for the season? *Canadian Psychology / Psychologie canadienne* 2017 May;58(2):168-177. [doi: [10.1037/cap0000090](https://doi.org/10.1037/cap0000090)]
27. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am Psychol* 2004;59(2):93-104. [doi: [10.1037/0003-066x.59.2.93](https://doi.org/10.1037/0003-066x.59.2.93)]
28. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci* 2010 Jun 15;33(2-3):61-83. [doi: [10.1017/s0140525x0999152x](https://doi.org/10.1017/s0140525x0999152x)]
29. Elicker JD, McConnell NL, Hall RJ. Research participation for course credit in introduction to psychology: why don't people participate? *Teaching Psychol* 2010 Jun 29;37(3):183-185. [doi: [10.1080/00986283.2010.488521](https://doi.org/10.1080/00986283.2010.488521)]
30. Rocchi M, Beaudry SG, Anderson, C, Pelletier LG. The perspective of undergraduate research participant pool nonparticipants. *Teaching Psychol* 2016 Aug 22;43(4):285-293. [doi: [10.1177/0098628316662756](https://doi.org/10.1177/0098628316662756)]
31. Antoun C, Zhang C, Conrad FG, Schober MF. Comparisons of online recruitment strategies for convenience samples. *Field Methods* 2015 Sep 16;28(3):231-246. [doi: [10.1177/1525822X15603149](https://doi.org/10.1177/1525822X15603149)]
32. Boas TC, Christenson DP, Glick DM. Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. *Political Sci Res Method* 2018 Aug 08;8(2):232-250. [doi: [10.1017/psrm.2018.28](https://doi.org/10.1017/psrm.2018.28)]
33. Zhang B, Gearhart S. *Surv Pract* 2020 Dec 03;13(1):1-10. [doi: [10.29115/sp-2020-0015](https://doi.org/10.29115/sp-2020-0015)]
34. Bauer EA, Braitman AL, Judah MR, Cigularov KP. Worry as a mediator between psychosocial stressors and emotional sequelae: moderation by contrast avoidance. *J Affect Disord* 2020 Apr 01;266:456-464. [doi: [10.1016/j.jad.2020.01.092](https://doi.org/10.1016/j.jad.2020.01.092)] [Medline: [32056913](https://pubmed.ncbi.nlm.nih.gov/32056913/)]
35. Braitman AL, Strowger M, Lau-Barraco C, Shipley JL, Kelley ML, Carey KB. Examining the added value of harm reduction strategies to emailed boosters to extend the effects of online interventions for college drinkers. *Psychol Addict Behav* 2022 Sep 03;36(6):635-647. [doi: [10.1037/adb0000755](https://doi.org/10.1037/adb0000755)] [Medline: [34081487](https://pubmed.ncbi.nlm.nih.gov/34081487/)]
36. Meyer T, Miller M, Metzger R, Borkovec TD. Development and validation of the penn state worry questionnaire. *Behav Res Ther* 1990 Jan;28(6):487-495. [doi: [10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)]
37. Lovibond P, Lovibond S. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995 Mar;33(3):335-343. [doi: [10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)]
38. Wilkinson L. Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist* 1999 Aug;54(8):594-604. [doi: [10.1037/0003-066x.54.8.594](https://doi.org/10.1037/0003-066x.54.8.594)]
39. Hayes AF, Coutts JJ. Use omega rather than Cronbach's alpha for estimating reliability. *But... Commun Method Measures* 2020 Feb 11;14(1):1-24. [doi: [10.1080/19312458.2020.1718629](https://doi.org/10.1080/19312458.2020.1718629)]
40. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Method* 2007 May;39(2):175-191. [doi: [10.3758/BF03193146](https://doi.org/10.3758/BF03193146)]
41. Collins RL, Parks GA, Marlatt GA. Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. *J Consulting Clin Psychol* 1985;53(2):189-200. [doi: [10.1037/0022-006X.53.2.189](https://doi.org/10.1037/0022-006X.53.2.189)]
42. Matthews DB, Miller WR. Estimating blood alcohol concentration: two computer programs and their applications in therapy and research. *Addictive Behav* 1979 Jan;4(1):55-60. [doi: [10.1016/0306-4603\(79\)90021-2](https://doi.org/10.1016/0306-4603(79)90021-2)]
43. Read JP, Kahler CW, Strong DR, Colder CR. Development and preliminary validation of the young adult alcohol consequences questionnaire. *J Stud Alcohol* 2006 Jan;67(1):169-177. [doi: [10.15288/jsa.2006.67.169](https://doi.org/10.15288/jsa.2006.67.169)] [Medline: [16536141](https://pubmed.ncbi.nlm.nih.gov/16536141/)]
44. Ward M, Pond SB. Using virtual presence and survey instructions to minimize careless responding on internet-based surveys. *Comput Human Behav* 2015 Jul;48:554-568. [doi: [10.1016/j.chb.2015.01.070](https://doi.org/10.1016/j.chb.2015.01.070)]
45. Hamby T, Taylor W. Survey satisficing inflates reliability and validity measures: an experimental comparison of college and amazon mechanical turk samples. *Educ Psychol Meas* 2016 Dec 29;76(6):912-932 [FREE Full text] [doi: [10.1177/0013164415627349](https://doi.org/10.1177/0013164415627349)] [Medline: [29795893](https://pubmed.ncbi.nlm.nih.gov/29795893/)]
46. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J Consult Clin Psychol* 2020 Jan;88(1):1-13. [doi: [10.1037/ccp0000459](https://doi.org/10.1037/ccp0000459)] [Medline: [31697093](https://pubmed.ncbi.nlm.nih.gov/31697093/)]
47. Fernandez E, Salem D, Swift JK, Ramtahal N. Meta-analysis of dropout from cognitive behavioral therapy: magnitude, timing, and moderators. *J Consult Clin Psychol* 2015 Dec;83(6):1108-1122. [doi: [10.1037/ccp0000044](https://doi.org/10.1037/ccp0000044)] [Medline: [26302248](https://pubmed.ncbi.nlm.nih.gov/26302248/)]
48. Carey KB, Scott-Sheldon LA, Carey MP, DeMartini KS. Individual-level interventions to reduce college student drinking: a meta-analytic review. *Addict Behav* 2007 Nov;32(11):2469-2494 [FREE Full text] [doi: [10.1016/j.addbeh.2007.05.004](https://doi.org/10.1016/j.addbeh.2007.05.004)] [Medline: [17590277](https://pubmed.ncbi.nlm.nih.gov/17590277/)]
49. King KM, Kim DS, McCabe CJ. Random responses inflate statistical estimates in heavily skewed addictions data. *Drug Alcohol Depend* 2018 Feb 01;183:102-110 [FREE Full text] [doi: [10.1016/j.drugalcdep.2017.10.033](https://doi.org/10.1016/j.drugalcdep.2017.10.033)] [Medline: [29245102](https://pubmed.ncbi.nlm.nih.gov/29245102/)]
50. Conrad F, Couper M, Tourangeau R, Zhang C. Reducing speeding in web surveys by providing immediate feedback. *Surv Res Methods* 2017 Apr 10;11(1):45-61 [FREE Full text] [doi: [10.18148/srm/2017.v11i1.6304](https://doi.org/10.18148/srm/2017.v11i1.6304)] [Medline: [31745400](https://pubmed.ncbi.nlm.nih.gov/31745400/)]
51. Shamon H, Berning C. Attention check items and instructions in online surveys with incentivized and non-incentivized samples: boon or bane for data quality? *SSRN J* 2020;14(1):55-77. [doi: [10.2139/ssrn.3549789](https://doi.org/10.2139/ssrn.3549789)]

52. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. *J Med Internet Res* 2009 May 14;11(2):e16 [FREE Full text] [doi: [10.2196/jmir.1138](https://doi.org/10.2196/jmir.1138)] [Medline: [19632970](https://pubmed.ncbi.nlm.nih.gov/19632970/)]
53. Ribisl KM, Walton MA, Mowbray CT, Luke DA, Davidson WS, Bootsmiller BJ. Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: review and recommendations. *Evaluation Program Planning* 1996 Feb;19(1):1-25. [doi: [10.1016/0149-7189\(95\)00037-2](https://doi.org/10.1016/0149-7189(95)00037-2)]
54. Buhrmester M, Kwang T, Gosling SD. Amazon's mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* 2011 Jan 03;6(1):3-5. [doi: [10.1177/1745691610393980](https://doi.org/10.1177/1745691610393980)] [Medline: [26162106](https://pubmed.ncbi.nlm.nih.gov/26162106/)]
55. Hauser DJ, Schwarz N. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav Res Methods* 2016 Mar 12;48(1):400-407. [doi: [10.3758/s13428-015-0578-z](https://doi.org/10.3758/s13428-015-0578-z)] [Medline: [25761395](https://pubmed.ncbi.nlm.nih.gov/25761395/)]
56. Lewis RJ, Dawson CA, Shappie AT, Braitman AL, Heron KE. Recruiting cisgender female couples for health disparity-focused daily diary research: challenges, successes, and lessons learned. *Psychol Sexuality* 2021 Jun 23;13(4):931-951. [doi: [10.1080/19419899.2021.1942177](https://doi.org/10.1080/19419899.2021.1942177)]

Abbreviations

eBAC: estimated blood alcohol concentration

MTurk: Mechanical Turk

Edited by A Mavragani; submitted 11.05.22; peer-reviewed by S Bidmon, T Sagi, D Oppenheimer; comments to author 29.08.22; revised version received 03.10.22; accepted 20.10.22; published 09.12.22.

Please cite as:

Braitman AL, Strowger M, Shipley JL, Ortman J, MacIntyre RI, Bauer EA

Data Quality and Study Compliance Among College Students Across 2 Recruitment Sources: Two Study Investigation

JMIR Form Res 2022;6(12):e39488

URL: <https://formative.jmir.org/2022/12/e39488>

doi: [10.2196/39488](https://doi.org/10.2196/39488)

PMID: [36485020](https://pubmed.ncbi.nlm.nih.gov/36485020/)

©Abby L Braitman, Megan Strowger, Jennifer L Shipley, Jordan Ortman, Rachel I MacIntyre, Elizabeth A Bauer. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 09.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Development of an mHealth App–Based Intervention for Depressive Rumination (RuminAid): Mixed Methods Focus Group Evaluation

Eve A Rosenfeld^{1,2}, PhD; Cassandra Lyman³, BS, BA; John E Roberts⁴, PhD

¹Dissemination and Training Division, National Center for PTSD, VA Palo Alto Healthcare System, Menlo Park, CA, United States

²Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States

³Department of Psychology, University of South Florida, Tampa, FL, United States

⁴Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY, United States

Corresponding Author:

Eve A Rosenfeld, PhD

Dissemination and Training Division, National Center for PTSD

VA Palo Alto Healthcare System

NCPTSD – 334

795 Willow Road

Menlo Park, CA, 94025

United States

Phone: 1 908 907 4135

Email: earosenf@stanford.edu

Abstract

Background: Depression is a common mental health condition that poses a significant public health burden. Effective treatments for depression exist; however, access to evidence-based care remains limited. Mobile health (mHealth) apps offer an avenue for improving access. However, few mHealth apps are informed by evidence-based treatments and even fewer are empirically evaluated before dissemination. To address this gap, we developed RuminAid, an mHealth app that uses evidence-based treatment components to reduce depression by targeting a single key depressogenic process—rumination.

Objective: The primary objective of this study was to collect qualitative and quantitative feedback that could be used to improve the design of RuminAid before the software development phase.

Methods: We reviewed empirically supported interventions for depression and rumination and used the key aspects of each to create a storyboard version of RuminAid. We distributed an audio-guided presentation of the RuminAid storyboard to 22 individuals for viewing and solicited user feedback on app content, design, and perceived functionality across 7 focus group sessions.

Results: The consumer-rated quality of the storyboard version of RuminAid was in the acceptable to good range. Indeed, most participants reported that they thought RuminAid would be an engaging, functional, and informational app. Likewise, they endorsed overwhelming positive beliefs about the perceived impact of RuminAid; specifically, 96% (21/22) believed that RuminAid will help depressed ruminators with depression and rumination. Nevertheless, the results highlighted the need for improved app aesthetics (eg, a more appealing color scheme and modern design).

Conclusions: Focus group members reported that the quality of information was quite good and had the potential to help adults who struggle with depression and rumination but expressed concern that poor aesthetics would interfere with users' desire to continue using the app. To address these comments, we hired a graphic designer and redesigned each screen to improve visual appeal. We also removed time gating from the app based on participant feedback and findings from related research. These changes helped elevate RuminAid and informed its initial software build for a pilot trial that focused on evaluating its feasibility and acceptability.

(*JMIR Form Res* 2022;6(12):e40045) doi:[10.2196/40045](https://doi.org/10.2196/40045)

KEYWORDS

depression; rumination; mobile health; mHealth; evidence-based treatment; focus group; mental health; mobile app; mobile phone

Introduction

Background

Depression is the leading cause of global disability [1] and poses a significant public health burden [2-7]. Following the COVID-19 outbreak, approximately one-third of Americans reported clinically significant symptoms of depression [8], highlighting the substantial need for effective treatment. Evidence-based treatments (EBTs) for depression exist; however, access to EBTs remains limited. Less than 35% of individuals with depression receive “minimally adequate” treatment [9-11], and the ratio is even lower for ethnic minority groups [9,10,12,13]. “Stay-at-home” orders during the peak of the COVID-19 pandemic disrupted traditional health care delivery, and our already-overburdened health care system struggled to meet patient needs [14]. Consequently, the pandemic not only increased the prevalence of depression but also reduced treatment accessibility.

Smartphone-based mobile health (mHealth) apps are well suited to address treatment barriers given the relative accessibility of smartphones. In the United States, 85% of adults own a smartphone, and 91% of users report that their smartphone is within arm’s reach 24 hours per day [15]. Consequently, mHealth apps can be accessed at any time, allowing for real-time intervention. In addition, 15% of Americans reported that their only internet-accessible device was their smartphone, and marginalized communities reported higher proportions of smartphone-only internet access [16]. Furthermore, emerging mHealth interventions—particularly self-guided interventions—alleviate the burden on the mental health care system by providing contactless, automated interventions and facilitating self-management.

Unfortunately, many mHealth apps purported to treat depression are not informed by evidence-based care. Specifically, only 10% of publicly available depression apps use evidence-based principles of cognitive behavioral therapy or behavioral activation (BA) [17]. Of those that are empirically informed (eg, MoodTools [18]), few have been empirically evaluated. A systematic review found that only 2% of publicly available mental health apps were empirically supported [19]. Thus, it is unclear to what extent specific mHealth apps are efficacious or effective treatments for depression.

A minority of depression apps have been informed by EBTs and subjected to at least an initial empirical evaluation. For example, Moodivate is a BA self-help app involving psychoeducation, identification of personal values, activity scheduling, and daily mood ratings [20]. A pilot trial found initial evidence of its feasibility and efficacy [21], and a large-scale randomized clinical trial is underway [22]. Nevertheless, too few app-based interventions are subjected to rigorous empirical testing before they reach the market.

In addition, existing evidence-informed mHealth apps, such as Moodivate, target depressive symptoms by delivering an app version of a traditional face-to-face psychotherapy protocol using a varied set of nonoverlapping skills (eg, values clarification, cognitive restructuring, and activity scheduling).

These tend to be presented in a nonstepwise fashion, with all psychoeducational material available at the outset in a handout format. This approach is inconsistent with how consumers use apps: in frequent, short bursts [23]. Treatment engagement may suffer if mHealth interventions require lengthy engagement, are overly didactic, or lack a clear sequential structure. Therefore, we argue that it might be most useful for an mHealth app to target a specific psychological process using a limited set of scaffolded skills [23]. Each targeted skill builds on the skills previously developed as the user progresses through the app content. The intervention would include minimal, targeted psychoeducation and narrowly focus on brief, sequential skill-oriented tasks that are “gamified.”

To date, no empirically supported mHealth apps have focused on rumination, despite it being a potentially promising treatment target. According to response styles theory [24], rumination is a pattern of behaviors and cognitions that focus attention on depressive symptoms, including their causes, consequences, and implications [25]. Rumination causes excessive focus on negative emotional states, inhibits mood-enhancing behaviors, and exacerbates and prolongs depression [24,25]. Although various theories of rumination conceptualize its content, antecedents, and functions somewhat differently [26,27], most converge on key features: rumination is a specific, depression-related form of repetitive negative thinking that occurs in response to a triggering event, and it is experienced as distressing and difficult to control [28-30].

Overwhelming evidence demonstrates a strong relationship between rumination and depression. Rumination prospectively predicts higher levels of depressive symptoms over time [31-33], is a trait of vulnerability to depression [34-37], contributes to the maintenance of depressive episodes [24,25], is a precursor to the onset of clinical depression [32,35,37,38], and is a risk factor for relapse and recurrence [39-42]. Experimental studies have causally linked rumination to the maintenance of depression [25,43,44] and have demonstrated its co-occurrence with other forms of psychopathology [45-52]. Rumination functions as an experiential avoidance strategy (ie, avoidance of uncomfortable private experiences [53]) wherein ruminators distract themselves from emotionally arousing material (eg, sadness) through repetitive thinking, which contributes to negative sequelae [54].

Several EBTs explicitly target rumination, including BA, mindfulness-based cognitive therapy, and rumination-focused cognitive behavioral therapy (RFCBT). Although these interventions are well supported by research [55-57], they are neither widely disseminated nor easily accessible. Considering that depression is so pervasive and problematic, rumination is strongly implicated in depression, and effective treatments for rumination exist but are not widely disseminated, rumination is a promising target for mHealth intervention. To address the need for an easily accessible evidence-based intervention targeting rumination, we designed RuminAid—a new mHealth app.

Treatment Development and Description

Overview

RuminAid skills were drawn from well-supported EBTs for depression and rumination (eg, BA, mindfulness-based cognitive therapy, and RFCBT) and mHealth research, which highlighted the importance of balancing EBT components with simplicity and user friendliness by distilling EBTs to key components. We identified five essential components: (1) brief psychoeducation, (2) recognizing ruminative episodes, (3) alternative behaviors to replace rumination, (4) counteracting rumination-related attentional deficits, and (5) gamification.

RuminAid integrates these elements across 5 brief lessons presented to users as sequential “quests,” with users completing 1 quest per day. Each quest includes brief psychoeducation and gamified elements. Quests 1 to 2 focus on identifying rumination. Quests 3 to 4 teach users to use alternative behaviors to combat rumination. Quest 5 teaches users mindfulness and behavioral skills to counteract the deleterious effects of rumination on attention.

Brief Psychoeducation

Brief psychoeducation is integrated into each quest ([Multimedia Appendix 1](#)). This material identifies rumination as a treatment target, normalizes rumination as an experience, addresses harmful meta-cognitive beliefs, and orients users to new skills. As overemphasis on didactic content is inconsistent with typical smartphone use [23], psychoeducation is not the primary focus. Instead, psychoeducation is intended to facilitate skills acquisition and is limited to a few minutes per quest, at the most.

Identifying Rumination

RuminAid teaches users to discriminate between rumination and nonpathological processes (eg, problem-solving and introspection) and familiarizes users with rumination warning signs ([Multimedia Appendix 2](#)). For example, RuminAid users explicitly label and log periods of rumination in real time, tracking associated content, triggers, and internal contexts (eg, thoughts and emotions). The “Two-Minute Rule for Recognizing Rumination” involves engaging in current patterns of thinking for 2 minutes and then answering specific questions to determine if the user is ruminating [58].

To accommodate the idiographic nature of rumination, users store their ruminative content, triggers, and contexts in a personalized, editable list of rumination “red flags” (ie, signs they are ruminating). RuminAid also includes a list of “common red flags” that users can save on their personalized list. When logging rumination, items from the personalized lists of users are available for tracking via a drop-down menu. Moreover, users can use text entries to enter novel red flags. As users familiarize themselves with their personal red flags, they learn to quickly identify and label periods of rumination, allowing for rapid application of later therapeutic skills.

Alternative Behaviors

RuminAid is behaviorally oriented, providing users with specific alternative behaviors to replace unhelpful rumination habits [59,60] and counteract its avoidance function [61]. RuminAid includes 2 forms of alternative behaviors: self-soothing and

approach-oriented behaviors ([Multimedia Appendix 3](#)). Both forms of alternative behaviors can be conceptualized as BA strategies, wherein self-soothing behaviors aim to increase pleasure and improve mood, whereas approach behaviors aim to increase mastery and promote problem-solving over avoidance.

Self-soothing behaviors provide distressed individuals with a sense of calmness or pleasure and are promoted in EBTs (eg, BA [57] and Dialectical Behavior Therapy [62,63]). In quest 3, users are instructed to engage in self-soothing whenever they ruminate and are automatically prompted to do so every time rumination is logged. Self-soothing helps disrupt the ruminative cycle and replace the rumination-avoidance association with a rumination-action association. Self-soothing orients users to use “rumination as a call to action,” preparing them for the greater challenge of implementing approach-oriented behaviors. Practicing self-soothing (rather than avoidance) introduces opportunities for positive reinforcement, potentially improving mood. Users are briefed on the distinction between avoidance and self-soothing behaviors and how to discriminate between them ([Multimedia Appendix 4](#)). Specifically, RuminAid emphasizes observing the functional consequences of a behavior on one’s mood and thoughts. If a given behavior results in improvements in mood and disrupts the ruminative thought cycle, it can be used again as an effective self-soothing behavior. In contrast, if a given behavior results in emotional numbness and temporary distraction from ruminative thoughts, the behavior is not an effective self-soothing behavior. Users are also taught that the same behavior (eg, watching a comedy show) might function as an effective self-soothing behavior for one person but as an avoidance behavior for another person. Users are encouraged to try new self-soothing behaviors to explore which options work best for them. They are able to edit their list of self-soothing behaviors at any time.

The second category of alternative behaviors is approach-oriented behaviors. Introduced in quest 4, approach-oriented behaviors are aimed at directly counteracting avoidance. Specifically, RuminAid helps users to create a new, adaptive habitual response to rumination and its triggers, a strategy drawn from EBTs (eg, BA and RFCBT). Users are taught to identify what they are avoiding during rumination by using their rumination topics. They are then instructed to generate and engage in alternative approach-oriented behaviors. These behaviors focus on addressing avoidance head-on and are framed as “facing your fears” (akin to exposure) or “doing the opposite” of the avoidance impulse (akin to opposite action in dialectical behavior therapy). Examples are provided to users to demonstrate these concepts ([Multimedia Appendix 5](#)). In addition, users are provided with tools to identify when self-soothing or approach-oriented behavior techniques may be more helpful in a given context.

Counteracting Rumination-Related Attentional Deficits

Rumination has negative effects on attention and concentration, particularly attention-switching [64-69], that is, the ability to flexibly attend to and adjust behavior in accordance with changes in task goals [70]. Resource allocation theory [71-73] suggests that depression-related thoughts consume cognitive

resources, making it difficult for ruminators to attend to task-relevant processes [74]. Because of this, ruminators may not be fully attentive to their experiences while executing alternative behaviors, which could reduce purported positive effects. RuminAid teaches users to use mindfulness (ie, purposeful, present-focused, and nonjudgmental awareness of internal experiences and external environment; [Multimedia Appendix 6](#)) to counteract these deficits during quest 5. This approach has effectively reduced rumination in other interventions [75-77]. RuminAid users are taught to choose active rather than passive alternative behaviors and to enhance present-moment awareness using the 5 senses. RuminAid also includes formal mindfulness exercises that can be used as alternative behaviors.

Gamification

RuminAid rewards treatment engagement through gamification ([Multimedia Appendix 7](#)). Gamification enhances user experience in mHealth apps by integrating gaming elements (eg, completing a quest map) into the intervention [19]. Although gamification research is limited, initial evidence has suggested that gamified elements reduce attrition and increase engagement by creating enjoyable, engaging, and reinforcing experiences [19,78,79]. Pleasurable gaming experiences trigger dopamine and endorphin release [80]. By triggering this response, game-like experiences in digital interventions may reinforce engagement [81]. A meta-analysis found moderate effect sizes for the effectiveness of gamified digital depression interventions [81].

RuminAid's quest structure and "map" gamifies psychoeducation and assignments. Progressing in a given quest is rewarded by earning a corresponding star on the map, unlocking new app features and map stages, and prompting celebratory messages of encouragement. These elements should facilitate more active and enjoyable user experiences and reinforce engagement.

Goal of This Study

An overwhelming number of mHealth apps claim to treat depression, but a minority are empirically supported [19]. To address this issue directly, we have and will continue to incorporate scientific inquiry into RuminAid from development to dissemination in an iterative, data-driven process. This study aimed to estimate the initial acceptability of RuminAid by evaluating its user-perceived quality and identifying the potential modifications required. We distributed a storyboard presentation of the initial version of RuminAid to potential consumers and collected quantitative and qualitative feedback via individual surveys and focus group interviews. This feedback was used to facilitate improvements in RuminAid before testing its feasibility, acceptability, and effectiveness.

Methods

Participants

We recruited participants from the University at Buffalo Psychology 101 courses. Individuals were eligible to participate if they were native English speakers and adults (age ≥ 18 years). Participants were oversampled for moderate, moderately severe,

or severe depression, but individuals with lower depression scores were also allowed to participate. A total of 22 individuals completed all the required surveys, viewed the complete storyboard presentation, passed the attention check items, and attended a focus group interview. We conducted a total of 7 focus group sessions with groups ranging in size from 1 to 7 participants (mean 3.14, SD 2.04; median 3).

Our sample comprised 77% (17/22) men and 23% (5/22) women. The mean age of the participants was approximately 19 (mean 18.86, SD 0.83) years. The sample was primarily heterosexual (19/22, 86%), and 14% (3/22) of the participants were bisexual. None of the participants endorsed any other sexual orientation. The sample consisted mainly of first- and second-year students: 68% (15/22) were first-year students, 23% (5/22) were second-year students, and 9% (2/22) were third-year students. In terms of marital status, most (20/22, 91%) participants were single and a minority (2/22, 9%) were married or partnered. In terms of religious background, 46% (10/22) identified as Catholic, 9% (2/22) as Protestant, 5% (1/22) as Muslim, and 9% (2/22) as some other religion (Christian nondenominational, $n=1$; spiritual, $n=1$), whereas 32% (7/22) of the participants reported that they did not identify with a religion. Participants were allowed to select all ethnic and racial identities with which they identified. The ethnic and racial makeup of the sample was as follows: 50% (11/22) White, 36% (8/22) Black, 18% (4/22) Latin American, and 9% (2/22) Asian American and Pacific Islander.

Regarding depressive symptoms, 41% (9/22) of the participants reported minimal symptoms of depression based on the Patient Health Questionnaire (PHQ; PHQ-8), 41% (9/22) reported mild depression, 9% (2/22) reported moderate depression, and 9% (2/22) reported moderately severe depression. None of the participants reported experiencing severe depression. The participants had an average PHQ-8 score of 6.55 (SD 5.12) and a median score of 5, which indicated that, overall, the participants experienced mild depression.

Measures

PHQ-8 Measure

The PHQ-8 [82-84] is an 8-item self-report measure of depressive symptoms and severity. Items are rated on a 4-point scale, with higher scores indicating more severe depression. The PHQ-8 omits the self-harm item from the PHQ-9 and is often used in research settings where interventions for suicidality or self-injury are difficult to coordinate [83]. The PHQ-8 has demonstrated reliability and validity as a measure of depression severity [83,84]. The PHQ-8 was used to oversample for moderate or worse depression (PHQ-8 score ≥ 10) at screening and readministered at baseline. Scores at readministration were used for all analyses.

Mobile Application Rating Scale: User Version

The Mobile Application Rating Scale–user version (uMARS [85]) is a 26-item measure of user-rated app quality consisting of the following scales: (1) engagement (degree of fun, interestingness, customizability, interactivity, whether it has prompts such as sending alerts, reminders, etc), (2) functionality (app functioning, ease of use, navigation, flow logic, and

gestural design), (3) aesthetics (graphic design, visual appeal, color scheme, and stylistic consistency), (4) information (contains high-quality information from a credible source), (5) app quality (average mean score of the 4 preceding scales), (6) subjective app quality (overall like or dislike), and (7) perceived impact (whether this app will help the target population with the target problem). Subjective app quality and perceived impact scales can be reported as individual items or mean scores. For our purposes, we used mean scores. To identify qualitative descriptors for each scale (1=inadequate, 2=poor, 3=acceptable, 4=good, and 5=excellent), the average mean scores were rounded to the nearest whole number [86]. The uMARS has demonstrated good test-retest reliability and excellent internal consistency [85].

Focus Group Interview

The semistructured focus group interview was designed specifically for this study to collect qualitative feedback through a standardized set of questions. Initial questions asked participants to elaborate on uMARS ratings and identify potential improvements for items rated <4 by at least one focus group member. Additional questions included, “what app features did you find most helpful?” and “based on your experience trying out RuminAid, do you think this app would help you to identify rumination in your day-to-day life?” The interviews lasted for 1 hour and were moderated by the first author through Zoom videoconferencing (Zoom Video Communications).

Qualitative content analysis [87] was used to systematically identify and describe themes within the participants’ focus group responses. Main categories were generated in a content-driven manner (ie, using theoretical models to derive categories [87]) derived from uMARS scales—engagement, functionality, aesthetics, information, subjective quality, and perceived impact; we added an “other” category for feedback that fell outside of these domains. Subcategories within these domains were generated in a data-driven manner (ie, using the data collected to derive categories [87]). We used subsummation (ie, an iterative process whereby relevant concepts were identified, compared with existing categories, added to existing categories if appropriate, or used to define new categories [87]) to add data-driven subcategories until saturation (ie, the inability to find additional concepts [87]) was achieved. For example, we identified that participants tended to comment on the entertainment value of the app, whether entertainment would impact engagement, and ideas they had about how a social media component could be added to the app to increase engagement. We created 3 corresponding data-driven subcategories within the engagement domain: “Not Entertaining,” “Entertainment not Important,” and “Social Media.”

Transcriptions of the recorded interviews were coded by 2 trained research assistants for the presence or absence of positive and negative statements on each content domain. “Positive statements” were operationalized as comments that were positive in nature, such as liking elements of the app, suggesting that existing elements be retained, recommendations to further capitalize on well-liked components, or explicit agreement with

another participant’s positive statement. “Negative statements” were operationalized as comments that were negative in nature, such as suggesting modifications to a disliked element, recommending the removal of elements, or explicit agreement with another participant’s negative statement. Within a given domain, positive and negative statements were not mutually exclusive, that is, an individual participant could have made both positive and negative statements. Responses were coded according to the presence or absence of statements related to each subcategory.

Procedures

Participants were screened through mass testing using the University at Buffalo’s Sona Systems to determine whether they met the eligibility criteria. Eligible participants signed up for this study on the University at Buffalo’s Sona Systems and attended a web-based focus group session. The sessions were conducted on a web-based videoconferencing platform (Zoom). Upon enrollment, the participants completed a demographics questionnaire and the PHQ-8 via Qualtrics. Next, participants were provided with a link to view the RuminAid storyboard using Panopto, which is a lecture recording and streaming platform. Participants’ progress in viewing and listening to the presentation was tracked using Panopto, which allowed study staff to view the percentage of the presentation the participants had completed. The participants had to complete 100% of the storyboard to proceed with the study. After viewing the storyboard presentation for 2 hours, 2 minutes, and 53 seconds, participants completed the follow-up Qualtrics survey, which consisted of the uMARS and attention-check questions that the participants needed to answer correctly to participate in the focus groups. The correct answers to the attention-check items were embedded within the audio of the storyboard presentation, with clear instructions to make a note of the information for the attention-check items. This ensured that the attention-check items would be easy to answer correctly for participants who attended to the presentation and quite difficult for those who did not. Next, participants attended a scheduled focus group session and provided feedback in a discussion-based format. Focus group sessions were video- and audio-recorded and then transcribed and coded.

Ethical Considerations

The University at Buffalo Institutional Review Board deemed this project exempt from review; all documents used were reviewed and approved by the institutional review board. Before participating, participants were informed that this research was being conducted to investigate the user-perceived quality of a new mHealth smartphone app called RuminAid and that their feedback would be used to improve the quality of the app. We told participants that their responses would be deidentified and stored on a password-protected server. We did not collect identifying information (eg, name, email, and phone number), but participants could provide their email address to opt in to receive a free download of RuminAid once it reached the marketplace; this was entirely optional. Similarly, participants could opt out of their feedback being anonymously quoted in publications about RuminAid. Participants were told that they had the right to end the study at any time and were compensated

with credits that partially fulfilled the course research requirement.

Results

Quantitative Feedback

In our sample, engagement (mean 3.58, SD 0.65), functionality (mean 3.86, SD 0.64), information (mean 4.07, SD 0.46), app quality (mean 3.63, SD 0.47), and perceived impact (mean 3.63, SD 0.47) were rated as “good.” In contrast, aesthetics (mean 3.02, SD 0.66) and subjective app quality (mean 2.89, SD 0.64) were rated as “acceptable.”

Qualitative Feedback

Engagement

During focus group sessions, 50% (11/22) of the participants made positive statements about engagement (ie, degree of fun, interestingness, customizability, interactivity, whether it has prompts such as sending alerts, reminders, etc) and 86% (19/22) made negative statements about engagement. This suggested potentially mixed feelings about how engaging RuminAid was for the participants. Within this domain, 64% (14/22) of the participants reported that they did not find the app entertaining. However, 36% (8/22) of the participants felt that the entertainment value of the app was unimportant, given that its purpose was related to mental health, not entertainment. In addition, 18% (4/22) of the participants suggested adding a social media element to RuminAid to improve engagement.

Functionality

In terms of app functionality (ie, app functioning, ease of use, navigation, flow logic, and gestural design), 55% (12/22) of the participants made positive statements and 41% (9/22) made negative statements. This suggested that there were mixed feelings about app functionality, although most participants identified positive aspects of app functionality. Specifically, 27% (6/22) of the participants noted that RuminAid might be difficult to navigate. In contrast, 50% (11/22) of the participants reported that the flow logic of the app (ie, how the content progresses from one screen to the next) made sense.

Aesthetics

Regarding aesthetics (ie, graphic design, visual appeal, color scheme, and stylistic consistency), 9% (2/22) of the participants made positive statements, whereas 96% (21/22) of the participants made negative statements. This suggested that app aesthetics was a major concern for the participants. Specifically, 91% (20/22) of the participants described the color scheme of RuminAid as boring, “drab,” “depressing,” or dull. In addition, 18% (4/22) of the participants expressed that this issue would interfere with app use and decrease the likelihood of initial or continued app use, although 5% (1/22) of the participants stated that the color scheme would not interfere. In terms of graphics, 45% (10/22) of the participants expressed that these appeared

amateur and 59% (13/22) stated that the graphics appeared outdated.

Information

Participants also commented on the information included in RuminAid (ie, containing high-quality information from a credible source): 77% (17/22) of the participants made positive statements and 32% (7/22) made negative statements. This indicated that, overall, information appeared to be a relative strength of the app. More specifically, 73% (16/22) of the participants stated that the amount of detail included in the information was appropriate and to their liking. A minority (4/22, 18%) of the participants found that there was excessive detail. None of the participants reported insufficient detail. Furthermore, 59% (13/22) of the participants reported learning something new from RuminAid. In terms of credibility, 18% (4/22) of the participants explicitly reported feeling that information came from a credible source; however, 23% (5/22) of the participants felt that credibility could be improved if the sources were cited within the app.

Subjective Quality

The subjective quality of the app (ie, overall like or dislike of the app) was overwhelmingly positive: 96% (21/22) of the participants made positive statements about the subjective quality of RuminAid. None of the participants made negative statements regarding the subjective app quality. This suggested that the participants felt positive about the app as a whole.

Perceived Impact

Participants endorsed overwhelmingly positive beliefs about the perceived impact of RuminAid: 96% (21/22) of the participants indicated that they believed RuminAid would help depressed ruminators with their depression and rumination. None of the participants made negative statements regarding the perceived impact of RuminAid. This indicated that the participants saw the app as potentially effective and helpful for depression and rumination.

Other

Approximately 18% (4/22) of the participants made positive statements and 36% (8/22) made negative statements that did not fit within the aforementioned domains. Specifically, 46% (10/22) of the participants reported that integration of measurement-based care features would improve the app, and 23% (5/22) of the participants raised concerns that “time gating” (ie, when users are prevented from accessing new app content until a certain amount of time has passed) quests with a 1-week delay between quests (as originally planned) would negatively impact app quality, frustrate users, and decrease engagement.

Example quotes from participants and the interrater reliability for each domain and subcategory mentioned above are included in [Tables 1](#) and [2](#). Examples of modifications based on focus group feedback can be seen in [Figures 1-3](#).

Table 1. Domains of qualitative focus group feedback^a.

Domain	Description	Quotes from participants	Cohen κ^b
Engagement			
Positive	App is fun, interesting, customizable, or interactive and has prompts (eg, sends alerts, messages, reminders, or feedback or enables sharing)	<ul style="list-style-type: none"> “I feel like [engagement] is kind of the whole point. And I think that it does it pretty well.” “I really like games. So, going through the levels—I really liked that and how it was kind of an incentive to keep going.” 	0.91
Negative	App is not fun, interesting, customizable, or interactive or does not have prompts (eg, sends alerts, messages, reminders, or feedback or enables sharing)	<ul style="list-style-type: none"> “I would say it could be a little more interactive, because the steps are just so repetitive and it’s kind of the same thing going through each task.” 	1
Functionality			
Positive	App functioning well; easy to learn; or good navigation, flow logic, or gestural design	<ul style="list-style-type: none"> “I thought the format where everything is laid out, like, the layout is good.” “My favorite part of the app personally was the quest system put in place.” 	0.47
Negative	App not functioning well; not easy to learn; or poor navigation, flow logic, and gestural design	<ul style="list-style-type: none"> “[You should] make the flow a little smoother.” 	0.38
Aesthetics			
Positive	Graphic design, overall visual appeal, color scheme, and stylistic consistency are good or appealing	<ul style="list-style-type: none"> “The brain graphic...was pretty good with the pink on blue brain contrast.” 	0.8
Negative	Graphic design, overall visual appeal, color scheme, and stylistic consistency are not good or unappealing	<ul style="list-style-type: none"> “I feel like a complete artistic overhaul of the app needs to be done.” 	0.38
Information			
Positive	App contains high-quality information (eg, text, feedback, measures, or references) from a credible source	<ul style="list-style-type: none"> “It had a lot of solid information that would be useful to someone struggling with depression.” 	1
Negative	App does not contain high-quality information (eg, text, feedback, measures, and references) or lacks credible source	<ul style="list-style-type: none"> “Some of the information seems sort of redundant.” 	0.51
Subjective quality			
Positive	Overall positive impression or liked the app	<ul style="list-style-type: none"> “Everything that was presented was presented clearly. So, I was able to retain it better, and actually learn about it...I also like some of the smaller, finer details...certain ways of rewarding you for staying on the app.” 	1
Negative	Overall negative impression or did not like the app	<ul style="list-style-type: none"> No participants made negative comments about subjective quality. 	1
Perceived impact			
Positive	This app will help the target population (ie, depressed ruminators) with the problem (ie, depression and rumination)	<ul style="list-style-type: none"> “I think it would help because it does have those red flag areas, you make it customizable for you...I think that it would definitely be helpful for people.” “If I were struggling with rumination and depression, it would be eye-opening in a way...If you didn’t realize you were doing that, or how it affected you, you’re going to get a new perspective to see why it’s affecting you and how you can stop ruminating.” 	1
Negative	This app will not help the target population (ie, depressed ruminators) with the problem (ie, depression and rumination)	<ul style="list-style-type: none"> No participants made negative comments about perceived impact. 	1
Other			

Domain	Description	Quotes from participants	Cohen κ^b
Positive	Additional themes that emerged but did not fit into the above categories; positive valence to this feedback	<ul style="list-style-type: none"> “When I first started looking at it, I thought it was going to be too short...I like how it’s a process, because at first I almost wrote it off. I felt like it was kind of just like, ‘this is a quick fix’ type thing. But a lot of mental health things aren’t a quick fix. So, I like how it was a process that you go through and you can plan out ahead of time and that’s how you can really make a change with it is by making it a process.” 	0.31
Negative	Additional themes that emerged but did not fit into the above categories; negative valence to this feedback	<ul style="list-style-type: none"> “I would say maybe like further on down the road, make sure you keep updating it, because if someone uses it frequently, they could very easily go through all quests very quick. And then after that, there’s really no use for it.” 	0.36

^aDescriptions and examples of categories identified in qualitative focus group feedback.

^bCohen κ represents an estimate of interrater reliability for qualitative items.

Table 2. Subcategories of qualitative focus group feedback^a.

Subcategories	Description	Quotes from participants	Cohen κ^b
Engagement			
Not entertaining	The app was not entertaining	<ul style="list-style-type: none"> “I saw a lot of stick figure-like images. And if they were just replaced with humans or something doing the same thing, it would have been more entertaining.” 	0.8
Entertainment not important	The entertainment value is not relevant to this app, is not an important factor, or is not the purpose of the app and would not impact the likelihood of using it	<ul style="list-style-type: none"> “I don’t really feel like it’s intended to be entertaining. It’s meant to help somebody.” “I wouldn’t, at least for me, you know, as somebody that would be downloading it, I don’t consider the entertainment aspect necessarily a priority.” 	0.56
Social media	Adding a social media component to the app would make it more entertaining or engaging	<ul style="list-style-type: none"> “I would [add] a feature where you can connect with other users and maybe talk about with another actual person, like what’s going on, and they can give you like feedback or something like that.” 	1
Functionality			
Difficult navigation	The app appears difficult to navigate	<ul style="list-style-type: none"> “I felt like...it jumped around a little bit. [I suggest] making the flow of it better.” 	0.38
Logical flow	The app flow makes sense	<ul style="list-style-type: none"> “I liked how it was kind of structured so that you aren’t just having a bunch of information dumped on you, it’s kind of separated into these five larger segments, that you slowly make your way through with demonstrations that you do yourself.” 	0.54
Aesthetics			
Boring color scheme	The color scheme was boring, drab, “depressing,” etc	<ul style="list-style-type: none"> “Other mental health apps that I’ve seen typically have a little bit better contrast, just things to make the other menu items pop out a little bit more.” “The color could use a little updating.” 	0.32
Color scheme interferes	The color scheme would interfere with app use or the likelihood of using the app	<ul style="list-style-type: none"> “Sometimes people’s emotions can be affected by the colors they see and all that. So just brighten it up.” 	0.54
Color scheme does not interfere	The color scheme would not interfere with app use or the likelihood of using the app	<ul style="list-style-type: none"> “When you need the help, you need the help. And I don’t think making a whole bunch of pretty colors and all that is really going to change what the app is doing.” 	1
Amateur graphics	Graphics looked amateur	<ul style="list-style-type: none"> “People will trust it more if it’s looks more professional.” “In terms of just the buttons, specifically, the little arrow icons, took up half the screen, almost. And it just kind of looked like out of place and too big. And proportions like that matter if you’re going to have a clear and concise experience with the app.” 	0.67
Outdated graphics	Graphics looked outdated	<ul style="list-style-type: none"> “It just looked a little outdated, especially the pictures. Like it’s not current.” 	0.35
Information			
Good detail	Liked the level of detail; found it appropriate	<ul style="list-style-type: none"> “Yeah, I think I think it does a really good job and gives very detailed information.” 	0.89
Insufficient detail	Insufficient detail; too little detail	<ul style="list-style-type: none"> No participants made comments about insufficient detail. 	1
Excessive detail	Felt overwhelmed by information; too much detail	<ul style="list-style-type: none"> “It’s just like, a lot of information, I guess, and a lot of things to do. I think, personally, I would get a little confused on the app.” 	1

Subcategories	Description	Quotes from participants	Cohen κ ^b
Learned	Learned new things	<ul style="list-style-type: none"> “And I wasn’t really like educated on the subject before. So that’s something I did like.” “I learned a lot.” 	0.81
Credible information	Information seemed credible	<ul style="list-style-type: none"> “I don’t have any reason to believe that it’s not a credible source.” 	0.62
Credible source needed	Including information such as references, expert videos, or other ways to make the credibility explicit would be useful	<ul style="list-style-type: none"> “I feel like there could be some kind of link to information that has credited sources, and not just put those facts up with no way to immediately check the background of it.” 	0.51
Other			
Measurement-based care	Measurement-based care would improve the app (eg, tracking ruminative episodes over time)	<ul style="list-style-type: none"> “I think the tracking is a good idea, especially to see how much you’ve progressed since starting the app.” 	0.06
Time gating	Restricting people’s access to parts of the quest based on time (ie, 1 week before you can move on to the next quest) might be frustrating to users	<ul style="list-style-type: none"> “People do not like time-gated content, especially artificially time gating.” 	0.64

^aDescriptions and examples of subcategories identified in qualitative focus group feedback.

^bCohen κ represents an estimate of interrater reliability for qualitative items.

Figure 1. Home screen redesign. RuminAid home screen before user-centered redesign (left) and after user-centered redesign (right).



Figure 2. Ask an expert screen redesign. RuminAid’s first set of “Ask an Expert” screens before user-centered redesign (left) and after user-centered redesign (right).

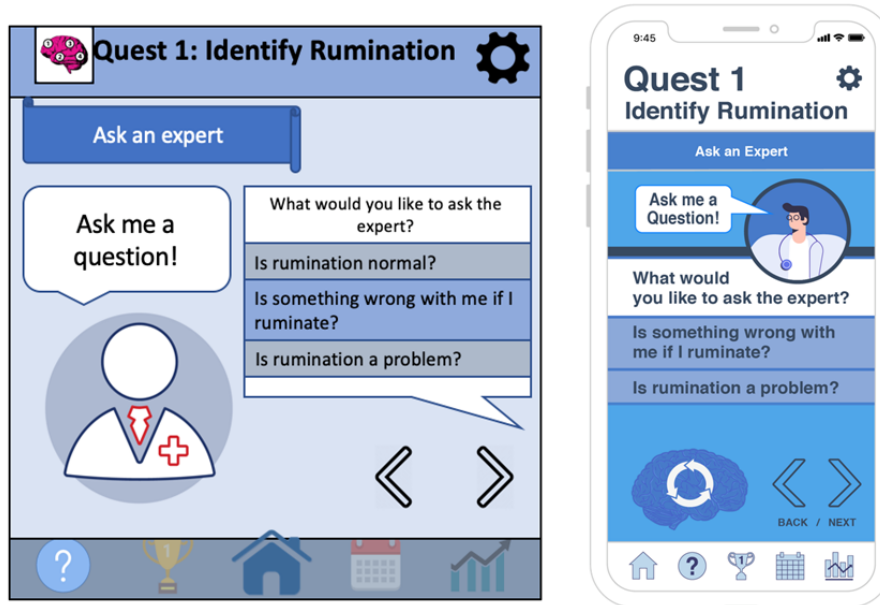
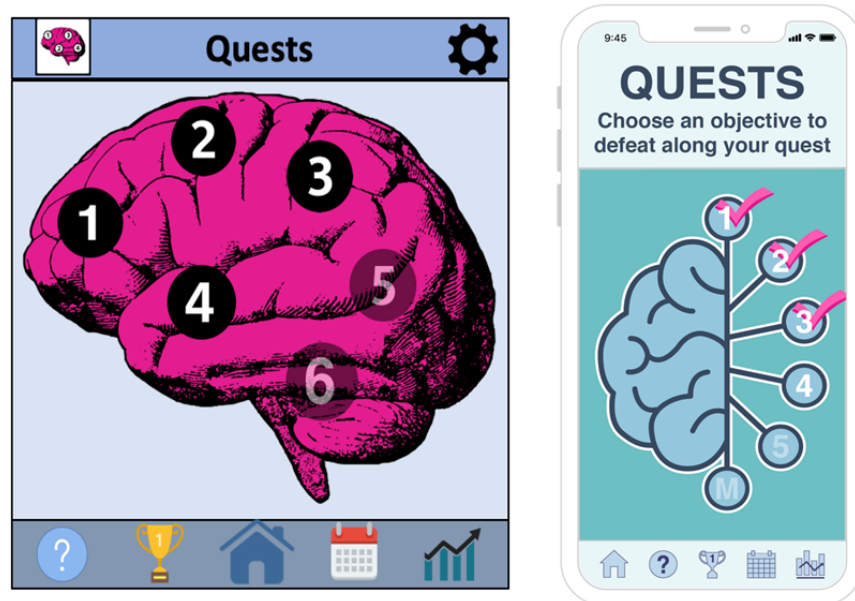


Figure 3. Quest launch screen redesign. Quest 4 launch screen before user-centered redesign (left) and after user-centered redesign (right).



Discussion

Principal Findings

This study used a mixed methods approach and aimed to estimate the initial acceptability of RuminAid by evaluating its user-perceived quality and identifying potential modifications needed. The results of this study suggested that, overall, RuminAid was perceived as “acceptable” to “good” by the focus group participants. In terms of strengths, participants highlighted the information and perceived impact of RuminAid: focus group members reported that the quality of information was quite good and had the potential to help adults who struggle with depression and rumination. In contrast, focus group members expressed concern that poor aesthetics could interfere with users’ desire to continue using the app. Indeed, focus group members highlighted RuminAid’s aesthetics as its primary weakness.

In particular, participants felt that the visual elements of the app—especially the color scheme and graphics—appeared amateur, outdated, and dull. Specific suggestions for improvement included a brighter color scheme, modern aesthetics, and professionally designed graphics. To address these concerns, we hired a graphic designer to redesign all RuminAid screens. The graphic designer was instructed to revise the app screens to look more professional, modern, and colorful, based on participant recommendations. Comparisons of the RuminAid app screens before and after focus group testing are displayed in [Figures 1-3](#) to demonstrate how the data collected in this study directly informed the modifications.

In addition to redesigning the app screens so that they were more aesthetically pleasing, we modified the RuminAid timeline. Although only a minority of focus group participants commented on the “time gating” involved in RuminAid (ie, restricting users

from moving from one quest to the next for a period of 1 week), this feedback highlighted potential frustration and loss of interest that might result from this restriction. This is reflected in the literature, which indicates that people use phone apps sporadically [23,88-90], and in lay commentary, which suggested that people dislike apps and games with time-gated content [91]. In addition, attrition tends to be high in mHealth interventions [92-94]. Thus, to prevent frustration and attrition, we reduced RuminAid from a 5-week intervention restricting users to completing 1 quest per week to a 5-day intervention with instructions (but no time gates) to complete 1 quest per day. Notably, we retained the calendar feature for scheduling quests but modified the instructions to fit the revised timeline.

Limitations

It is possible that demand characteristics and group factors may have played a role in participants reporting that they liked the app in a focus group setting led by a researcher [95]. During the interviews, positive statements about the subjective app quality were made by all but 1 participant. In contrast, the subjective app quality scores obtained via uMARS suggested less satisfaction. However, the uMARS subjective quality scale includes items such as, "would you pay for this app?" The sample of focus group participants consisted entirely of college students, who may be unlikely to pay for any app. More broadly, monetary commitment may not indicate an individual's true feelings regarding app quality. For example, some individuals may never feel comfortable paying for a smartphone app, regardless of what the app is. Notably, the app quality domain, based on mean scores of engagement, information, functionality, and aesthetics, was more consistent with the feedback provided during focus group interviews. Nevertheless, it is possible that demand characteristics and group factors may have played a role in participants reporting that they liked the app in a focus group setting led by a researcher. In the future, it might be helpful to have the research assistant running the focus group explicitly state that they were not involved in app development. Furthermore, it may be more effective to collect supplemental qualitative feedback via one-on-one interviews to help reduce feedback biases owing to group factors, such as groupthink or reluctance to dissent [96].

Another potential limitation of our study was the use of a sample of college students. We intentionally selected a sample of college students because this population tends to be young and highly digitally literate. As such, college students often have extensive experience using a variety of high-quality smartphone apps with which they could meaningfully compare RuminAid to provide detailed qualitative feedback informed by the current standards of app quality. However, the use of a college student sample may have potentially limited the amount of critical feedback related to app functionality, particularly flow logic. For example, older adults with less digital literacy may have identified more potential difficulties with app navigation and functionality. As we plan for the next stages of this research, it is important to seek feedback from stakeholders of all ages to identify a broader spectrum of potential modifications before making RuminAid available on the public marketplace.

Finally, participants interacted with a story board rather than a beta version of the app itself, and there were limitations to seeking feedback on an app that participants did not have the opportunity to use. Without having first-hand experience of using RuminAid, the participants may have relied on personal biases, expectations, and perceptions to provide feedback on the app's usability.

Future Directions

In the next phase of this research, we are most interested in evaluating the core therapeutic content contained within RuminAid as an intervention. As such, we will conduct a pilot trial of RuminAid as an intervention for depression and rumination, which will allow us to examine the app's feasibility and acceptability and provide initial estimates of its effectiveness among a community sample of adults with depression and rumination. Subsequently, we may conduct a case series study of a beta version of RuminAid. To do so, we may have a community sample of adults with depression install RuminAid and use it for a week before interviewing them individually about their experience. The rich qualitative data collected in such a study would inform any additional modifications made to RuminAid before it is made available to the public.

Once the feasibility, acceptability, and effectiveness of the basic RuminAid approach are established with support for its core therapeutic content, additional features could be incorporated. For example, a substantial percentage of the participants stated that integration of measurement-based care features would be favorable. Measurement-based care has been shown to improve outcomes of treatment for depression [97,98]. Therefore, a longer-term future direction would be to develop measurement-based care features and evaluate whether outcomes are enhanced. Likewise, several participants mentioned wanting a social media component to help them feel more engaged. The incorporation of social media has been shown to improve app engagement [98] and may help to improve long-term app use [99]; therefore, we may also develop and include social features in future iterations of RuminAid.

Conclusions

RuminAid was intended to address the need for accessible EBT for depression by targeting a key risk factor and maintenance factor—rumination. Although the overwhelming majority of mHealth apps are neither empirically tested nor empirically informed, RuminAid incorporates recent research findings and elements of EBTs for rumination and depression. Should research support its feasibility, acceptability, and effectiveness, RuminAid stands to increase access to treatment for rumination, depression, and their negative sequelae. Importantly, this intervention involved no therapist contact, making it inexpensive and timely, given the recent COVID-19 pandemic and the need for remote service delivery options. The self-guided nature of RuminAid may also minimize stigma-related concerns and "better engage individuals who have limited access to or interest in traditional face-to-face interventions.

In summary, the overall quality and treatment approach used in RuminAid was acceptable to potential users. The feedback obtained in this study directly informed the modifications to

RuminAid, and the usability, acceptability, and feasibility of this modified version will be tested in a pilot trial. In terms of its potential impact, RuminAid has broad applications. It is a contactless intervention, based on EBTs, and could reduce the burden on the mental health care system by offsetting care to automated service delivery. If effective, RuminAid could be implemented and tested in a variety of settings such as primary care, behavioral health integrated care, stepped care facilities, or remote care clinics or offered to individuals on clinic waitlists.

Potential future augmentations to RuminAid could include peer coaching, therapist coaching, or group facilitation. Consequently, pending empirical evaluation, RuminAid may be an accessible and effective intervention for depression and rumination with positive public health consequences, and its development process could serve as a road map for developing evidence-based mHealth apps and empirically supported mHealth apps.

Acknowledgments

The authors would like to thank Steven Redler, Brendon Jones, and Daniel Lesniak for their contributions to the development of RuminAid, without whom this project would not have been possible. In addition, the authors thank Ian Lechevet for his dedicated contributions as a research assistant on this project.

Data Availability

All quantitative data sets generated and analyzed during this study are available at Open Science Framework [100]. Full access to the qualitative data collected during focus group feedback sessions is available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Psychoeducational content screens. These screens from quests 1, 2, and 5 display psychoeducational content to users to enhance their understanding before skills application. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study.

[[PNG File , 272 KB - formative_v6i12e40045_app1.png](#)]

Multimedia Appendix 2

Identification and labeling of rumination screens. These screens display the in-app use of skills related to identifying and labeling rumination. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study.

[[PNG File , 239 KB - formative_v6i12e40045_app2.png](#)]

Multimedia Appendix 3

Alternative behavior screens. These screens from quests 3 and 4 display psychoeducation about alternative behaviors as well as in-app skills use. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study.

[[PNG File , 236 KB - formative_v6i12e40045_app3.png](#)]

Multimedia Appendix 4

Self-soothing behavior screens. These screens from quest 3 demonstrate how users learn to discriminate between self-soothing and avoidance behaviors. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study.

[[PNG File , 319 KB - formative_v6i12e40045_app4.png](#)]

Multimedia Appendix 5

Approach behavior screens. These screens from quest 4 display examples of identification of avoidance (left) and generating approach behaviors (right). These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study.

[[PNG File , 148 KB - formative_v6i12e40045_app5.png](#)]

Multimedia Appendix 6

Mindfulness screens. These screens from quest 5 display psychoeducation, in-app skills use, and tips for mindfulness practice. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study. [\[PNG File , 245 KB - formative_v6i12e40045_app6.png \]](#)

Multimedia Appendix 7

Gamified elements screens. These screens display gamified elements of RuminAid, including the quest 1 map in progress and at completion. These images are screenshots of RuminAid after the user-centered redesign, which was informed by the results of this study. [\[PNG File , 219 KB - formative_v6i12e40045_app7.png \]](#)

References

1. Depression. World Health Organization. 2021 Sep 13. URL: <https://www.who.int/news-room/fact-sheets/detail/depression> [accessed 2022-09-16]
2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003 Jun 18;289(23):3095-3105. [doi: [10.1001/jama.289.23.3095](https://doi.org/10.1001/jama.289.23.3095)] [Medline: [12813115](https://pubmed.ncbi.nlm.nih.gov/12813115/)]
3. McLaughlin KA. The public health impact of major depression: a call for interdisciplinary prevention efforts. *Prev Sci* 2011 Dec 6;12(4):361-371 [FREE Full text] [doi: [10.1007/s11121-011-0231-8](https://doi.org/10.1007/s11121-011-0231-8)] [Medline: [21732121](https://pubmed.ncbi.nlm.nih.gov/21732121/)]
4. McKenna MT, Michaud CM, Murray CJ, Marks JS. Assessing the burden of disease in the United States using disability-adjusted life years. *Am J Prev Med* 2005 Jun;28(5):415-423. [doi: [10.1016/j.amepre.2005.02.009](https://doi.org/10.1016/j.amepre.2005.02.009)] [Medline: [15894144](https://pubmed.ncbi.nlm.nih.gov/15894144/)]
5. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. *JAMA Psychiatry* 2015 Apr;72(4):334-341 [FREE Full text] [doi: [10.1001/jamapsychiatry.2014.2502](https://doi.org/10.1001/jamapsychiatry.2014.2502)] [Medline: [25671328](https://pubmed.ncbi.nlm.nih.gov/25671328/)]
6. Angst J, Angst F, Stassen HH. Suicide risk in patients with major depressive disorder. *J Clin Psychiatry* 1999;60 Suppl 2:57-62; discussion 75. [Medline: [10073389](https://pubmed.ncbi.nlm.nih.gov/10073389/)]
7. Lépine JP, Briley M. The increasing burden of depression. *Neuropsychiatr Dis Treat* 2011;7(Suppl 1):3-7 [FREE Full text] [doi: [10.2147/NDT.S19617](https://doi.org/10.2147/NDT.S19617)] [Medline: [21750622](https://pubmed.ncbi.nlm.nih.gov/21750622/)]
8. Anxiety and Depression. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm> [accessed 2020-10-28]
9. González HM, Vega WA, Williams DR, Tarraf W, West BT, Neighbors HW. Depression care in the United States: too little for too few. *Arch Gen Psychiatry* 2010 Jan 01;67(1):37-46 [FREE Full text] [doi: [10.1001/archgenpsychiatry.2009.168](https://doi.org/10.1001/archgenpsychiatry.2009.168)] [Medline: [20048221](https://pubmed.ncbi.nlm.nih.gov/20048221/)]
10. Hahn HC, Cook BL, Ault-Brutus A, Alegría M. Intersection of race-ethnicity and gender in depression care: screening, access, and minimally adequate treatment. *Psychiatr Serv* 2015 Mar 01;66(3):258-264 [FREE Full text] [doi: [10.1176/appi.ps.201400116](https://doi.org/10.1176/appi.ps.201400116)] [Medline: [25727113](https://pubmed.ncbi.nlm.nih.gov/25727113/)]
11. Thornicroft G, Chatterji S, Evans-Lacko S, Gruber M, Sampson N, Aguilar-Gaxiola S, et al. Undertreatment of people with major depressive disorder in 21 countries. *Br J Psychiatry* 2017 Feb;210(2):119-124 [FREE Full text] [doi: [10.1192/bjp.bp.116.188078](https://doi.org/10.1192/bjp.bp.116.188078)] [Medline: [27908899](https://pubmed.ncbi.nlm.nih.gov/27908899/)]
12. Alegría M, Chatterji P, Wells K, Cao Z, Chen C, Takeuchi D, et al. Disparity in depression treatment among racial and ethnic minority populations in the United States. *Psychiatr Serv* 2008 Nov;59(11):1264-1272 [FREE Full text] [doi: [10.1176/ps.2008.59.11.1264](https://doi.org/10.1176/ps.2008.59.11.1264)] [Medline: [18971402](https://pubmed.ncbi.nlm.nih.gov/18971402/)]
13. Bailey R, Patel M, Barker N, Ali S, Jabeen S. Major depressive disorder in the African American population. *J National Med Assoc* 2011 Jul;103(7):548-559 [FREE Full text] [doi: [10.1016/s0027-9684\(15\)30380-1](https://doi.org/10.1016/s0027-9684(15)30380-1)]
14. Bojdani E, Rajagopalan A, Chen A, Gearin P, Olcott W, Shankar V, et al. COVID-19 pandemic: impact on psychiatric care in the United States. *Psychiatry Res* 2020 Jul;289:113069 [FREE Full text] [doi: [10.1016/j.psychres.2020.113069](https://doi.org/10.1016/j.psychres.2020.113069)] [Medline: [32413707](https://pubmed.ncbi.nlm.nih.gov/32413707/)]
15. Conner C. Fifty essential mobile marketing facts. *Forbes*. 2013 Nov 12. URL: <https://www.forbes.com/sites/cherylsnappconner/2013/11/12/fifty-essential-mobile-marketing-facts/> [accessed 2021-01-03]
16. Mobile Fact Sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2022-04-28]
17. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS One* 2016;11(5):e0154248 [FREE Full text] [doi: [10.1371/journal.pone.0154248](https://doi.org/10.1371/journal.pone.0154248)] [Medline: [27135410](https://pubmed.ncbi.nlm.nih.gov/27135410/)]
18. MoodTools Review. One Mind PsyberGuide. URL: <https://onemindpsyberguide.org/apps/moodtools/> [accessed 2022-04-28]
19. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. *Front Psychiatry* 2016;7:215 [FREE Full text] [doi: [10.3389/fpsy.2016.00215](https://doi.org/10.3389/fpsy.2016.00215)] [Medline: [28119636](https://pubmed.ncbi.nlm.nih.gov/28119636/)]

20. Dahne J, Lejuez CW, Kustanowitz J, Felton JW, Diaz VA, Player MS, et al. Moodivate: a self-help behavioral activation mobile app for utilization in primary care-Development and clinical considerations. *Int J Psychiatry Med* 2017 Mar;52(2):160-175 [FREE Full text] [doi: [10.1177/0091217417720899](https://doi.org/10.1177/0091217417720899)] [Medline: [28792292](https://pubmed.ncbi.nlm.nih.gov/28792292/)]
21. Dahne J, Lejuez C, Diaz VA, Player MS, Kustanowitz J, Felton JW, et al. Pilot randomized trial of a self-help behavioral activation mobile app for utilization in primary care. *Behav Ther* 2019 Jul;50(4):817-827 [FREE Full text] [doi: [10.1016/j.beth.2018.12.003](https://doi.org/10.1016/j.beth.2018.12.003)] [Medline: [31208690](https://pubmed.ncbi.nlm.nih.gov/31208690/)]
22. Dahne J. Evaluation of a mobile therapy for elevated depressive symptoms. *ClinicalTrials.gov*. URL: <https://clinicaltrials.gov/ct2/show/NCT04463914> [accessed 2022-04-27]
23. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. Intellicare: an eclectic, skills-based app suite for the treatment of depression and anxiety. *J Med Internet Res* 2017 Jan 05;19(1):e10 [FREE Full text] [doi: [10.2196/jmir.6645](https://doi.org/10.2196/jmir.6645)] [Medline: [28057609](https://pubmed.ncbi.nlm.nih.gov/28057609/)]
24. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S. Rethinking rumination. *Perspect Psychol Sci* 2008 Sep;3(5):400-424. [doi: [10.1111/j.1745-6924.2008.00088.x](https://doi.org/10.1111/j.1745-6924.2008.00088.x)] [Medline: [26158958](https://pubmed.ncbi.nlm.nih.gov/26158958/)]
25. Nolen-Hoeksema S. Responses to depression and their effects on the duration of depressive episodes. *J Abnormal Psychol* 1991 Nov;100(4):569-582. [doi: [10.1037/0021-843x.100.4.569](https://doi.org/10.1037/0021-843x.100.4.569)]
26. Smith JM, Alloy LB. A roadmap to rumination: a review of the definition, assessment, and conceptualization of this multifaceted construct. *Clin Psychol Rev* 2009 Mar;29(2):116-128 [FREE Full text] [doi: [10.1016/j.cpr.2008.10.003](https://doi.org/10.1016/j.cpr.2008.10.003)] [Medline: [19128864](https://pubmed.ncbi.nlm.nih.gov/19128864/)]
27. Smith ER, Henry S. An in-group becomes part of the self: response time evidence. *Pers Soc Psychol Bull* 2016 Jul 02;22(6):635-642. [doi: [10.1177/0146167296226008](https://doi.org/10.1177/0146167296226008)]
28. Martin L, Tesser A. Some ruminative thoughts. In: *Ruminative Thoughts Advances in Social Cognition, Volume IX*. New York: Psychology Press; 1996.
29. Morrison R, O'Connor RC. A systematic review of the relationship between rumination and suicidality. *Suicide Life Threat Behav* 2008 Oct;38(5):523-538. [doi: [10.1521/suli.2008.38.5.523](https://doi.org/10.1521/suli.2008.38.5.523)] [Medline: [19014305](https://pubmed.ncbi.nlm.nih.gov/19014305/)]
30. Watkins E. *Rumination-Focused Cognitive-Behavioral Therapy for Depression*. New York: Guilford Publications; 2018.
31. Just N, Alloy LB. The response styles theory of depression: tests and an extension of the theory. *J Abnormal Psychol* 1997 May;106(2):221-229. [doi: [10.1037/0021-843x.106.2.221](https://doi.org/10.1037/0021-843x.106.2.221)]
32. Nolan S, Roberts J, Gotlib I. Neuroticism and ruminative response style as predictors of change in depressive symptomatology. *Cognit Ther Res* 1998;22(5):445-455. [doi: [10.1023/A:1018769531641](https://doi.org/10.1023/A:1018769531641)]
33. Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: a psychometric analysis. *Cognit Ther Res* 2003;27:247-259. [doi: [10.1023/A:1023910315561](https://doi.org/10.1023/A:1023910315561)]
34. Nolen-Hoeksema S. The response styles theory. In: *Depressive Rumination: Nature, Theory and Treatment*. Hoboken, New Jersey: Wiley; 2003.
35. Roberts J, Gilboa E, Gotlib I. Ruminative response style and vulnerability to episodes of dysphoria: gender, neuroticism, and episode duration. *Cognit Ther Res* 1998;22:401-423. [doi: [10.1023/A:1018713313894](https://doi.org/10.1023/A:1018713313894)]
36. Spasojević J, Alloy L, Abramson L, Maccoon D, Robinson M. Reactive rumination: outcomes, mechanisms, and developmental antecedents. In: *Depressive Rumination: Nature, Theory and Treatment*. Hoboken, New Jersey, United States: Wiley; 2003.
37. Martell C, Addis M, Jacobson N. *Depression in Context: Strategies for Guided Action*. US: W.W. Norton; 2001.
38. De Raedt R, Koster EH. Understanding vulnerability for depression from a cognitive neuroscience perspective: a reappraisal of attentional factors and a new conceptual framework. *Cogn Affect Behav Neurosci* 2010 Mar;10(1):50-70. [doi: [10.3758/CABN.10.1.50](https://doi.org/10.3758/CABN.10.1.50)] [Medline: [20233955](https://pubmed.ncbi.nlm.nih.gov/20233955/)]
39. Buckman J, Underwood A, Clarke K, Saunders R, Hollon S, Fearon P, et al. Risk factors for relapse and recurrence of depression in adults and how they operate: a four-phase systematic review and meta-synthesis. *Clin Psychol Rev* 2018 Aug;64:13-38 [FREE Full text] [doi: [10.1016/j.cpr.2018.07.005](https://doi.org/10.1016/j.cpr.2018.07.005)] [Medline: [30075313](https://pubmed.ncbi.nlm.nih.gov/30075313/)]
40. Hick SF, Chan L. Mindfulness-based cognitive therapy for depression: effectiveness and limitations. *Social Work Mental Health* 2010 Apr 13;8(3):225-237. [doi: [10.1080/15332980903405330](https://doi.org/10.1080/15332980903405330)]
41. Michalak J, Hölz A, Teismann T. Rumination as a predictor of relapse in mindfulness-based cognitive therapy for depression. *Psychol Psychother* 2011 Jun;84(2):230-236. [doi: [10.1348/147608310X520166](https://doi.org/10.1348/147608310X520166)] [Medline: [22903859](https://pubmed.ncbi.nlm.nih.gov/22903859/)]
42. Spinhoven P, Drost J, de Rooij M, van Hemert AM, Penninx BW. Is experiential avoidance a mediating, moderating, independent, overlapping, or proxy risk factor in the onset, relapse and maintenance of depressive disorders? *Cognit Ther Res* 2016 Dec 30;40(2):150-163 [FREE Full text] [doi: [10.1007/s10608-015-9747-8](https://doi.org/10.1007/s10608-015-9747-8)] [Medline: [27069286](https://pubmed.ncbi.nlm.nih.gov/27069286/)]
43. Watkins ER. Constructive and unconstructive repetitive thought. *Psychol Bull* 2008 Mar;134(2):163-206 [FREE Full text] [doi: [10.1037/0033-2909.134.2.163](https://doi.org/10.1037/0033-2909.134.2.163)] [Medline: [18298268](https://pubmed.ncbi.nlm.nih.gov/18298268/)]
44. Watkins ER. Depressive rumination: investigating mechanisms to improve cognitive behavioural treatments. *Cogn Behav Ther* 2009 Jan;38 Suppl 1(sup1):8-14 [FREE Full text] [doi: [10.1080/16506070902980695](https://doi.org/10.1080/16506070902980695)] [Medline: [19697180](https://pubmed.ncbi.nlm.nih.gov/19697180/)]
45. Joormann J, Dkane M, Gotlib I. Adaptive and maladaptive components of rumination? Diagnostic specificity and relation to depressive biases. *Behav Ther* 2006 Sep;37(3):269-280 [FREE Full text] [doi: [10.1016/j.beth.2006.01.002](https://doi.org/10.1016/j.beth.2006.01.002)] [Medline: [16942978](https://pubmed.ncbi.nlm.nih.gov/16942978/)]

46. Watkins ER. Depressive rumination and co-morbidity: evidence for brooding as a transdiagnostic process. *J Ration Emot Cogn Behav Ther* 2009 Sep 7;27(3):160-175 [[FREE Full text](#)] [doi: [10.1007/s10942-009-0098-9](https://doi.org/10.1007/s10942-009-0098-9)] [Medline: [19718267](https://pubmed.ncbi.nlm.nih.gov/19718267/)]
47. Wong Q, Moulds M. Impact of rumination versus distraction on anxiety and maladaptive self-beliefs in socially anxious individuals. *Behav Res Ther* 2009 Oct;47(10):861-867 [[FREE Full text](#)] [doi: [10.1016/j.brat.2009.06.014](https://doi.org/10.1016/j.brat.2009.06.014)] [Medline: [19608157](https://pubmed.ncbi.nlm.nih.gov/19608157/)]
48. Jenness JL, Jager-Hyman S, Heleniak C, Beck AT, Sheridan MA, McLaughlin KA. Catastrophizing, rumination, and reappraisal prospectively predict adolescent PTSD symptom onset following a terrorist attack. *Depress Anxiety* 2016 Nov 24;33(11):1039-1047 [[FREE Full text](#)] [doi: [10.1002/da.22548](https://doi.org/10.1002/da.22548)] [Medline: [27557454](https://pubmed.ncbi.nlm.nih.gov/27557454/)]
49. Naumann E, Tuschen-Caffier B, Voderholzer U, Caffier D, Svaldi J. Rumination but not distraction increases eating-related symptoms in anorexia and bulimia nervosa. *J Abnorm Psychol* 2015 May;124(2):412-420. [doi: [10.1037/abn0000046](https://doi.org/10.1037/abn0000046)] [Medline: [25730516](https://pubmed.ncbi.nlm.nih.gov/25730516/)]
50. Startup H, Lavender A, Oldershaw A, Stott R, Tchanturia K, Treasure J, et al. Worry and rumination in anorexia nervosa. *Behav Cogn Psychother* 2012 Oct 29;41(3):301-316. [doi: [10.1017/s1352465812000847](https://doi.org/10.1017/s1352465812000847)]
51. Nolen-Hoeksema S, Harrell ZA. Rumination, depression, and alcohol use: tests of gender differences. *J Cogn Psychother* 2002 Oct 01;16(4):391-403. [doi: [10.1891/jcop.16.4.391.52526](https://doi.org/10.1891/jcop.16.4.391.52526)]
52. Selby EA, Joiner TE. Cascades of emotion: the emergence of borderline personality disorder from emotional and behavioral dysregulation. *Rev Gen Psychol* 2009 Sep 01;13(3):219-229 [[FREE Full text](#)] [doi: [10.1037/a0015687](https://doi.org/10.1037/a0015687)] [Medline: [20161656](https://pubmed.ncbi.nlm.nih.gov/20161656/)]
53. Hayes SC, Wilson KG, Gifford EV, Follette VM, Strosahl K. Experimental avoidance and behavioral disorders: a functional dimensional approach to diagnosis and treatment. *J Consult Clin Psychol* 1996 Dec;64(6):1152-1168. [doi: [10.1037//0022-006x.64.6.1152](https://doi.org/10.1037//0022-006x.64.6.1152)] [Medline: [8991302](https://pubmed.ncbi.nlm.nih.gov/8991302/)]
54. Giorgio JM, Sanflippo J, Kleiman E, Reilly D, Bender RE, Wagner CA, et al. An experiential avoidance conceptualization of depressive rumination: three tests of the model. *Behav Res Ther* 2010 Oct;48(10):1021-1031 [[FREE Full text](#)] [doi: [10.1016/j.brat.2010.07.004](https://doi.org/10.1016/j.brat.2010.07.004)] [Medline: [20691426](https://pubmed.ncbi.nlm.nih.gov/20691426/)]
55. Cook L, Watkins E. Guided, internet-based, rumination-focused cognitive behavioural therapy (i-RFCBT) versus a no-intervention control to prevent depression in high-ruminating young adults, along with an adjunct assessment of the feasibility of unguided i-RFCBT, in the REducing Stress and Preventing Depression trial (RESPOND): study protocol for a phase III randomised controlled trial. *Trials* 2016 Jan 4;17(1). [doi: [10.1186/s13063-015-1128-9](https://doi.org/10.1186/s13063-015-1128-9)]
56. Hvenegaard M, Watkins ER, Poulsen S, Rosenberg NK, Gondan M, Grafton B, et al. Rumination-focused cognitive behaviour therapy vs. cognitive behaviour therapy for depression: study protocol for a randomised controlled superiority trial. *Trials* 2015 Aug 11;16(1):344 [[FREE Full text](#)] [doi: [10.1186/s13063-015-0875-y](https://doi.org/10.1186/s13063-015-0875-y)] [Medline: [26260780](https://pubmed.ncbi.nlm.nih.gov/26260780/)]
57. Perestelo-Perez L, Barraca J, Peñate W, Rivero-Santana A, Alvarez-Perez Y. Mindfulness-based interventions for the treatment of depressive rumination: systematic review and meta-analysis. *Int J Clin Health Psychol* 2017;17(3):282-295 [[FREE Full text](#)] [doi: [10.1016/j.ijchp.2017.07.004](https://doi.org/10.1016/j.ijchp.2017.07.004)] [Medline: [30487903](https://pubmed.ncbi.nlm.nih.gov/30487903/)]
58. Addis M, Martell C. *Overcoming Depression One Step at a Time: The New Behavioral Activation Approach to Getting Your Life Back*. Oakland, California, United States: New Harbinger Publications; 2004.
59. Hertel P, Mor N, Ferrari C, Hunt O, Agrawal N. Looking on the dark side: rumination and cognitive-bias modification. *Clin Psychol Sci* 2014 Apr 16;2(6):714-726. [doi: [10.1177/2167702614529111](https://doi.org/10.1177/2167702614529111)]
60. Watkins ER, Nolen-Hoeksema S. A habit-goal framework of depressive rumination. *J Abnorm Psychol* 2014 Feb;123(1):24-34. [doi: [10.1037/a0035540](https://doi.org/10.1037/a0035540)] [Medline: [24661156](https://pubmed.ncbi.nlm.nih.gov/24661156/)]
61. Thomas J, Raynor M, Ribott D. Depressive rumination and experiential avoidance: a task based exploration. *Personal Ment Health* 2015 Feb;9(1):58-65 [[FREE Full text](#)] [doi: [10.1002/pmh.1276](https://doi.org/10.1002/pmh.1276)] [Medline: [25273612](https://pubmed.ncbi.nlm.nih.gov/25273612/)]
62. Linehan M. *DBT Skills Training Manual, Second Edition*. New York, United States: Guilford Publications; 2014.
63. McKay M, Wood J, Brantley J. *The Dialectical Behavior Therapy Skills Workbook: Practical DBT Exercises for Learning Mindfulness, Interpersonal Effectiveness, Emotion Regulation, and Distress Tolerance*. Oakland, California, United States: New Harbinger Publications; 2019.
64. Chuen Yee Lo B, Lau S, Cheung S, Allen NB. The impact of rumination on internal attention switching. *Cogn Emot* 2012 Feb;26(2):209-223. [doi: [10.1080/02699931.2011.574997](https://doi.org/10.1080/02699931.2011.574997)] [Medline: [21614702](https://pubmed.ncbi.nlm.nih.gov/21614702/)]
65. Davis R, Nolen-Hoeksema S. Cognitive inflexibility among ruminators and nonruminators. *Cognit Ther Res* 2000;24(6):699-711.
66. Koster EH, De Lissnyder E, De Raedt R. Rumination is characterized by valence-specific impairments in switching of attention. *Acta Psychol (Amst)* 2013 Nov;144(3):563-570. [doi: [10.1016/j.actpsy.2013.09.008](https://doi.org/10.1016/j.actpsy.2013.09.008)] [Medline: [24140824](https://pubmed.ncbi.nlm.nih.gov/24140824/)]
67. Philippot P, Brutoux F. Induced rumination dampens executive processes in dysphoric young adults. *J Behav Ther Exp Psychiatry* 2008 Sep;39(3):219-227 [[FREE Full text](#)] [doi: [10.1016/j.jbtep.2007.07.001](https://doi.org/10.1016/j.jbtep.2007.07.001)] [Medline: [17698028](https://pubmed.ncbi.nlm.nih.gov/17698028/)]
68. Whitmer AJ, Gotlib IH. Switching and backward inhibition in major depressive disorder: the role of rumination. *J Abnorm Psychol* 2012 Aug;121(3):570-578. [doi: [10.1037/a0027474](https://doi.org/10.1037/a0027474)] [Medline: [22468767](https://pubmed.ncbi.nlm.nih.gov/22468767/)]
69. Wilkinson PO, Goodyer IM. Attention difficulties and mood-related ruminative response style in adolescents with unipolar depression. *J Child Psychol Psychiatry* 2006 Dec 07;47(12):1284-1291. [doi: [10.1111/j.1469-7610.2006.01660.x](https://doi.org/10.1111/j.1469-7610.2006.01660.x)] [Medline: [17176383](https://pubmed.ncbi.nlm.nih.gov/17176383/)]

70. Hanania R, Smith L. Selective attention and attention switching: towards a unified developmental approach. *Dev Sci* 2010 Jul;13(4):622-635 [FREE Full text] [doi: [10.1111/j.1467-7687.2009.00921.x](https://doi.org/10.1111/j.1467-7687.2009.00921.x)] [Medline: [20590726](https://pubmed.ncbi.nlm.nih.gov/20590726/)]
71. Ellis H, Ashbrook P. Resource allocation model of the effects of depressed mood states on memory. *BibSonomy*. 1988. URL: <https://www.bibsonomy.org/publication/22058cc28abec9a4c57dcc07a46e391c/brian.mingus> [accessed 2022-11-29]
72. Ellis H, Ashbrook P. The "state" of mood and memory research: a selective review. *J Social Behav Personal* 1989;4(2):1-21.
73. Ellis HC, Thomas RL, McFarland AD, Lane JW. Emotional mood states and retrieval in episodic memory. *J Experiment Psychol Learn Memory Cognit* 1985;11(2):363-370. [doi: [10.1037/0278-7393.11.2.363](https://doi.org/10.1037/0278-7393.11.2.363)]
74. Rosenfeld EA, Kennedy K, Farchione TJ, Roberts JE. Cleansing the attentional palate: a preliminary test of a novel approach to facilitate disengagement from rumination. *J Cogn Psychother* 2019 May 01;33(2):128-139. [doi: [10.1891/0889-8391.33.2.128](https://doi.org/10.1891/0889-8391.33.2.128)] [Medline: [32746388](https://pubmed.ncbi.nlm.nih.gov/32746388/)]
75. Deyo M, Wilson K, Ong J, Koopman C. Mindfulness and rumination: does mindfulness training lead to reductions in the ruminative thinking associated with depression? *Explore (NY)* 2009;5(5):265-271 [FREE Full text] [doi: [10.1016/j.explore.2009.06.005](https://doi.org/10.1016/j.explore.2009.06.005)] [Medline: [19733812](https://pubmed.ncbi.nlm.nih.gov/19733812/)]
76. Hawley LL, Schwartz D, Bieling PJ, Irving J, Corcoran K, Farb NA, et al. Mindfulness practice, rumination and clinical outcome in mindfulness-based treatment. *Cogn Ther Res* 2013 Oct 8;38(1):1-9. [doi: [10.1007/s10608-013-9586-4](https://doi.org/10.1007/s10608-013-9586-4)]
77. Sipe WE, Eisendrath SJ. Mindfulness-based cognitive therapy: theory and practice. *Can J Psychiatry* 2012 Feb;57(2):63-69. [doi: [10.1177/070674371205700202](https://doi.org/10.1177/070674371205700202)] [Medline: [22340145](https://pubmed.ncbi.nlm.nih.gov/22340145/)]
78. Fleming T, Dixon R, Frampton C, Merry S. A pragmatic randomized controlled trial of computerized CBT (SPARX) for symptoms of depression among adolescents excluded from mainstream education. *Behav Cogn Psychother* 2012 Oct;40(5):529-541. [doi: [10.1017/S1352465811000695](https://doi.org/10.1017/S1352465811000695)] [Medline: [22137185](https://pubmed.ncbi.nlm.nih.gov/22137185/)]
79. Batterham PJ, Neil AL, Bennett K, Griffiths KM, Christensen H. Predictors of adherence among community users of a cognitive behavior therapy website. *Patient Prefer Adherence* 2008 Feb 02;2:97-105 [FREE Full text] [Medline: [19920949](https://pubmed.ncbi.nlm.nih.gov/19920949/)]
80. Koeppe MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, Jones T, et al. Evidence for striatal dopamine release during a video game. *Nature* 1998 May 21;393(6682):266-268. [doi: [10.1038/30498](https://doi.org/10.1038/30498)] [Medline: [9607763](https://pubmed.ncbi.nlm.nih.gov/9607763/)]
81. Li J, Theng Y, Foo S. Game-based digital interventions for depression therapy: a systematic review and meta-analysis. *Cyberpsychol Behav Soc Netw* 2014 Aug;17(8):519-527 [FREE Full text] [doi: [10.1089/cyber.2013.0481](https://doi.org/10.1089/cyber.2013.0481)] [Medline: [24810933](https://pubmed.ncbi.nlm.nih.gov/24810933/)]
82. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
83. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009 Apr;114(1-3):163-173. [doi: [10.1016/j.jad.2008.06.026](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
84. Kroenke K, Spitzer R. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals* 2002 Sep;32(9):509-515. [doi: [10.3928/0048-5713-20020901-06](https://doi.org/10.3928/0048-5713-20020901-06)]
85. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth* 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
86. LeBeau K, Huey LG, Hart M. Assessing the quality of mobile apps used by occupational therapists: evaluation using the user version of the mobile application rating scale. *JMIR Mhealth Uhealth* 2019 May 01;7(5):e13019 [FREE Full text] [doi: [10.2196/13019](https://doi.org/10.2196/13019)] [Medline: [31066712](https://pubmed.ncbi.nlm.nih.gov/31066712/)]
87. Schreier M. *The SAGE Handbook of Qualitative Data Analysis*. Thousand Oaks, California, United States: SAGE Publications; 2014.
88. Oulasvirta A, Tamminen S, Roto V, Kuorelahti J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2005 Presented at: CHI05: CHI 2005 Conference on Human Factors in Computing Systems; Apr 2 - 7, 2005; Portland Oregon USA. [doi: [10.1145/1054972.1055101](https://doi.org/10.1145/1054972.1055101)]
89. Smith AC, Fowler LA, Graham AK, Jaworski BK, Firebaugh M, Monterubio GE, et al. Digital overload among college students: implications for mental health app use. *Social Sci* 2021 Jul 21;10(8):279. [doi: [10.3390/socsci10080279](https://doi.org/10.3390/socsci10080279)]
90. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res* 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: [10.2196/14567](https://doi.org/10.2196/14567)] [Medline: [31573916](https://pubmed.ncbi.nlm.nih.gov/31573916/)]
91. Walker N. Time Gating in Video Games. YouTube. 2018. URL: <https://www.youtube.com/watch?v=b1phy6TpgXk> [accessed 2022-11-15]
92. Meyer T, Miller M, Metzger R, Borkovec T. Development and validation of the penn state worry questionnaire. *Behav Res Ther* 1990;28(6):487-495 [FREE Full text] [doi: [10.1016/0005-7967\(90\)90135-6](https://doi.org/10.1016/0005-7967(90)90135-6)]
93. Grand View Research. *MarketResearch.com*. URL: <https://www.marketresearch.com/Grand-View-Research-v4060/> [accessed 2022-09-28]
94. Mobile app retention rate: what's a good retention rate? Upland. URL: <https://uplandsoftware.com/localytics/resources/blog/mobile-apps-whats-a-good-retention-rate/> [accessed 2022-09-28]

95. Wooffitt R. Communication and laboratory performance in parapsychology experiments: demand characteristics and the social organization of interaction. *Br J Soc Psychol* 2007 Sep;46(Pt 3):477-498. [doi: [10.1348/014466606X152667](https://doi.org/10.1348/014466606X152667)] [Medline: [17877849](https://pubmed.ncbi.nlm.nih.gov/17877849/)]
96. Brew K, Clark T, Feingold-Link J, Barth H. Do demand characteristics contribute to minimal ingroup preferences? *J Exp Child Psychol* 2021 Apr;204:105043. [doi: [10.1016/j.jecp.2020.105043](https://doi.org/10.1016/j.jecp.2020.105043)] [Medline: [33360283](https://pubmed.ncbi.nlm.nih.gov/33360283/)]
97. Guo T, Xiang Y, Xiao L, Hu C, Chiu HF, Ungvari GS, et al. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. *Am J Psychiatry* 2015 Oct;172(10):1004-1013. [doi: [10.1176/appi.ajp.2015.14050652](https://doi.org/10.1176/appi.ajp.2015.14050652)] [Medline: [26315978](https://pubmed.ncbi.nlm.nih.gov/26315978/)]
98. Alqahtani F, Winn A, Orji R. Co-designing a mobile app to improve mental health and well-being: focus group study. *JMIR Form Res* 2021 Feb 26;5(2):e18172 [FREE Full text] [doi: [10.2196/18172](https://doi.org/10.2196/18172)] [Medline: [33635281](https://pubmed.ncbi.nlm.nih.gov/33635281/)]
99. Kaveladze BT, Wasil AR, Bunyi JB, Ramirez V, Schueller SM. User experience, engagement, and popularity in mental health apps: secondary analysis of app analytics and expert app reviews. *JMIR Hum Factors* 2022 Jan 31;9(1):e30766 [FREE Full text] [doi: [10.2196/30766](https://doi.org/10.2196/30766)] [Medline: [35099398](https://pubmed.ncbi.nlm.nih.gov/35099398/)]
100. Lyman C, Rosenfeld E, Roberts JE. RuminAid: A mHealth App for Depressive Rumination. *Open Science Framework*. 2022 Sep 18. URL: <https://osf.io/wjct4/> [accessed 2022-12-02]

Abbreviations

- BA:** behavioral activation
- EBT:** evidence-based treatment
- mHealth:** mobile health
- PHQ:** Patient Health Questionnaire
- RFCBT:** rumination-focused cognitive behavioral therapy
- uMARS:** Mobile Application Rating Scale–user version

Edited by A Mavragani; submitted 03.06.22; peer-reviewed by K Stawarz, P Anderson; comments to author 09.09.22; revised version received 05.10.22; accepted 16.10.22; published 13.12.22.

Please cite as:

Rosenfeld EA, Lyman C, Roberts JE

Development of an mHealth App–Based Intervention for Depressive Rumination (RuminAid): Mixed Methods Focus Group Evaluation
JMIR Form Res 2022;6(12):e40045

URL: <https://formative.jmir.org/2022/12/e40045>

doi: [10.2196/40045](https://doi.org/10.2196/40045)

PMID: [36512400](https://pubmed.ncbi.nlm.nih.gov/36512400/)

©Eve A Rosenfeld, Cassondra Lyman, John E Roberts. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 13.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Detecting Elevated Air Pollution Levels by Monitoring Web Search Queries: Algorithm Development and Validation

Chen Lin¹, BSc, MSc; Safoora Yousefi¹, DPhil; Elvis Kahoro², BSc; Payam Karisani¹, DPhil; Donghai Liang³, DPhil; Jeremy Sarnat³, DPhil; Eugene Agichtein¹, DPhil

¹Department of Computer Science, Emory University, Atlanta, GA, United States

²Department of Computer Science, Pomona College, Claremont, CA, United States

³Department of Environmental Health, Emory University, Atlanta, GA, United States

Corresponding Author:

Chen Lin, BSc, MSc

Department of Computer Science

Emory University

201 Dowman Drive

W302

Atlanta, GA, 30322

United States

Phone: 1 404 395 0266

Email: chen.lin@emory.edu

Abstract

Background: Real-time air pollution monitoring is a valuable tool for public health and environmental surveillance. In recent years, there has been a dramatic increase in air pollution forecasting and monitoring research using artificial neural networks. Most prior work relied on modeling pollutant concentrations collected from ground-based monitors and meteorological data for long-term forecasting of outdoor ozone (O₃), oxides of nitrogen, and fine particulate matter (PM_{2.5}). Given that traditional, highly sophisticated air quality monitors are expensive and not universally available, these models cannot adequately serve those not living near pollutant monitoring sites. Furthermore, because prior models were built based on physical measurement data collected from sensors, they may not be suitable for predicting the public health effects of pollution exposure.

Objective: This study aimed to develop and validate models to *nowcast* the observed pollution levels using web search data, which are publicly available in near real time from major search engines.

Methods: We developed novel machine learning-based models using both traditional supervised classification methods and state-of-the-art deep learning methods to detect elevated air pollution levels at the US city level by using generally available meteorological data and aggregate web-based search volume data derived from Google Trends. We validated the performance of these methods by predicting 3 critical air pollutants (O₃, nitrogen dioxide, and PM_{2.5}) across 10 major US metropolitan statistical areas in 2017 and 2018. We also explore different variations of the long short-term memory model and propose a novel search term dictionary learner-long short-term memory model to learn sequential patterns across multiple search terms for prediction.

Results: The top-performing model was a deep neural sequence model long short-term memory, using meteorological and web search data, and reached an accuracy of 0.82 (F_1 -score 0.51) for O₃, 0.74 (F_1 -score 0.41) for nitrogen dioxide, and 0.85 (F_1 -score 0.27) for PM_{2.5}, when used for detecting elevated pollution levels. Compared with using only meteorological data, the proposed method achieved superior accuracy by incorporating web search data.

Conclusions: The results show that incorporating web search data with meteorological data improves the nowcasting performance for all 3 pollutants and suggest promising novel applications for tracking global physical phenomena using web search data.

(JMIR Form Res 2022;6(12):e23422) doi:[10.2196/23422](https://doi.org/10.2196/23422)

KEYWORDS

nowcasting of air pollution; web-based public health surveillance; neural network sequence modeling; search engine log analysis; air pollution exposure assessment; mobile phone

Introduction

Background

Web-based crowd surveillance has been used to track emergent risks to public health [1-3]. Most commonly, these efforts involve the collection of web-based search queries to document acute changes in the incidence or symptom occurrence of primary infectious disease agents, such as influenza [4-7], Ebola [8], dengue fever [9], and COVID-19 [10]. These methods have the potential to provide public health and medical professionals with benefits over traditional health surveillance and environmental epidemiology in their ability to capture both personal exposures and response dynamics at more sensitive spatial and temporal scales [2].

Despite the promise of these approaches for infectious diseases, only a limited number of studies have examined how crowd surveillance approaches can be used to track environmental exposures and, less frequently, responses to noninfectious environment-mediated disease processes [11-13]. The global burden of disease attributable to outdoor and indoor air pollution has been quantified by recent efforts and has increased public awareness of the severity of this public health crisis worldwide [14]. Therefore, urban air pollution provides a key test case for the evaluation of web-based surveillance approaches for noninfectious environmental risks. The web-based surveillance approach is distinct from traditional approaches for measuring urban air pollution exposure. Therefore, it could possibly serve as a substitute to or complement the existing approaches. Traditional indicators of air pollution exposure, namely, concentrations measured at ambient monitoring sites, are widely used to assess the health effects associated with air pollution in epidemiological studies. However, the use of ambient monitoring measurements as surrogates of exposure may result in the misclassification of health responses and potential risks, especially for those not living near pollutant monitoring sites [15-17]. Moreover, ambient monitoring, by design, provides information on measured outdoor pollutant concentrations and may not necessarily reflect accurate personal exposures for individuals spending most of their time indoors or for those with preexisting biological susceptibility to air pollution. Several recent studies have focused on using smartphones within distributed air pollution sensing networks, where users record and upload local air pollution conditions to crowd-generated, geospatially refined pollution maps [11-13]. These studies demonstrate the feasibility of web-based crowd-generated participation in projects predicted on urban air pollution awareness.

To the best of our knowledge, few studies have investigated the feasibility of using web search data to produce accurate “nowcasts” of urban air pollution levels in real time. Conducting accurate predictions using web search data is a challenging task with 2 major challenges. The first is the selection of search terms to comprehensively capture people’s responses. Several approaches have been proposed to select search terms. For example, some studies preliminarily prepare keywords related to the target disease and then use these keywords to filter the search terms, which is often difficult because finding related

keywords could be difficult for some diseases or be costly when conducting for multiple diseases. The second is the selection of the appropriate models. Although the literature on data-driven nowcasting methods for estimating infectious disease activity is well developed from an epidemiological standpoint, the machine learning methods used lag behind the state-of-the-art methods. The nowcasting models introduced to date mainly use variations of regularized linear regressions or, less often, random forests (RFs) or support vector machines. From a machine learning perspective, the problem of disease activity estimation is most suited to a more sophisticated and time series-specific model architecture. Because of the growing volume of recorded environment-mediated disease data, the use of recurrent neural networks (RNNs) and, more specifically, their variants long short-term memory (LSTM) and gated recurrent unit networks is increasingly feasible. The vanilla LSTM model makes predictions solely relying on the time series of the search activity while ignoring the semantic information in the search query phrases. Previous studies have pointed out that search queries could be semantically related, and ignoring their correlation would lead to a decrease in model performance [18,19]. Recent advances in natural language processing have led to the development of a technique called word embeddings to represent the semantic information in phrases, and fine-tuning of word embeddings has been encouraged for downstream tasks (Wu, Y, unpublished data, September 2016) [20-22]. However, there is still a lack of knowledge on incorporating both the semantic information of search queries and time series of search activities to make predictions.

Objectives

In this study, we investigate web search data as an important source of a web-based crowd-based indicator. As web search data are free and broadly accessible, we posit that they could serve as a scalable means of tracking urban air pollution exposures and corresponding population-level health responses. To measure search interest, we used the freely accessible Google Trends service, which reports aggregate search volume data at a city-level geographical resolution. For this analysis, we use known health end point terms and topics, such as “difficulty breathing,” and observations (eg, “haze”) suggested by public health researchers, augmented by automatic term expansion based on semantic and temporal correlations, to estimate the levels of search activities related to air pollution, and ultimately to predict whether the pollution levels were elevated [23,24].

Compared with existing air pollution classification models, this study explores the use of web search anomalies as an auxiliary signal to detect air pollution. We compared our approach with the state-of-the-art physical sensor-based models that incorporate various pollutant covariates such as historical pollutant concentrations and meteorological data [25]. Using web search data for prediction introduces several challenges, including an unclear relationship between search interest and pollution levels and the trade-off between model complexity and convergence for the inclusion of web search data in a data-deficient scenario.

In summary, our contributions are as follows:

- We proposed a novel search term dictionary learner-LSTM (DL-LSTM) model to learn sequential patterns from broad historical records of web search data for air pollution nowcasting.
- We compared the DL-LSTM models with a variety of baseline models on the efficacy of using web search data to indicate exposure to a noninfectious environmental stressor (ie, air pollution) and demonstrate that the proposed models are effective across different experimental settings.
- We evaluated the efficacy of combining web search data and meteorological data for air pollution prediction and showed that the inclusion of web search data improves the prediction accuracy and provides a promising substitute when historical pollutant data are unavailable.

Methods

We now describe the methodology. First, we formalize our problem setting, then describe the data, and then introduce our modeling approaches.

Problem Statement

We formalized this task as a classification problem and adapted state-of-the-art machine learning models. We constructed a multivariate autoregressive model and an RF model fit on historical air pollutant concentrations as well as search and meteorological data as baseline models. We evaluated the performance of our proposed models (described below) in comparison with the baselines in terms of prediction accuracy and other standard classification prediction metrics.

Ethical Considerations

The data available to the public are not individually identifiable and therefore analysis does not involve human subjects. The International Review Board (IRB) recognizes that the analysis of de-identified, publicly available data does not constitute human subjects research and therefore does not require IRB review.

Data Collection

We collected daily air pollutant concentration data as well as temperature and relative humidity in the 10 largest US metropolitan statistical areas (MSAs) from January 2007 to December 2018. We focused on 3 air pollutants: ozone (O₃), nitrogen dioxide (NO₂), and fine particulate matter (PM_{2.5}). The in-situ pollutant concentrations and meteorological data such as temperature, relative humidity, and dew point temperature were retrieved from the US Environmental Protection Agency, Air Quality System, and AirNow database. To create a single daily pollutant concentration for each city, we used the median pollutant concentration from all available monitoring sites within each city to avoid outlier bias.

We collected the daily search frequency of pollution-related terms from Google Trends for the same 12-year period and cities. We created a curated list of 152 pollution-related terms based on our previous air pollution epidemiology studies and in reviewing the environmental health literature [14,26-30], and we downloaded the reports of trending results terms using PyTrends [31]. For each PyTrends request, we downloaded the search history of pollution-related terms over a 6-month window with 1 overlapping month for calibration. PyTrends provided us with a search frequency scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics. Because of the PyTrends restriction, we downloaded the reports of trending results multiple times, and the search frequencies were scaled separately in each 6-month window, which required us to calibrate the search frequency for the 12-year period. We calibrated the search frequencies by joining the search logs on the overlapping periods (1 out of 6 months) for intercalibration [32].

We investigated the available input features from meteorological data (temperature and relative humidity), historical pollutant concentrations, and web search data (Table 1).

Table 1. Input features calculated per time step in the input sequence.

Input feature	Feature transformation
Meteorological data (Met ^a)	<ul style="list-style-type: none"> • Maximum temperature (Temp_max^b) • Mean temperature (Temp_mean^c) • Relative humidity (humidity) • Square of Temp_mean • Cube of Temp_mean • Square of humidity • Cube of humidity • Dew point temperature
Pollutant concentration (Pol ^d)	<ul style="list-style-type: none"> • Concentration on day t-7^e • Concentration on day t-6^e • Concentration on day t-5^e • Concentration on day t-4^e • Concentration on day t-3^e • Concentration on day t-2^e • Concentration on day t-1^e
Search	<ul style="list-style-type: none"> • Search volumes of search terms

^aMet, meteorological data.

^bTemp_max: maximum temperature

^cTemp_mean: mean temperature

^dPol: pollutant concentration.

^eDay t-7,..., t-1: days preceding the prediction day t.

Missing Data Imputation and Normalization

Smoothing and interpolation are simple and efficient data imputation methods [33], and we applied linear interpolation to fill the missing data in historical pollutant concentration, temperature, and humidity, with a rolling window size of 3. To fill in the missing data in infrequent search terms for which Google Trends does not return a count, we used random numbers close to 0 ($e^{-10} \sim e^{-5}$). We normalized all the input features to standard scores by subtracting their mean values and dividing them by the respective SDs.

Search Term Expansion

As web-based search queries may reflect individual exposure to ambient air pollution, the seed terms were mostly related to symptoms, observations, and emission sources (Table S1 in Multimedia Appendix 1). However, because an exhaustive list of user queries was not available, reliance on only expert-generated seed words may result in poor prediction because of the high mismatch rate between the user queries and our expected search words.

Query expansion is a common approach for resolving this discrepancy. A recent study [18] showed that the initial set of seed words could be effectively expanded through semantic and temporal correlations. Thus, for each seed word, we used Google Correlate [34] to retrieve the top 100 correlated query terms. Then, we used the pretrained word2vec model [21] to retrieve the vector representation of each query; phrases were mapped to the centroid of the constituent terms. A utility score was calculated for each candidate query by measuring the maximum cosine similarity between the query and seed words. Queries

with a high utility score were retained, and the remaining queries were eliminated, and we empirically set the utility cutoff to 0.55. This method expanded the set of search terms for the 152 search terms to track (Table S2 in Multimedia Appendix 1).

Modeling and Evaluation

Problem Definition

Given sequences of physical sensor data $P = [p_{t-L}, \dots, p_{t-1}]^T$ with the dimension of L times d_p , and search interest data $S = [s_{t-L+2}, \dots, s_{t+1}]^T$ with the dimension of L times d_s , the task is to classify day t as *polluted* or not, where a positive class label indicates that the air pollution was above a predefined threshold. L denotes the sequence length, and d_p and d_s are the number of physical sensor features and the number of search-related terms, respectively.

Autoregressive and RF Classification Models

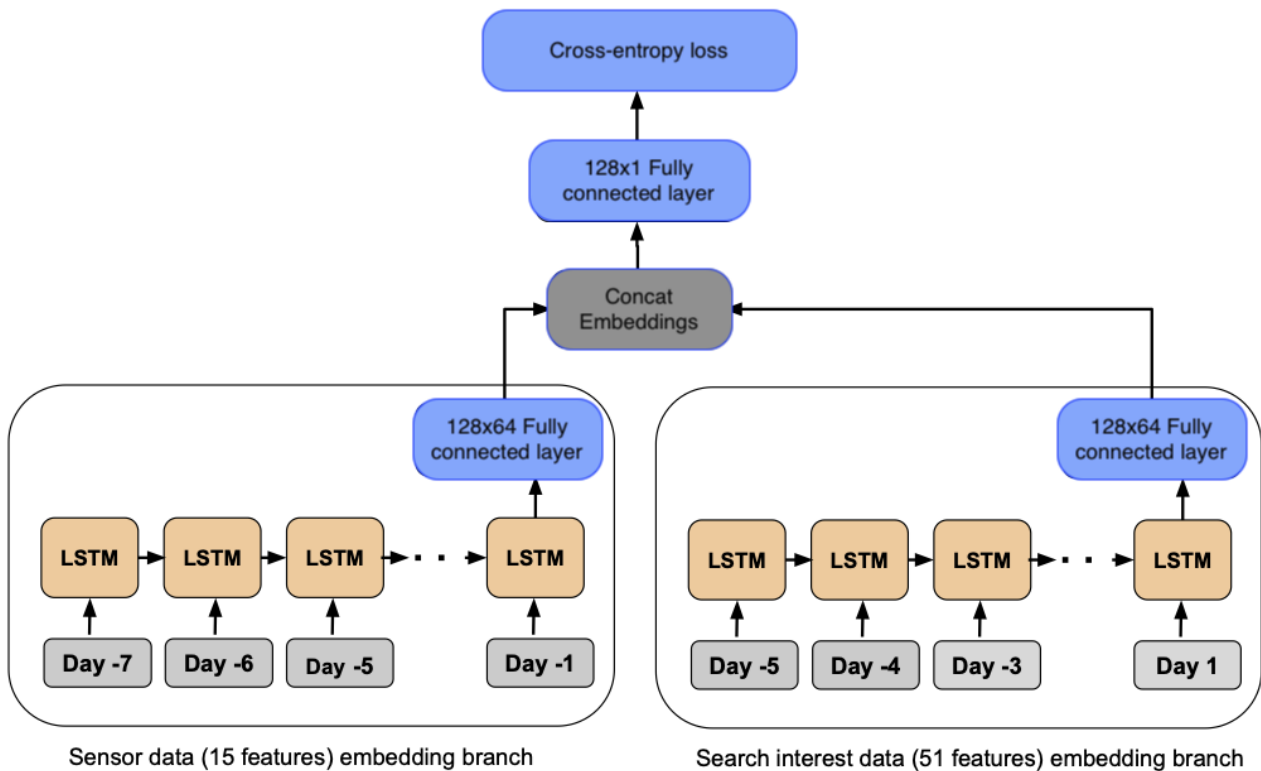
Previous work has shown that simple autoregressive models using web search data can generate nowcast estimates for influenza-like illnesses at the US national level [19]. We adapted autoregressive models with a logistic regression (LR) classifier for classification purposes. Furthermore, we applied elastic net regularization, which is a linear combination of l_1 and l_2 regularization, as proposed in previous studies [18,19]. LR+Elastic Net was implemented using the Python *scikit-learn* package, using cross-validation to set the model's hyperparameters to maximize the F_1 -score on the validation set, with `class_weight` set to "balanced."

RF is an ensemble learning model that is robust against overfitting and provides a strong baseline for the development of nonlinear predictive models [35]. We used the *scikit-learn* implementation of RFs. The number of trees and maximum depth of individual trees were selected to maximize the F_1 -score on the validation set, with balanced class_weight for positive and negative samples.

LSTM and Its Variants

LSTM units [36] are RNN models designed for sequence modeling, which can learn nonlinear relationships in time series data [37]. First, we describe a baseline LSTM model with 2 subnetworks to separate the search data and meteorological data. As shown in Figure 1, there are 4 layers in the model, that is, the sequence embedding layer, LSTM layer, fully connected hidden layer, and output layer [38].

Figure 1. The architecture of the long short-term memory (LSTM) model.



In the left subnetwork of the LSTM model with search data as input, we propose 2 methods for capturing semantic information in search terms. The first is the LSTM semantic model (GloVe [Global Vectors for Word Representation]; LSTM-GloVe). As a variant of the vanilla LSTM model, for the sequence embedding layer of the right subnetwork in Figure 1, we introduce the matrix multiplication operation to project the search values of search terms to their semantic embedding space (GloVe embeddings), as shown in equation 1.

Given the search interest data $S = [s_1, \dots, s_7]^T$ with the dimension of 7 times d_s , and their GloVe embedding $G = [g_1, \dots, g_{d_g}]$ with the dimension of d_s times d_g , where $d_g = 50$ (GloVe 50-dimensional word vectors trained on tweets [22]). The matrix multiplication operation is defined as

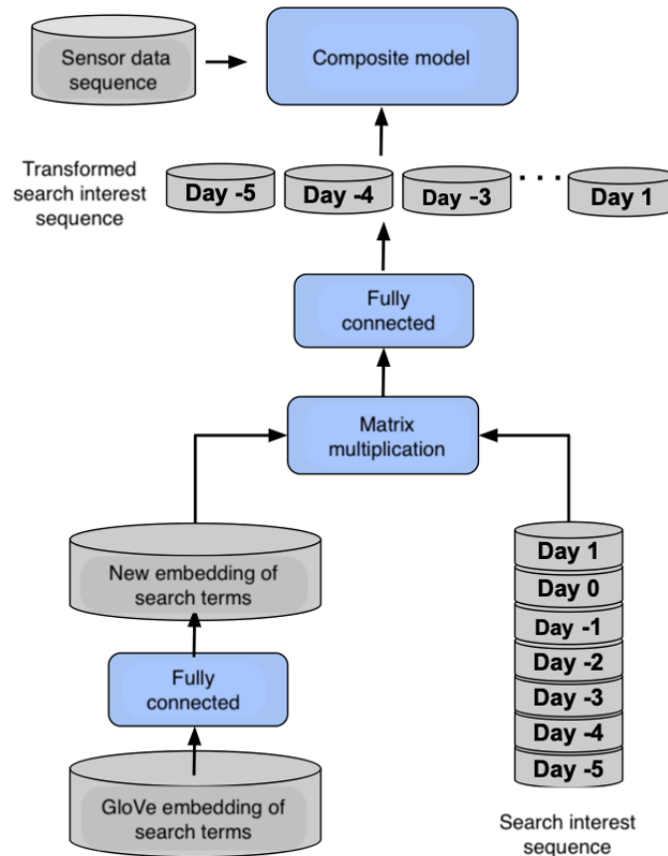
$$S \times G$$

Specifically, the tensor generated by the matrix multiplication operation was then fed into the LSTM layer for further

calculations. This matrix multiplication is designed specifically for the model consistency problem when introducing collinear predictors after search term expansion (STE).

The second variation of the LSTM model is the DL-LSTM model, which is theoretically based on the idea of matrix multiplication, as shown in LSTM-GloVe. However, instead of directly applying the GloVe embedding for matrix multiplication, it introduces the fine-tuning of the word embeddings via a d_g by d_e rectified linear unit-activated fully connected layer. As shown in Figure 2, the rectified linear unit-activated fully connected layer was applied to the initial GloVe embedding, where $d_e=100$ is the size of the new embedding. In this architecture, the GloVe 50-dimensional word vectors are used to initialize the search term embedding dictionary, and the matrix multiplication operation is used to transform the input embedding of search terms into the semantic embedding space [39].

Figure 2. The architecture of the dictionary learner-long short-term memory model.



In summary, we evaluate the following models in this paper:

- LR: LR is LR classifier with elastic net regularization.
- RF: RF is RF classifier with the number of trees and maximum depth tuned for prediction.
- LSTM: The baseline LSTM model, as shown in Figure 1, combines physical sensor features, if available, with the search interest volume data directly, providing a direct adaptation of RNNs to this problem without any problem-specific extensions.
- LSTM-GloVe: LSTM semantic model is a variant of the LSTM model as described in equation 1, where we control the input of search interest data (ie, 51 seed search terms vs 152 terms after STE) in this model. We refer to the variants as *LSTM-GloVe* and *LSTM-GloVe with [w/] STE*, respectively.
- DL-LSTM: The DL-LSTM model is shown in Figure 2. We control the input of the search interest data (ie, 51 seed search terms vs 152 terms after STE) in this model and refer

to the variants as *DL-LSTM* and *DL-LSTM w/STE*, respectively.

Validation

To tune the model parameters and validate the model performance, we split the available data into training (from January 2007 to December 2014), validation (from January 2015 to December 2016), and testing (from January 2017 to December 2018) sets. This 8-year training period provides a broad history for learning the relationship between input and output variables, and the predictive models are evaluated based on their ability to make predictions for completely unseen periods. For evaluating our model, we made predictions for each day from January 2017 to December 2018 in the test data set. The distribution of the classes in the training, validation, and test data sets is presented in Table 2. Note that the positive and negative classes are heavily imbalanced, with positive classes comprising, for instance, only 16% of the training samples when $PM_{2.5}$ is the target pollutant.

Table 2. The distribution of classes in the training, validation, and test sets.

Pollutant	Negative samples			Positive samples		
	Training	Validation	Test	Training	Validation	Test
O ₃ ^a	24,322	6269	6311	4896	1038	982
NO ₂ ^b	23,926	6119	6332	5292	1188	961
PM _{2.5} ^c	24,297	6745	6757	4921	562	536

^aO₃: ozone.

^bNO₂: nitrogen dioxide.

^cPM_{2.5}: fine particulate matter.

Evaluation Metrics

As we defined this task as a classification problem, we used the standard classification evaluation metrics. We report the accuracy and F_1 -score of the positive class (the harmonic mean of precision and recall) of the predictions as evaluation metrics for all models. Although accuracy measures the total fraction of correct predictions and could misrepresent model performance in the presence of heavily imbalanced classes, the F_1 -score considers class imbalance and is, therefore, a more appropriate metric for our problem.



Where TP , TN , FP , and FN are the number of true positive samples, true negative samples, false positive samples, and false negative samples, respectively.

Results

Overview

In this section, we first present the findings of the data exploration. Next, we present the principal findings of this study.

Insights From Collected Data

In this section, we describe the thresholds of abnormal air pollutant concentrations and present the lag between the search anomalies and air pollution.

Thresholds of Abnormal Air Pollutant Concentrations

The major MSAs chosen for this study have different distributions of pollutant concentrations over time and almost always fall below the Environmental Protection Agency standard 24-hour threshold (Figure 3). However, multiple studies have shown that even at low concentrations, chronic exposure to air pollution negatively affects human health [26,27]. Therefore, calibrating a meaningful threshold for each city, especially those with generally lower levels of air pollution (eg, Miami), may be critical for adequately protecting population health. A natural way to do this may be to set the threshold to 1 SD above the mean daily pollutant concentration within each city, which was adopted in this study. The input predictors were also normalized within each city to reflect the city-level dynamics. The resulting thresholds for the 3 pollutants and cities under investigation are reported in Table 3.

Figure 3. Distribution of pollution values for Atlanta, Los Angeles, Philadelphia, and Miami, with city-specific elevated pollution level (dashed line) and the general Environmental Protection Agency–mandated standard (dotted line), for ozone (O₃; left column), nitrogen dioxide (NO₂; middle column), and fine particulate matter (PM_{2.5}; right column). EPA: Environmental Protection Agency.

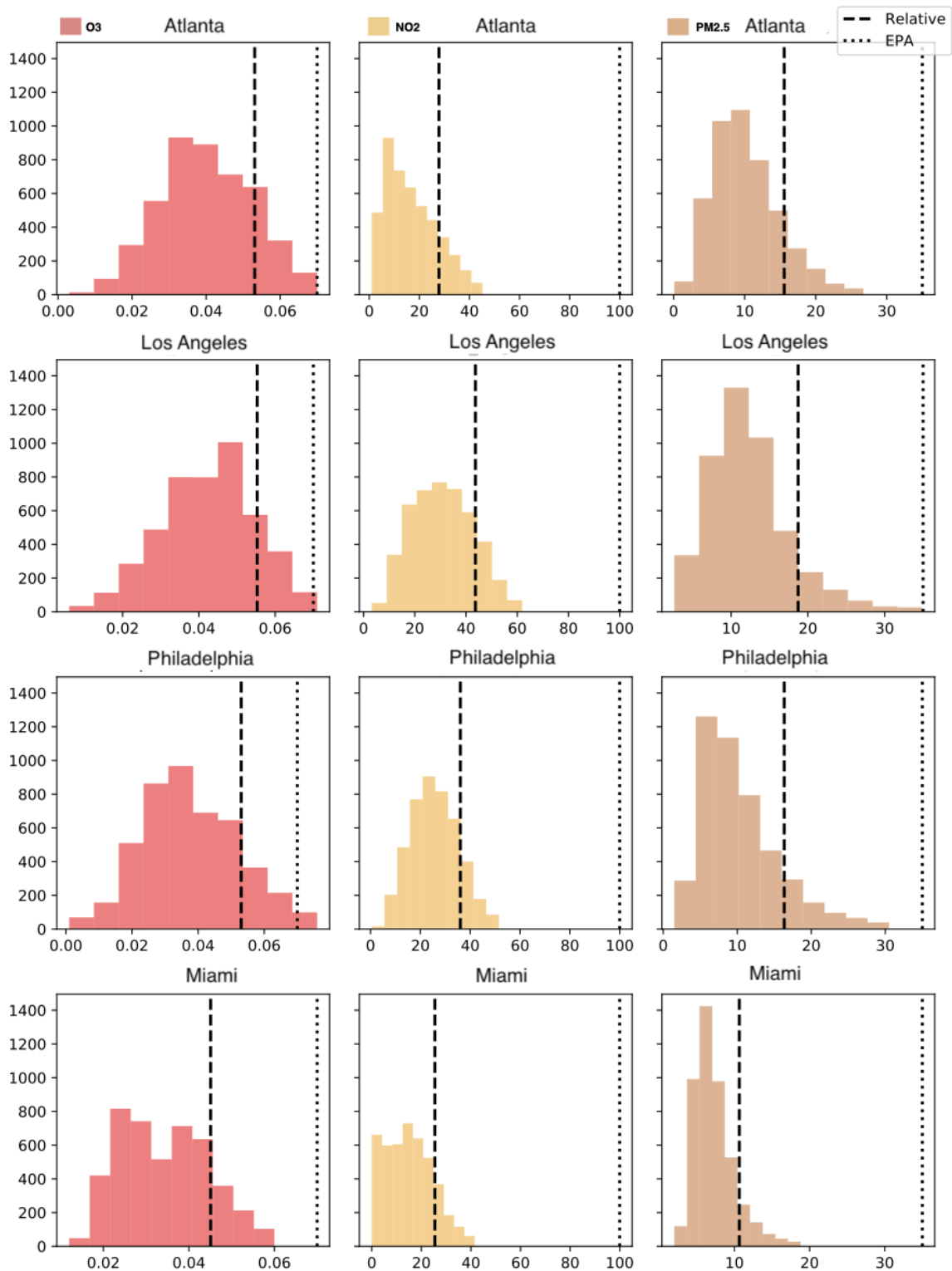


Table 3. Classification thresholds for 3 pollutants across 10 major metropolitan statistical areas in the United States.

Pollutant	Los Angeles	District of Columbia	Philadelphia	Dallas	Atlanta	Boston	New York	Miami	Chicago	Houston
O ₃ ^a (ppb ^b)	55	54	53	53	53	48	49	45	49	49
NO ₂ ^c (ppb)	43.7	38.1	36	25.2	27.8	30.7	45.3	25.5	43.7	27.7
PM _{2.5} ^d (µg/m ³)	18.7	15.1	16.4	13.1	15.6	12.4	13.9	10.6	16.2	14.4

^aO₃: ozone.

^bppb: parts per billion.

^cNO₂: nitrogen dioxide.

^dPM_{2.5}: fine particulate matter.

Lag Between Search Anomalies and Air Pollution

A previous study showed that there could be a lag between incident occurrence and Google search activity [40]. As shown in Figure 4, the normalized search frequency of the term “cough” is correlated with the concentration of NO₂ in Atlanta with a certain lag of time. To determine the lag between elevated pollution levels and consequent pollution-related searches, the mean absolute Spearman correlation between pollutant

concentrations and search interest data was calculated and shifted forward in time for 0, 1, 2, and 3 days. As shown in Table 4, for O₃ and PM_{2.5}, the mean absolute Spearman correlation increased with an increase in the shifted days. Considering that the task aimed to detect elevated pollution levels as soon as possible, a lag of 1 day was applied to search data. In other words, the search interest data from the current day were used to estimate whether air pollution was elevated on the previous day.

Figure 4. Daily nitrogen dioxide (NO₂) levels and search interest for the term “cough” in October 2016 in Atlanta.

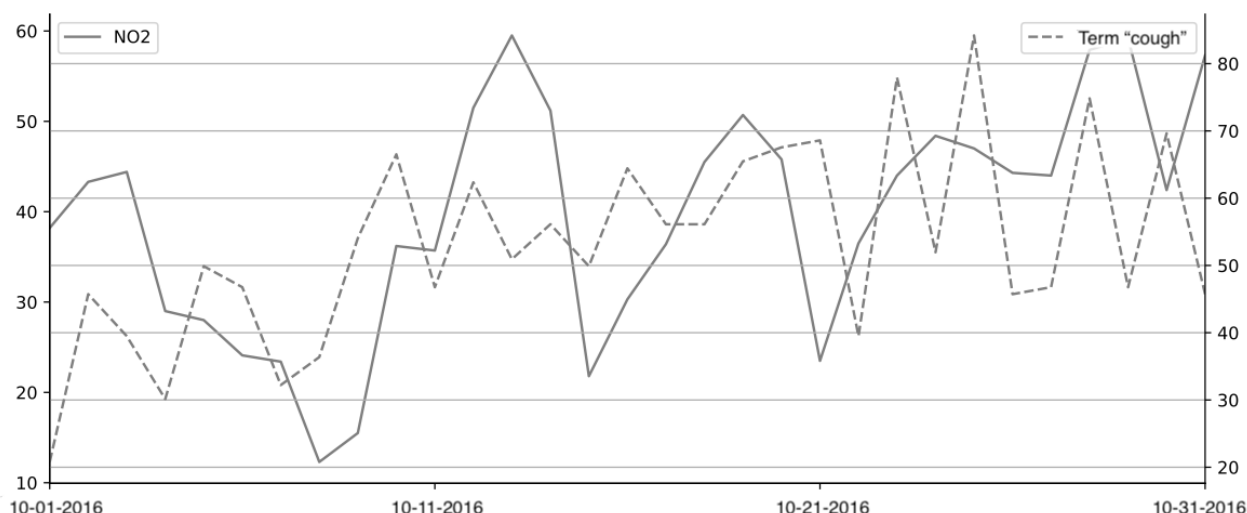


Table 4. Cross-correlation of top 5 search terms with different lags for 3 pollutants in the Atlanta metropolitan area in 2016 (N=366).

Pollutant	Lag=0; search term (Spearman correlation)	P value	Lag=1; search term (Spearman correlation)	P value	Lag=2; search term (Spearman correlation)	P value	Lag=3; search term (Spearman correlation)	P value
O₃^a								
	Cough (-0.34)	<.001	Cough (-0.38)	<.001	Cough (-0.41)	<.001	Cough (-0.41)	<.001
	Bronchitis (-0.31)	<.001	Bronchitis (-0.32)	<.001	Bronchitis (-0.33)	<.001	Bronchitis (-0.35)	<.001
	Traffic (0.26)	<.001	Traffic (0.27)	<.001	Traffic (0.26)	<.001	Smoke (0.24)	<.001
	Smoke (0.23)	<.001	Chest pain (-0.23)	<.001	Chest pain (-0.23)	<.001	Traffic (0.23)	<.001
	Snoring (0.22)	<.001	Snoring (0.22)	<.001	Smoke (0.22)	<.001	Chest pain (-0.22)	<.001
NO₂^b								
	Asthma (0.20)	<.001	Sulfate (0.20)	<.001	Sulfate (0.16)	.002	Cough (0.16)	.002
	Sulfate (0.19)	<.001	Bronchitis (0.16)	.002	Bronchitis (0.15)	.005	COPD ^c (-0.16)	.003
	Cough (0.17)	<.001	Inhaler (0.15)	.005	Cough (0.14)	.008	Bronchitis (0.14)	.008
	Bronchitis (0.17)	.001	Cough (0.14)	.006	Inhaler (0.11)	.03	Wheezing (-0.12)	.02
	Inhaler (0.16)	.002	Difficulty breathing (-0.12)	.02	Headache (-0.11)	.03	Headache (-0.10)	.04
PM_{2.5}^d								
	Wildfires (0.14)	.009	COPD (-0.15)	.005	Air pollution (0.19)	<.001	Air pollution (0.18)	<.001
	COPD (-0.11)	.03	Wildfires (0.14)	.007	COPD (-0.17)	.001	COPD (-0.18)	<.001
	Snoring (0.11)	.03	Air pollution (0.14)	.008	Wildfires (0.14)	.009	Wildfires (0.15)	.004
	Inhaler (0.10)	.06	Asthma attack (0.11)	.04	Respiratory illness (0.10)	.05	Sulfate (-0.11)	.03
	Difficulty breathing (-0.09)	.08	Respiratory illness (0.10)	.05	Traffic (0.10)	.06	Traffic (0.11)	.04

^aO₃: ozone.^bNO₂: nitrogen dioxide.^cCOPD: chronic obstructive pulmonary disease.^dPM_{2.5}: fine particulate matter.

Evaluation Outcomes

In this section, we consider 3 conditions to evaluate the performance of using web search data to detect elevated pollution, that is, using only search data, using search data as auxiliary data for meteorological data, and using search data as auxiliary data for meteorological data and historical pollutant concentrations.

Using Only Search Data

For areas where ambient pollution monitoring is unavailable, investigating whether web search data can be used as the only

signal for nowcasting elevated air pollution is a vital question. When relying only on search data for air pollution prediction, both the proposed DL-LSTM architecture and STE contribute to the improvement of prediction accuracy. As shown in the “Search” section of Table 5, the LSTM-based models exhibited superior accuracy over the baseline LR and RF models for O₃ and NO₂. For PM_{2.5}, the proposed models did not perform better than the baseline LR or LSTM model because the validation and test data sets were heavily imbalanced (Table 5). The proposed DL-LSTM w/STE model achieved the highest F₁-score (32.44% for O₃ and 27.70% for NO₂) for detecting O₃ and NO₂ pollution.

Table 5. Accuracy and F_1 -score of the logistic regression, random forest, and long short-term memory models for detecting elevated pollution across 10 major US cities, for varying input feature combinations: no prior knowledge, search data only (Search), meteorological data only (Met), meteorological data and search data (Met+Search), meteorological data and historical pollutant concentration (Met+Pol) and all input features (Met+Pol+Search).

Features and model	O ₃ ^a , accuracy (F_1 -score; %)	NO ₂ ^b , accuracy (F_1 -score; %)	PM _{2.5} ^c , accuracy (F_1 -score; %)
No prior knowledge			
All positives	13.46 (23.73)	13.18 (23.28)	7.35 (13.69)
All negatives	86.54 (0.0)	86.82 (0.0)	92.65 (0.0)
Random (prob of positive=0.5)	50.29 (20.63)	50.56 (20.68)	50.65 (12.67)
Search			
LR ^d	36.93 (17.77)	53.97 (24.17)	78.29 (10.72)
RF ^e	33.53 (23.36)	55.22 (18.1)	92.65 ^f (0.0)
LSTM ^g	46.73 (23.63)	69.68 (21.62)	89.96 (7.58)
LSTM-GloVe ^h	53.23 (28.45)	63.44 (27.4)	90.09 (3.73)
LSTM-GloVe w/STE ⁱ	69.17 (28.04)	46.85 (26.51)	91.73 (1.31)
DL-LSTM ^j	62.46 (30.4)	65.99 (26.19)	88.61 (7.97)
DL-LSTM w/STE	69.61 (32.44)	56.84 (27.7)	87.59 (6.99)
Met			
LR	62.57 (39.81)	63.64 (37.25)	58.58 (22)
RF	78.76 (50.59)	71.77 (39.88)	73.78 (24.67)
LSTM	76.54 (48.29)	72.52 (41.27)	67.89 (24.69)
Met+search			
LR	55.99 (36.56)	62 (36.25)	61.25 (21.5)
RF	81.39 (45.35)	73.77 (38.71)	87.96 (23.78)
LSTM	78.18 (47.65)	77.75 (40.31)	88.14 (21.29)
LSTM-GloVe	80.04 (49.37)	72.75 (40.35)	85.38 (26.99)
LSTM-GloVe w/STE	81.85 (50.71)	74.21 (41.49)	85.42 (26.13)
DL-LSTM	77.97 (48.94)	74.81 (40.53)	84.94 (24.07)
DL-LSTM w/STE	80.16 (49.32)	72.99 (40.34)	87.04 (21.32)
Met+pol			
LR	67.38 (44.61)	70.05 (44.09)	74.45 (32.82)
RF	82.81 (57.23)	80.35 (51.24)	86.45 (40.63)
LSTM	86.97 (63.01)	84.64 (55.59)	85.25 (43.19)
Met+pol+search			
LR	66.91 (43.71)	69.13 (43.6)	74.45 (32.82)
RF	82.76 (55.91)	78.91 (47.72)	89.43 (37.57)
LSTM	87.11 (61.54)	84.71 (54.02)	90.74 (44.81)
LSTM-GloVe	87.94 (63.81)	82.98 (53.78)	88.19 (46.55)
LSTM-GloVe w/STE	87.63 (63.83)	83.81 (54.59)	88.24 (46.51)
DL-LSTM	87.30 (63.02)	82.65 (53.65)	89.66 (47.35)
DL-LSTM w/STE	87.60 (63.61)	83.40 (53.58)	89.25 (46.59)

^aO₃: ozone.^bNO₂: nitrogen dioxide.^cPM_{2.5}: fine particulate matter.

^dLR: logistic regression.

^eRF: random forest.

^fThis high accuracy is simply due to class imbalance; this model always predicts negative class, and the corresponding F_1 -score is 0.

^gLSTM: long short-term memory.

^hGloVe: Global Vectors for Word Representation.

ⁱSTE: search term expansion.

^jDL-LSTM: dictionary learner-long short-term memory.

Using Search Data and Meteorological Data

When meteorological data were available, we investigated the feasibility of using meteorological data with or without search activity data to nowcast air pollution under this condition. As shown in the “Met” and “Met+Search” sections of [Table 5](#), the inclusion of web search data improves the nowcasting accuracy for all 3 pollutants. In addition, the LSTM-GloVe w/STE model achieved the highest F_1 -score (50.71% for O₃ and 41.49% for NO₂) for the detection of O₃ and NO₂ pollution. The LSTM-GloVe without STE model achieved the highest F_1 -score (26.99%) for detecting PM_{2.5} pollution.

Using Search Data, Meteorological Data, and Historical Pollutant Concentration

When historical pollution concentration is available, search activity data are added as auxiliary data to both meteorological data and historical pollution data. As shown in the “Met+Pol” and “Met+Pol+Search” sections of [Table 5](#), the inclusion of web search data improves the nowcasting accuracy for O₃ and PM_{2.5}. However, for NO₂, the inclusion of web search data does not improve the nowcasting accuracy, which indicates that

increases in NO₂ concentrations may not be directly noticeable by people sufficiently to increase their search interest. This difference in the performance for different pollutants and locations merits further investigation.

City-Level Analysis of O₃ Pollution Prediction

We investigated the potential of using search interest and meteorological data to replace ground-based O₃ sensor data for predicting O₃ pollution in individual cities. As shown in [Table 6](#), including search interest data (Met+Search) to augment purely meteorological data (Met) increases both the accuracy and F_1 -score metrics for most cities. Although these metrics do not reach performance when ground-level pollution sensors are available (Met+Pol), at least for two of the major MSAs (Philadelphia and Houston), search volume data indeed provides a useful alternative to pollution monitors, with only 1.6% and 0.14% degradation in accuracy, respectively. In addition, the differences in model performance across different cities indicate that web-based search patterns could vary from city to city. As shown in [Table 7](#), the top 5 correlated terms differ across US cities over 10 years. The variation in search patterns could lead to degraded prediction performance in certain areas, leaving promising directions for improvement.

Table 6. City-level accuracy and F_1 -score for detecting elevated ozone pollution in 10 US cities, with Met (long short-term memory model), Met+Search (dictionary learner-long short-term memory w/search term expansion) and Met+Pol (long short-term memory model) as features.

Features	Los Angeles	District of Columbia	Philadelphia	Dallas	Atlanta	Boston	New York	Miami	Chicago	Houston
Accuracy, %										
Met ^a	72.6	77.4	83.29	83.42	83.56	75.62	68.36	58.09	76.71	85.89
Met+search	76.71	80.68	87.4	79.86	83.84	78.63	74.93	69.29	80	90.14
Met+pol ^b	85.89	86.99	89.04	89.04	88.22	84.66	86.85	82.02	86.85	90
F_1-score, %										
Met	51.69	48.28	53.79	53.28	48.72	46.06	44.07	32.52	56.19	57.26
Met+search	54.3	50.53	58.56	41.9	42.72	48	47.86	35.84	57.56	59.09
Met+pol	68.11	60.58	64.29	64.6	56.12	55.56	63.64	55.48	70.73	67.26

^aMet: meteorological data.

^bPol: pollution data.

Table 7. Top 5 correlated search terms for ozone pollution in 10 US cities: January 1, 2010, to December 31, 2019.

City and search term	Spearman correlation (lag=1)
Los Angeles	
Cough	-0.40
Bronchitis	-0.33
Wildfires	0.24
Traffic	0.14
Respiratory infection	-0.12
District of Columbia	
Bronchitis	-0.25
Cough	-0.25
Coughing	-0.19
Headache	-0.14
Wildfires	0.13
Philadelphia	
Cough	-0.33
Traffic	0.27
Bronchitis	-0.20
Organic carbon	-0.10
Respiratory infection	-0.09
Dallas	
Cough	-0.25
Bronchitis	-0.24
Ozone	0.17
Wildfires	0.15
Coughing	-0.14
Atlanta	
Bronchitis	-0.14
Cough	-0.11
Chest pain	-0.10
Respiratory infection	-0.09
Wheezing	-0.07
Boston	
Smoke	-0.11
Haze	-0.07
Code red	-0.06
Coughing	0.06
Smog	0.05
New York	
Bronchitis	-0.31
Traffic	0.29
Cough	-0.25
Wildfires	0.19
Wheezing	-0.15

City and search term	Spearman correlation (lag=1)
Miami	
Bronchitis	0.14
Air pollution	0.13
Cough	0.13
Power plants	0.09
Nitrogen dioxide	0.08
Chicago	
Wildfires	0.18
Smoke	0.08
Shortness of breath	0.04
Heart murmur	0.04
Tail pipe	0.04
Houston	
Ozone	0.12
Air pollution	0.12
Asthma	0.06
Organic carbon	0.05
Wildfires	0.05

Sensitivity Analysis of Air Pollution Thresholds

Classification thresholds play an important role in our model. In this study, an SD threshold from the mean of the corresponding pollutants was used as a “probability threshold” to detect air pollution at a spatial-temporal resolution. However, the proposed method is sensitive to this threshold. We further

investigated the performance of the proposed method using a variety of fixed classification thresholds. As shown in Figures 5-7, we fixed the classification thresholds for all 10 cities to detect O₃, NO₂, and PM_{2.5} pollutions. The results show that the meteorological and search data are complementary, and combining the search and meteorological data leads to better prediction performance for all classification thresholds.

Figure 5. Accuracy (left figure) and F1-score (right figure) for detecting ozone (O₃) pollution on various classification thresholds, with Met (long short-term memory model) and Met+Search (dictionary learner-long short-term memory w/search term expansion) as features. Met: meteorological data; ppb: parts per billion.

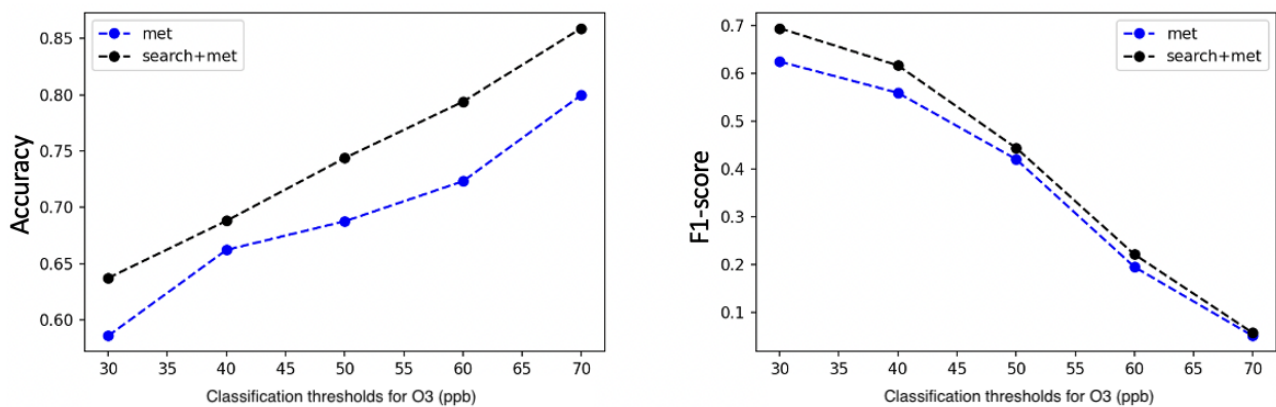


Figure 6. Accuracy (left figure) and F_1 -score (right figure) for detecting nitrogen dioxide (NO_2) pollution on various classification thresholds, with Met (long short-term memory model) and Met+Search (dictionary learner-long short-term memory w/search term expansion) as features. Met: meteorological data; ppb: parts per billion.

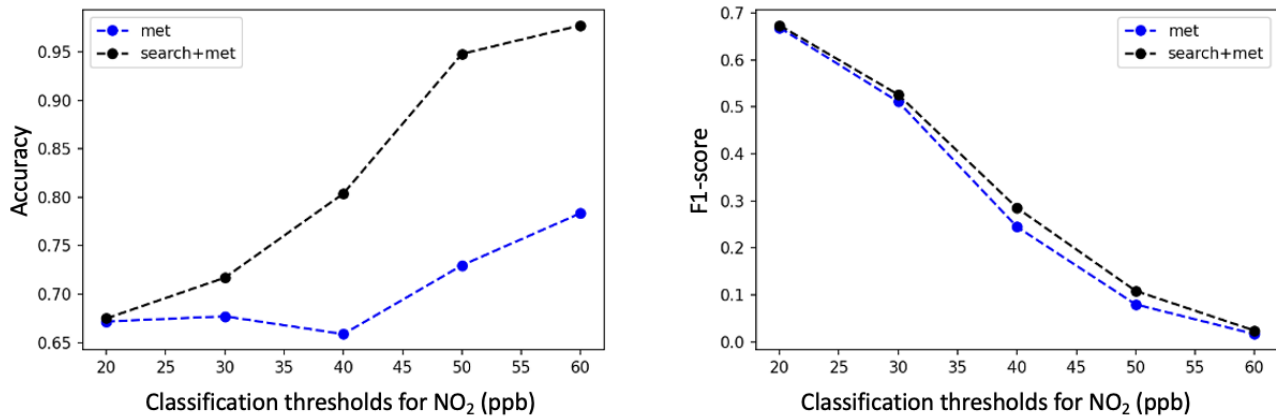
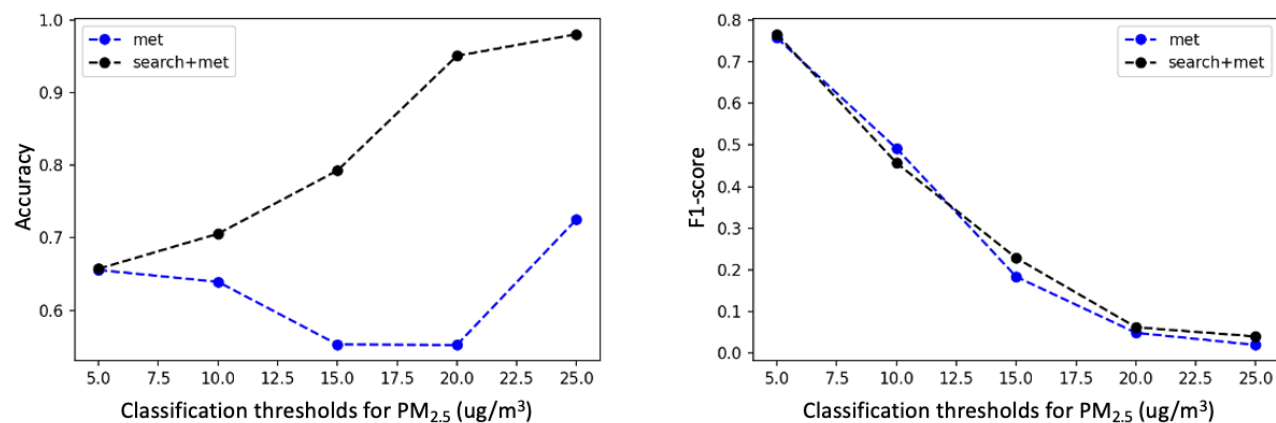


Figure 7. Accuracy (left figure) and F_1 -score (right figure) for detecting fine particulate matter ($\text{PM}_{2.5}$) pollution on various classification thresholds, with Met (long short-term memory model) and Met+Search (dictionary learner-long short-term memory w/search term expansion) as features. Met: meteorological data.



Discussion

Principal Findings

In this study, we explored various existing air pollution prediction models and found that the use of a time series neural network approach achieved the highest predictive accuracy in most of our experiments. The results showed that the LSTM-based models achieved superior accuracy for the 3 air pollutants when both meteorological data and web search data were available. Furthermore, our results on the inclusion of web search data with meteorological data indicate that under short reporting delays, the LSTM models could provide highly accurate predictions compared with baseline models using meteorological and historical pollution concentration data.

Compared with existing studies that predict urban air pollution concentrations using linear and nonlinear machine learning models [25,41-47], our proposed method can predict air pollution when source emissions and remotely sensed satellite data are infeasible (eg, sensed satellite data often suffer from a high missing rate owing to frequent cloud cover [48]). Previous studies using web-based search behavior have emphasized the use of Google Trends [40,49] and applied regularized linear regression to collinear web search queries to estimate disease

rates from social media or web-based search data [18,19,50-54]. Our research further explored the possibility of using LSTM models with semantic embeddings of search queries to predict air pollution. As shown in Figures 8 and 9, the semantic embeddings of search terms fine-tuned by the DL-LSTM model are less correlated compared with their initial GloVe embeddings, which shows that the collinearity between search terms is reduced during the training process.

We also explored various combinations of search terms and found that a comprehensive set of user queries was critical for accurately capturing people's responses to urban air pollution. In this study, we expanded the initial set of seed terms using semantic and temporal correlations with search queries from Google Correlate. We investigated the contribution of different search term groups by manually classifying the search terms into 4 categories, where the unclassified category includes terms with ambiguous meanings. Table 8 shows the accuracy and F_1 -score when we removed search terms by categories for predicting O_3 , NO_2 , and $\text{PM}_{2.5}$ pollution. Removing the search terms in the symptom, observation, and source categories led to a decrease in the accuracy score for detecting at least two pollutants. At the same time, removing the search terms with ambiguous meaning only led to a slightly higher accuracy score for all 3 pollutants.

Figure 8. Cosine similarity between GloVe embeddings of seed search terms. GloVe: Global Vectors for Word Representation.

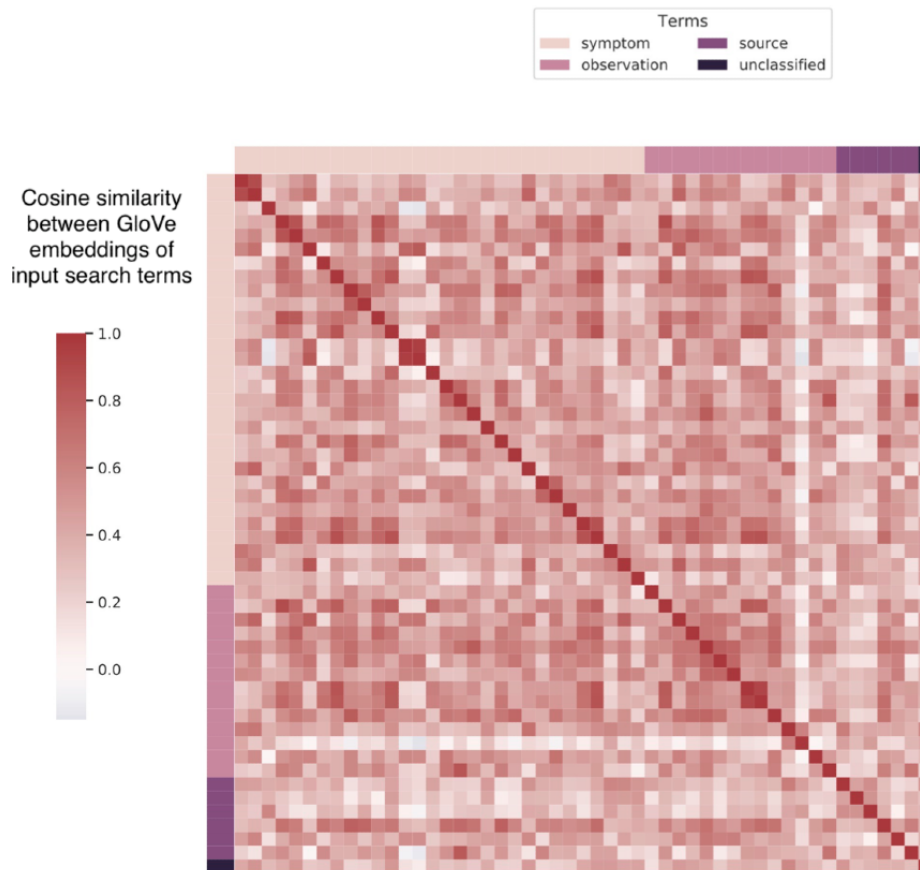


Figure 9. Cosine similarity between trained embeddings of seed search terms.

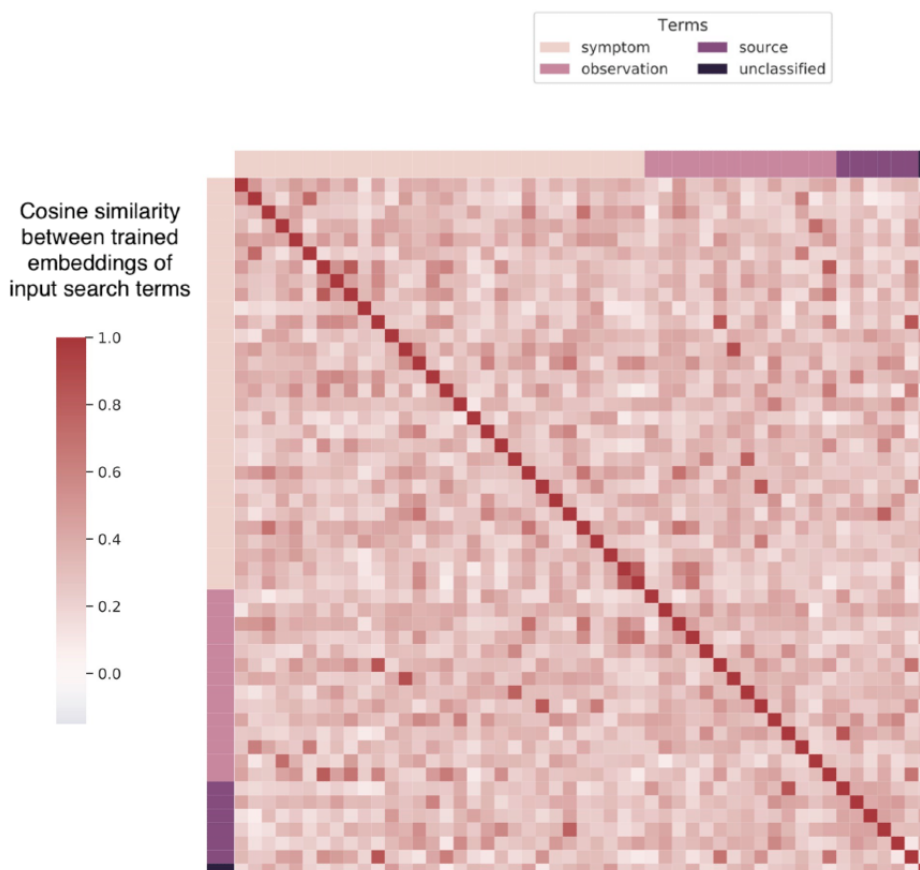


Table 8. Accuracy and F_1 -score of removing different categories of search terms for detecting ozone, nitrogen dioxide, and fine particulate matter pollution using search (dictionary learner-long short-term memory w/search term expansion) as features.

Pollutant and terms	Accuracy (change; %)	F_1 -score (change; %)
O₃^a		
All	0.6961	0.3244
All wo ^b symptom	0.647 (-7.1)	0.3024 (-6.8)
All wo observation	0.622 (-10.6)	0.3264 (+0.6)
All wo source	0.6712 (-3.6)	0.3033 (-6.5)
All wo unclassified	0.7057 (+1.4)	0.3273 (+0.9)
NO₂^c		
All	0.5684	0.2770
All wo symptom	0.4452 (-22.0)	0.2418 (-12.7)
All wo observation	0.6125 (+7.8)	0.2480 (-10.5)
All wo source	0.5452 (-4.1)	0.2647 (-4.4)
All wo unclassified	0.6534 (+15.0)	0.2134 (-23.0)
PM_{2.5}^d		
All	0.8759	0.0699
All wo symptom	0.7897 (-9.8)	0.1029 (+47.2)
All wo observation	0.7496 (-14.4)	0.1049 (+50.1)
All wo source	0.8994 (+2.7)	0.0393 (-43.8)
All wo unclassified	0.8991 (+2.6)	0.0264 (-62.2)

^aO₃: ozone.

^bwo: without.

^cNO₂: nitrogen dioxide.

^dPM_{2.5}: fine particulate matter.

By analyzing the coefficients of each search term, the results show that several search terms contribute more than other search terms. The average feature importance of the seed search terms was calculated using the RF model. As shown in Figure S1, Figure S2, and Figure S3 in [Multimedia Appendix 2](#), search terms including “particular matter,” “rapid breathing,” and “throat irritation” have relatively high feature importance for detecting O₃, NO₂, and PM_{2.5} pollution, respectively. The results also indicated that no search terms worked best for all 3 pollutants.

Limitations

A key limitation of this study is the tuning of the neural network model. First, the performance of neural network models is sensitive to several hyperparameters, including optimization choices, depth, width, and regularization. Owing to computational limitations, we adopted a simple LSTM architecture with a single 128-unit hidden layer and tuned the model using validation data sets for other hyperparameters. In addition, we noticed that stochastic components such as the random seed for the RF model and the randomness in the optimization process of LSTM models influenced the interpretation of the results. Therefore, we repeated the experiments 10 times with different random seeds for the RF

and LSTM models. As the time cost of repeating LSTM models is high, we only repeated the RF, LSTM, and DL-LSTM models 10 times to predict O₃ pollution with all input features. The accuracy of the DL-LSTM model is mean 0.8744 (SD 0.0046). Compared with the LSTM model (mean 0.8714, SD 0.0036), the improvement was not significant ($P=.11$). Compared with the RF model (mean 0.8273, SD 0.0017), the improvement was significant ($P<.001$). The F_1 -score for the DL-LSTM model is mean 0.6314 (SD 0.0058). Compared with both the LSTM (mean 0.6019, SD 0.0096) and RF models (mean 0.5588, SD 0.0024), the improvements are significant ($P<.001$), which shows that the results of the LSTM models are stable. There is room for further exploration of more sophisticated neural network model architectures for noninfectious disease prediction [55-57]. We leave the exploration of deeper and wider architectures to future work.

Another limitation relates to the biases introduced by relying on search data, which may not reflect the underlying population demographics or experiences. Although some of these issues are alleviated automatically by training a model against ground sensor pollution levels, understanding and correcting these data biases requires further study. In the future, we plan to investigate other sources of crowd-based surveillance data, such as

self-reports on social media, to augment traditional physical sensor methods, thus providing a more direct, human-centered measure of how people experience elevated air pollution levels.

Conclusions

In this study, we posit that although web search data cannot yet completely replace ground-based pollution monitors, it may already serve as a valuable additional signal to augment ground-based pollution data, providing significant accuracy

improvements for detecting unusual spikes in air pollution. We also found that the correlation between search terms and pollution concentration varies at the city level. Therefore, the model must be fine-tuned when applied to specific cities. For model and search term selection, we used the simplest LSTM architecture with a dictionary learner module and found that no search terms worked best for all the 3 pollutants. We propose the use of our model to learn the semantic correlations between available search terms to obtain better prediction results.

Acknowledgments

This work was supported by grants from the National Institutes of Health National Library of Medicine (R21LM013014). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The descriptions of search terms, data source, and model hyperparameters.

[[PDF File \(Adobe PDF File\), 107 KB - formative_v6i12e23422_app1.pdf](#)]

Multimedia Appendix 2

Average feature importance for detecting ozone, nitrogen dioxide, and fine particulate matter pollution using random forest models.

[[PDF File \(Adobe PDF File\), 385 KB - formative_v6i12e23422_app2.pdf](#)]

References

1. Brynjolfsson E, Geva T, Reichman S. Crowd-squared: amplifying the predictive power of search trend data. *MIS Q* 2016 Apr 4;40(4):941-961. [doi: [10.25300/MISQ/2016/40.4.07](https://doi.org/10.25300/MISQ/2016/40.4.07)]
2. Fung IC, Tse ZT, Fu KW. The use of social media in public health surveillance. *Western Pac Surveill Response J* 2015 Jun 26;6(2):3-6 [FREE Full text] [doi: [10.5365/WPSAR.2015.6.1.019](https://doi.org/10.5365/WPSAR.2015.6.1.019)] [Medline: [26306208](https://pubmed.ncbi.nlm.nih.gov/26306208/)]
3. Hill S, Merchant R, Ungar L. Lessons learned about public health from online crowd surveillance. *Big Data* 2013 Oct 10;1(3):160-167 [FREE Full text] [doi: [10.1089/big.2013.0020](https://doi.org/10.1089/big.2013.0020)] [Medline: [25045598](https://pubmed.ncbi.nlm.nih.gov/25045598/)]
4. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. *PLoS One* 2013 Dec 9;8(12):e83672 [FREE Full text] [doi: [10.1371/journal.pone.0083672](https://doi.org/10.1371/journal.pone.0083672)] [Medline: [24349542](https://pubmed.ncbi.nlm.nih.gov/24349542/)]
5. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol* 2015 Oct;11(10):e1004513 [FREE Full text] [doi: [10.1371/journal.pcbi.1004513](https://doi.org/10.1371/journal.pcbi.1004513)] [Medline: [26513245](https://pubmed.ncbi.nlm.nih.gov/26513245/)]
6. Kandula S, Hsu D, Shaman J. Subregional nowcasts of seasonal influenza using search trends. *J Med Internet Res* 2017 Nov 06;19(11):e370 [FREE Full text] [doi: [10.2196/jmir.7486](https://doi.org/10.2196/jmir.7486)] [Medline: [29109069](https://pubmed.ncbi.nlm.nih.gov/29109069/)]
7. Ning S, Yang S, Kou SC. Accurate regional influenza epidemics tracking using Internet search data. *Sci Rep* 2019 Mar 27;9(1):5238 [FREE Full text] [doi: [10.1038/s41598-019-41559-6](https://doi.org/10.1038/s41598-019-41559-6)] [Medline: [30918276](https://pubmed.ncbi.nlm.nih.gov/30918276/)]
8. Fung IC, Tse ZT, Cheung CN, Miu AS, Fu KW. Ebola and the social media. *Lancet* 2014 Dec 20;384(9961):2207. [doi: [10.1016/S0140-6736\(14\)62418-1](https://doi.org/10.1016/S0140-6736(14)62418-1)] [Medline: [25625391](https://pubmed.ncbi.nlm.nih.gov/25625391/)]
9. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 2011 May;5(5):e1206 [FREE Full text] [doi: [10.1371/journal.pntd.0001206](https://doi.org/10.1371/journal.pntd.0001206)] [Medline: [21647308](https://pubmed.ncbi.nlm.nih.gov/21647308/)]
10. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Niakan Kalhori SR. Predicting COVID-19 incidence through analysis of Google trends data in Iran: data mining and deep learning pilot study. *JMIR Public Health Surveill* 2020 Apr 14;6(2):e18828 [FREE Full text] [doi: [10.2196/18828](https://doi.org/10.2196/18828)] [Medline: [32234709](https://pubmed.ncbi.nlm.nih.gov/32234709/)]
11. de Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, et al. Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environ Pollut* 2013 May;176:92-99 [FREE Full text] [doi: [10.1016/j.envpol.2012.12.032](https://doi.org/10.1016/j.envpol.2012.12.032)] [Medline: [23416743](https://pubmed.ncbi.nlm.nih.gov/23416743/)]

12. Devarakonda S, Sevusu P, Liu H, Iftode L, Nath B. Real-time air quality monitoring through mobile sensing in metropolitan areas. In: Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing. 2013 Aug Presented at: UrbComp '13; August 11, 2013; Chicago, IL, USA p. 1-8. [doi: [10.1145/2505821.2505834](https://doi.org/10.1145/2505821.2505834)]
13. Snik F, Rietjens JH, Apituley A, Volten H, Mijling B, Di Noia A, et al. Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters. *Geophys Res Lett* 2014 Oct 27;41(20):7351-7358. [doi: [10.1002/2014gl061462](https://doi.org/10.1002/2014gl061462)]
14. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 2017 May 13;389(10082):1907-1918 [FREE Full text] [doi: [10.1016/S0140-6736\(17\)30505-6](https://doi.org/10.1016/S0140-6736(17)30505-6)] [Medline: [28408086](https://pubmed.ncbi.nlm.nih.gov/28408086/)]
15. Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000 May;108(5):419-426 [FREE Full text] [doi: [10.1289/ehp.00108419](https://doi.org/10.1289/ehp.00108419)] [Medline: [10811568](https://pubmed.ncbi.nlm.nih.gov/10811568/)]
16. Sarnat SE, Sarnat JA, Mulholland J, Isakov V, Özkaynak H, Chang HH, et al. Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *J Expo Sci Environ Epidemiol* 2013;23(6):593-605. [doi: [10.1038/jes.2013.41](https://doi.org/10.1038/jes.2013.41)] [Medline: [23963512](https://pubmed.ncbi.nlm.nih.gov/23963512/)]
17. Liang D, Golan R, Moutinho JL, Chang HH, Greenwald R, Sarnat SE, et al. Errors associated with the use of roadside monitoring in the estimation of acute traffic pollutant-related health effects. *Environ Res* 2018 Aug;165:210-219. [doi: [10.1016/j.envres.2018.04.013](https://doi.org/10.1016/j.envres.2018.04.013)] [Medline: [29727821](https://pubmed.ncbi.nlm.nih.gov/29727821/)]
18. Zou B, Lampos V, Cox I. Transfer learning for unsupervised influenza-like illness models from online search data. In: Proceedings of the 2019 World Wide Web Conference. 2019 Presented at: WWW '19; May 13-17, 2019; San Francisco, CA, USA p. 2505-2516. [doi: [10.1145/3308558.3313477](https://doi.org/10.1145/3308558.3313477)]
19. Lampos V, Miller AC, Crossan S, Stefansen C. Advances in nowcasting influenza-like illness rates using search query logs. *Sci Rep* 2015 Aug 03;5:12760 [FREE Full text] [doi: [10.1038/srep12760](https://doi.org/10.1038/srep12760)] [Medline: [26234783](https://pubmed.ncbi.nlm.nih.gov/26234783/)]
20. Graves A, Jaitly N. Towards end-to-end speech recognition with recurrent neural networks. In: Proceedings of the 31st International Conference on International Conference on Machine Learning. 2014 Presented at: ICML '14; June 21-26, 2014; Beijing, China p. II-1764-II-1772.
21. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013 Presented at: NeurIPS '13; December 5-10, 2013; Lake Tahoe, NV, USA p. 3111-3119. [doi: [10.4324/9780203776506-14](https://doi.org/10.4324/9780203776506-14)]
22. Pennington J, Socher R, Manning C. Glove: global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014 Presented at: EMNLP '14; October 25-29, 2014; Doha, Qatar p. 1532-1543. [doi: [10.3115/v1/d14-1162](https://doi.org/10.3115/v1/d14-1162)]
23. Pilotto LS, Douglas RM, Attewell RG, Wilson SR. Respiratory effects associated with indoor nitrogen dioxide exposure in children. *Int J Epidemiol* 1997 Aug;26(4):788-796. [doi: [10.1093/ije/26.4.788](https://doi.org/10.1093/ije/26.4.788)] [Medline: [9279611](https://pubmed.ncbi.nlm.nih.gov/9279611/)]
24. Chauhan AJ, Inskip HM, Linaker CH, Smith S, Schreiber J, Johnston SL, et al. Personal exposure to nitrogen dioxide (NO₂) and the severity of virus-induced asthma in children. *Lancet* 2003 Jul 07;361(9373):1939-1944 [FREE Full text] [doi: [10.1016/s0140-6736\(03\)13582-9](https://doi.org/10.1016/s0140-6736(03)13582-9)] [Medline: [12801737](https://pubmed.ncbi.nlm.nih.gov/12801737/)]
25. Rybarczyk Y, Zalakeviciute R. Machine learning approaches for outdoor air quality modelling: a systematic review. *Appl Sci* 2018 Dec 11;8(12):2570. [doi: [10.3390/app8122570](https://doi.org/10.3390/app8122570)]
26. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, et al. Air pollution and mortality in the Medicare population. *N Engl J Med* 2017 Jun 29;376(26):2513-2522 [FREE Full text] [doi: [10.1056/NEJMoa1702747](https://doi.org/10.1056/NEJMoa1702747)] [Medline: [28657878](https://pubmed.ncbi.nlm.nih.gov/28657878/)]
27. Vedal S, Brauer M, White R, Petkau J. Air pollution and daily mortality in a city with low levels of pollution. *Environ Health Perspect* 2003 Jan;111(1):45-52 [FREE Full text] [doi: [10.1289/ehp.5276](https://doi.org/10.1289/ehp.5276)] [Medline: [12515678](https://pubmed.ncbi.nlm.nih.gov/12515678/)]
28. Sarnat JA, Sarnat SE, Flanders WD, Chang HH, Mulholland J, Baxter L, et al. Spatiotemporally resolved air exchange rate as a modifier of acute air pollution-related morbidity in Atlanta. *J Expo Sci Environ Epidemiol* 2013;23(6):606-615. [doi: [10.1038/jes.2013.32](https://doi.org/10.1038/jes.2013.32)] [Medline: [23778234](https://pubmed.ncbi.nlm.nih.gov/23778234/)]
29. Kelly FJ, Fussell JC. Air pollution and public health: emerging hazards and improved understanding of risk. *Environ Geochem Health* 2015 Aug;37(4):631-649 [FREE Full text] [doi: [10.1007/s10653-015-9720-1](https://doi.org/10.1007/s10653-015-9720-1)] [Medline: [26040976](https://pubmed.ncbi.nlm.nih.gov/26040976/)]
30. Sarnat JA, Russell A, Liang D, Moutinho JL, Golan R, Weber RJ, et al. Developing multipollutant exposure indicators of traffic pollution: the dorm room inhalation to vehicle emissions (DRIVE) study. *Res Rep Health Eff Inst* 2018 Apr(196):3-75. [Medline: [31872750](https://pubmed.ncbi.nlm.nih.gov/31872750/)]
31. Google Trends. Google. URL: <https://support.google.com/trends/answer/4365533?hl=en> [accessed 2019-08-31]
32. Challet D, Bel Hadj Ayed A. Do Google Trend Data Contain More Predictability than Price Returns? SSRN J 2014 Mar 7. [doi: [10.2139/ssrn.2405804](https://doi.org/10.2139/ssrn.2405804)]
33. Kreindler DM, Lumsden CJ. The effects of the irregular sample and missing data in time series analysis. In: Guastello SJ, Gregson RA, editors. *Nonlinear Dynamical Systems Analysis for the Behavioral Sciences Using Real Data*. Boca Raton, FL, USA: CRC Press; 2006.
34. Google correlate. Google. URL: <https://searchengineland.com/google-correlate-more-search-data-to-mine-78560> [accessed 2019-08-31]

35. Kane MJ, Price N, Scotch M, Rabinowitz P. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics* 2014 Aug 13;15(1):276 [FREE Full text] [doi: [10.1186/1471-2105-15-276](https://doi.org/10.1186/1471-2105-15-276)] [Medline: [25123979](https://pubmed.ncbi.nlm.nih.gov/25123979/)]
36. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997 Dec 15;9(8):1735-1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)] [Medline: [9377276](https://pubmed.ncbi.nlm.nih.gov/9377276/)]
37. Elman JL. Distributed representations, simple recurrent networks, and grammatical structure. *Mach Learn* 1991 Sep;7(2-3):195-225. [doi: [10.1007/bf00114844](https://doi.org/10.1007/bf00114844)]
38. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision*. 2015 Presented at: ICCV '15; December 7-13, 2015; Santiago, Chile p. 1026-1034. [doi: [10.1109/iccv.2015.123](https://doi.org/10.1109/iccv.2015.123)]
39. CGAP Project-Nowcasting Air Pollution. URL: <https://github.com/emory-irlab/airpollutionnowcast> [accessed 2021-12-10]
40. Carrière-Swallow Y, Labbé F. Nowcasting with Google trends in an emerging market. *J Forecast* 2013 Jul;32(4):289-298. [doi: [10.1002/for.1252](https://doi.org/10.1002/for.1252)]
41. Chen S, Kan G, Li J, Liang K, Hong Y. Investigating China's urban air quality using big data, information theory, and machine learning. *Pol J Environ Stud* 2018;27(2):565-578. [doi: [10.15244/pjoes/75159](https://doi.org/10.15244/pjoes/75159)]
42. Zhao X, Zhang R, Wu JL, Chang PC. A deep recurrent neural network for air quality classification. *J Inf Hiding Multimed Signal Process* 2018 Mar;9(2):346-354.
43. Lin Y, Mago N, Gao Y, Li Y, Chiang YY, Shahabi C, et al. Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning. In: *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 2018 Presented at: SIGSPATIAL '18; November 6-9, 2018; Seattle, WA, USA p. 359-368. [doi: [10.1145/3274895.3274907](https://doi.org/10.1145/3274895.3274907)]
44. Zhu W, Wang J, Zhang W, Sun D. Short-term effects of air pollution on lower respiratory diseases and forecasting by the group method of data handling. *Atmos Environ* 2012 May;51:29-38. [doi: [10.1016/j.atmosenv.2012.01.051](https://doi.org/10.1016/j.atmosenv.2012.01.051)]
45. Zhang Y, Bocquet M, Mallet V, Seigneur C, Baklanov A. Real-time air quality forecasting, part I: history, techniques, and current status. *Atmos Environ* 2012 Dec;60:632-655. [doi: [10.1016/j.atmosenv.2012.06.031](https://doi.org/10.1016/j.atmosenv.2012.06.031)]
46. Brokamp C, Jandarov R, Rao MB, LeMasters G, Ryan P. Exposure assessment models for elemental components of particulate matter in an urban environment: a comparison of regression and random forest approaches. *Atmos Environ* (1994) 2017 Mar;151:1-11 [FREE Full text] [doi: [10.1016/j.atmosenv.2016.11.066](https://doi.org/10.1016/j.atmosenv.2016.11.066)] [Medline: [28959135](https://pubmed.ncbi.nlm.nih.gov/28959135/)]
47. Cabaneros SM, Calautit JK, Hughes BR. A review of artificial neural network models for ambient air pollution prediction. *Environ Model Soft* 2019 Sep;119:285-304. [doi: [10.1016/j.envsoft.2019.06.014](https://doi.org/10.1016/j.envsoft.2019.06.014)]
48. Misra P, Takeuchi W. Assessing population sensitivity to urban air pollution using google trends and remote sensing datasets. *Int Arch Photogramm Remote Sens Spatial Inf Sci* 2020 Feb 14;XLII-3/W11:93-100. [doi: [10.5194/isprs-archives-xxii-3-w11-93-2020](https://doi.org/10.5194/isprs-archives-xxii-3-w11-93-2020)]
49. Jun SP, Yoo HS, Choi S. Ten years of research change using Google Trends: from the perspective of big data utilizations and applications. *Technol Forecast Soc Change* 2018 May;130:69-87. [doi: [10.1016/j.techfore.2017.11.009](https://doi.org/10.1016/j.techfore.2017.11.009)]
50. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc Natl Acad Sci U S A* 2015 Dec 24;112(47):14473-14478 [FREE Full text] [doi: [10.1073/pnas.1515373112](https://doi.org/10.1073/pnas.1515373112)] [Medline: [26553980](https://pubmed.ncbi.nlm.nih.gov/26553980/)]
51. Lamos V, Cristianini N. Tracking the flu pandemic by monitoring the social web. In: *Proceedings of the 2nd International Workshop on Cognitive Information Processing*. 2010 Presented at: CIP '10; June 14-16, 2010; Elba, Italy p. 411-416. [doi: [10.1109/cip.2010.5604088](https://doi.org/10.1109/cip.2010.5604088)]
52. Lamos V, De Bie T, Cristianini N. Flu detector - tracking epidemics on Twitter. In: *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases*. 2010 Presented at: ECML PKDD '10; September 20-24, 2010; Barcelona, Spain p. 599-602. [doi: [10.1007/978-3-642-15939-8_42](https://doi.org/10.1007/978-3-642-15939-8_42)]
53. Lamos V, Preoțiuc-Pietro D, Cohn T. A user-centric model of voting intention from Social Media. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. 2013 Presented at: ACL '13; August 4-9, 2013; Sofia, Bulgaria p. 993-1003.
54. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Statistical Soc B* 2005 Apr;67(2):301-320. [doi: [10.1111/j.1467-9868.2005.00503.x](https://doi.org/10.1111/j.1467-9868.2005.00503.x)]
55. Deng S, Wang S, Rangwala H, Wang L, Ning Y. Cola-GNN: cross-location attention based graph neural networks for long-term ILI prediction. In: *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2020 Oct 19 Presented at: CIKM '20; October 19-23, 2020; Virtual p. 245-254.
56. Zou B, Lamos V, Cox I. Multi-task learning improves disease models from web search. In: *Proceedings of the 2018 World Wide Web Conference*. 2018 Presented at: WWW '18; April 23-27, 2018; Lyon, France p. 87-96. [doi: [10.1145/3178876.3186050](https://doi.org/10.1145/3178876.3186050)]
57. Zhang Y, Yakob L, Bonsall MB, Hu W. Predicting seasonal influenza epidemics using cross-hemisphere influenza surveillance data and local internet query data. *Sci Rep* 2019 Mar 01;9(1):3262 [FREE Full text] [doi: [10.1038/s41598-019-39871-2](https://doi.org/10.1038/s41598-019-39871-2)] [Medline: [30824756](https://pubmed.ncbi.nlm.nih.gov/30824756/)]

Abbreviations

DL-LSTM: dictionary learner-long short-term memory

GloVe: Global Vectors for Word Representation

LR: logistic regression

LSTM: long short-term memory

MSA: metropolitan statistical area

NO₂: nitrogen dioxide

O₃: ozone

PM_{2.5}: fine particulate matter

RF: random forest

RNN: recurrent neural network

STE: search term expansion

Edited by A Mavragani; submitted 30.03.22; peer-reviewed by C Zhao, A Staffini, W Ceron; comments to author 21.06.22; revised version received 06.10.22; accepted 25.10.22; published 19.12.22.

Please cite as:

Lin C, Yousefi S, Kahoro E, Karisani P, Liang D, Sarnat J, Agichtein E

Detecting Elevated Air Pollution Levels by Monitoring Web Search Queries: Algorithm Development and Validation

JMIR Form Res 2022;6(12):e23422

URL: <https://formative.jmir.org/2022/12/e23422>

doi: [10.2196/23422](https://doi.org/10.2196/23422)

PMID: [36534457](https://pubmed.ncbi.nlm.nih.gov/36534457/)

©Chen Lin, Safoora Yousefi, Elvis Kahoro, Payam Karisani, Donghai Liang, Jeremy Sarnat, Eugene Agichtein. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 19.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Social Media to Engage Justice-Involved Young Adults in Digital Health Interventions for Substance Use: Pilot Feasibility Survey Study

Anna Harrison^{1,2}, PhD; Johanna Folk², PhD; Christopher Rodriguez², MSc; Amanda Wallace^{2,3}, MD; Marina Tolou-Shams², PhD

¹Mental Health Service, San Francisco Veterans Affairs Health Care System, San Francisco, CA, United States

²Department of Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, United States

³Department of Child and Adolescent Psychiatry, Weill Cornell Medical College and Columbia University College of Physicians and Surgeons, New York, NY, United States

Corresponding Author:

Anna Harrison, PhD

Mental Health Service

San Francisco Veterans Affairs Health Care System

4150 Clement Street

San Francisco, CA, 94121

United States

Phone: 1 415 221 4810

Email: anna.harrison@ucsf.edu

Abstract

Background: Young adults involved in the justice system have high rates of substance use disorders and low rates of treatment engagement. Most justice-involved young adults are supervised in the community—not incarcerated in jail or prison—where they have ongoing access to substances and experience significant barriers to care. When they do engage in treatment, they tend to have worse outcomes than justice-involved adolescents and older adults. Despite the need to develop targeted treatments, there are unique challenges in recruiting this population into clinical research. Digital health technology offers many novel avenues for recruiting justice-involved young adults into clinical research studies and disseminating substance use disorder treatments to justice-involved young adults. Because the vast majority of young adults regularly use one or more social media platforms, social media may offer a cost-effective and efficient way to achieve these goals.

Objective: This study aimed to describe the process and feasibility of using social media platforms (Facebook and Reddit) to recruit justice-involved young adults into clinical research. Justice-involved young adults recruited from these platforms completed a survey assessing the acceptability of digital health interventions to address substance use in this population.

Methods: Justice-involved young adults (aged 18-24 years) were recruited through paid advertisements placed on Facebook and Reddit. Participants responded to a web-based survey focused on their substance use, treatment use history, and acceptability of various digital health interventions focused on substance use.

Results: A national sample of justice-involved young adults were successfully enrolled and completed the survey (N=131). Participants were racially diverse (8/131, 6.1% American Indian individuals; 27/131, 20.6% Asian individuals; 23/131, 17.6% Black individuals; 26/131, 19.8% Latinx individuals; 8/131, 6.1% Pacific Islander individuals; 49/131, 37.4% White individuals; and 2/131, 1.5% individuals who identified as “other” race and ethnicity). Advertisements were cost-effective (US \$0.66 per click on Facebook and US \$0.47 per click on Reddit). More than half (72/131, 54.9%) of the participants were on probation or parole in the past year and reported hazardous alcohol (54/131, 51.9%) or drug (66/131, 57.4%) use. Most of the participants (103/131, 78.6%) were not currently participating in substance use treatment. Nearly two-third (82/131, 62.6%) of the participants were willing to participate in one or more hypothetical digital health interventions.

Conclusions: Social media is a feasible and cost-effective method for reaching justice-involved young adults to participate in substance use research trials. With limited budgets, researchers can reach a broad audience, many of whom could benefit from treatment but are not currently engaged in care. Proposed digital health interventions focusing on reducing substance use, such

as private Facebook groups, SMS text message–based appointment reminders, and coaching, had high acceptability. Future work will build on these findings to develop substance use treatment interventions for this population.

(*JMIR Form Res* 2022;6(12):e37609) doi:[10.2196/37609](https://doi.org/10.2196/37609)

KEYWORDS

substance use; young adult; social media; digital health technology; mobile phone

Introduction

Background

Young adults exhibit high rates of substance use disorders (SUDs), yet few receive necessary SUD treatment. In the general population, 5.2 million (15%) young adults aged 18- 25 years reported SUD symptoms severe enough to warrant treatment [1]. Despite being more likely than older adults to meet the clinical criteria for a SUD, young adults are far less likely to receive treatment: only 331,000 (ie, 6.3% of those with an identified need) received specialty SUD services in the prior year [1]. Engaging young adults in SUD treatment can be challenging, as existing treatments often overlook the unique context and circumstances they face [2]. For example, young adults often initiate substance use at an earlier age and experience more psychiatric comorbidities than older adults [2-4]. They also face unique psychosocial barriers to treatment, such as unstable housing and finances [5]. Young adults also tend to have lower self-efficacy for abstinence and fewer coping skills, both of which are key to SUD treatment success [6]. As many young adults have peers who are also using substances [1], their social context tends to be less conducive to reducing use or abstaining. Moreover, popular mutual help organizations, such as Alcoholics Anonymous, may be less appealing to young adults: only 13% of Alcoholics Anonymous members are aged ≤30 years [7].

Substance use problems are particularly common among young adults involved with the justice system [8], with recent national estimates showing that 39% of young adults (aged 18-25 years) on probation and 41% on parole or supervised release meet the clinical criteria for a SUD [9]. Justice-involved young adults have greater difficulty accessing health care than their older counterparts (eg, because they are more likely to remain uninsured [10]). Involvement in the justice system may exacerbate challenges young adults already face with treatment access and engagement. Of the approximately 6 million adults involved with the US adult correctional systems on any given day, approximately 4.5 million live in the community either on probation or parole [11] with ongoing access to substances. Justice-involved young adults are also more likely to live in urban neighborhoods with a disproportionately high density of outlets selling substances (eg, liquor stores and marijuana dispensaries), which may trigger cravings or lead to increased substance use and relapse [12,13]. In response, community supervision agencies across the country have increasingly incorporated SUD treatment into community corrections [14]. Diversion programs offering alternatives to incarceration for those charged with drug offenses are becoming more common, and evidence suggests they are effective at reducing both substance use and recidivism [15]. Specialty drug courts, a

subset of collaborative courts that typically mandate participation in SUD treatment, have also proliferated in recent years [16], thus increasing access to SUD treatment.

Despite the criminal justice system's embrace of these practices, uptake of SUD treatment for the broader community-supervised population has been slow [17]. Only approximately one-third of justice-involved adults with SUD receive treatment in any given year [10]. Significant barriers to treatment engagement may include stigma as well as a failure to identify substance use as a problem behavior [10]. When referred to SUD treatment by probation, young adults are less likely to identify their substance use as a problem behavior that requires change, despite experiencing associated legal consequences, compared with older adults [4]. Furthermore, justice-involved young adults tend to have worse treatment outcomes than both adolescents and older adults [2].

New strategies are needed to reach justice-involved young adults and improve their access to SUD treatment tailored to their unique circumstances. Digital health technology offers numerous ways to adapt SUD treatment to the specific needs of justice-involved young adults. Ranging from web-based to SMS text messaging to smartphone app, digital health technology can facilitate rapid delivery of evidence-based interventions [18]. For example, SMS text messaging interventions have led to tobacco use reductions in young adults [19]. A recent meta-analysis revealed that in the primary care setting, SMS text messaging interventions improve retention in treatment and medication adherence and promote reductions in alcohol, methamphetamine, and opioid use. Web-based interventions are comparable with in-person interventions in terms of their success in improving treatment engagement and increasing abstinence [20].

Despite its promise, to our knowledge, no studies have examined digital health interventions developed specifically for justice-involved young adults. Although digital health technology has shown promise among the adult probation population [21,22], to our knowledge, the unique needs of young adults have not been specifically explored or addressed. More research is needed about how to test and scale substance use interventions using technology with justice-involved young adults. Little is known about their preferences for the format of digital health interventions to address substance use. The first step is to see if and how to recruit justice-involved young adults into such research.

In total, 84% of all young adults (aged 18-29) in the United States use one or more social media platforms [23]. Even populations of young adults without reliable internet access, such as those experiencing homelessness, report using social media regularly [24]. Social media, particularly Facebook, has

been effective for recruiting other subpopulations of young adults (such as smokers and military veterans) into digital health interventions [25-27]. Social media has also been shown to be an effective platform to recruit hard-to-reach populations and traditionally marginalized populations, such as HIV-positive men who have sex with men, Spanish-speaking Latino gay couples, and others [28-32]. Facebook is also an effective medium to deliver interventions [33]. Young adults who use Facebook are diverse in regard to racial and ethnic backgrounds, socioeconomic statuses, and geographic locations. Thus, Facebook may be one way to both recruit justice-involved young adults and deliver digital health SUD interventions specific to substance use. Other social media platforms have similar user bases and promise for recruitment.

Objectives

This pilot feasibility study seeks to understand (1) if and how social media might be used to effectively reach justice-involved young adults for digital health substance use intervention trials and (2) what types of digital health interventions targeting substance use would justice-involved young adults be willing to participate in when such interventions are eventually brought to scale. This study lays the groundwork for future research adapting digital health interventions for this underserved population, thus expanding access and engagement in SUD treatment.

Methods

Participants

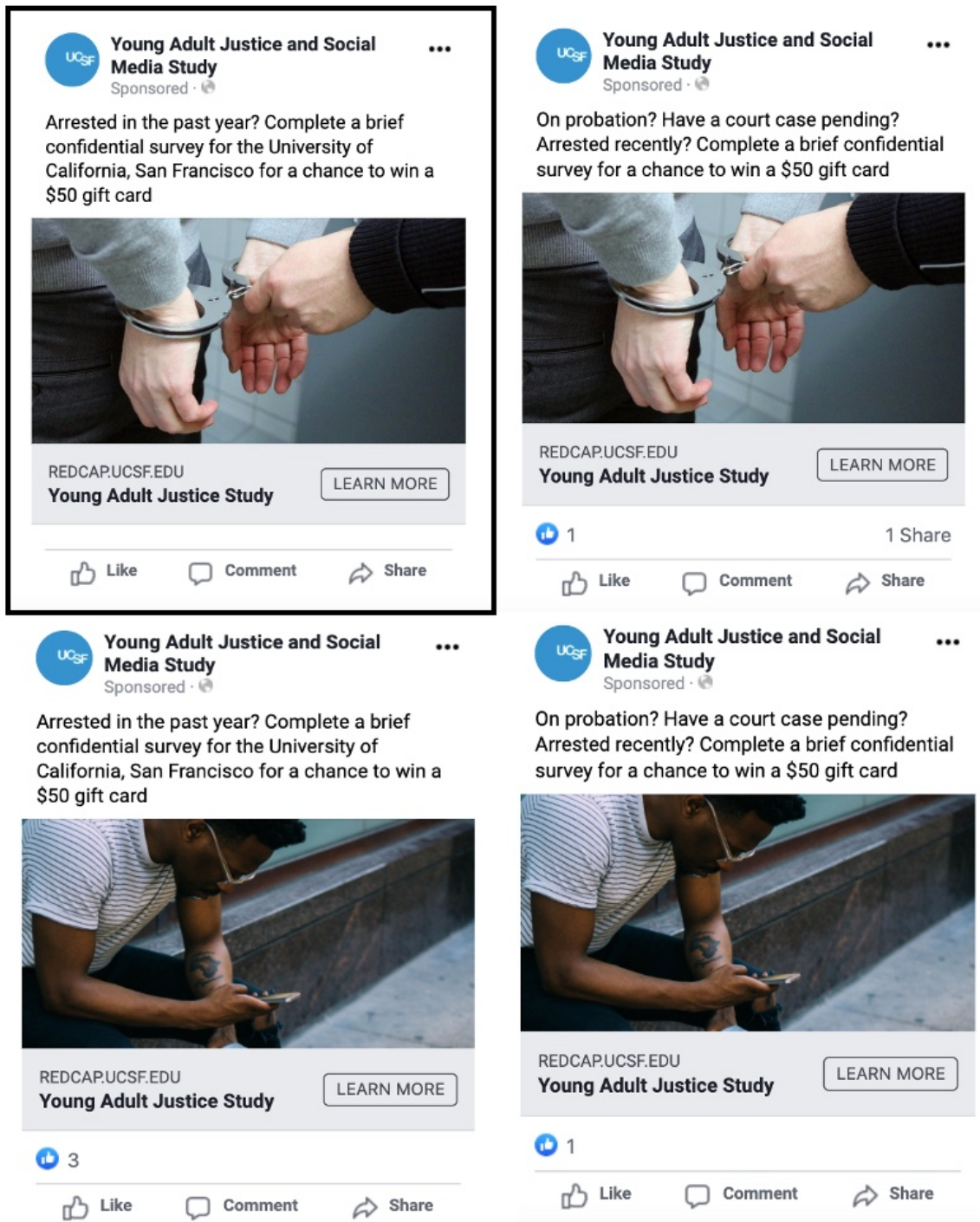
Participants were 131 young adults from across the United States. To be eligible, participants needed to (1) be aged between 18 and 24 years, (2) have been arrested within the past year, (3) have access to a computer or mobile device with the internet,

and (4) be proficient in English. Respondents were not eligible if they were currently incarcerated.

Procedures

Participants were recruited through Facebook and Reddit paid advertisement campaigns (Figure 1). These platforms were selected because of their popularity among young adults: according to Pew Research Center, 70% of adults aged 18-29 years use Facebook and 36% use Reddit [23]. We included both these platforms to minimize sampling bias, as their intended uses and user bases tend to be somewhat different. Moreover, Reddit has specific groups (“subreddits”) for users involved in the justice system, increasing the advertisement’s visibility to the study’s target audience. All study advertisements were direct promotions for the survey website, such that individuals who clicked on the advertisement would be immediately directed to an external website (REDCap [Research Electronic Data Capture; Vanderbilt University]) containing a screening questionnaire for the web-based survey. Eligible participants were then directed to a study information page outlining the purpose of the study, contact information for the principal investigator, and the consent form. We also maintained a Facebook page where participants could review this material, which was reachable from the advertisements themselves. After providing informed consent, respondents were directed to the web-based survey. Participants who submitted their survey and who met eligibility criteria received the incentive. A total of 145 respondents were deemed eligible and completed the consent document. Of the 145 participants, 8 (5.5%) subsequently indicated they had no criminal justice contact in the past year (despite responding “yes” on the screening questionnaire) and 6 (4.1%) did not respond to any questions on the survey. These participants were removed from the final analytic sample. In 2 cases, participants entered the same email address for reimbursement; in these cases, the second response was removed.

Figure 1. Facebook advertisements used for pilot testing. Black outline signifies the most cost-effective advertisement.



Advertising Pilot-Testing

Overview

Study advertising was conducted in an iterative manner. We used the Facebook A/B split testing feature in the advertising manager to compare 2 different images paired with 2 different text options. The A/B split testing feature shows different advertisements to similar types of Facebook users to understand which advertisements garner the most attention. We purposefully chose 2 images with very different stimulus values and paired each of the images with 2 different text options with the aim to better understand the unique effects of both the image and text.

These 4 unique advertisements were initially deployed in the advertisement manager to understand which combination of image and text generated the most engagement with our target population (Figure 1).

The 4 different test advertisements ran for 3 days, from August 30, 2018, to September 2, 2018. At the end of the 3-day trial period, 1 advertisement was determined to be the most “successful” (ie, received the most unique link clicks at the lowest cost per click). After the initial A/B split testing was complete, advertisement campaigns using the most successful advertisement were launched through both the Facebook and Reddit paid advertising platforms.

In addition to the paid advertising campaigns, we posted the images and text in relevant Facebook groups (eg, Probation Researcher Network, SF Adult Probation, and Surviving Probation) and on subreddit threads (eg, Probation, Criminal Justice, Jail, Prison, and Paid Studies) at no cost. These appeared as regular posts and were not displayed as advertisements. As the autofiltering function in Reddit flagged our posts in subreddits as spam, the text of these advertisements was posted with no accompanying image.

During this initial advertisement launch, the incentive was an entry into a raffle for a US \$50 gift card. Although these advertisements had satisfactory engagement on social media (on Facebook, the initial launch garnered 56,441 impressions and 244 clicks over the course of 25 days, which came to US \$0.67 per click; on Reddit, there were 165,988 impressions with 733 clicks over the same period for US \$0.42 per click), they yielded few participants. For example, the original Facebook advertisements yielded 24 people who completed the screener, with 11 eligible, 4 consenting, but only 3 completing the survey. From the original Reddit advertisements, 50 people completed the screener, with 10 eligible, 7 consenting, but only 6 completing the survey. Given these low enrollment numbers, we streamlined our study information sheet and increased our incentive to be a guaranteed US \$10 gift card. Using the same winning advertisement from the original A/B split test, with the text updated to reflect the new incentive, advertisements were relaunched and posted in the same relevant Facebook groups. The sample for this study included only participants from the final advertisement strategy approach (ie, updated study information sheet and increased incentive).

Facebook Advertising Strategy

Paid advertisements were optimized for link clicks and presented to adults aged 18-24 years who used the platforms in English in the 10 most populous metropolitan areas in the United States (New York, Newark, or Jersey City; Los Angeles, Long Beach, or Anaheim; Chicago, Naperville, or Elgin; Dallas, Fort Worth, or Arlington; Washington, Arlington, or Alexandria; Houston, The Woodlands, or Sugar Land; San Francisco, Oakland, or Hayward; Philadelphia, Camden, or Wilmington; Boston, Cambridge, or Newton; and Atlanta, Sandy Springs, or Roswell). As our advertising budget was limited, we chose to limit advertisements to densely populated urban areas where community-supervised young adults are the most likely to reside. Advertisements ran over 12 days (November 11, 2018, to November 23, 2018) with a budget of approximately US \$55 per day for a total of US \$587.74.

Reddit Advertising Strategy

The advertising strategy for Reddit was as similar as possible to the strategy deployed for Facebook; however, there were some alterations because of inherent differences in the advertisement managers. Advertisements were optimized for link clicks and presented to users in the following 11 metropolitan areas, which corresponded as closely as possible to options in the Facebook advertisement manager (noted earlier): New York, Los Angeles, Chicago, Dallas, Washington, District of Columbia, Houston, San Francisco/Oakland/San Jose, Philadelphia/Wilmington, Boston/Manchester, and Atlanta.

The advertisement manager did not allow for targeting of specific age demographics. Although it does allow for targeting of specific groups, or subreddits, the subreddits of potential interest (eg, criminal justice) to our population were not able to be targeted because of their small sizes. Advertisements ran over 14 days (November 11, 2018, to November 26, 2018) with a budget of approximately US \$55 per day and a bid cap per 1000 impressions of US \$1.00 for a total of US \$786.26.

Measures

Demographics and Criminal Justice Characteristics

Participants completed a brief demographics questionnaire to identify their gender, race, ethnicity, marital status, and current criminal justice involvement (eg, currently on probation or parole)

Substance Use

The *Alcohol Use Disorders Identification Test* (AUDIT [34]), a 10-item self-report screening tool, was used to assess past-year alcohol consumption, drinking behaviors, and alcohol-related problems. Responses are rated on a 0-4 scale (anchors differ for each question), with a maximum score of 40 and higher values representing more problematic alcohol use. Individuals who score ≥ 8 are considered at risk for or experiencing alcohol problems [35]. There is considerable evidence for the AUDIT's internal consistency [36,37] and validity [34,38,39]. In this study, internal consistency was 0.89.

The *Drug Use Disorders Identification Test* (DUDIT [40]), an 11-item self-report assessment designed to parallel the AUDIT, was used to assess drug-related problems. Responses are rated on a 0-4 scale (anchors differ for each question), with a maximum score of 44; males with scores of ≥ 6 and females with scores of ≥ 2 are considered to probably have drug-related problems [40]. The DUDIT is psychometrically sound, with high internal, convergent, and discriminant validity [41]. In this study, internal consistency was 0.95.

Substance Use Treatment

Participants self-reported whether they were currently participating in substance use treatment and if so, what type of treatment (eg, residential treatment, outpatient treatment, or peer support groups), whether it was court mandated, and the proportion of scheduled sessions they typically attend. Participants also rated their level of motivation to participate in treatment (1=*not motivated at all* to 5=*very motivated*).

Acceptability of Digital Health Interventions

Participants reported their access to electronic devices (eg, smartphone and computer) and their willingness to participate in digital health interventions focused on substance use (1=*I would definitely participate* to 5=*I would definitely not participate*). Participants were asked to identify whether they would participate in any of the following programs (yes or no) and to rate which top 3 sound most interesting to them: (1) receiving appointment reminders via SMS text messages, (2) web-based peer support community, (3) private groups on Facebook focusing on reducing substance use where you can connect with peers, (4) private groups on Facebook focusing on reducing substance use where you can connect with peers

and chat or contact a provider, (5) motivational posts on social media (eg, Instagram or Snapchat), (6) sessions with providers through video chat (eg, Facetime or Skype), (7) coaching by providers through SMS text messaging or secure messaging, and (8) others (please specify).

Informed Consent and Participation Incentive

The consent form outlined the risks and benefits of study participation and confidentiality procedures. Participants were informed that the information would not be provided to court staff, police, or other justice system staff and that participation did not have any bearing on court-, probation-, or parole-related program requirements. As an incentive for participation, survey respondents were offered a US \$10 Amazon gift card delivered via email. Participants were also provided with contact information of the principal investigator. After completing the survey, participants entered an email address, which was disconnected from survey responses. All study data used in analyses were anonymous.

Ethics Approval

All study procedures were approved by the University of California, San Francisco, Institutional Review Board (IRB approval number: 18-24899).

Plan of Analysis

We first describe findings from pilot advertisement and discuss how these data informed final advertisements to recruit the analysis sample. We then present demographic characteristics, criminal justice involvement, substance use, and treatment characteristics for participants from each recruitment platform. Participants who did not respond to all questions on the AUDIT ($n=27$) or the DUDIT ($n=16$) were excluded from those analyses. There were 6 participants who did not respond to items about current treatment and 9 who did not respond to items concerning the acceptability of digital health interventions. As this was a pilot study focused on the feasibility of recruiting a sample using social media, we did not conduct hypothesis testing to examine demographic or other differences for each platform or use advanced statistical methods to impute missing data. Differences in treatment preferences for those with and without hazardous substance use were assessed using independent sample t tests (2-tailed).

Results

Cost-effectiveness of Facebook and Reddit Advertisements

During the initial A/B split testing phase, the 4 “test” advertisements shown in [Figure 1](#) were shown to a total of 44,168 people, resulting in 246 unique link clicks and costing US \$164.96. Cost per click varied by advertisement. The most successful advertisement was an image of a White person’s hands being handcuffed with accompanying text, “Arrested in the past year? Complete a brief confidential survey for the University of California, San Francisco for a chance to win a \$50 gift card.” This advertisement cost US \$0.57 per click. The least successful advertisement, an image of a young Black man

using a cellphone in an urban environment with the same text, cost US \$0.76 per click.

Facebook advertisements were shown to a total of 90,431 Facebook users, with a total number of 818 unique link clicks. The cost per click was US \$0.66. In total, we recruited 37 young adults through paid Facebook advertisements, which translated to an advertising cost of US \$15.88 per participant. No participants were successfully enrolled from Facebook group posts (which were posted at no cost).

Reddit advertisements were shown to a total of 401,895 Reddit users, with a total number of 1618 advertisement clicks. The cost per click was US \$0.47. In total, we recruited 94 young adults and spent a total of US \$768.26 on Reddit advertisements, which translated to an advertising cost of US \$8.17 per participant. A total of 3 participants were successfully enrolled from subreddit posts (which were posted at no cost).

Sample Characteristics

In total, 37.4% (49/131) of the sample identified as White, and the majority of the sample (72/131, 59%) identified as cisgender men ([Table 1](#)). The median (IQR) age was 21 (20-23) years. Participants were largely single (94/131, 77.7%) and reported a wide range of experiences with the criminal justice system. In the past year, most (84/131, 64.1%) participants spent one or more days in jail, and 20% (26/131) of the participants spent one or more days in prison ([Table 2](#)). Approximately 40.5% (53/131) of the participants had been on probation in the past year, and 16.9% (22/131) of the participants had been on parole. Moreover, 12% (15/131) of the sample reported involvement with collaborative courts in the past year.

Both alcohol and drug use were common. The mean AUDIT score was in the “hazardous” range at 10.8 (SD 10.1), and just more than half (54/104, 51.9%) of the participants scored in the “hazardous” range. Similarly, the mean DUDIT score was 12.3 (SD 12.5), and more than half (66/115, 57.4%) of the participants scored in the hazardous range. Nearly one-fourth (28/125, 22.4%) of participants reported current participation in substance use treatment, and the vast majority (26/28, 96%) of the participants reported treatment that was mandated by the court. The most common treatment modalities were peer support groups (such as Alcoholics Anonymous, Narcotics Anonymous, and Life Ring; $n=16$); individual counseling ($n=14$); and group counseling ($n=12$). Of those who were engaged in formal individual or group therapy, most (22/28, 79%) reported attending regularly (more than 75% of the sessions). Those who reported alcohol or drug use in the hazardous range ($n=85$) placed greater importance on reducing substance use compared with those without hazardous substance use ($n=18$; $t_{101}=-7.31$; $P<.001$). There were no significant differences between groups in participants’ confidence in their ability to reduce substance use.

We did not conduct hypothesis testing to detect differences in participant characteristics between recruiting platforms. It is notable, however, that Facebook appeared to be somewhat more efficient in recruiting Hispanic or Latinx participants: approximately 30% (11/37) of the participants recruited through Facebook identified as Hispanic or Latinx, whereas only 16%

(15/94) of the participants recruited through Reddit identified as such. Regarding recent incarceration history, approximately 90% (33/37) of the participants recruited via Facebook had been incarcerated in jail in the past year, whereas only a little more

than half (51/94, 54%) of the Reddit users reported jail time in the past year. Electronic monitoring appeared to be more common among Reddit users (32/94, 34%) than Facebook users (5/37, 14%).

Table 1. Study sample demographics.

	Total sample	Facebook	Reddit
Gender (total sample: n=122; Facebook: n=34; Reddit: n=88), n (%)			
Cisgender woman	36 (29.5)	8 (23.5)	28 (31.8)
Cisgender man	72 (59)	22 (64.7)	50 (56.8)
Transgender woman	5 (4.1)	0 (0)	5 (5.7)
Transgender man	3 (2.5)	2 (5.9)	1 (1.1)
Nonbinary	5 (4.1)	2 (5.9)	3 (3.4)
Other	1 (0.8)	0 (0)	1 (1.1)
Age (years), median (IQR)	21 (20-23)	21 (20-23)	22 (21-23)
Age (years; total sample: N=131; Facebook: n=37; Reddit: n=94), n (%)			
18-19	23 (17.6)	12 (32.4)	11 (11.7)
20-21	44 (33.6)	10 (27)	34 (36.2)
22-24	64 (48.9)	15 (40.5)	49 (52.1)
Race and ethnicity^a (total sample: N=131; Facebook: n=37; Reddit: n=94), n (%)			
Latinx or Hispanic	26 (19.8)	11 (29.7)	15 (16)
American Indian	8 (6.1)	0 (0)	8 (8.5)
Asian	27 (20.6)	6 (16.2)	21 (22.3)
Black or African American	23 (17.6)	6 (16.2)	17 (18.1)
Pacific Islander or Native Hawaiian	8 (6.1)	2 (5.4)	6 (6.4)
White	49 (37.4)	12 (32.4)	37 (39.5)
Other	2 (1.5)	1 (2.7)	1 (1.1)
Declined to answer	3 (2.3)	1 (2.7)	2 (2.1)
Marital status (total sample: n=121; Facebook: n=33; Reddit: n=88), n (%)			
Single	94 (77.7)	25 (75.8)	69 (78.4)
Married	6 (5)	1 (3)	5 (5.7)
Separated or divorced	7 (5.8)	2 (6.1)	5 (5.7)
Living with partner	13 (10.7)	5 (15.2)	8 (9.1)
Widow or widower	1 (0.8)	0 (0)	1 (1.1)

^aPercentages do not add to 100 because participants were allowed to select more than one option.

Table 2. Justice system involvement, substance use, and treatment history.

	Total sample, n (%)	Facebook, n (%)	Reddit, n (%)
Justice system involvement (past year^a; total sample: N=131; Facebook: n=37; Reddit: n=94)			
Arrested	126 (96.9)	36 (97.3)	90 (96.8)
≥1 day in jail	84 (64.1)	33 (89.2)	51 (54.3)
≥1 day in prison	26 (20)	12 (33.3)	14 (14.9)
Probation	53 (40.5)	17 (45.9)	36 (38.3)
Electronic monitoring	37 (28.5)	5 (13.5)	32 (34.4)
Court-mandated living arrangement	17 (13.2)	3 (8.1)	14 (15.2)
Parole	22 (16.9)	8 (21.6)	14 (15.1)
Awaiting court proceeding	17 (13)	7 (18.9)	10 (10.6)
Recently convicted and waiting to serve sentence	6 (4.6)	2 (5.4)	4 (4.3)
Involved in collaborative court	15 (11.5)	5 (13.9)	10 (10.6)
Hazardous alcohol use (total sample: n=104; Facebook: n=27; Reddit: n=77)	54 (51.9)	14 (51.9)	40 (51.9)
Hazardous drug use (total sample: n=115; Facebook: n=29; Reddit: n=86)	66 (57.4)	16 (55.2)	50 (58.1)
Current substance use treatment (total sample: n=125; Facebook: n=35; Reddit: n=90)	28 (22.4)	3 (8.6)	25 (27.8)
Current substance use treatment, mandated by court (total sample: n=28; Facebook: n=3; Reddit: n=25)	26 (96.3)	3 (100)	23 (95.8)
Treatment type (total sample: N=131; Facebook: n=37; Reddit: n=94)			
Inpatient or residential	3 (2.3)	0 (0)	3 (3.2)
Partial hospitalization	5 (3.8)	0 (0)	5 (5.3)
Opioid replacement	9 (6.9)	1 (2.7)	8 (8.5)
Narcan or naloxone prescription	1 (0.8)	0 (0)	1 (1.1)
Narcan or naloxone use	0 (0)	0 (0)	0 (0)
Other medications	1 (0.8)	1 (2.7)	0 (0)
Individual counseling	14 (10.7)	3 (8.1)	11 (11.7)
Group counseling	12 (9.2)	2 (5.4)	10 (10.6)
Peer support	16 (12.2)	3 (8.1)	13 (13.8)
Other treatment	0 (0)	0 (0)	0 (0)

^aPercentages do not add to 100 because participants were allowed to select more than one option.

Interest in Digital Health Interventions

Nearly all participants reported owning a smartphone (115/122, 94.3%) or having regular access to a computer (113/122, 92.6%). In total, 62.6% (82/131) of the participants expressed interest in participating in at least one digital health substance use intervention listed (asked as dichotomous yes or no). Each option presented to participants had relatively similar levels of interest (between 28/131, 24%, and 35/131, 27%, of the participants reported that they would consider participation). When asked to rank digital health substance use interventions, the three most popular were (1) receiving appointment reminders via SMS text messages, (2) web-based peer support community, and (3) private groups on Facebook focusing on reducing substance use. Those who reported hazardous substance use ([Multimedia Appendix 1](#)) reported more openness to

motivational posts on social media ($t_{101}=-1.96$; $P=.05$) and coaching by providers via SMS text messaging or secure messaging ($t_{101}=-2.55$; $P=.01$) than their peers without hazardous substance use.

Discussion

Principal Findings

This study demonstrated the feasibility of recruiting a national sample of justice-involved young adults via social media. With some minor adjustments to initial advertisements and incentives, recruitment was efficient and cost-effective. Over the course of approximately 2 weeks, we surpassed our recruitment goal of 100 participants (the final sample included 131 participants) and spent US \$1356 on advertisements. Reddit was somewhat

more cost-effective than Facebook in recruiting participants (US \$8.17 per participant vs US \$15.88 per participant). Despite this cost difference, placing advertisements on multiple social media platforms was helpful to recruit a diverse sample of justice-involved young adults.

Although our sample was fairly heterogeneous in terms of race and ethnicity, there were key racial and ethnic differences between the sample included in this study and the overall population of justice-involved adults in the United States. When compared with the overall population of adults who have been arrested in the past year, this study included a higher proportion of Asian, Pacific Islander or Native Hawaiian, and American Indian participants than expected. For example, although approximately 1.3% of the population of adults arrested in the past year are Asian [42], 20.6% (27/131) of the participants of this study identified themselves as Asian. Similarly, Native American or American Indian people comprise approximately 2.4% of the total population of arrestees [42] but represented 6.1% (8/131) of our sample. These data are consistent with prior research demonstrating that using social media to recruit participants may be an effective way to recruit specific populations of minoritized groups that are underrepresented in clinical research [28,29].

Notably, Black-identified people were underrepresented in our sample. Although nationally, 26.1% of adults arrested in the past year identify as Black or African American [42], less than 18% (23/131) of our sample did. Just under 20% (26/131) of our sample identified themselves as Latinx or Hispanic, which roughly corresponds to national estimates: 18.8% of adults arrested in the past year are reported to be Hispanic or Latino [42]. Only 37.4% (49/131) of our sample identified as White, which is substantially lower than expected. Across the United States, 69.9% of adults arrested in the past year are White. This finding may be a result of our recruitment strategy, which was focused on advertising to participants from more diverse, urban locations and did not include rural communities.

We found that an image of a White-appearing person in handcuffs received the most engagement and was the most cost-effective for recruitment (US \$0.57 per click). Despite the fact that most people arrested in any given year are White [42], stereotypes in the United States associating people of color with the criminal justice system remain pervasive. A White-appearing person in handcuffs may have been unexpected and the most eye-catching of the images we trialed. However, our sample included fewer than expected Black participants, thus raising the possibility that this image may have been less engaging to Black young adults. Although academic research regarding structural racism in advertising remains somewhat limited, this finding is consistent with prior research from the field of marketing showing that advertisements featuring images of Black people are significantly more likely to engage Black young adults than images of people from other racial and ethnic groups [43].

Digital health SUD interventions were broadly acceptable to survey respondents, with nearly two-thirds (82/131, 62.6%) of the sample expressing interest in one or more potential programs. The most broadly acceptable intervention was SMS

text reminders for appointments. Interventions focused on fostering supportive communities on the web, either with peer support or via private Facebook groups, were also popular.

Notably, this study was conducted before the COVID-19 pandemic. When social distancing protocols were widely implemented across the United States in March 2020, access to in-person SUD treatment and peer support communities (eg, Alcoholics Anonymous) drastically reduced [44,45]. Providing most behavioral health services via telehealth is feasible, particularly with regulatory changes made during the pandemic [46,47]; however, many SUD treatment models rely on groups and thus are more challenging to adapt to web-based formats and often have less group cohesion and treatment alliance [48]. Surveillance data suggest substance use has increased during the COVID-19 pandemic [49], particularly among young adults, thus widening the already-considerable gap between the number of people who would benefit from treatment and the number of people who receive it. Digital health interventions for SUDs may therefore be even more acceptable than reported in this study, and swift adoption by providers and systems is more important than ever before.

Consistent with prior studies [9,10,50], the results of this survey demonstrated a need for SUD treatment among justice-involved young adults. More than half of the sample reported alcohol use (54/104, 51.9%) or drug use (66/115, 57.4%) in the hazardous range. Less than one-fourth (28/125, 22.4%) of the sample were currently receiving treatment, the vast majority of which was court ordered. Notably, young adults who reported hazardous substance use expressed preferences for active engagement with providers in digital spaces, such as receiving motivational posts and coaching by providers via SMS text messaging or secure messaging. More justice-involved young adults may, therefore, benefit from substance treatment. Reaching justice-involved young adults through social media, with the ultimate goal of expanding available digital health interventions, may reduce barriers to entry for care, especially for young adults who have difficulty attending in-person sessions or have ambivalent engagement in treatment.

Strengths and Limitations

This study has several notable strengths and limitations that can inform future research. The strengths include nationwide sampling and the heterogeneity of young adults recruited in terms of gender, race and ethnicity, and substance use. Obtaining perspectives from a variety of young adults allowed a more comprehensive (ie, nationwide survey) investigation of young adults' perspectives regarding the potential feasibility and acceptability of substance use digital health interventions. The study supports the use of social media to recruit justice-involved young adults for research studies and provides insight into several potentially acceptable avenues for digital health interventions. The use of these approaches has the potential to expand access to treatment and promote health equity among justice-involved young adults.

A key limitation of this study, similar to most internet-based data collection, involves verifying respondent identity. We relied on self-report of justice involvement, so it is possible that some participants mischaracterized their experiences to gain study

entry. Given the minimal incentive and that only participants who completed the full survey received the incentive, false identification of oneself as a justice-involved person seems less likely. To protect our respondents' privacy, we did not track IP addresses, so it is possible that some respondents attempted to take the survey more than once; in any cases where the email was identical (this occurred in 2 cases), the duplicate (second response) was removed. When possible, tracking IP addresses in future studies could provide an additional method of preventing multiple entries from a single individual.

Our study relied on recruitment from only 2 social media platforms, so we did not reach justice-involved young adults who use other forms of social media. Facebook and Reddit are used by many young adults: approximately 23% of Facebook users are aged between 18 and 24 years [51], and 58% of Reddit users are aged between 18 and 29 years [52]. Furthermore, Facebook and Reddit both offer easy-to-use advertisement platforms for conducting research. Future studies should consider expanding recruitment to other social media platforms (eg, YouTube and Twitter, which are used by 95% and 42% of young adults, respectively [23]) to reach a wider range of justice-involved youth adults. Justice-involved young adults who do not have internet access were unable to participate in this study. However, as 96% of young adults in the United States own smartphones [53], the vast majority of young adults have some form of internet access.

Although our sample was mixed in terms of gender, race, ethnicity, and substance use history, we only enrolled English-speaking young adults. This excludes a portion of the justice-involved population in the United States and decreases generalizability. The inclusion of individuals who speak languages other than English may provide insight into different preferences for digital health interventions, which should be considered in future studies and in design of interventions to promote health equity.

Finally, we did not collect data on concerns that young adults may have with engaging in digital health interventions or in internet-based SUD treatment or whether young adults would prefer traditional services over digital health. Such concerns will be critical to understand and address in future work developing digital health interventions.

Future Directions

Despite limitations, this pilot study suggests several promising directions for future research. First, future studies should continue to hone recruitment methods using social media to reach larger samples of justice-involved young adults. Expanding this research to new platforms such as YouTube and

Twitter, in addition to refining advertisements, will be critical for reaching a representative sample of justice-involved young adults. Future work should also test the efficacy of advertisements using a wider variety of text options and images featuring young adults from other racial and ethnic backgrounds, particularly Black young adults.

This study demonstrates that social media is an effective tool to recruit justice-involved young adults into clinical research; future work can build on this finding to develop robust empirical support for substance use treatment interventions for this population for whom a dearth of interventions are currently available and used [6,10]. Young adults indicated that digital health interventions were broadly acceptable. Adjuncts to existing treatment, such as receiving appointment reminders via SMS text messages, are already broadly implemented in many medical systems and have been effective in enhancing connection with care [54]. Systems serving young adults in the justice system, such as community mental health centers and courts, should consider adopting this technology as well.

Young adults who reported hazardous substance use were also open to coaching via secure messages or texting treatment providers may wish to implement this into existing practices to increase treatment engagement and improve outcomes. Although there are barriers to widely implementing this mode of communication, particularly with regard to suicide risk management and difficulty receiving reimbursement [49,55], existing structures of treatment are not effective for many justice-involved young adults [4]. Adapting treatment to the modalities by which young adults typically communicate may enhance treatment engagement and outcomes.

Private Facebook groups, with daily posts and discussion moderated by trained clinical staff, have been shown to reduce cigarette smoking and alcohol use among diverse populations of young adults [26,33,56]. Young adults found such groups to be convenient and expressed that social support is particularly beneficial [57]. Aspects of this model may be adapted for young adults involved in the justice system who seek to reduce their own substance use.

Conclusions

This pilot feasibility study established the utility of social media in recruiting a broad sample of young adults involved in the criminal justice system. These young adults reported high rates of hazardous substance use; however, few were in treatment voluntarily. As this high-need, underserved population reported strong interest in digital health interventions, future work will leverage social media to create programs to engage justice-involved young adults in substance use treatment.

Acknowledgments

The authors extend their gratitude to the young adults who participated in this study, without whom this work would not have been possible. This work was supported by the National Institute of Mental Health (grant number T32MH018261); the National Institute on Drug Abuse (grant numbers K24DA046569 and T32DA007250); and the University of California, San Francisco, Department of Psychiatry and Behavioral Sciences Clinical Research Start-up Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Acceptable digital mental health interventions for participants with and without hazardous drug and alcohol use. * $P=.05$. ** $P=.01$.
[PNG File , 262 KB - formative_v6i12e37609_app1.png]

References

1. Lipari RN, Park-Lee E. Key Substance Use and Mental Health Indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. 2018. URL: <https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHNationalFindingsReport2018/NSDUHNationalFindingsReport2018.pdf> [accessed 2020-03-30]
2. Smith DC. Emerging Adults and Substance Use Disorder Treatment: Developmental Considerations and Innovative Approaches. Oxford, UK: Oxford University Press; 2017.
3. Mason MJ, Luckey B. Young adults in alcohol-other drug treatment: an understudied population. *Alcohol Treat Q* 2003 May 05;21(1):17-32. [doi: [10.1300/j020v21n01_02](https://doi.org/10.1300/j020v21n01_02)]
4. Sinha R, Easton C, Kemp K. Substance abuse treatment characteristics of probation-referred young adults in a community-based outpatient program. *Am J Drug Alcohol Abuse* 2003 Aug;29(3):585-597. [doi: [10.1081/ada-120023460](https://doi.org/10.1081/ada-120023460)] [Medline: [14510042](https://pubmed.ncbi.nlm.nih.gov/14510042/)]
5. Elswick A, Fallin-Bennett A, Ashford K, Werner-Wilson R. Emerging adults and recovery capital: barriers and facilitators to recovery. *J Addict Nurs* 2018;29(2):78-83. [doi: [10.1097/JAN.0000000000000218](https://doi.org/10.1097/JAN.0000000000000218)] [Medline: [29864054](https://pubmed.ncbi.nlm.nih.gov/29864054/)]
6. Kelly JF, Urbanoski KA, Hoepfner BB, Slaymaker V. "Ready, willing, and (not) able" to change: young adults' response to residential treatment. *Drug Alcohol Depend* 2012 Mar 01;121(3):224-230 [FREE Full text] [doi: [10.1016/j.drugalcdep.2011.09.003](https://doi.org/10.1016/j.drugalcdep.2011.09.003)] [Medline: [21955361](https://pubmed.ncbi.nlm.nih.gov/21955361/)]
7. 2014 Membership Survey. Alcoholics Anonymous. 2015. URL: https://www.aa.org/sites/default/files/literature/assets/p-48_membershipsurvey.pdf [accessed 2022-09-05]
8. Mumola CJ, Karberg JC. Drug use and dependence, state and federal prisoners, 2004. US Department of Justice, Office of Justice Programs, Bureau of Justice. 2007. URL: <https://bjs.ojp.gov/content/pub/pdf/dudsfp04.pdf> [accessed 2020-03-30]
9. 2018 National Survey of Drug Use and Health (NSDUH): Detailed Tables. Substance Abuse and Mental Health Services Administration. 2018 Aug 20. URL: <https://www.samhsa.gov/data/release/2018-national-survey-drug-use-and-health-nsduh-releases> [accessed 2022-09-05]
10. Saloner B, Bandara SN, McGinty EE, Barry CL. Justice-involved adults with substance use disorders: coverage increased but rates of treatment did not in 2014. *Health Aff (Millwood)* 2016 Jun 01;35(6):1058-1066. [doi: [10.1377/hlthaff.2016.0005](https://doi.org/10.1377/hlthaff.2016.0005)] [Medline: [27269023](https://pubmed.ncbi.nlm.nih.gov/27269023/)]
11. Kaebler D, Cowhig M. Correctional populations in the United States, 2016. U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. 2018 Apr. URL: <https://bjs.ojp.gov/content/pub/pdf/cpus16.pdf> [accessed 2020-03-30]
12. Farahmand P, Arshed A, Bradley MV. Systemic racism and substance use disorders. *Psychiatr Ann* 2020 Nov;50(11):494-498. [doi: [10.3928/00485713-20201008-01](https://doi.org/10.3928/00485713-20201008-01)]
13. Mennis J, Stahler GJ, Mason MJ. Risky substance use environments and addiction: a new frontier for environmental justice research. *Int J Environ Res Public Health* 2016 Jun 18;13(6):607 [FREE Full text] [doi: [10.3390/ijerph13060607](https://doi.org/10.3390/ijerph13060607)] [Medline: [27322303](https://pubmed.ncbi.nlm.nih.gov/27322303/)]
14. Guevara M, Solomon E. Implementing evidence-based policy and practice in community corrections. 2nd edition. National Institute of Corrections, U.S. Department of Justice. 2009 Oct. URL: https://www.crj.org/assets/2017/07/Community_Corrections_BoxSet_Oct09.pdf [accessed 2022-09-05]
15. Belenko S, Hiller M, Hamilton L. Treating substance use disorders in the criminal justice system. *Curr Psychiatry Rep* 2013 Nov;15(11):414 [FREE Full text] [doi: [10.1007/s11920-013-0414-z](https://doi.org/10.1007/s11920-013-0414-z)] [Medline: [24132733](https://pubmed.ncbi.nlm.nih.gov/24132733/)]
16. Marlowe DB, Hardin CD, Fox CL. Painting the current Picture: A national report on drug courts and other problem-solving court programs in the United States. National Drug Court Institute. Washington, DC, USA: National Drug Court Institute; 2016. URL: <https://www.ndci.org/wp-content/uploads/2016/05/Painting-the-Current-Picture-2016.pdf> [accessed 2022-09-05]
17. Belenko S, Johnson ID, Taxman FS, Rieckmann T. Probation staff attitudes toward substance abuse treatment and evidence-based practices. *Int J Offender Ther Comp Criminol* 2018 Feb;62(2):313-333. [doi: [10.1177/0306624X16650679](https://doi.org/10.1177/0306624X16650679)] [Medline: [27220361](https://pubmed.ncbi.nlm.nih.gov/27220361/)]
18. Aguilera A. Digital technology and mental health interventions: opportunities and challenges. *Arbor* 2015 May 12;191(771):a210. [doi: [10.3989/arbor.2015.771n1012](https://doi.org/10.3989/arbor.2015.771n1012)]
19. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. *Prev Sci* 2015 Feb;16(2):181-188. [doi: [10.1007/s11121-014-0498-7](https://doi.org/10.1007/s11121-014-0498-7)] [Medline: [24930386](https://pubmed.ncbi.nlm.nih.gov/24930386/)]

20. Tofighi B, Abrantes A, Stein MD. The role of technology-based interventions for substance use disorders in primary care: a review of the literature. *Med Clin North Am* 2018 Jul;102(4):715-731 [FREE Full text] [doi: [10.1016/j.mcna.2018.02.011](https://doi.org/10.1016/j.mcna.2018.02.011)] [Medline: [29933825](https://pubmed.ncbi.nlm.nih.gov/29933825/)]
21. Cowell AJ, Zarkin GA, Wedehase BJ, Lerch J, Walters ST, Taxman FS. Cost and cost-effectiveness of computerized vs. in-person motivational interventions in the criminal justice system. *J Subst Abuse Treat* 2018 Apr;87:42-49 [FREE Full text] [doi: [10.1016/j.jsat.2018.01.013](https://doi.org/10.1016/j.jsat.2018.01.013)] [Medline: [29471925](https://pubmed.ncbi.nlm.nih.gov/29471925/)]
22. Walters ST, Ondersma SJ, Ingersoll KS, Rodriguez M, Lerch J, Rossheim ME, et al. MAPIT: development of a web-based intervention targeting substance abuse treatment in the criminal justice system. *J Subst Abuse Treat* 2014 Jan;46(1):60-65 [FREE Full text] [doi: [10.1016/j.jsat.2013.07.003](https://doi.org/10.1016/j.jsat.2013.07.003)] [Medline: [23954392](https://pubmed.ncbi.nlm.nih.gov/23954392/)]
23. Social Media Fact Sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/> [accessed 2022-09-04]
24. Adkins EC, Zalta AK, Boley RA, Glover A, Karnik NS, Schueller SM. Exploring the potential of technology-based mental health services for homeless youth: a qualitative study. *Psychol Serv* 2017 May;14(2):238-245 [FREE Full text] [doi: [10.1037/ser0000120](https://doi.org/10.1037/ser0000120)] [Medline: [28481610](https://pubmed.ncbi.nlm.nih.gov/28481610/)]
25. Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. *JMIR Res Protoc* 2015 Jun 01;4(2):e63 [FREE Full text] [doi: [10.2196/resprot.3996](https://doi.org/10.2196/resprot.3996)] [Medline: [26033209](https://pubmed.ncbi.nlm.nih.gov/26033209/)]
26. Ramo DE, Rodriguez TM, Chavez K, Sommer MJ, Prochaska JJ. Facebook recruitment of young adult smokers for a cessation trial: methods, metrics, and lessons learned. *Internet Interv* 2014 Apr;1(2):58-64 [FREE Full text] [doi: [10.1016/j.invent.2014.05.001](https://doi.org/10.1016/j.invent.2014.05.001)] [Medline: [25045624](https://pubmed.ncbi.nlm.nih.gov/25045624/)]
27. Teo AR, Liebow SB, Chan B, Dobscha SK, Graham AL. Reaching those at risk for psychiatric disorders and suicidal ideation: Facebook advertisements to recruit military veterans. *JMIR Ment Health* 2018 Jul 05;5(3):e10078 [FREE Full text] [doi: [10.2196/10078](https://doi.org/10.2196/10078)] [Medline: [29980498](https://pubmed.ncbi.nlm.nih.gov/29980498/)]
28. Topolovec-Vranic J, Natarajan K. The use of social media in recruitment for medical research studies: a scoping review. *J Med Internet Res* 2016 Nov 07;18(11):e286 [FREE Full text] [doi: [10.2196/jmir.5698](https://doi.org/10.2196/jmir.5698)] [Medline: [27821383](https://pubmed.ncbi.nlm.nih.gov/27821383/)]
29. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: systematic review. *Internet Interv* 2016 May;4:72-81 [FREE Full text] [doi: [10.1016/j.invent.2016.02.001](https://doi.org/10.1016/j.invent.2016.02.001)] [Medline: [30135792](https://pubmed.ncbi.nlm.nih.gov/30135792/)]
30. Hernandez-Romieu AC, Sullivan PS, Sanchez TH, Kelley CF, Peterson JL, Del Rio C, et al. The comparability of men who have sex with men recruited from venue-time-space sampling and Facebook: a cohort study. *JMIR Res Protoc* 2014 Jul 17;3(3):e37 [FREE Full text] [doi: [10.2196/resprot.3342](https://doi.org/10.2196/resprot.3342)] [Medline: [25048694](https://pubmed.ncbi.nlm.nih.gov/25048694/)]
31. Kayrouz R, Dear BF, Karin E, Titov N. Facebook as an effective recruitment strategy for mental health research of hard to reach populations. *Internet Interv* 2016 May;4:1-10 [FREE Full text] [doi: [10.1016/j.invent.2016.01.001](https://doi.org/10.1016/j.invent.2016.01.001)] [Medline: [30135786](https://pubmed.ncbi.nlm.nih.gov/30135786/)]
32. Martinez O, Wu E, Shultz AZ, Capote J, López Rios J, Sandfort T, et al. Still a hard-to-reach population? Using social media to recruit Latino gay couples for an HIV intervention adaptation study. *J Med Internet Res* 2014 Apr 24;16(4):e113 [FREE Full text] [doi: [10.2196/jmir.3311](https://doi.org/10.2196/jmir.3311)] [Medline: [24763130](https://pubmed.ncbi.nlm.nih.gov/24763130/)]
33. Ramo DE, Thrul J, Delucchi KL, Hall S, Ling PM, Belohlavek A, et al. A randomized controlled evaluation of the tobacco status project, a Facebook intervention for young adults. *Addiction* 2018 May 24;113(9):1683-1695 [FREE Full text] [doi: [10.1111/add.14245](https://doi.org/10.1111/add.14245)] [Medline: [29797621](https://pubmed.ncbi.nlm.nih.gov/29797621/)]
34. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on early detection of persons with harmful alcohol consumption--II. *Addiction* 1993 Jun;88(6):791-804. [doi: [10.1111/j.1360-0443.1993.tb02093.x](https://doi.org/10.1111/j.1360-0443.1993.tb02093.x)] [Medline: [8329970](https://pubmed.ncbi.nlm.nih.gov/8329970/)]
35. Conigrave KM, Hall WD, Saunders JB. The AUDIT questionnaire: choosing a cut-off score. *Alcohol Use Disorder Identification Test*. *Addiction* 1995 Oct;90(10):1349-1356. [doi: [10.1046/j.1360-0443.1995.901013496.x](https://doi.org/10.1046/j.1360-0443.1995.901013496.x)] [Medline: [8616463](https://pubmed.ncbi.nlm.nih.gov/8616463/)]
36. Barry KL, Fleming MF. The Alcohol Use Disorders Identification Test (AUDIT) and the SMAST-13: predictive validity in a rural primary care sample. *Alcohol Alcohol* 1993 Jan;28(1):33-42. [Medline: [8471085](https://pubmed.ncbi.nlm.nih.gov/8471085/)]
37. Doyle SR, Donovan DM, Kivlahan DR. The factor structure of the Alcohol Use Disorders Identification Test (AUDIT). *J Stud Alcohol Drugs* 2007 May;68(3):474-479. [doi: [10.15288/jsad.2007.68.474](https://doi.org/10.15288/jsad.2007.68.474)] [Medline: [17446988](https://pubmed.ncbi.nlm.nih.gov/17446988/)]
38. Bohn MJ, Babor TF, Kranzler HR. The Alcohol Use Disorders Identification Test (AUDIT): validation of a screening instrument for use in medical settings. *J Stud Alcohol* 1995 Jul;56(4):423-432. [doi: [10.15288/jsa.1995.56.423](https://doi.org/10.15288/jsa.1995.56.423)] [Medline: [7674678](https://pubmed.ncbi.nlm.nih.gov/7674678/)]
39. Hays RD, Merz JF, Nicholas R. Response burden, reliability, and validity of the CAGE, Short MAST, and AUDIT alcohol screening measures. *Behav Res Methods Instruments Comput* 1995 Jun;27(2):277-280. [doi: [10.3758/bf03204745](https://doi.org/10.3758/bf03204745)]
40. Berman AH, Bergman H, Palmstierna T, Schlyter F. Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *Eur Addict Res* 2005;11(1):22-31. [doi: [10.1159/000081413](https://doi.org/10.1159/000081413)] [Medline: [15608468](https://pubmed.ncbi.nlm.nih.gov/15608468/)]

41. Voluse AC, Gioia CJ, Sobell LC, Dum M, Sobell MB, Simco ER. Psychometric properties of the Drug Use Disorders Identification Test (DUDIT) with substance abusers in outpatient and residential treatment. *Addict Behav* 2012 Jan;37(1):36-41. [doi: [10.1016/j.addbeh.2011.07.030](https://doi.org/10.1016/j.addbeh.2011.07.030)] [Medline: [21937169](https://pubmed.ncbi.nlm.nih.gov/21937169/)]
42. Table 43: Arrests by race and ethnicity. Federal Bureau of Investigation. Washington, DC, USA: U.S. Department of Justice; 2019. URL: <https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/tables/table-43> [accessed 2022-09-21]
43. Jeanine Meyers Y, Janeice Morgan A. Targeted marketing and African American millennial consumers. *J Res Interact Market* 2013;7(1):6-17. [doi: [10.1108/17505931311316716](https://doi.org/10.1108/17505931311316716)]
44. Huskamp HA, Busch AB, Uscher-Pines L, Barnett ML, Riedel L, Mehrotra A. Treatment of opioid use disorder among commercially insured patients in the context of the COVID-19 pandemic. *JAMA* 2020 Dec 15;324(23):2440-2442 [FREE Full text] [doi: [10.1001/jama.2020.21512](https://doi.org/10.1001/jama.2020.21512)] [Medline: [33320214](https://pubmed.ncbi.nlm.nih.gov/33320214/)]
45. Liese BS, Monley CM. Providing addiction services during a pandemic: lessons learned from COVID-19. *J Subst Abuse Treat* 2021 Jan;120:108156 [FREE Full text] [doi: [10.1016/j.jsat.2020.108156](https://doi.org/10.1016/j.jsat.2020.108156)] [Medline: [33036878](https://pubmed.ncbi.nlm.nih.gov/33036878/)]
46. Torous J, Wykes T. Opportunities from the coronavirus disease 2019 pandemic for transforming psychiatric care with telehealth. *JAMA Psychiatry* 2020 Dec 01;77(12):1205-1206. [doi: [10.1001/jamapsychiatry.2020.1640](https://doi.org/10.1001/jamapsychiatry.2020.1640)] [Medline: [32391857](https://pubmed.ncbi.nlm.nih.gov/32391857/)]
47. Frank HE, Grumbach NM, Conrad SM, Wheeler J, Wolff J. Mental health services in primary care: evidence for the feasibility of telehealth during the COVID-19 pandemic. *J Affect Disord Rep* 2021 Jul;5:100146 [FREE Full text] [doi: [10.1016/j.jadr.2021.100146](https://doi.org/10.1016/j.jadr.2021.100146)] [Medline: [33870261](https://pubmed.ncbi.nlm.nih.gov/33870261/)]
48. Oesterle TS, Kolla B, Risma CJ, Breiting SA, Rakocevic DB, Loukianova LL, et al. Substance use disorders and telehealth in the COVID-19 pandemic era: a new outlook. *Mayo Clin Proc* 2020 Dec;95(12):2709-2718 [FREE Full text] [doi: [10.1016/j.mayocp.2020.10.011](https://doi.org/10.1016/j.mayocp.2020.10.011)] [Medline: [33276843](https://pubmed.ncbi.nlm.nih.gov/33276843/)]
49. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: [10.15585/mmwr.mm6932a1](https://doi.org/10.15585/mmwr.mm6932a1)] [Medline: [32790653](https://pubmed.ncbi.nlm.nih.gov/32790653/)]
50. Taxman FS, Perdoni ML, Harrison LD. Drug treatment services for adult offenders: the state of the state. *J Subst Abuse Treat* 2007 Apr;32(3):239-254 [FREE Full text] [doi: [10.1016/j.jsat.2006.12.019](https://doi.org/10.1016/j.jsat.2006.12.019)] [Medline: [17383549](https://pubmed.ncbi.nlm.nih.gov/17383549/)]
51. Distribution of Facebook users worldwide as of January 2022, by age and gender. Statista. URL: <https://www.statista.com/statistics/376128/facebook-global-user-age-distribution/> [accessed 2022-09-05]
52. Distribution of Reddit users in the United States as of February 2016, by age group. Statista. 2016. URL: <https://www.statista.com/statistics/517218/reddit-user-distribution-usa-age/> [accessed 2022-09-05]
53. Mobile fact sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2022-09-12]
54. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. *J Med Internet Res* 2014 Oct 17;16(10):e222 [FREE Full text] [doi: [10.2196/jmir.3442](https://doi.org/10.2196/jmir.3442)] [Medline: [25326646](https://pubmed.ncbi.nlm.nih.gov/25326646/)]
55. Bunnell BE, Barrera JF, Paige SR, Turner D, Welch BM. Acceptability of telemedicine features to promote its uptake in practice: a survey of community telemental health providers. *Int J Environ Res Public Health* 2020 Nov 17;17(22):8525 [FREE Full text] [doi: [10.3390/ijerph17228525](https://doi.org/10.3390/ijerph17228525)] [Medline: [33212979](https://pubmed.ncbi.nlm.nih.gov/33212979/)]
56. Meacham MC, Ramo DE, Prochaska JJ, Maier LJ, Delucchi KL, Kaur M, et al. A Facebook intervention to address cigarette smoking and heavy episodic drinking: a pilot randomized controlled trial. *J Subst Abuse Treat* 2021 Mar;122:108211 [FREE Full text] [doi: [10.1016/j.jsat.2020.108211](https://doi.org/10.1016/j.jsat.2020.108211)] [Medline: [33509414](https://pubmed.ncbi.nlm.nih.gov/33509414/)]
57. Ramo DE, Liu H, Prochaska JJ. A mixed-methods study of young adults' receptivity to using Facebook for smoking cessation: if you build it, will they come? *Am J Health Promot* 2015;29(4):e126-e135 [FREE Full text] [doi: [10.4278/ajhp.130326-QUAL-128](https://doi.org/10.4278/ajhp.130326-QUAL-128)] [Medline: [24575728](https://pubmed.ncbi.nlm.nih.gov/24575728/)]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
DUDIT: Drug Use Disorders Identification Test
REDCap: Research Electronic Data Capture
SUD: substance use disorder

Edited by A Mavragani; submitted 28.02.22; peer-reviewed by W Buente, C García, K Winskell, N Karnik; comments to author 19.08.22; revised version received 30.09.22; accepted 13.10.22; published 02.12.22.

Please cite as:

Harrison A, Folk J, Rodriguez C, Wallace A, Tolou-Shams M

Using Social Media to Engage Justice-Involved Young Adults in Digital Health Interventions for Substance Use: Pilot Feasibility Survey Study

JMIR Form Res 2022;6(12):e37609

URL: <https://formative.jmir.org/2022/12/e37609>

doi: [10.2196/37609](https://doi.org/10.2196/37609)

PMID: [36459404](https://pubmed.ncbi.nlm.nih.gov/36459404/)

©Anna Harrison, Johanna Folk, Christopher Rodriguez, Amanda Wallace, Marina Tolou-Shams. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Continuous Glucose Monitoring to Detect and Intervene on Dietary Restriction in Individuals With Binge Eating: The SenseSupport Withdrawal Design Study

Adrienne S Juarascio^{1,2}, PhD; Paakhi Srivastava¹, PhD; Emily K Presseller^{1,2}, MS; Mandy Lin¹, MSED, MPhil; Anna G G Patarinski¹, BS; Stephanie M Manasse¹, PhD; Evan M Forman^{1,2}, PhD

¹Center for Weight, Eating, and Lifestyle Science, Drexel University, Philadelphia, PA, United States

²Department of Psychological and Brain Sciences, Drexel University, Philadelphia, PA, United States

Corresponding Author:

Adrienne S Juarascio, PhD

Center for Weight, Eating, and Lifestyle Science

Drexel University

3201 Chestnut St.

Philadelphia, PA, 19104

United States

Phone: 1 215 553 7154

Email: asj32@drexel.edu

Abstract

Background: Dietary restraint is a key factor for maintaining engagement in binge eating among individuals with binge eating disorder (BED) and bulimia nervosa (BN). Reducing dietary restraint is a mechanism of change in cognitive behavioral therapy (CBT) for individuals with BN and BED. However, many individuals who undergo CBT fail to adequately reduce dietary restraint during treatment, perhaps owing to difficulty in using treatment skills (eg, regular eating) to reduce dietary restraint during their daily lives. The SenseSupport system, a novel just-in-time, adaptive intervention (JITAI) system that uses continuous glucose monitoring to detect periods of dietary restraint, may improve CBT to reduce dietary restraint during treatment by providing real-time interventions.

Objective: This study aimed to describe the feasibility, acceptability, and initial evaluation of SenseSupport. We presented feasibility, acceptability, target engagement, and initial treatment outcome data from a small trial using an ABAB (A=continuous glucose monitoring data sharing and JITAI-Off, B=continuous glucose monitoring data sharing and JITAI-On) design (in which JITAI were turned on for 2 weeks and then turned off for 2 weeks throughout the treatment).

Methods: Participants (N=30) were individuals with BED or BN engaging in ≥ 3 episodes of ≥ 5 hours without eating per week at baseline. Participants received 12 sessions of CBT and wore continuous glucose monitors to detect eating behaviors and inform the delivery of JITAI. Participants completed 4 assessments and reported eating disorder behaviors, dietary restraint, and barriers to app use weekly throughout treatment.

Results: Retention was high (25/30, 83% after treatment). However, the rates of continuous glucose monitoring data collection were low (67.4% of expected glucose data were collected), and therapists and participants reported frequent app-related issues. Participants reported that the SenseSupport system was comfortable, minimally disruptive, and easy to use. The only form of dietary restraint that decreased significantly more rapidly during JITAI-On periods relative to JITAI-Off periods was the desire for an empty stomach ($t_{43}=1.69$; $P=.049$; Cohen $d=0.25$). There was also a trend toward greater decrease in overall restraint during JITAI-On periods compared with JITAI-Off periods, but these results were not statistically significant ($t_{43}=1.60$; $P=.06$; Cohen $d=0.24$). There was no significant difference in change in the frequency of binge eating during JITAI-On periods compared with JITAI-Off periods ($P=.23$). Participants demonstrated clinically significant, large decreases in binge eating ($t_{24}=10.36$; $P<.001$; Cohen $d=2.07$), compensatory behaviors ($t_{24}=3.40$; $P=.001$; Cohen $d=0.68$), and global eating pathology ($t_{24}=6.25$; $P<.001$; Cohen $d=1.25$) from pre- to posttreatment.

Conclusions: This study describes the successful development and implementation of the first intervention system combining passive continuous glucose monitors and JITAI to augment CBT for binge-spectrum eating disorders. Despite the

lower-than-anticipated collection of glucose data, the high acceptability and promising treatment outcomes suggest that the SenseSupport system warrants additional investigation via future, fully powered clinical trials.

Trial Registration: ClinicalTrials.gov NCT04126694; <https://clinicaltrials.gov/ct2/show/NCT04126694>

(*JMIR Form Res* 2022;6(12):e38479) doi:[10.2196/38479](https://doi.org/10.2196/38479)

KEYWORDS

binge eating; loss-of-control eating; continuous glucose monitoring; mobile phone

Introduction

Background

Binge eating (ie, eating a large amount of food within a discrete time period accompanied by a sense of loss of control overeating) is a key symptom of several eating disorders (EDs), including bulimia nervosa (BN) and binge ED (BED). According to a recent study using a nationally representative sample, as many as 3.7 million Americans will have a lifetime BN or BED diagnosis [1] and as many as 42.2 million Americans will experience clinically significant binge eating [2]. EDs are considered critical public health issues, are associated with significant negative physical and psychosocial consequences [3-10], and place a substantial burden on health care services [11].

Dietary restriction (ie, deliberate attempts to drastically reduce the overall amount of food eaten or the types of food eaten) is a key maintenance factor for binge eating in BN and BED, and reducing dietary restriction is the most well-established mechanism of existing treatments. Dietary restriction increases the vulnerability to binge eating episodes because it leaves patients in a state of physical or psychological deprivation or both [12]. Moreover, reducing dietary restriction is one of the only treatment mechanisms that has been empirically supported [13-19]. Furthermore, the adoption of a regular eating schedule (eg, eating 3 meals and 1 or 2 snacks per day and not going >4 waking hours without eating) is one of the strongest predictors of treatment success for both BN and BED [14,20-22]. Thus, dietary restriction has been identified as an essential clinical target for the treatment of BN and BED.

Cognitive behavioral therapy (CBT), including an enhanced transdiagnostic version, is the current frontline treatment approach for both BN and BED [23-25]. Most CBT manuals focus >50% of the session content on the reduction of dietary restriction and recommend achieving a regular eating schedule before moving on to other treatment content [25-27]. However, many patients with BN and BED continue to engage in restrictive eating behaviors until the end of treatment, suggesting that CBT fails to sufficiently improve this clinical target [28]. For example, one trial of CBT found that <40% of participants achieved regular eating (defined as eating at least three meals and one snack per day) in the month immediately before the posttreatment assessment [20]. In addition, recent systematic reviews and meta-analyses have found that 40% to 50% of patients with BED [29] and nearly 70% of patients with BN [13] remain symptomatic after a full course of CBT, likely owing in part to insufficient amelioration of dietary restriction.

CBT may fail to adequately improve dietary restriction because the typical method of intervention delivery (eg, weekly in-person therapy sessions) limits the ability to intervene during the moments when an intervention may be most needed. The inability of CBT to intervene sufficiently in dietary restriction may be due to 3 main reasons. First, dietary restriction can fluctuate significantly within a day and between days in individuals with BN and BED [30]. For example, within any given day, restrictive eating behaviors can re-emerge following binge eating episodes, as individuals attempt to control weight via restriction [25]. Indeed, research has shown that individuals with BN and BED often engage in *chaotic* eating, ie, profound within- and between-day fluctuations in meal and calorie patterning [31]. Given the within- and between-day fluctuations in restrictive intent and behavior, it is not surprising that weekly treatment approaches may fail to alter this constantly moving target. Second, CBT relies on patients' self-reporting (typically through a review of self-monitoring records that involves tracking eating and ED behaviors) to identify continued engagement in restrictive eating behaviors. Even among individuals without eating pathologies, accurate tracking of dietary intake is notoriously difficult [32-34]. In patients with eating pathology, many individuals are unable or unwilling to accurately record their eating behaviors, including restrictive eating, which limits a CBT clinician's ability to intervene appropriately to prevent dietary restriction. Finally, obtaining adequate compliance with self-monitoring records during CBT can be difficult because of the perceived burden during a long-term treatment [14]. Thus, the limitations in both the quality of data used to guide interventions and the frequency and immediacy with which interventions can be provided may explain why CBT results in inadequate amelioration of dietary restriction.

Objectives

The limitations described above for traditional in-person CBT suggest that patients may need additional real-time support and accountability to facilitate sufficient amelioration of dietary restriction. To overcome the aforementioned limitations inherent in the standard CBT treatment delivery approach to address fluctuations in dietary restriction, we developed an intervention system called SenseSupport to augment standard CBT for binge eating. SenseSupport is a state-of-the-art intervention system that combines passive sensing technologies for continuous and unobtrusive collection of data on eating behaviors and analysis to detect dietary restriction with a just-in-time, adaptive intervention (JITAI) system to intervene in this behavior in real time. SenseSupport has 3 key capabilities.

First, SenseSupport uses continuous glucose monitoring (CGM; ie, continuous and passive measurement of blood glucose levels)

device to detect dietary restriction in real time. CGM can detect the within-day fluctuations in the dietary restriction because meal intake is associated with characteristic patterns of changes in blood glucose levels [35-40]. When individuals are not eating, their blood glucose levels are maintained at remarkably constant levels. In response to a meal, glucose levels fluctuate, with simple carbohydrates producing a quick and distinct rise in glucose and meals heavy in protein, fat, or fiber producing slower, less steep, and longer-lasting increases in glucose [41-43]. SenseSupport transfers the real-time CGM data to a smartphone and analyzes these data using an embedded meal detection algorithm to accurately detect meal consumption and estimate the size and macronutrient content of a meal. The meal detection app is based on the parameter-invariant algorithm developed by Weimer et al [40] in 2016. The parameter-invariant algorithm is *invariant* to individual physiological parameters and therefore achieves near-constant accuracy across the population without individual tuning (as required in previous algorithms).

Second, SenseSupport uses a JITAI system (a smartphone app-based system that uses real-time analysis of data to deliver momentary interventions at identified times of need) that delivers real-time brief CBT-based interventions when dietary restriction and other ED symptoms, including binge eating and purging, are detected. These interventions are designed to provide in-the-moment reminders to patients to eat regularly as they go about their daily lives, thereby augmenting the therapeutic content delivered in-session during CBT. These reminders are hypothesized to improve treatment outcomes as patients often report engaging in habitual dietary restriction, of which they may not be consciously aware, during treatment, and increased awareness of problematic behavior is essential to changing one's behavior. For example, when the meal detection algorithm detects fasting behavior, SenseSupport's JITAI system delivers an intervention encouraging patients to eat regularly throughout the day to prevent future binge episodes or suggesting use of problem-solving skills to address barriers to regular eating. JITAI as augmentation of treatment-as-usual have demonstrated promise for improving outcomes for a variety of mental health conditions, including substance use disorders, schizophrenia, and affective disorders [44].

Third, SenseSupport includes a clinician portal that displays objective data on dietary restriction and ED behaviors from the CGM data to the treating clinician. Clinicians can quickly and easily view patients' eating behaviors and use this information to guide treatment planning and implementation. For example, if clinicians observe continued engagement in dietary restriction between sessions, they may have an in-depth discussion to spur motivation to reduce restrictive eating and encourage the between-session practice of regular eating. The objective and real-time detection of dietary restriction from CGM data also allows for accurate, less burdensome, and more complete data collection during standard CBT for binge eating to further guide therapeutic work during CBT.

In the remainder of this paper, we present data from a small open clinical trial (N=30) in which patients with clinically significant binge eating and restrictive eating received 12 weeks of CBT treatment and completed electronic self-monitoring of

eating and ED behaviors on a smartphone app (the SenseSupport app). The SenseSupport app also collected and analyzed the CGM data using a meal detection algorithm. The app had two additional features that could be turned on and off: (1) sharing CGM data with the study clinicians and (2) a JITAI system to detect and intervene in dietary restriction. We used an ABAB design (A=CGM data sharing and JITAI-Off, B=CGM data sharing and JITAI-On) to test the feasibility, acceptability, and target engagement of SenseSupport when paired with a CBT treatment program. The primary aims of this study were (1) to test the hypothesis that SenseSupport will be a feasible and acceptable system for use during a 12-week CBT treatment protocol; (2) to test the hypothesis that larger decreases in dietary restriction (ie, episodes of fasting for ≥ 5 hours, limiting the overall amount of food consumed, number of days when ≥ 8 hours of fasting was observed, excluding specific foods, following specific dietary rules, and desire for empty stomach) will be observed during JITAI-On phases compared with JITAI-Off phases; (3) to test the hypothesis that CBT augmented by the SenseSupport system will yield large decreases in binge eating, compensatory behaviors, and global eating pathology in pre- to posttreatment results; and (4) to test the hypothesis that larger decreases in binge eating and compensatory behaviors will be observed during JITAI-On phases compared with JITAI-Off phases.

Methods

Participants

Participants were recruited through professional referrals and radio, newspaper, and web-based (social media) advertisements. Advertisements called for individuals who engaged in fasting and binge eating to participate in clinical treatment studies. Participants included adults who met the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition criteria for BED or BN (ie, BED: experienced at least 12 episodes of binge eating in the past 3 months, accompanied by marked distress about binge eating and at least three binge eating features [eating rapidly; eating until excessively full; eating large amounts of food in the absence of hunger; eating alone because of embarrassment; or feeling disgusted, depressed, or very guilty]; BN: experienced at least 12 episodes of binge eating and at least 12 episodes of inappropriate compensatory behavior in the past 3 months, accompanied by excessive influence of body shape and weight on self-evaluation) and had 3 or more episodes of fasting for >5 waking hours per week in the last 4 weeks. Individuals were excluded if they (1) were receiving treatment for an ED or behavioral weight loss, (2) required immediate treatment for medical complications because of the ED, (3) were experiencing other severe psychopathology that would limit the participants' ability to comply with this study (eg, severe depression with suicidal intent), (4) were not stable on psychiatric medications for at least 1 month, (5) had diabetes, (6) were taking a medication known to impact insulin or glucose levels, (7) had a history of bariatric surgery, (8) were pregnant or nursing, or (9) had a BMI <17.5 or >40 .

Procedure

Participants attended a Zoom-based baseline assessment in which they completed the Eating Disorder Examination (EDE) and self-report questionnaires (demographics and BMI) to confirm eligibility. Upon eligibility, participants were sent a Dexcom G6 CGM and a MiFit smart band for tracking activity, heart rate, and sleep. They were asked to trial the devices for 3 days before committing to the study for 12 weeks. The participants met the research coordinator via Zoom videoconference to arrange the trial period. They were instructed on how to download and use the SenseSupport phone app and MiFit smart band app. They were also instructed on how to use the Dexcom G6 CGM system. Participants were told to monitor all instances of food intake, ED behaviors, and mood on the SenseSupport app as close to real time as possible. The trial period occurred immediately following the baseline assessment. After completing the trial period and agreeing to commit to the study obligations, participants were sent additional CGM sensors and transmitters to last the 12 weeks of the study.

After the trial period, the participants began 12 weeks of therapy sessions. Participants completed telehealth treatment sessions with a study therapist once a week. In weeks 1 to 2, participants wore the CGM devices and tracked their eating behaviors and mood in the SenseSupport app but did not receive JITAIs, and therapists did not have access to the clinician portal (eg, an *A* period of the ABAB design). To track eating behaviors, including ED behaviors and mood, participants opened the smartphone app, initiated an entry, and answered a series of questions related to their eating behavior, including the type of eating episode, the context (ie, time and location) of eating episodes, and whether loss of control was experienced and the participant engaged in compensatory behaviors. In the same entry, participants rated their current experience of anxiety or worry, sadness or depression, and other emotions. Throughout the study, participants could view previous entries within the SenseSupport app. During weeks 3 to 4, JITAIs were *turned on* such that patients began to receive push notifications based on CGM data, and therapists had access to the clinician portal and were instructed to review CGM data during their session with patients (eg, a *B* period of the ABAB design). The *A/B* periods were repeated (weeks 5-6, and 9-10 were *A* periods; weeks 7-8, and 11-12 were *B* periods) to test whether the amelioration of dietary restriction was due to the use of the SenseSupport and not simply to the effect of time in treatment. Weekly data collected at each therapy session included the therapists' and participants' weekly questionnaires on issues reported with the CGM or SenseSupport app, a modified version of the EDE Questionnaire (EDE-Q) restraint subscale that asked about the preceding 7-day period (rather than the usual 28-day period), weekly participant-reported episodes of fasting for ≥ 5 hours, and finally, weekly participant-reported ED behaviors (binge episodes and compensatory behaviors).

Assessments of feasibility, acceptability, target engagement, and clinical outcomes were completed between sessions 4 and 5, between sessions 8 and 9, and after session 12 (posttreatment assessment) by the study team. Feasibility data included CGM data, weekly therapist and participant questionnaires, and retention calculations. Acceptability data included participant

feedback and questionnaires. Target engagement data included the full EDE, the modified EDE-Q restraint subscale, participant-reported episodes of fasting for ≥ 5 hours, and participant-reported ED behaviors.

Ethics Approval

This study was approved by the Institutional Review Board of Drexel University (protocol number 1907007293). Informed consent was obtained from all the participants before any study procedures.

Measures

Baseline Characteristics

Demographics Questionnaire

The participants were asked to provide information such as age, sex, ethnicity, and socioeconomic status.

BMI Measures

Height and weight data were collected before the treatment was started and at each assessment point.

Feasibility

Retention

To assess feasibility, the percentage of participants enrolled in this study retained at each assessment point was collected. We also examined the percentage of participants who completed the 3-day trial period but declined to continue participation in the study based on this trial.

CGM Data

The Dexcom G6 CGM sensor provided blood glucose readings (in mg/dL) every 5 minutes. Participants wore CGM sensors, which used 0.5-inch flexible wires inserted into the abdomen to collect blood glucose readings. A transmitter attached to the sensor wirelessly transmitted the blood glucose readings to the SenseSupport app portal. The percentage of total expected CGM data that were collected was computed for each participant as a measure of the feasibility of the SenseSupport system.

Weekly Therapist and Participant Questionnaires

Therapists and participants reported problems with the CGM sensors and the SenseSupport app in web-based questionnaires at each therapy session.

Acceptability

Participant Feedback Questionnaire

Participants completed a feedback questionnaire at session 12, which was used to obtain qualitative ratings regarding the acceptability of and compliance with the SenseSupport system. The feedback questionnaire included questions about the comfort, pain, disruption of daily life, ease of use, and helpfulness of the SenseSupport system.

Participant Feedback Interviews

Participants completed interviews after sessions 4, 8, and 12 to provide qualitative feedback regarding their experiences with the SenseSupport system.

Target Engagement

EDE-Q Restraint Subscale

The EDE-Q is a self-report version of the EDE. The restraint subscale consists of 5 items that measure dietary restraint, including overall dietary restraint, avoidance of eating for ≥ 8 hours, desire for an empty stomach, food avoidance, and dietary rules. Participants completed each of these items, which were modified to ask about the past 7 days (rather than the past 28 days) at each therapy session. These items were examined as measures of dietary restraint (ie, attempts to limit food consumption).

Weekly Participant-Reported Episodes Fasting for ≥ 5 Hours

Participants reported the frequency with which they went ≥ 5 hours without eating anything over the past 7 days at each therapy session. This was examined as a measure of dietary restriction (ie, the actual limitation of food consumption).

Treatment Outcomes

EDE-Measured ED Symptoms

The EDE [12] is a semistructured clinician-administered interview that includes 4 subscales (restraint, eating concern, weight concern, and shape concern). Treatment outcomes were assessed through reductions in loss-of-control episodes, compensatory behaviors, and global EDE severity scores. The EDE global score can range from 1 to 5, with a higher number indicating a more severe pathology. The score was calculated by summing the subscales and dividing the total score by 4. A score of 4 or higher was considered clinically significant. A shortened version of the interview was conducted between sessions 4 and 5 and sessions 8 and 9, while the full interview was administered at baseline and after session 12. The shortened interview ended after the compensatory behaviors section of the EDE (the shape and weight concerns sections were excluded).

Weekly Participant-Reported ED Behaviors

At each therapy session, participants reported the number of episodes of binge eating and compensatory behaviors in which they engaged over the previous 7 days (eg, “In the past 7 days, how many binge episodes did you have?” and “In the past 7 days, how many times did you vomit to compensate for a binge episode?”).

Statistical Analyses

Feasibility

Feasibility was characterized using percent retention during treatment. We considered retention $\geq 80\%$ following the 3-day

trial period and at each assessment point to be adequate feasibility. The percentage of data obtained from the CGM sensors was also quantified, with $\geq 80\%$ of the expected data being considered adequate. Finally, the issues reported with the SenseSupport system by therapists and patients were also summarized.

Acceptability

Acceptability was characterized using patient-reported acceptability ratings at session 12, with $>80\%$ ratings of 4 or 5 (on 1 to 5 scales wherein higher scores indicate greater acceptability) and with $>80\%$ ratings of 1 or 2 (on 1 to 5 scales wherein lower scores indicate greater acceptability) being considered acceptable.

Target Engagement

Target engagement variables included participant-reported episodes of fasting for ≥ 5 hours, and responses to the EDE-Q restraint subscale items measured each week of treatment. The slope of change in target engagement variables (weekly episodes of fasting for ≥ 5 hours, days of limiting overall amount consumed, number of days when ≥ 8 hours of fasting was observed, days excluding foods from diet, days following specific dietary rules, and days with desire for empty stomach) was computed for each *A* period and *B* period. Consecutive *A* periods and *B* periods were paired and compared using 1-tailed paired samples *t* tests for each variable of interest (ie, the period from session to 1-3 was compared with the period from session 3-5; the period from session 5-7 was compared with the period from session 7-9).

Treatment Outcomes

Treatment outcomes (past month binge episodes; compensatory behaviors; and EDE global and subscale scores, which was measured by the EDE) were assessed by comparing pre- and posttreatment measures using paired sample 1-tailed *t* tests. In addition, slopes of change in treatment outcome variables (weekly binge episodes and weekly compensatory behaviors) were computed for each *A* period and *B* period. Consecutive *A* and *B* periods were paired and then compared via paired samples *t* tests.

Results

Overview

Demographic information is shown in Table 1. The feasibility, acceptability, treatment outcomes, and target engagement results are summarized in Table 2, Table 3, Table 4, and Table 5, respectively.

Table 1. Participant demographics (N=30).

	Participants
Gender, n (%)	
Man	3 (10)
Woman	26 (87)
Genderqueer or gender nonconforming	1 (3)
Race, n (%)	
White	27 (90)
Black or African American	3 (10)
Asian	2 (7)
More than one race	2 (7)
Ethnicity, n (%)	
Non-Hispanic or Latino	28 (93)
Hispanic or Latino	2 (7)
Baseline diagnosis, n (%)	
BN ^a	20 (67)
BED ^b	10 (33)
Employment status^c, n (%)	
Full time	17 (57)
Part time	6 (20)
Full-time student	4 (13)
Part-time student	1 (3)
Disability or social security	1 (3)
No income	6 (20)
Relationship status, n (%)	
Married	8 (27)
Divorced	4 (13)
Single	12 (40)
Living with partner, but not married	3 (10)
In relationship but not living with partner	2 (7)
Widowed	1 (3)
Household income (US \$), n (%)	
0-10,000	1 (3)
10,000-24,999	1 (3)
25,000-34,999	1 (3)
35,000-49,999	6 (20)
50,000-74,999	6 (20)
75,000-99,999	3 (10)
≥100,000	10 (33)
Unknown	2 (7)
Age (years), mean (SD)	37.10 (12.19)
BMI (kg/m ²), mean (SD)	29.51 (5.20)

^aBN: bulimia nervosa.

^bBED: binge eating disorder.

^cFor employment status, if a participant indicated that they were a student while also working, both responses would be reflected in the data above.

Table 2. Feasibility of SenseSupport.

	Values
Retention (n=30), n (%)	
Time of assessment	
Week 4	29 (96.7)
Week 8	27 (90)
Posttreatment	25 (83.3)
Reported problems with SenseSupport app (n=333 therapy sessions), n (%)	
By participants	140 (42)
By therapists	48 (14.4)
Reported problems with CGM^a sensors (n=333 therapy sessions), n (%)	
By participants	16 (4.8)
By therapists	7 (2.1)
Compliance with CGM sensors (%), mean (SD; range)	67.4 (18.3; 16.8-92.2)

^aCGM: continuous glucose monitoring.

Table 3. Acceptability of SenseSupport (n=24).

	Values, mean (SD; range)	Value, n (%)
In the past 4 weeks, on average, how comfortable did you feel wearing the CGM ^a sensor? (1=Completely uncomfortable to 5=Completely comfortable)	4.50 (0.93; 2-5)	19 (79) rated 4 or 5
In the past 4 weeks, on average, how painful was the insertion of the CGM sensor? (1=Not at all painful to 5=Extremely painful)	1.38 (0.71; 1-4)	23 (96) rated 1 or 2
How much did the CGM sensor disrupt your daily life? (1=It was not at all disruptive to 5=It was extremely disruptive)	1.33 (0.48; 1-2)	24 (100) rated 1 or 2
In the past 4 weeks, on average, how easy was it to learn how to navigate the full SenseSupport system? (1=Not at all easy to learn to 5=Extremely easy to learn)	4.58 (0.65; 3-5)	22 (92) rated 4 or 5
How helpful were the push notifications? (1=Not at all helpful to 5=Extremely helpful)	4.74 (0.45; 4-5)	24 (100) rated 4 or 5
How much did you like using the SenseSupport system with your therapist? (1=Not at all to 5=Extremely; n=6)	4.00 (0.63; 3-5)	6 (83) rated 4 or 5
Overall, how likely would you be to use the SenseSupport system in the future as part of treatment for an eating disorder? (1=Not at all likely to 5=Extremely likely)	4.08 (0.83; 2-5)	19 (79) rated 4 or 5
Overall, how likely would you be to recommend the SenseSupport system in the future as part of treatment for an eating disorder? (1=Not at all likely to recommend to 5=Extremely likely to recommend)	4.21 (1.02; 1-5)	19 (79) rated 4 or 5

^aCGM: continuous glucose monitoring.

Table 4. Treatment outcomes of SenseSupport^a.

	Baseline, mean (SD)	Posttreatment, mean (SD)	Slope of change during data sharing- and JITAIs ^b -On weeks, mean (SD)	Slope of change during data sharing- and JITAIs-Off weeks, mean (SD)	<i>t</i> test (<i>df</i>)	<i>P</i> value ^c	Cohen <i>d</i>
Past month total binge episodes	23.88 (11.23)	4.60 (6.25)	N/A ^d	N/A	10.36 (24)	<.001	2.07
Past month total compensatory behaviors	15.96 (19.41)	3.60 (14.87)	N/A	N/A	3.40 (24)	.001	0.68
EDE ^e restraint subscale score	2.37 (1.38)	0.93 (1.40)	N/A	N/A	4.34 (24)	<.001	0.87
EDE eating concern subscale score	1.97 (1.10)	0.97 (1.20)	N/A	N/A	3.43 (24)	.001	0.69
EDE weight concern subscale score	3.75 (1.01)	2.08 (1.22)	N/A	N/A	6.14 (24)	<.001	1.23
EDE shape concern subscale score	4.22 (0.87)	2.79 (1.14)	N/A	N/A	5.83 (24)	<.001	1.17
EDE global score	3.08 (0.76)	1.69 (1.10)	N/A	N/A	6.25 (24)	<.001	1.25
Weekly binge eating episodes	N/A	N/A	0.06 (.11)	0.03 (.15)	0.73 (24)	.23	0.11
Weekly total compensatory behaviors	N/A	N/A	-.01 (.16)	0.06 (.23)	1.78 (24)	.04	0.25

^aPositive values for change in eating disorder behaviors indicate decreases in the frequency of behavior during the period of interest.

^bJITAI: just-in-time, adaptive intervention.

^cItalicized *P* values indicate statistical significance.

^dN/A: not applicable.

^eEDE: Eating Disorder Examination.

Table 5. Target engagement of SenseSupport^a.

	Slope of change during data sharing- and JITAIs ^b -On weeks, mean (SD)	Slope of change during data sharing- and JITAIs-Off weeks, mean (SD)	<i>t</i> test (<i>df</i>)	<i>P</i> value ^c	Cohen <i>d</i>
Weekly episodes of fasting for ≥5 hours	0.04 (0.14)	0.07 (0.20)	-0.82 (44)	.21	-0.13
On how many days have you been deliberately trying to limit the amount of food you eat to influence your shape or weight (whether or not you have succeeded)?	0.03 (0.08)	-.01 (0.10)	1.60 (44)	.06	0.24
On how many days have you gone for long periods (8 waking hours or more) without eating anything at all to influence your shape or weight?	0.01 (0.06)	0.02 (0.08)	-1.05 (44)	.15	-0.16
On how many days have you tried to exclude from your diet any foods that you like to influence your shape or weight (whether or not you have succeeded)?	-0.003 (0.11)	.02 (0.11)	-0.84 (44)	.20	-0.13
On how many days have you tried to follow definite rules regarding your eating (eg, a calorie limit) to influence your shape or weight (whether or not you have succeeded)?	0.02 (0.11)	0.01 (0.08)	0.77 (44)	.22	0.12
On how many days have you had a definite desire to have an empty stomach with the aim of influencing your shape or weight?	0.04 (0.09)	-0.002 (0.08)	1.69 (44)	.049	0.25

^aPositive values for change in eating disorder behaviors indicate decreases in the frequency of behavior during the period of interest.

^bJITAI: just-in-time, adaptive intervention.

^cItalicized *P* values indicate statistical significance.

Feasibility

The participants' retention was high in this study. Only 3% (1/30) of participants completed the 3-day trial period and then declined to participate in the study (constituting 3.2% of the sample who completed the trial period). Furthermore, 83% (25/30) of the participants completed all treatment sessions and assessments. The mean percentage of expected CGM data that were collected was 67.4%, and the range was broad (16.8%-92.2%). Notably, only 10 participants (33% of the sample) met the threshold for the successful collection of >80% of the expected CGM data. Participants (28/30, 93%) and therapists (7/8, 88%) also frequently reported app-related problems, including frequent disconnections (reported by patients at 123 sessions and therapists at 35 sessions), nondelivery of JITAIs when expected (reported by patients at 5 sessions and therapists at 5 sessions), or delivery of JITAIs without the occurrence of ≥ 5 hours without eating (reported by a patient at 1 session), slow loading of app content (reported by a patient at 1 session), unspecified bugs (reported by patients at 10 sessions and therapists at 4 sessions), difficulty connecting new sensors to the app (reported by patients at 3 sessions and a therapist at 1 session), and failure to save electronic self-monitoring records (reported by patients at 5 sessions and therapists at 5 sessions). Participants (11/30, 37%) and therapists (2/8, 25%) also reported issues with the sensors, including sensor insertion being more uncomfortable than usual (reported by patients at 2 sessions), location of the sensor relative to clothing impacting data collection (reported by a patient at 1 session and a therapist at 1 session), sensors falling off (reported by a patient at 1 session), and nonspecific sensor or transmitter failures (reported by patients at 11 sessions and by therapists at 3 sessions).

Acceptability

Participants rated the use of the SenseSupport system as highly acceptable. Of the 24 participants with complete acceptability

data, 19 (79%) and 24 (100%) participants rated wearing the CGM sensors as comfortable and minimally disruptive to their lives, respectively. More than 95% (23/24) of the participants rated the insertion of the CGM sensors as minimally painful. The participants also reported that the SenseSupport system was easy to learn (22/24, 92%) and useful (24/24, 100%). Approximately 83% (5/6; this question had substantial missing data) of participants also reported that they liked using the system, and 79% (19/24) reported that they would use the system or would recommend the system to others as part of ED treatment. Participants reported that they enjoyed the supportive accountability provided by their clinician's ability to view their glucose levels, the reminders to eat regularly when JITAIs were turned on, and the opportunity to observe how their bodies processed meals and snacks they had consumed (refer to [Textbox 1](#) for more details). The established benchmarks for acceptability were met (>80% ratings of >4 or 5 and >80% ratings of <1 or 2, depending on the question) for insertion pain, disruptiveness, ease of learning the system, helpfulness of push notifications, and enjoyment of using the system (see [Table 3](#) for mean ratings). However, the a priori benchmarks were not met for the likelihood of using the system in future ED treatment or the likelihood of recommending the system to others for use during ED treatment.

Qualitative participant feedback particularly emphasized the ease of using the SenseSupport system and CGM sensors, the unobtrusiveness of the CGM sensors, the benefits of supportive accountability facilitated by observable patterns in blood glucose associated with meals, the helpful push notifications reminding them to eat regularly, and an appreciation of learning about their physiological responses to eating ([Textbox 1](#)). Several participants endorsed that they wished that they had received more push notifications to remind them to eat regularly.

Textbox 1. Qualitative participant feedback on continuous glucose monitoring (CGM) sensors and SenseSupport system.

Unobtrusiveness of CGM sensors

- “I like that [the CGM] is so small so you can’t really see it through my clothing.”
- “I didn’t really notice [the CGM] because it was small.”
- “There wasn’t any pain. [The CGM] didn’t interfere with activities.”
- “[I liked that the CGM] is relatively discrete, it doesn’t come off, it sticks for the most part. The insertion is easy.”
- “[The CGM] is very easy to put on and forget it’s there.”

Ease of use of SenseSupport system

- “I thought [SenseSupport] was pretty easy to use.”
- “It was easy to use and understand.”
- “It was pretty easy to navigate.”
- “I thought that [the SenseSupport system] was really easy to use and I found it really helpful.”
- “[The SenseSupport system] was super simple.”

Supportive accountability

- “It helped keep me accountable to what I was recording. There was data to match up with what I was recording.”
- “It made me more likely to stick to regular eating and avoid binging because I knew what the glucose levels would show up as and I wanted that to look good.”
- “There was no lying with the glucose monitoring. I knew if it was going in my mouth it was going on paper.”
- “[The system] gives you accountability because your therapist is also using your portal to monitor your treatment.”
- “[The system] helps you see your eating patterns and just create better eating habits. It helps you 24 hours of the day. There’s always an app helping you. It’s the hardest afterwards now there’s nobody watching. Going from 100% to 0% accountability post-treatment is hard.”
- “I liked that [my therapist] could see what I was eating and noticed changes I couldn’t notice. The glucose graphs were really interesting.”
- “[Using SenseSupport] was very productive and very positive. [I liked] the external structure and accountability that came with the system as a whole.”
- “I liked the involved process and self-checks to have accountability with and without my therapist. [The SenseSupport system] made me check myself to eat regularly. [There was] a lot more consistency with all the parts operating together.”

Push notifications as reminders to eat regularly

- “[The push notifications] were encouraging in the moment.”
- “[The notifications] would come when I didn’t eat frequently. They were helpful. I wish there were more.”
- “If [the SenseSupport system] felt like I had loss of control or overeating, it helped me get back to the structure that we had been on previously.”
- “I thought [the reminder push notifications] were really helpful.”
- “[The notifications] would remind me when I wasn’t eating. That was good.”

Learning about physiological responses to food

- “It was helpful to be able to record my food and have [my therapist] tell me that what I was eating wasn’t adequate according to my blood glucose levels and whether or not I should continue eating like that.”
- “[The system] shows you how your body responds to things and made me want to eat regularly so I can elicit a specific response from my body...I believe my metabolism is all messed up. I spent a lot of time fasting and not eating but seeing how my body responds to food made me more comfortable with food.”
- “I really liked having the specific information about checking glucose levels and comparing what I was eating and when.”
- “[SenseSupport] was helpful because it helped me to keep consistently eating throughout the day and making sure to eat enough at each meal. It helped me align my hunger cues more and organize, get into a routine.”
- “[SenseSupport] helped keep me accountable and connect what I was eating to my body.”
- “It helped to understand the importance of regular eating and what a normal glucose pattern should look like versus what my pattern looked like.”
- “I really enjoyed going over the glucose data with [my therapist]. Seeing all of that really helped me want to regulate my body. [It] made me want to eat for my body’s wellbeing.”

Target Engagement

The frequency of desiring an empty stomach to influence shape or weight decreased significantly more rapidly during JITAI-On weeks compared with JITAI-Off weeks with a small effect size. There was also a trend-level effect for the frequency of deliberately trying to limit the amount of food consumed to influence shape or weight, with greater decreases during periods during the JITAI-On weeks. There were no significant differences among episodes of fasting for ≥ 5 hours, going long periods without eating, excluding foods from diet, or following definite rules regarding eating by notification delivery.

Treatment Outcomes

Participants showed significant decreases in the frequency of binge eating and compensatory behaviors, and in the severity of eating pathology from pre- to posttreatment with medium (compensatory behaviors and EDE eating concern subscale scores) and large (binge eating episodes; EDE restraint, weight concern, and shape concern subscale scores; and EDE global scores) effect sizes. Contrary to the hypotheses, the change in the frequency of weekly compensatory behaviors was significantly more rapid during JITAI-Off periods compared with JITAI-On. Changes in the frequency of weekly binge eating episodes were not significantly different during JITAI-On weeks compared with JITAI-Off weeks.

Discussion

This study is the first to evaluate a novel intervention system that uses CGM sensors to augment CBT for binge eating to improve dietary restriction and treatment outcomes.

Principal Findings

The high retention rates for the study assessments and sessions demonstrated the feasibility of the SenseSupport system. However, the benchmarks for successful data collection from CGM were not met. The most notable barrier to achieving this benchmark was the connectivity issues between the CGM sensors and the SenseSupport app owing to technical problems in the software development kit provided by Dexcom during the study. Because the publicly available application programming interface for Dexcom devices has a built-in 3-hour time delay and thus would not provide data quickly enough for use in the current system, we relied on the donated software development kit for this initial proof-of-concept development phase. These findings may suggest that although it may be feasible to use SenseSupport as an augmentation to in-person CBT, future development work is needed to ensure that technological problems owing to Bluetooth connection dropping between the CGM sensor and the SenseSupport system are resolved before additional testing is indicated. Future iterations of SenseSupport may involve either using newer-generation CGM sensors that may better integrate with the SenseSupport app or enabling the software development kit from Dexcom to support more robust data collection.

Despite the concerns regarding missing CGM data, SenseSupport demonstrated high acceptability and achieved nearly all a priori benchmarks for the acceptability of the system. Qualitatively, the patients indicated that the CGM sensors were

comfortable to wear and minimally disruptive to their lives. A majority of patients perceived SenseSupport to be useful in learning about patterns in blood glucose associated with meals and reminding them to eat regularly, and they expressed a desire to receive more frequent JITAI during treatment. SenseSupport fell marginally short of meeting the benchmarks for the likelihood of using or recommending the system to others as part of the ED treatments. When probed further, this appeared to be largely due to the burdensome nature of manually reconnecting the sensor and app after the Bluetooth connections failed, as discussed earlier. These interruptions may have reduced the perceived value of SenseSupport to augment ED treatments despite the participants rating the overall SenseSupport system favorably.

Compared with JITAI-Off weeks, JITAI-On weeks had small effect size improvements in desire for empty stomach (ie, intent that maintains dietary restriction) and reductions in the frequency of deliberately trying to limit the amount of food consumed (ie, actual restriction) to influence shape or weight at the trend-level, suggesting the possible benefit of the SenseSupport system in reducing both intent and actual dietary restriction. However, contrary to the hypotheses, the observed differences among episodes of fasting for ≥ 5 hours, going long periods without eating, excluding foods from diet, or following definite rules regarding eating during JITAI-On compared with JITAI-Off weeks were small. One possible reason for the small effect size differences between *A* and *B* periods may be that during the full duration of the study, participants were receiving CBT treatment with a specific focus on reducing dietary restriction. In other words, it may be plausible that the changes through CBT treatment were so notable that there was little opportunity for the JITAI to have a significant impact on dietary restriction. In addition, in *A* and *B* periods, participants continued to wear the sensors and track eating behavior and ED symptoms in the SenseSupport app. Thus, it is possible that the high levels of overlap between the JITAI-On and JITAI-Off weeks made it difficult to observe differences between the conditions. In addition, a common challenge of ABAB study designs is that carryover effects may be observed that extend beyond the periods being measured. For example, in this study, it is possible that JITAI delivered in *B* weeks led to general amelioration of dietary restriction that persisted into the next *A* period. Future research using alternative study designs is needed to determine the additive value of the SenseSupport system above and beyond CBT.

We observed medium effect size improvements in the frequency of compensatory episodes and large effect size improvements in binge eating frequency and cognitive symptoms of ED pathology in pre- to posttreatment results. However, contrary to the hypotheses, the reduction in the frequency of binge eating during JITAI-On weeks was only marginally higher than the improvements in binge eating frequency during JITAI-Off weeks, perhaps owing to many of the limitations described for detecting changes in our target engagement variables as a result of the ABAB design described earlier.

Comparison With Prior Work

The medium effect size improvements observed in the frequency of compensatory episodes and large effect size improvements in binge eating frequency and cognitive symptoms of ED pathology in pre- to posttreatment results are notable when compared with those reported in meta-analyses and systematic reviews testing the efficacy of CBT for BN and BED (ie, an average small effect size improvement in compensatory behaviors and an average medium size improvement in cognitive symptoms) [45]. The strong outcomes observed over the full duration of treatment suggest the promise of the SenseSupport system as an augmentation of CBT for EDs characterized by binge eating and indicate that additional research is warranted. Although these results are preliminary and replication is needed in large-scale clinical trials, the SenseSupport system may also be a useful augmentation for other EDs, such as anorexia nervosa, where individuals are at risk of overreporting their eating episodes and could benefit from receiving a treatment that uses the SenseSupport system to collect objective data on dietary restriction and ED behaviors to guide treatment planning and implementation. In addition, SenseSupport system could also augment treatments for comorbid medical conditions with binge eating that are associated with high glucose variability. For example, the SenseSupport system could provide objective data on blood glucose levels in individuals with comorbid type 2 diabetes mellitus and binge eating to guide their eating-related decisions on a day-to-day basis that may further help stabilize blood sugar levels.

Limitations

This study has several limitations. First, we used an ABAB design, which may have introduced the carryover effects of JITAI-On weeks to JITAI-Off weeks. Thus, this design substantially reduces our confidence in the causal effect of JITAI in improving dietary restriction or behavioral clinical outcomes such as binge eating and compensatory behaviors. A randomized controlled trial is the logical next step in determining whether SenseSupport is the causal factor in improving dietary restriction and eating pathology. Second, it is possible that the participants' eating decisions were influenced

merely by wearing CGMs, in addition to those guided by the SenseSupport app and interventions, as reactivity to these devices has been documented in other populations [46]. Our inability to clearly delineate the effects of the CGM sensors from those of SenseSupport interventions is a limitation of the study although the effects of wearing CGMs could partially explain the changes in eating behavior observed during JITAI-Off periods. Similarly, the participants were compensated for the completion of the assessments and received free treatment by participating in the study. These incentives may have influenced the participants' engagement in the study, thereby affecting the feasibility data collected. Third, the study did not include a follow-up assessment that precluded the knowledge of the effect of SenseSupport-augmented CBT in maintaining long-term gains in dietary restriction and eating pathology. Future research would benefit from assessing the long-term effects of SenseSupport-augmented CBT for binge eating. Fourth, in this study, the treatment included multiple intervention components (eg, digital self-monitoring, CGM data sharing with study clinicians, and the JITAI system), which precluded the understanding of the unique contribution of CGM data sharing with clinicians and the JITAI system as an augmentation of CBT treatment. Future research should attempt to isolate these different technological intervention components that may impact dietary restriction and clinical outcomes, as well as test the use of JITAI to deliver more complete CBT interventions as an augmentation of CBT. Fifth, our methods preclude determining the percentage of missing CGM data that was due to CGM app malfunctioning. Sixth, similar to the previous ED studies, most patients were White women. Future research should attempt to replicate our findings in samples that are more diverse in race, ethnicity, and gender.

Conclusions

In summary, this study successfully developed and deployed the first ever intervention system combining passive CGM sensors and the JITAI system as an augmentation of CBT to improve dietary restriction and clinical outcomes in EDs characterized by binge eating. Our findings suggest that the SenseSupport system is worthy of additional study in future through fully powered clinical trials.

Acknowledgments

This study was funded by a grant from the National Institute of Health awarded to ASJ (R43MH121205).

Data Availability

Data from this project will be available upon reasonable request.

Conflicts of Interest

EMF is a member of the scientific advisory board for Nutrisystem. The authors have no further conflicts to declare.

References

1. Udo T, Grilo CM. Prevalence and correlates of DSM-5-defined eating disorders in a nationally representative sample of U.S. Adults. *Biol Psychiatry* 2018 Sep 01;84(5):345-354 [FREE Full text] [doi: [10.1016/j.biopsych.2018.03.014](https://doi.org/10.1016/j.biopsych.2018.03.014)] [Medline: [29859631](https://pubmed.ncbi.nlm.nih.gov/29859631/)]

2. Mitchison D, Touyz S, González-Chica DA, Stocks N, Hay P. How abnormal is binge eating? 18-Year time trends in population prevalence and burden. *Acta Psychiatr Scand* 2017 Aug;136(2):147-155. [doi: [10.1111/acps.12735](https://doi.org/10.1111/acps.12735)] [Medline: [28419425](https://pubmed.ncbi.nlm.nih.gov/28419425/)]
3. Wilfley D, Wilson GT, Agras WS. The clinical significance of binge eating disorder. *Int J Eat Disord* 2003;34 Suppl:S96-106. [doi: [10.1002/eat.10209](https://doi.org/10.1002/eat.10209)] [Medline: [12900990](https://pubmed.ncbi.nlm.nih.gov/12900990/)]
4. Kessler RC, Berglund PA, Chiu WT, Deitz AC, Hudson JI, Shahly V, et al. The prevalence and correlates of binge eating disorder in the World Health Organization World Mental Health Surveys. *Biol Psychiatry* 2013 May 01;73(9):904-914 [FREE Full text] [doi: [10.1016/j.biopsych.2012.11.020](https://doi.org/10.1016/j.biopsych.2012.11.020)] [Medline: [23290497](https://pubmed.ncbi.nlm.nih.gov/23290497/)]
5. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007 Feb 01;61(3):348-358 [FREE Full text] [doi: [10.1016/j.biopsych.2006.03.040](https://doi.org/10.1016/j.biopsych.2006.03.040)] [Medline: [16815322](https://pubmed.ncbi.nlm.nih.gov/16815322/)]
6. Mehler PS, Rylander M. Bulimia Nervosa - medical complications. *J Eat Disord* 2015;3:12 [FREE Full text] [doi: [10.1186/s40337-015-0044-4](https://doi.org/10.1186/s40337-015-0044-4)] [Medline: [25866627](https://pubmed.ncbi.nlm.nih.gov/25866627/)]
7. Mitchell JE, Crow S. Medical complications of anorexia nervosa and bulimia nervosa. *Curr Opin Psychiatry* 2006 Jul;19(4):438-443. [doi: [10.1097/01.yco.0000228768.79097.3e](https://doi.org/10.1097/01.yco.0000228768.79097.3e)] [Medline: [16721178](https://pubmed.ncbi.nlm.nih.gov/16721178/)]
8. American Psychiatric Association. *Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition*. Virginia, United States: American Psychiatric Association; 2013.
9. Bulik CM. Eating disorders in adolescents and young adults. *Child Adolesc Psychiatr Clin N Am* 2002 Apr;11(2):201-218. [doi: [10.1016/s1056-4993\(01\)00004-9](https://doi.org/10.1016/s1056-4993(01)00004-9)] [Medline: [12109318](https://pubmed.ncbi.nlm.nih.gov/12109318/)]
10. Wonderlich S, Mitchell JE. Eating disorders and comorbidity: empirical, conceptual, and clinical implications. *Psychopharmacol Bull* 1997;33(3):381-390. [Medline: [9550882](https://pubmed.ncbi.nlm.nih.gov/9550882/)]
11. Roberto CA, Grilo CM, Masheb RM, White MA. Binge eating, purging, or both: eating disorder psychopathology findings from an internet community survey. *Int J Eat Disord* 2010 Dec 27;43(8):724-731 [FREE Full text] [doi: [10.1002/eat.20770](https://doi.org/10.1002/eat.20770)] [Medline: [19862702](https://pubmed.ncbi.nlm.nih.gov/19862702/)]
12. Murphy R, Straebl S, Cooper Z, Fairburn CG. Cognitive behavioral therapy for eating disorders. *Psychiatr Clin North Am* 2010 Sep;33(3):611-627 [FREE Full text] [doi: [10.1016/j.psc.2010.04.004](https://doi.org/10.1016/j.psc.2010.04.004)] [Medline: [20599136](https://pubmed.ncbi.nlm.nih.gov/20599136/)]
13. Linardon J. Meta-analysis of the effects of cognitive-behavioral therapy on the core eating disorder maintaining mechanisms: implications for mechanisms of therapeutic change. *Cogn Behav Ther* 2018 Mar 30;47(2):107-125. [doi: [10.1080/16506073.2018.1427785](https://doi.org/10.1080/16506073.2018.1427785)] [Medline: [29378481](https://pubmed.ncbi.nlm.nih.gov/29378481/)]
14. Barakat S, Maguire S, Surgenor L, Donnelly B, Miceska B, Fromholtz K, et al. The role of regular eating and self-monitoring in the treatment of bulimia nervosa: a pilot study of an online guided self-help CBT program. *Behav Sci (Basel)* 2017 Jun 26;7(3):39 [FREE Full text] [doi: [10.3390/bs7030039](https://doi.org/10.3390/bs7030039)] [Medline: [28672851](https://pubmed.ncbi.nlm.nih.gov/28672851/)]
15. Wilson GT, Fairburn CC, Agras WS, Walsh BT, Kraemer H. Cognitive-behavioral therapy for bulimia nervosa: time course and mechanisms of change. *J Consult Clin Psychol* 2002;70(2):267-274. [doi: [10.1037/0022-006x.70.2.267](https://doi.org/10.1037/0022-006x.70.2.267)]
16. Steel Z, Jones J, Adcock S, Clancy R, Bridgford-West L, Austin J. Why the high rate of dropout from individualized cognitive-behavior therapy for bulimia nervosa? *Int J Eat Disord* 2000 Sep;28(2):209-214. [doi: [10.1002/1098-108x\(200009\)28:2<209::aid-eat10>3.0.co;2-1](https://doi.org/10.1002/1098-108x(200009)28:2<209::aid-eat10>3.0.co;2-1)]
17. Towell D, Woodford S, Reid S, Rooney B, Towell A. Compliance and outcome in treatment-resistant anorexia and bulimia: a retrospective study. *Br J Clin Psychol* 2001 Jun;40(2):189-195. [doi: [10.1348/014466501163634](https://doi.org/10.1348/014466501163634)] [Medline: [11446240](https://pubmed.ncbi.nlm.nih.gov/11446240/)]
18. Troop N, Schmidt U, Tiller J, Todd G, Keilen M, Treasure J. Compliance with a self-care manual for bulimia nervosa: predictors and outcome. *Br J Clin Psychol* 1996 Sep;35(3):435-438. [doi: [10.1111/j.2044-8260.1996.tb01197.x](https://doi.org/10.1111/j.2044-8260.1996.tb01197.x)] [Medline: [8889084](https://pubmed.ncbi.nlm.nih.gov/8889084/)]
19. Thiels C, Schmidt U, Troop N, Treasure J, Garthe R. Compliance with a self-care manual in guided self-change for bulimia nervosa. *Eur Eat Disorders Rev* 2001 Jul;9(4):255-262. [doi: [10.1002/erv.395](https://doi.org/10.1002/erv.395)]
20. Shah N, Passi V, Bryson S, Agras WS. Patterns of eating and abstinence in women treated for bulimia nervosa. *Int J Eat Disord* 2005 Dec;38(4):330-334. [doi: [10.1002/eat.20204](https://doi.org/10.1002/eat.20204)] [Medline: [16231339](https://pubmed.ncbi.nlm.nih.gov/16231339/)]
21. Zendegui EA, West JA, Zandberg LJ. Binge eating frequency and regular eating adherence: the role of eating pattern in cognitive behavioral guided self-help. *Eat Behav* 2014 Apr;15(2):241-243. [doi: [10.1016/j.eatbeh.2014.03.002](https://doi.org/10.1016/j.eatbeh.2014.03.002)] [Medline: [24854811](https://pubmed.ncbi.nlm.nih.gov/24854811/)]
22. Ellison JM, Simonich HK, Wonderlich SA, Crosby RD, Cao L, Mitchell JE, et al. Meal patterning in the treatment of bulimia nervosa. *Eat Behav* 2016 Jan;20:39-42 [FREE Full text] [doi: [10.1016/j.eatbeh.2015.11.008](https://doi.org/10.1016/j.eatbeh.2015.11.008)] [Medline: [26630618](https://pubmed.ncbi.nlm.nih.gov/26630618/)]
23. Wilson GT. Psychological treatment of eating disorders. *Annu Rev Clin Psychol* 2005 Apr 01;1(1):439-465. [doi: [10.1146/annurev.clinpsy.1.102803.144250](https://doi.org/10.1146/annurev.clinpsy.1.102803.144250)] [Medline: [17716095](https://pubmed.ncbi.nlm.nih.gov/17716095/)]
24. National Guideline Alliance (UK). *Eating Disorders: Recognition and Treatment*. London: National Institute for Health and Care Excellence (NICE); 2017.
25. Fairburn CG. *Cognitive Behavior Therapy and Eating Disorders*. New York, United States: Guilford Publications; 2008.
26. Mitchell J, Devlin M, Zwaan M, Peterson C, Crow S. *Binge-Eating Disorder Clinical Foundations and Treatment*. New York, United States: Guilford Press; 2007.
27. Fairburn C. *Overcoming Binge Eating, First Edition*. New York, United States: Guilford Press; 1995.

28. Hay P, Bacaltchuk J, Stefano S, Kashyap P. Psychological treatments for bulimia nervosa and bingeing. *Cochrane Database Syst Rev* 2009 Oct 07(4):CD000562 [FREE Full text] [doi: [10.1002/14651858.CD000562.pub3](https://doi.org/10.1002/14651858.CD000562.pub3)] [Medline: [19821271](https://pubmed.ncbi.nlm.nih.gov/19821271/)]
29. Peat CM, Berkman ND, Lohr KN, Brownley KA, Bann CM, Cullen K, et al. Comparative effectiveness of treatments for binge-eating disorder: systematic review and network meta-analysis. *Eur Eat Disord Rev* 2017 Sep 03;25(5):317-328. [doi: [10.1002/erv.2517](https://doi.org/10.1002/erv.2517)] [Medline: [28467032](https://pubmed.ncbi.nlm.nih.gov/28467032/)]
30. Shingleton RM, Palfai TP. Technology-delivered adaptations of motivational interviewing for health-related behaviors: a systematic review of the current research. *Patient Educ Couns* 2016 Jan;99(1):17-35 [FREE Full text] [doi: [10.1016/j.pec.2015.08.005](https://doi.org/10.1016/j.pec.2015.08.005)] [Medline: [26298219](https://pubmed.ncbi.nlm.nih.gov/26298219/)]
31. Masheb RM, Grilo CM, White MA. An examination of eating patterns in community women with bulimia nervosa and binge eating disorder. *Int J Eat Disord* 2011 Nov 26;44(7):618-624 [FREE Full text] [doi: [10.1002/eat.20853](https://doi.org/10.1002/eat.20853)] [Medline: [21997425](https://pubmed.ncbi.nlm.nih.gov/21997425/)]
32. Macdiarmid J, Blundell J. Assessing dietary intake: who, what and why of under-reporting. *Nutr Res Rev* 2007 Dec 14;11(2):231-253. [doi: [10.1079/nrr19980017](https://doi.org/10.1079/nrr19980017)]
33. Schoeller D. How accurate is self-reported dietary energy intake? *Nutr Rev* 1990 Oct;48(10):373-379. [doi: [10.1111/j.1753-4887.1990.tb02882.x](https://doi.org/10.1111/j.1753-4887.1990.tb02882.x)] [Medline: [2082216](https://pubmed.ncbi.nlm.nih.gov/2082216/)]
34. CHAMPAGNE CM, BRAY GA, KURTZ AA, MONTEIRO JB, TUCKER E, VOLAUFOVA J, et al. Energy intake and energy expenditure. *J Am Dietetic Assoc* 2002 Oct;102(10):1428-1432. [doi: [10.1016/s0002-8223\(02\)90316-0](https://doi.org/10.1016/s0002-8223(02)90316-0)]
35. Cameron F, Niemeyer G, Buckingham BA. Probabilistic evolving meal detection and estimation of meal total glucose appearance. *J Diabetes Sci Technol* 2009 Sep 01;3(5):1022-1030 [FREE Full text] [doi: [10.1177/193229680900300505](https://doi.org/10.1177/193229680900300505)] [Medline: [20144415](https://pubmed.ncbi.nlm.nih.gov/20144415/)]
36. Chen S, Weimer J, Rickels MR, Peleckis A, Lee I. Towards a model-based meal detector for type I diabetics. In: Proceedings of the 6th Workshop on Medical Cyber-Physical Systems (MedicalCPS 2015). 2015 Presented at: 6th Workshop on Medical Cyber-Physical Systems (MedicalCPS 2015); Apr, 2015; Seattle, WA.
37. Dassau E, Bequette BW, Buckingham BA, Doyle FJ. Detection of a meal using continuous glucose monitoring: implications for an artificial beta-cell. *Diabetes Care* 2008 Feb;31(2):295-300. [doi: [10.2337/dc07-1293](https://doi.org/10.2337/dc07-1293)] [Medline: [17977934](https://pubmed.ncbi.nlm.nih.gov/17977934/)]
38. Harvey RA, Dassau E, Zisser H, Seborg DE, Doyle FJ. Design of the glucose rate increase detector: a meal detection module for the health monitoring system. *J Diabetes Sci Technol* 2014 Mar 13;8(2):307-320 [FREE Full text] [doi: [10.1177/1932296814523881](https://doi.org/10.1177/1932296814523881)] [Medline: [24876583](https://pubmed.ncbi.nlm.nih.gov/24876583/)]
39. Turksoy K, Samadi S, Feng J, Littlejohn E, Quinn L, Cinar A. Meal detection in patients with type 1 diabetes: a new module for the multivariable adaptive artificial pancreas control system. *IEEE J Biomed Health Inform* 2016 Jan;20(1):47-54. [doi: [10.1109/jbhi.2015.2446413](https://doi.org/10.1109/jbhi.2015.2446413)]
40. Weimer J, Chen S, Peleckis A, Rickels MR, Lee I. Physiology-invariant meal detection for type 1 diabetes. *Diabetes Technol Ther* 2016 Oct;18(10):616-624 [FREE Full text] [doi: [10.1089/dia.2015.0266](https://doi.org/10.1089/dia.2015.0266)] [Medline: [27704875](https://pubmed.ncbi.nlm.nih.gov/27704875/)]
41. Smart C, Evans M, O'Connell SM, McElduff P, Lopez PE, Jones TW, et al. Both dietary protein and fat increase postprandial glucose excursions in children with type 1 diabetes, and the effect is additive. *Diabetes Care* 2013 Dec;36(12):3897-3902 [FREE Full text] [doi: [10.2337/dc13-1195](https://doi.org/10.2337/dc13-1195)] [Medline: [24170749](https://pubmed.ncbi.nlm.nih.gov/24170749/)]
42. Pańkowska E, Błazik M, Groele L. Does the fat-protein meal increase postprandial glucose level in type 1 diabetes patients on insulin pump: the conclusion of a randomized study. *Diabetes Technol Ther* 2012 Jan;14(1):16-22. [doi: [10.1089/dia.2011.0083](https://doi.org/10.1089/dia.2011.0083)] [Medline: [22013887](https://pubmed.ncbi.nlm.nih.gov/22013887/)]
43. Holt S, Miller JC, Petocz P. An insulin index of foods: the insulin demand generated by 1000-kJ portions of common foods. *Am J Clin Nutr* 1997 Nov;66(5):1264-1276. [doi: [10.1093/ajcn/66.5.1264](https://doi.org/10.1093/ajcn/66.5.1264)] [Medline: [9356547](https://pubmed.ncbi.nlm.nih.gov/9356547/)]
44. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. *J Med Internet Res* 2016 Jun 10;18(6):e135 [FREE Full text] [doi: [10.2196/jmir.5066](https://doi.org/10.2196/jmir.5066)] [Medline: [27287668](https://pubmed.ncbi.nlm.nih.gov/27287668/)]
45. Linardon J, Wade TD, de la Piedad Garcia X, Brennan L. The efficacy of cognitive-behavioral therapy for eating disorders: a systematic review and meta-analysis. *J Consult Clin Psychol* 2017 Nov;85(11):1080-1094. [doi: [10.1037/ccp0000245](https://doi.org/10.1037/ccp0000245)] [Medline: [29083223](https://pubmed.ncbi.nlm.nih.gov/29083223/)]
46. Rodbard D. Continuous glucose monitoring: a review of recent studies demonstrating improved glycemic outcomes. *Diabetes Technol Ther* 2017 Jun;19(S3):S25-S37 [FREE Full text] [doi: [10.1089/dia.2017.0035](https://doi.org/10.1089/dia.2017.0035)] [Medline: [28585879](https://pubmed.ncbi.nlm.nih.gov/28585879/)]

Abbreviations

- ABAB:** A, CGM data sharing and JITAIs-Off; B, CGM data sharing and JITAIs-On
BED: binge eating disorder
BN: bulimia nervosa
CBT: cognitive behavioral therapy
CGM: continuous glucose monitoring
ED: eating disorder
EDE: Eating Disorder Examination

EDE-Q: Eating Disorder Examination Questionnaire
JITAI: just-in-time, adaptive intervention

Edited by A Mavragani; submitted 04.04.22; peer-reviewed by YC Kato-Lin, D Estrin; comments to author 01.08.22; revised version received 26.09.22; accepted 03.10.22; published 14.12.22.

Please cite as:

Juarascio AS, Srivastava P, Presseller EK, Lin M, Patarinski AGG, Manasse SM, Forman EM

Using Continuous Glucose Monitoring to Detect and Intervene on Dietary Restriction in Individuals With Binge Eating: The SenseSupport Withdrawal Design Study

JMIR Form Res 2022;6(12):e38479

URL: <https://formative.jmir.org/2022/12/e38479>

doi: [10.2196/38479](https://doi.org/10.2196/38479)

PMID: [36515992](https://pubmed.ncbi.nlm.nih.gov/36515992/)

©Adrienne S Juarascio, Paakhi Srivastava, Emily K Presseller, Mandy Lin, Anna G G Patarinski, Stephanie M Manasse, Evan M Forman. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Preferences of Older Adult Veterans With Heart Failure for Engaging With Mobile Health Technology to Support Self-care: Qualitative Interview Study Among Patients With Heart Failure and Content Analysis

Marva Foster^{1,2*}, PhD; Wei Xiong^{3*}, MA, MS; Lisa Quintiliani², PhD; Christine W Hartmann^{4,5}, PhD; Stephan Gaehde^{6,7}, MD, MPH

¹VA Boston Healthcare System, Center for Healthcare Organization and Implementation Research, Boston, MA, United States

²Department of General Internal Medicine, School of Medicine, Boston University, Boston, MA, United States

³Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

⁴VA Bedford Healthcare System, Center for Healthcare Organization and Implementation Research, Bedford, MA, United States

⁵Department of Public Health, Zuckerberg College of Health Sciences, University of Massachusetts, Lowell, MA, United States

⁶VA Boston Healthcare System, Department of Medicine, Section of Emergency Services, Boston, MA, United States

⁷Department of Medicine, School of Medicine, Boston University, Boston, MA, United States

*these authors contributed equally

Corresponding Author:

Marva Foster, PhD

VA Boston Healthcare System

Center for Healthcare Organization and Implementation Research

150 S Huntington Avenue

Boston, MA, 02130

United States

Phone: 1 857 203 6671

Fax: 1 857 364 4511

Email: marva.foster@va.gov

Abstract

Background: Heart failure (HF) affects approximately 6.5 million adults in the United States, disproportionately afflicting older adults. Mobile health (mHealth) has emerged as a promising tool to empower older adults in HF self-care. However, little is known about the use of this approach among older adult veterans.

Objective: The goal of this study was to explore which features of an app were prioritized for older adult veterans with HF.

Methods: Between January and July 2021, we conducted semistructured interviews with patients with heart failure aged 65 years and older at a single facility in an integrated health care system (the Veterans Health Administration). We performed content analysis and derived themes based on the middle-range theory of chronic illness, generating findings both deductively and inductively. The qualitative questions captured data on the 3 key themes of the theory: self-care maintenance, self-care monitoring, and self-care management. Qualitative responses were analyzed using a qualitative data management platform, and descriptive statistics were used to analyze demographic data.

Results: Among patients interviewed (n=9), most agreed that a smartphone app for supporting HF self-care was desirable. In addition to 3 a priori themes, we identified 7 subthemes: education on daily HF care, how often to get education on HF, support of medication adherence, dietary restriction support, goal setting for exercises, stress reduction strategies, and prompts of when to call a provider. In addition, we identified 3 inductive themes related to veteran preferences for app components: simplicity, ability to share data with caregivers, and positive framing of HF language.

Conclusions: We identified educational and tracking app features that can guide the development of HF self-care for an older adult veteran population. Future research needs to be done to extend these findings and assess the feasibility of and test an app with these features.

KEYWORDS

qualitative research; heart failure; self-care; mobile health; mobile health technology; older adults; elderly; perceptions; mhealth intervention; veteran health; mHealth technology; elderly health care; elderly self-care

Introduction

Heart failure (HF) is a major health problem in the United States [1], with an estimated 6.5 million adults living with this condition. Its prevalence is highest among older adults [2-4]. Approximately 20% of all hospital discharges in older adults are associated with HF, and it is one of the major readmission diagnoses among Medicare beneficiaries [5]. Among all older adults in the United States, veterans older than 65 years represent a specifically vulnerable population, as they have a consistently higher readmission rate for HF than nonveterans [5,6]. Within the Veterans Health Administration (VHA), HF is the second most common diagnosis as well as one of the most expensive diagnoses to treat annually [6-9]. The most common reason for HF-related hospitalizations is symptom exacerbations. Engaging in self-care [10]—self-monitoring of physiologic changes and symptom recognition—is among the most critical preventative efforts to decrease HF readmission rates among patients, as early detection of potentially serious symptoms enables patients and their caregivers to intervene before hospitalization is needed [11]. HF prevention efforts, therefore, ideally target improving patients' ability to identify and attribute meaning to early symptoms [12]. Yet, patients with HF—who are usually older persons and frequently have comorbidities—often struggle to recognize the signs of exacerbation owing to difficulty discriminating HF symptoms from other comorbidities [13,14]. They can also find it challenging to adhere to complicated medication regimens and lifestyle advice related to diet and exercise [12,15]. The use of mobile health (mHealth) technologies (eg, Fitbit and Apple Heart Study) with HF symptom-tracking features, which enable older adults to monitor changes in their conditions on a daily basis and determine when treatment is needed, may improve HF health outcomes and thereby decrease health care usage.

mHealth technology is an effective platform to support changes in health behavior (physical activity, diet, etc) because of its ease of use, consistent connectivity to information, and quick upgrades that lead to ever-increasing sophistication [16,17]. Furthermore, mHealth has the potential to be useful for symptom management among all populations including older adults because it can include behavioral prompts, reminders, illness monitoring, and self-care management programs that extend beyond clinic walls. The use of mHealth technologies among older adults is increasing [18]. Over 62% of adults aged 70 years and older use smartphones [19]. Over 30% of older adults report using a smartphone app to manage an aspect of their health [4,20]. Older veterans also report technology ownership rates and interest in using mHealth, which is in line with that of the general population [21,22]. The increased availability of mHealth tools and the increasing technological engagement of older adults offer a potentially cost-effective solution to support HF self-care.

Nevertheless, there are also a number of specific barriers for implementation in this patient population. These include older patients' physical limitations from existing health conditions, such as sensory impairments, cognitive changes, arthritis, and vision impairments. These age- and disease-related physical limitations can inhibit the ability of older adults to use the functions of the technology and receive maximum benefit [16,23]. Other perceived challenges are the burden or workload associated with device use [24]. One suggested way to address these barriers is to involve older adults in the design process from inception through development [16,25,26].

In addition, mHealth developers should consider getting input specifically on features that will support health behavior change. Health behavior refers to any behaviors (physical activity, healthy diet, etc) that can impact a person's physical and mental health and quality of life [27]. Successful HF health behavior change leads to better management of the disease and resultant reduction in the frequency of hospital readmissions [28,29]. Embedding health behavior change strategies into cardiovascular interventions have shown a sustainable change in self-management behaviors [30,31]. Strategies or features to support behavior change include instruction on how to perform a behavior, self-monitoring, goal setting, problem-solving, etc [32,33]. There are established national and international guidelines that support the use of behavior change features to improve cardiovascular disease, including HF [34,35]. However, key questions remain in terms of how mHealth interventions should be optimally designed for older adult veterans with HF.

A recent review of the functionalities of commercially available apps and their ability to support HF symptom monitoring and self-care management was conducted by Creber et al [36]. The authors searched 3 web-based app stores for apps that provided self-management to patients with HF. They then rated the apps using the Mobile Application Rating Scale [37] and the Heart Failure Society of America guidelines for nonpharmacologic management [38]. The authors reviewed 34 apps that met inclusion criteria and found that many apps were designed to support healthy living rather than chronic disease management and did not effectively support change in health behavior. This highlights the need for improving the ability of mHealth apps to support HF. It also opens the door to involving end users in helping identify effective and engaging mHealth interventions to improve patient self-care [39,40]. Therefore, in this preliminary study, we examined the behavior change features older adult veterans with HF would find important to include in an HF mHealth intervention.

Methods

Design

The study used a descriptive qualitative method [41,42]. We conducted individual, web-based, semistructured in-depth

interviews among older veterans with HF. We used a qualitative content analysis approach [41] to provide a rich description of older adult veteran perspectives and preferences for a mHealth intervention to support HF self-care.

Ethical Considerations

Ethical approval for this study was obtained from the VA Boston Healthcare System Institutional Review Board (#3216-X). The local institutional review board approved this study with a waiver of informed consent. Participants received a US \$20 cash voucher for their participation. Randomly generated ID numbers were assigned to participants to ensure confidentiality. No personal identifying information was used in the audio recordings.

Participant Sampling and Recruitment

We used purposive sampling [41] (a nonprobability selection of participants) to ensure the richness of data for this preliminary study. We recruited older veteran adults aged 70-80 years who used smartphones. Between October 2020 and May 2021, we used administrative data to obtain a list of 70 potential participants who were patients at VA Boston HF clinic. We excluded 15 patients who had cognitive impairment or a psychotic disorder diagnosis. We mailed opt-out letters to the remaining 55 patients and called potentially eligible participants who did not opt out or otherwise contact the study staff. Interested and eligible patients (n=9 for a response rate of 16%) were enrolled in the study if they (1) lived in their own house or apartment, (2) were aged 65 years or older, (3) had an HF diagnosis, (4) owned an Android or iOS platform smartphone, and (5) used any apps on their smartphone more than once in the preceding 30 days. All participants were enrolled after an introductory conversation with the researchers who explained to them the details and the purpose of the study.

Data Collection

Data were collected between January and July 2021. We conducted semistructured interviews using a web-based platform, WebEx (WebEx Video Communications, Inc) [43]. Participants were given the option to use the video function or call in via telephone. At the start of the interview, participants answered a short quantitative questionnaire to gather demographic information (age, race and ethnicity, level of education, and marital status), and usage of smartphone apps (see *Demographic Questionnaire* in [Multimedia Appendix 1](#)).

We developed the semistructured interview guide based on key concepts from the middle-range theory of self-care in chronic illness [10]. The theory addresses the process of maintaining health through health-promoting practices and managing illness. The 3 key concepts are self-care maintenance, self-care monitoring, and self-care management. Self-care maintenance refers to those behaviors performed to improve well-being, preserve health, or to maintain physical and emotional stability [10]. These behaviors can be related to lifestyle (eg, exercise, preparing healthy food, and coping with stress) or the medical regimen (eg, taking medication as prescribed and attending medical appointments). Self-care monitoring is a process of routine, vigilant body monitoring, surveillance, or body listening [10]. This type of monitoring is a common behavior. For

example, people may monitor weight or blood pressure regularly to follow changes. Self-care management involves an evaluation of physical and emotional signs and symptoms to determine if the change is present and action is needed [10]. It requires attention to the effectiveness of a treatment and evaluation of whether that approach should be tried again in the future. The theory was chosen to guide this study because its structure focused research efforts on results that could be readily translated into practice. In addition, the interview guide included items related to preselected behavior change features (feedback on behavior, self-monitoring of the behavior, reducing negative emotions, instruction on how to perform the behavior, goal setting, social support, problem solving, and action planning) that were rated as important in the literature and our previous research [31,44-47] (see *Interview Guide* in [Multimedia Appendix 1](#)).

During the interviews, participants were asked open-ended questions as well as directed questions to explore their perceptions and preferences for receiving app-based support to help them self-manage their HF. Participants were asked to report on features that stood out as most useful or least useful, how to brand apps so they would be recognizable to those with HF, and if there were any other features that should be included. Interview prompts centered on the preselected behavior change app features. In addition, we also asked participants their perspectives on the term "heart failure." This question was posed to determine whether the title of the diagnosis affected the veterans' self-efficacy, as self-efficacy is an antecedent to self-care and has been found to independently predict HF self-care [10,48,49].

The HF diagnosis of all participants was confirmed through the patient's electronic health record by confirming an HF (International Statistical Classification of Diseases and Related Problems, Tenth Revision, Clinical Modification) code. We also collected data on the type of HF, HF preserved ejection fraction, or HF reduced ejection fraction. Interviews were conducted by MF, were audio recorded, and lasted approximately 45 minutes each. They were transcribed verbatim, and transcripts were checked for accuracy.

Data Analysis

Quantitative Data Analyses

Demographic data were entered into an Excel (Microsoft Corp) spreadsheet and analyzed to generate descriptive information. The participants' ages were described using mean and SD values, and the descriptive data were summarized using frequencies and percentages.

Qualitative Data Analyses

We took a combined approach to analysis. We used a directed content analysis approach [42] with predetermined codes based on the middle-range theory of self-care in chronic illness followed by inductive coding to capture the accounts (experiences and views) of research participants. Interview transcripts were uploaded in the qualitative data management program NVivo (version 12; QSR International) [50]. Analysis began with 2 researchers (MF and WX) each reading 3 transcripts in their entirety to become familiar with the data.

They then highlighted all text that on first impression appeared to represent the predetermined codes (see *Coding Framework* in [Multimedia Appendix 1](#)). Next, they coded all the highlighted passages using the predetermined codes. Any relevant text that could not be categorized with the initial coding scheme was given a new code. After MF and WX reached a consensus on the final coding framework, the remaining transcripts were analyzed by MF, followed by discussions and consensus generating with WX. Explanatory verbatim quotes were selected cautiously to maintain data validity and follow themes.

Results

Sample Characteristics

In total, 9 men opted in to participate in an interview session. Among these 9 participants, the mean age was 73.4 (range

70-80) years; the majority were White (n=7), most had completed postsecondary education (diploma or degree, n=6), and had HF with reduced ejection fraction (n=6). Five patients reported mild to moderate HF symptoms (shortness of breath with exertion and tiredness), while 4 others reported having no HF symptoms. As it was an eligibility criterion, all had either iPhones or Android-based smartphones. Most used their smartphone for email or internet at least once a day (n=8) and had experience with downloading an app (n=4), but fewer had used their phones to search for health-related information (n=3). However, the majority (n=6) reported using their phones for looking up sports information or the news and (n=5) used their phones to play games. Sociodemographic data are presented in [Table 1](#). None of the participants had previously downloaded an app to help them manage their HF.

Table 1. Sociodemographic characteristics of all participants (N=9).

Demographic information	Values
Age (years), mean (SD)	73.4 (1.4)
Gender (male), n (%)	9 (100)
Race, n (%)	
Black or African American	2 (22)
White	7 (78)
Ethnicity, n (%)	
Not Hispanic or Latino	9 (100)
Level of education, n (%)	
Less than high school	1 (11)
High school	2 (22)
Some college	2 (22)
Associates' degree	3 (33)
Bachelor's degree	1 (11)
Marital status, n (%)	
Married	4 (44)
Widowed	2 (22)
Divorced	2 (22)
Never married	1 (11)
Living alone, n (%)	3 (33)
Type of heart failure, n (%)	
Heart failure preserved ejection fraction	3 (33)
Heart failure reduced ejection fraction	6 (67)
Had or currently used home telehealth, n (%)	3 (33)
Ever used a heart failure self-management app, n (%)	
Never	9 (100)

Preferences for App Features and Content

Below we discuss the participants' opinions of the app features grouped according to the 3 key concepts of the middle-range theory of chronic illness (self-care maintenance, self-care

monitoring, and self-care management; [Table 2](#)). After that, we describe three iteratively developed categories related to app preference: (1) simplicity, (2) ability to engage their caregiver, and (3) positively framed language.

Table 2. Themes and illustrative quotes of participants.

Theme or category	Illustrative quote
Self-care maintenance	<ul style="list-style-type: none"> Overall, most participants thought that having an app “would be helpful.” “The app probably would be good. You could look, not only myself, [but] others that have different problems with their heart problem.” [Vet #8]
Education on daily HF ^a care	<ul style="list-style-type: none"> Most participants thought that an app would help them gain knowledge about their condition. “That would probably make my life easier. You know, I’m curious about a lot of things and I would probably read it. I would look at it and then if I had [or] I thought I had something, I’d probably look it up.” [Vet #4] A few were not sure how useful an app would be: “I don’t think it would do much, to tell you the truth. Not right now, anyway.” [Vet #2]
How often to get education on HF	<ul style="list-style-type: none"> Participants thought the frequency of receiving education should be determined by each person: “I think the person who was getting the App should have a choice to how often they want to get the information, how often they think they need the information. Do you need it every day or should it be every day or should it be weekly or should it be done monthly, bi-weekly or whatever? I think it’s gonna be an individual choice, not just a blanket App.” [Vet #4]
Support of medication adherence	<ul style="list-style-type: none"> Some participants thought having an app could not only provide medication reminders but also provide education about medications: “Yes, that would be worthwhile, especially if you could use that app, [and say] ‘tell me about [a medication]’ and it would tell you.” [Vet #9] “If I had an app to do [learn about medications], the doctor would be thrilled, I’m sure.” [Vet #7] On the other hand, some participants thought medication reminders “... would get me somewhat upset.” Some who were being seen frequently by their cardiologist did not see the importance of using an app to support medication adherence. For these individuals, their frequent medical appointments were enough: “No, I doubt [the app would be useful]. I see the cardiologist like once a month.” [Vet #1]
Dietary restriction support	<ul style="list-style-type: none"> The need for support to adhere to dietary restrictions (limiting sodium consumption and weight management) varied based on the participant’s level of involvement with health care services (eg, home telehealth and, nutrition courses offered at the HF clinic) and whether they were the primary cooks in their homes. Those involved in services preferred less support: “I have that Telehealth thing, ...and every day that thing gives me messages about what to do about medications, about your diet, you know.” [Vet #1] Those who were not the primary cooks in their homes thought it might not help: “[M]y wife watch[es] everything that I eat.” [Vet #5] Some others stated the inclusion of dietary support, “Would be helpful. Especially like I said, I’m not educated at all in nutrition.” [Vet #9]
Goal setting for exercises	<ul style="list-style-type: none"> When asked about setting goals for exercise, some participants mentioned that motivation played a major factor in their desire to set goals for exercise, saying things such as, “I’ve done reading, I just need the willpower to get through it,” and, “I just don’t have the motivation to do it.” However, for some the desire to set goals was dependent on others: “I would welcome that [setting goals for exercise]... We started out great here, and then my wife started having problems with her back and sciatica. But I’m anticipating that that will get back on track when she completes a course of rehab and so on, but we don’t know about that, But it does inhibit my enthusiasm for getting out and around.” [Vet #7]
Stress reduction strategies (eg, meditation and breathing exercises)	<ul style="list-style-type: none"> When asked about how useful it would be for an app to offer methods that can be used to relax or reduce stress, most participants did not find this feature helpful: “I think it’s definitely something that the individual should ... be able to choose themselves, because I don’t want to be told what do to do relax.” [Vet #4]
Self-care monitoring	<ul style="list-style-type: none"> Some participants thought there was utility in using an app to monitor and track symptoms: “I think that’s a good idea [to be able to track symptoms].” [Vet #7] Others pointed to the issue of having to manually input data into an app as being a deterrent to self-monitoring: “Oh, that’s just more work for myself, isn’t it? I mean I would have to concentrate every day and put a) my weight in, b) my blood sugar, c) my blood pressure... So that’s 15 minutes [that] would be just fooling around with that stuff.” [Vet #4]

Theme or category	Illustrative quote
Self-care management	<ul style="list-style-type: none"> When asked about the ability to review previous symptoms, one participant remarked "... your memory plays tricks on you." Having the ability to review previous changes in symptoms was acceptable because: "if you get the same symptom again and it happened a month ago, it would be good to find out what you did, instead of writing it down and trying to look for it through paperwork, [which is] what I did last time." [Vet #3] Others were skeptical about how previous information could support decision-making and actions on future changes in symptoms: "I don't think my phone can tell me what's wrong with my heart. I really don't. What kind of App can you put on your phone that tells me that my... blood pressure is running high?" [Vet #4] Almost all participants did not think that it was important to be able to contact medical help from within an app: "Well my phone already has that. It will call 9-1-1 automatically if you want." [Vet #1]
Prompts of when to call a provider	<ul style="list-style-type: none"> Another important facet about self-care management is that the treatment indicated might require consultation with a health care provider. When asked if receiving prompts to notify a provider based on changes in their HF patterns that were similar to those that led to a hospitalization in the past would be helpful, almost half of the participants did not think that it would be helpful to have this feature, and one participant stated, "Actually, it'd probably scare me out of my wits, that I was gonna check out." Yet another participant responded, "That would be helpful. To me it wouldn't have to be a cardiologist, just somebody that's knowledgeable." [Vet #9]

^aHF: heart failure.

Inductive Themes Related to Preferences for App Features and Content

Simplicity

Participants expressed that "you need to make the app easy for the old people" and develop it at "common man's level." This preference was clearly a requirement for most:

"You know, some apps...I'll just delete...because I can't figure it out." [Vet #3]

Ability to Share Data With Caregivers

When asked who else should be given access their information in an app, most wanted to share data with others.

"My son and his wife, because they're the ones [who] keep an eye on me." [Vet #8]

"Let's put it this way. [As long as my wife] can access it through her computer ... to retrieve it, all well and good. I'll go along with that." [Vet #5]

Positively Frame HF Language

When asked about how they feel about the term "heart failure" and its effect on them, one person said, "I don't get that it's really a failure. Anomaly might be a better, but even that's a scary word." The term HF echoed even more negativity for another participant:

"I don't want to talk about how I'm suffering from heart failure because I'm not, you see. And somebody who is suffering from heart failure that echoes failure, failure, failure, failure in their mind. Or if it does, how have you helped them?" [Vet #4]

Discussion

Principal Findings

In this study, we aimed to identify the mHealth design preferences of behavior change features of older adult veterans with HF to inform the design of a future mHealth intervention. Our analysis of interviews with 9 older adult Veterans demonstrated that older adults are engaged and willing to use mobile technology to support their self-care. Participants' accounts not only identified features and characteristics important to them to support HF self-care maintenance, monitoring, and management, but also highlighted the challenge of designing mHealth for older adults with varying levels of caregiver support.

In terms of self-maintenance, veterans perceived mHealth to be helpful to gain knowledge on their HF condition but wanted the ability to choose how often they get that information. This finding adds to the literature regarding including patient perspectives in education. It has been found that educational activities that involve shared decision-making improve self-care practices [51], and user preferences have a significant impact on the use and effect of the education [52]. Participants also expressed the need for medication support to help them understand their medications, but they had mixed thoughts regarding medication reminders. Previous research has shown high satisfaction when patients receive medication reminders that are delivered at times they select and correlate with their medication schedule [53]. Gaining patients' perspectives on the optimal number of messages and alerts is warranted to decrease psychological stress. Dietary support was a component that some participants also indicated mixed feelings about. Responses were based on the amount of caregiver support participants already had. The differing viewpoints corroborate with existing literature that encourages mHealth designers to

allow users to customize mHealth for their individual case [54]. Some participants were also interested in using goal setting features for exercises and expressed the need for “willpower” to exercise. This finding reinforced the need for live coaching to facilitate exercise goal setting and attainment, in addition to app-based content (eg, self-monitoring, information modules, and exercises) [55,56]. Furthermore, the inclusion of behavior change features (ie, plan or goals and mobile diary) in mHealth correlates with a higher incidence of statistically significant outcomes [57]. Like many chronic illnesses, HF places great physical and psychological stress on patients. Participants endorsed the helpfulness of stress reduction strategies, suggesting that individuals should be able to choose the type of strategy that appeals to them. This adds to the literature that has found that stress reduction through mind/body interventions not only has benefits for healthy individuals but has many health benefits for patients in all stages of HF [58].

Monitoring of biometric measurements (weight, blood pressure, and pulse) was deemed useful by almost all participants, as long as they did not have to manually input data into an app. Their interest in self-care monitoring is in line with previous research that patients are interested in monitoring their symptoms as long as it does not intrude in their lifestyle [59]. Manual input of data has been cited as a deterrent to symptom self-care monitoring [60,61]. Our finding corroborates the findings of an existing body of literature on the importance of incorporating automatic uploading of biometric measurements to a mHealth app [57,62]. It should be noted that although Bluetooth technology (which supports automatic uploading of data) is widely available, it is not frequently used in mHealth research [57]. Participants also endorsed a feature that would allow them to evaluate past symptoms. Having a single source to record symptoms and actions taken to remedy symptoms can provide information to patients to identify worsening of HF. This finding is consistent with research that suggests the use of an electronic heart diary [63] to provide a consistent location for patients to track and review physical and psycho-emotional changes to determine if action is needed. Although receiving prompts as to when to call their provider was not seen as helpful by some participants, it was for others. This contributes to the literature on the desirability of using mHealth to guide patients to seek higher levels of care when needed [52].

In addition to feature preferences for health behavior change (reminders, monitoring, etc), veterans also had a preference for simple, easy-to-use technology. Ease of use is just one factor that can affect adherence among older adults [64]. They also wanted the ability to share data with family and caregivers. This preference was similar to previous research that explored the attitudes and preferences of older adults on warfarin therapy regarding the use of mHealth technology and health games to gain skills for self-management [65]. Another preference was the ability to customize the app once it is downloaded. Several participants indicated they wanted to have their family’s or caregiver’s support in conducting this task. These findings substantiate previous research indicating that some older adults say they need help setting up or using technology [18,66].

This study’s qualitative approach allowed for a better understanding of not only veterans’ mHealth preferences but also their feelings when told that they have HF. It is known that language can potentially transmit bias and affect the quality of care that patients subsequently receive [67,68]. Language, whether spoken or written, can also affect the attitudes of others. Our finding that participants had a negative reaction to the term “heart failure” indicates that there may be a need to reframe language surrounding a HF diagnosis.

Limitations

This study focused solely on VHA patients, which represent a population not typically studied in mHealth contexts (eg, older adults). This may present a limit to generalizability to other populations [56,69] but also represents a needed addition to the literature on veterans. Another limitation was that most participants were White, potentially limiting the transferability of these findings to other ethnic groups. This may mean that our participants may be overly similar in certain ways, having similar backgrounds and preferences. Another potential source of bias is the small sample size. However, this study used the concept of information power rather than saturation to determine the adequacy of the sample size. Information power indicates that the more information the sample holds, relevant for the actual study, the lower number of participants is needed [70,71]. Based on our use of a specific, vulnerable population, the use of an established theory, the quality of the dialogue, and our analysis strategy, we determined the sample size to be sufficient. Although the use of a theory to guide our work was a strength of our study, there is the possibility that our use of the specific use of the middle-range theory of self-care in chronic illness may have limited the scope of the results, and future studies should take this into consideration. In addition, our focus was to examine the perceptions of features older adult veterans with HF would find important to include in a mHealth intervention; more background on demographics and other participant characteristics would have strengthened the interpretation of our findings. Finally, 3 of our participants were enrolled in or had previous experience with home telehealth. This may have caused some bias regarding their intention to use mHealth. It is also important to acknowledge that some older adults, as a result of strong personal preferences or other barriers, may never adopt mHealth [23].

Conclusions

In conclusion, despite the proliferation of mHealth apps to manage HF [36], a dearth of information exists concerning the usage needs of older adult veterans with HF. Some needs uncovered in our study are relatively new findings, such as the potential need to positively frame HF information and the patients’ desire for their caregivers to have access to the patient’s self-care data. Other needs identified correlate with those of HF patients in general (eg, goal setting and dietary restriction). App features should facilitate addressing these needs and consider incorporating, for example, a simple interface accessible to older adults with little or no technological literacy. Future research needs to be done to extend these findings and assess the feasibility of and test an app with these features.

Acknowledgments

The authors would like to thank Jacob Joseph for his expertise during our study. Statements contained in this article reflect the views of the authors and do not represent the official positions of the US Department of Veterans Affairs or other authors' affiliate organizations. This research was supported by the Veterans Health Administration and the Fellowship Office of Academic Affiliations. MF is supported by the VA Office of Research and Development Research Supplements to Promote Diversity Award (IIR 18-244, P.I. McInnes).

Data Availability

The data sets presented in this article are not readily available because they will need to be approved by the VA IRB, requiring a modification. Requests to access the datasets should be directed to MF: Marva.Foster@va.gov.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demographic questionnaire, interview guide and coding framework.

[\[DOCX File, 22 KB - formative_v6i12e41317_app1.docx\]](#)

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee Stroke Statistics Subcommittee. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *Circulation* 2017 Mar 07;135(10):e146-e603 [FREE Full text] [doi: [10.1161/CIR.0000000000000485](https://doi.org/10.1161/CIR.0000000000000485)] [Medline: [28122885](https://pubmed.ncbi.nlm.nih.gov/28122885/)]
2. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Heart disease and stroke statistics—2016 update: a report from the American Heart Association. *Circulation* 2016 Jan 26;133(4):e38-e360. [doi: [10.1161/CIR.0000000000000350](https://doi.org/10.1161/CIR.0000000000000350)] [Medline: [26673558](https://pubmed.ncbi.nlm.nih.gov/26673558/)]
3. Savarese G, Lund LH. Global public health burden of heart failure. *Card Fail Rev* 2017 Apr;3(1):7-11 [FREE Full text] [doi: [10.15420/cfr.2016:25:2](https://doi.org/10.15420/cfr.2016:25:2)] [Medline: [28785469](https://pubmed.ncbi.nlm.nih.gov/28785469/)]
4. Schorr EN, Gepner AD, Dolansky MA, Forman DE, Park LG, Petersen KS, American Heart Association Cardiovascular Disease in Older Populations Committee of the Council on Clinical Cardiology Council on Cardiovascular Stroke Nursing; Council on Arteriosclerosis, Thrombosis Vascular Biology; Council on Lifestyle Cardiometabolic Health. Harnessing mobile health technology for secondary cardiovascular disease prevention in older adults: a scientific statement from the American Heart Association. *Circ Cardiovasc Qual Outcomes* 2021 May;14(5):e000103 [FREE Full text] [doi: [10.1161/HCQ.0000000000000103](https://doi.org/10.1161/HCQ.0000000000000103)] [Medline: [33793309](https://pubmed.ncbi.nlm.nih.gov/33793309/)]
5. Gorst SL, Armitage CJ, Brownsell S, Hawley MS. Home telehealth uptake and continued use among heart failure and chronic obstructive pulmonary disease patients: a systematic review. *Ann Behav Med* 2014 Dec;48(3):323-336 [FREE Full text] [doi: [10.1007/s12160-014-9607-x](https://doi.org/10.1007/s12160-014-9607-x)] [Medline: [24763972](https://pubmed.ncbi.nlm.nih.gov/24763972/)]
6. Nuti SV, Qin L, Rumsfeld JS, Ross JS, Masoudi FA, Normand ST, et al. Association of admission to veterans affairs hospitals vs non-veterans affairs hospitals with mortality and readmission rates among older men hospitalized with acute myocardial infarction, heart failure, or pneumonia. *JAMA* 2016 Feb 09;315(6):582-592 [FREE Full text] [doi: [10.1001/jama.2016.0278](https://doi.org/10.1001/jama.2016.0278)] [Medline: [26864412](https://pubmed.ncbi.nlm.nih.gov/26864412/)]
7. Zulman DM, Pal Chee C, Wagner TH, Yoon J, Cohen DM, Holmes TH, et al. Multimorbidity and healthcare utilisation among high-cost patients in the US Veterans Affairs Health Care System. *BMJ Open* 2015 Apr 16;5(4):e007771 [FREE Full text] [doi: [10.1136/bmjopen-2015-007771](https://doi.org/10.1136/bmjopen-2015-007771)] [Medline: [25882486](https://pubmed.ncbi.nlm.nih.gov/25882486/)]
8. Chapko MK, Liu C, Perkins M, Li Y, Fortney JC, Maciejewski ML. Equivalence of two healthcare costing methods: bottom-up and top-down. *Health Econ* 2009 Oct;18(10):1188-1201. [doi: [10.1002/hec.1422](https://doi.org/10.1002/hec.1422)] [Medline: [19097041](https://pubmed.ncbi.nlm.nih.gov/19097041/)]
9. Krishnamurthi N, Francis J, Fihn SD, Meyer CS, Whooley MA. Leading causes of cardiovascular hospitalization in 8.45 million US veterans. *PLoS One* 2018 Mar 22;13(3):e0193996 [FREE Full text] [doi: [10.1371/journal.pone.0193996](https://doi.org/10.1371/journal.pone.0193996)] [Medline: [29566396](https://pubmed.ncbi.nlm.nih.gov/29566396/)]
10. Riegel B, Jaarsma T, Strömberg A. A middle-range theory of self-care of chronic illness. *ANS Adv Nurs Sci* 2012;35(3):194-204. [doi: [10.1097/ANS.0b013e318261b1ba](https://doi.org/10.1097/ANS.0b013e318261b1ba)] [Medline: [22739426](https://pubmed.ncbi.nlm.nih.gov/22739426/)]
11. Groeneveld PW, Medvedeva EL, Walker L, Segal AG, Menno DM, Epstein AJ. Association between spending and survival of chronic heart failure across veterans affairs medical centers. *JAMA Netw Open* 2019 Jul 03;2(7):e197238 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.7238](https://doi.org/10.1001/jamanetworkopen.2019.7238)] [Medline: [31322689](https://pubmed.ncbi.nlm.nih.gov/31322689/)]
12. Guzman-Clark JRS, van Servellen G, Chang B, Mentis J, Hahn TJ. Predictors and outcomes of early adherence to the use of a home telehealth device by older veterans with heart failure. *Telemed J E Health* 2013 Mar;19(3):217-223. [doi: [10.1089/tmj.2012.0096](https://doi.org/10.1089/tmj.2012.0096)] [Medline: [23268695](https://pubmed.ncbi.nlm.nih.gov/23268695/)]

13. Yancy C, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, et al. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2013 Oct 15;128(16):1810-1852. [doi: [10.1161/CIR.0b013e31829e8807](https://doi.org/10.1161/CIR.0b013e31829e8807)] [Medline: [23741057](https://pubmed.ncbi.nlm.nih.gov/23741057/)]
14. Toback M, Clark N. Strategies to improve self-management in heart failure patients. *Contemp Nurse* 2017 Feb 19;53(1):105-120. [doi: [10.1080/10376178.2017.1290537](https://doi.org/10.1080/10376178.2017.1290537)] [Medline: [28151071](https://pubmed.ncbi.nlm.nih.gov/28151071/)]
15. Foster M. A mobile application for patients with heart failure: theory- and evidence-based design and testing. *Comput Inform Nurs* 2018 Nov;36(11):540-549. [doi: [10.1097/CIN.0000000000000465](https://doi.org/10.1097/CIN.0000000000000465)] [Medline: [30045131](https://pubmed.ncbi.nlm.nih.gov/30045131/)]
16. Masterson Creber RM, Hickey KT, Maurer MS. Gerontechnologies for older patients with heart failure: what is the role of smartphones, tablets, and remote monitoring devices in improving symptom monitoring and self-care management? *Curr Cardiovasc Risk Rep* 2016 Oct;10(10):30 [FREE Full text] [doi: [10.1007/s12170-016-0511-8](https://doi.org/10.1007/s12170-016-0511-8)] [Medline: [28713481](https://pubmed.ncbi.nlm.nih.gov/28713481/)]
17. Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. *JMIR Mhealth Uhealth* 2015 Sep 14;3(3):e87 [FREE Full text] [doi: [10.2196/mhealth.4532](https://doi.org/10.2196/mhealth.4532)] [Medline: [26369346](https://pubmed.ncbi.nlm.nih.gov/26369346/)]
18. Anderson MP, Perrin A. Tech adoption climbs among older adults. Pew Research Center. 2017. URL: <https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/> [accessed 2022-12-08]
19. Kakulla BN. 2020 Tech Trends of the 50+. AARP Research. Washington, DC: AARP Research; 2020 Jan. URL: <https://doi.org/10.26419/res.00329.001> [accessed 2022-04-21]
20. Fox S, Duggan M. Main findings. Pew Research Center. URL: <https://www.pewresearch.org/internet/2012/11/08/main-findings-6/> [accessed 2022-05-25]
21. Gould CE, Loup J, Kuhn E, Beaudreau SA, Ma F, Goldstein MK, et al. Technology use and preferences for mental health self-management interventions among older veterans. *Int J Geriatr Psychiatry* 2020 Mar 29;35(3):321-330. [doi: [10.1002/gps.5252](https://doi.org/10.1002/gps.5252)] [Medline: [31854029](https://pubmed.ncbi.nlm.nih.gov/31854029/)]
22. Lipschitz J, Miller CJ, Hogan TP, Burdick KE, Lippin-Foster R, Simon SR, et al. Adoption of mobile apps for depression and anxiety: cross-sectional survey study on patient interest and barriers to engagement. *JMIR Ment Health* 2019 Jan 25;6(1):e11334 [FREE Full text] [doi: [10.2196/11334](https://doi.org/10.2196/11334)] [Medline: [30681968](https://pubmed.ncbi.nlm.nih.gov/30681968/)]
23. Krishnaswami A, Beavers C, Dorsch MP, Dodson JA, Masterson Creber R, Kitsiou S, Innovations, Cardiovascular Teamthe Geriatric Cardiology Councils, American College of Cardiology. Gerontechnology for older adults with cardiovascular diseases: JACC state-of-the-art review. *J Am Coll Cardiol* 2020 Dec 01;76(22):2650-2670 [FREE Full text] [doi: [10.1016/j.jacc.2020.09.606](https://doi.org/10.1016/j.jacc.2020.09.606)] [Medline: [33243384](https://pubmed.ncbi.nlm.nih.gov/33243384/)]
24. Dharmarajan K, Chaudhry SI. New approaches to reduce readmissions in patients with heart failure. *JAMA Intern Med* 2016 Mar 01;176(3):318-320. [doi: [10.1001/jamainternmed.2015.7993](https://doi.org/10.1001/jamainternmed.2015.7993)] [Medline: [26857025](https://pubmed.ncbi.nlm.nih.gov/26857025/)]
25. Piau A, Campo E, Rumeau P, Vellas B, Nourhashemi F. Aging society and gerontechnology: a solution for an independent living? *J Nutr Health Aging* 2014 Jan 16;18(1):97-112. [doi: [10.1007/s12603-013-0356-5](https://doi.org/10.1007/s12603-013-0356-5)] [Medline: [24402399](https://pubmed.ncbi.nlm.nih.gov/24402399/)]
26. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. *J Med Internet Res* 2014 Apr 09;16(4):e104 [FREE Full text] [doi: [10.2196/jmir.2968](https://doi.org/10.2196/jmir.2968)] [Medline: [24718852](https://pubmed.ncbi.nlm.nih.gov/24718852/)]
27. Tombor I, Michie S. Methods of health behavior change. Oxford Research Encyclopedia of Psychology. Oxford University Press: New York; 2017 Jul 27. URL: <https://oxfordre.com/psychology/view/10.1093/acrefore/9780190236557.001.0001/acrefore-9780190236557-e-125> [accessed 2022-04-21]
28. Paul S, Sneed NV. Strategies for behavior change in patients with heart failure. *Am J Crit Care* 2004 Jul;13(4):305-313. [Medline: [15293582](https://pubmed.ncbi.nlm.nih.gov/15293582/)]
29. Laddu D, Ma J, Kaar J, Ozemek C, Durant RW, Campbell T, et al. Health behavior change programs in primary care and community practices for cardiovascular disease prevention and risk factor management among midlife and older adults: a scientific statement from the American Heart Association. *Circulation* 2021 Dec 14;144(24):e533-e549 [FREE Full text] [doi: [10.1161/CIR.0000000000001026](https://doi.org/10.1161/CIR.0000000000001026)] [Medline: [34732063](https://pubmed.ncbi.nlm.nih.gov/34732063/)]
30. Valente TW. Network interventions. *Science* 2012 Jul 06;337(6090):49-53. [doi: [10.1126/science.1217330](https://doi.org/10.1126/science.1217330)] [Medline: [22767921](https://pubmed.ncbi.nlm.nih.gov/22767921/)]
31. Akinosun AS, Polson R, Diaz-Skeete Y, De Kock JH, Carragher L, Leslie S, et al. Digital technology interventions for risk factor modification in patients with cardiovascular disease: systematic review and meta-analysis. *JMIR Mhealth Uhealth* 2021 Mar 03;9(3):e21061 [FREE Full text] [doi: [10.2196/21061](https://doi.org/10.2196/21061)] [Medline: [33656444](https://pubmed.ncbi.nlm.nih.gov/33656444/)]
32. Suls J, Mogavero JN, Falzon L, Pescatello LS, Hennessy EA, Davidson KW. Health behaviour change in cardiovascular disease prevention and management: meta-review of behaviour change techniques to affect self-regulation. *Health Psychol Rev* 2020 Mar 29;14(1):43-65 [FREE Full text] [doi: [10.1080/17437199.2019.1691622](https://doi.org/10.1080/17437199.2019.1691622)] [Medline: [31707938](https://pubmed.ncbi.nlm.nih.gov/31707938/)]
33. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013 Aug 20;46(1):81-95 [FREE Full text] [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]

34. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, ESC Scientific Document Group. Eur Heart J 2016 Aug 01;37(29):2315-2381 [FREE Full text] [doi: [10.1093/eurheartj/ehw106](https://doi.org/10.1093/eurheartj/ehw106)] [Medline: [27222591](https://pubmed.ncbi.nlm.nih.gov/27222591/)]
35. LeFevre ML, U.S. Preventive Services Task Force. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2014 Oct 21;161(8):587-593 [FREE Full text] [doi: [10.7326/M14-1796](https://doi.org/10.7326/M14-1796)] [Medline: [25155419](https://pubmed.ncbi.nlm.nih.gov/25155419/)]
36. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the mobile application rating scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74 [FREE Full text] [doi: [10.2196/mhealth.5882](https://doi.org/10.2196/mhealth.5882)] [Medline: [27302310](https://pubmed.ncbi.nlm.nih.gov/27302310/)]
37. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
38. Heart Failure Society of America, Lindenfeld J, Albert NM, Boehmer JP, Collins SP, Ezekowitz JA, et al. HFSA 2010 comprehensive heart failure practice guideline. J Card Fail 2010 Jun;16(6):e1-194. [doi: [10.1016/j.cardfail.2010.04.004](https://doi.org/10.1016/j.cardfail.2010.04.004)] [Medline: [20610207](https://pubmed.ncbi.nlm.nih.gov/20610207/)]
39. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
40. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
41. Patton M. Qualitative Research and Evaluation Methods, 4th ed. Thousand Oaks: Sage Publications; 2015.
42. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
43. Webex by Cisco. Cisco. 2022. URL: <https://www.webex.com/> [accessed 2022-05-17]
44. Quintiliani LM, Foster M, Oshry LJ. Preferences of mHealth app features for weight management among breast cancer survivors from underserved populations. Psychooncology 2019 Oct 15;28(10):2101-2104 [FREE Full text] [doi: [10.1002/pon.5190](https://doi.org/10.1002/pon.5190)] [Medline: [31368197](https://pubmed.ncbi.nlm.nih.gov/31368197/)]
45. Patterson K, Davey R, Keegan R, Kunstler B, Woodward A, Freene N. Behaviour change techniques in cardiovascular disease smartphone apps to improve physical activity and sedentary behaviour: systematic review and meta-regression. Int J Behav Nutr Phys Act 2022 Jul 07;19(1):81 [FREE Full text] [doi: [10.1186/s12966-022-01319-8](https://doi.org/10.1186/s12966-022-01319-8)] [Medline: [35799263](https://pubmed.ncbi.nlm.nih.gov/35799263/)]
46. Duff OM, Walsh DM, Furlong BA, O'Connor NE, Moran KA, Woods CB. Behavior change techniques in physical activity eHealth interventions for people with cardiovascular disease: systematic review. J Med Internet Res 2017 Aug 02;19(8):e281 [FREE Full text] [doi: [10.2196/jmir.7782](https://doi.org/10.2196/jmir.7782)] [Medline: [28768610](https://pubmed.ncbi.nlm.nih.gov/28768610/)]
47. Riegel B, Westland H, Iovino P, Barelds I, Bruins Slot J, Stawnychy MA, et al. Characteristics of self-care interventions for patients with a chronic condition: a scoping review. Int J Nurs Stud 2021 Apr;116:103713 [FREE Full text] [doi: [10.1016/j.ijnurstu.2020.103713](https://doi.org/10.1016/j.ijnurstu.2020.103713)] [Medline: [32768137](https://pubmed.ncbi.nlm.nih.gov/32768137/)]
48. Schweitzer RD, Head K, Dwyer JW. Psychological factors and treatment adherence behavior in patients with chronic heart failure. J Cardiovasc Nurs 2007;22(1):76-83. [doi: [10.1097/00005082-200701000-00012](https://doi.org/10.1097/00005082-200701000-00012)] [Medline: [17224702](https://pubmed.ncbi.nlm.nih.gov/17224702/)]
49. Schnell-Hoehn K, Naimark BJ, Tate RB. Determinants of self-care behaviors in community-dwelling patients with heart failure. J Cardiovasc Nurs 2009;24(1):40-47. [doi: [10.1097/01.JCN.0000317470.58048.7b](https://doi.org/10.1097/01.JCN.0000317470.58048.7b)] [Medline: [19114800](https://pubmed.ncbi.nlm.nih.gov/19114800/)]
50. Powerful research simplified. QSR International. 2022. URL: <https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/about/nvivo> [accessed 2022-12-08]
51. Athilingam P, Osorio RE, Kaplan H, Oliver D, O'neachtain T, Rogal PJ. Embedding patient education in mobile platform for patients with heart failure: theory-based development and beta testing. Comput Inform Nurs 2016 Feb;34(2):92-98. [doi: [10.1097/CIN.0000000000000216](https://doi.org/10.1097/CIN.0000000000000216)] [Medline: [26765655](https://pubmed.ncbi.nlm.nih.gov/26765655/)]
52. Allida S, Du H, Xu X, Prichard R, Chang S, Hickman LD, et al. mHealth education interventions in heart failure. Cochrane Database Syst Rev 2020 Jul 02;7(7):CD011845 [FREE Full text] [doi: [10.1002/14651858.CD011845.pub2](https://doi.org/10.1002/14651858.CD011845.pub2)] [Medline: [32613635](https://pubmed.ncbi.nlm.nih.gov/32613635/)]
53. Park LG, Howie-Esquivel J, Chung ML, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: a randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: [10.1016/j.pec.2013.10.027](https://doi.org/10.1016/j.pec.2013.10.027)] [Medline: [24321403](https://pubmed.ncbi.nlm.nih.gov/24321403/)]
54. Alnosayan N, Chatterjee S, Alluhaidan A, Lee E, Houston Feenstra L. Design and usability of a heart failure mHealth system: a pilot study. JMIR Hum Factors 2017 Mar 24;4(1):e9 [FREE Full text] [doi: [10.2196/humanfactors.6481](https://doi.org/10.2196/humanfactors.6481)] [Medline: [28341615](https://pubmed.ncbi.nlm.nih.gov/28341615/)]
55. Park LG, Ng F, K Shim J, Elnaggar A, Villero O. Perceptions and experiences of using mobile technology for medication adherence among older adults with coronary heart disease: a qualitative study. Digit Health 2020 May 20;6:2055207620926844 [FREE Full text] [doi: [10.1177/2055207620926844](https://doi.org/10.1177/2055207620926844)] [Medline: [32489672](https://pubmed.ncbi.nlm.nih.gov/32489672/)]

56. Breland JY, Agha K, Mohankumar R. Adoption and appropriateness of mHealth for weight management in the real world: a qualitative investigation of patient perspectives. *JMIR Form Res* 2021 Dec 08;5(12):e29916 [FREE Full text] [doi: [10.2196/29916](https://doi.org/10.2196/29916)] [Medline: [34889761](https://pubmed.ncbi.nlm.nih.gov/34889761/)]
57. Donevant S, Estrada RD, Culley JM, Habing B, Adams SA. Exploring app features with outcomes in mHealth studies involving chronic respiratory diseases, diabetes, and hypertension: a targeted exploration of the literature. *J Am Med Inform Assoc* 2018 Oct 01;25(10):1407-1418 [FREE Full text] [doi: [10.1093/jamia/ocy104](https://doi.org/10.1093/jamia/ocy104)] [Medline: [30137383](https://pubmed.ncbi.nlm.nih.gov/30137383/)]
58. Aggarwal M, Bozkurt B, Panjath G, Aggarwal B, Ostfeld RJ, Barnard ND, American College of Cardiology's NutritionLifestyle Committee of the Prevention of Cardiovascular Disease Council. Lifestyle modifications for preventing and treating heart failure. *J Am Coll Cardiol* 2018 Nov 06;72(19):2391-2405 [FREE Full text] [doi: [10.1016/j.jacc.2018.08.2160](https://doi.org/10.1016/j.jacc.2018.08.2160)] [Medline: [30384895](https://pubmed.ncbi.nlm.nih.gov/30384895/)]
59. Sohn A, Speier W, Lan E, Aoki K, Fonarow G, Ong M, et al. Assessment of heart failure patients' interest in mobile health apps for self-care: survey study. *JMIR Cardiol* 2019 Oct 29;3(2):e14332 [FREE Full text] [doi: [10.2196/14332](https://doi.org/10.2196/14332)] [Medline: [31758788](https://pubmed.ncbi.nlm.nih.gov/31758788/)]
60. Lu X, Yang H, Xia X, Lu X, Lin J, Liu F, et al. Interactive mobile health intervention and blood pressure management in adults. *Hypertension* 2019 Sep;74(3):697-704. [doi: [10.1161/HYPERTENSIONAHA.119.13273](https://doi.org/10.1161/HYPERTENSIONAHA.119.13273)] [Medline: [31327259](https://pubmed.ncbi.nlm.nih.gov/31327259/)]
61. Alessa T, Abdi S, Hawley MS, de Witte L. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, usability, and user satisfaction. *JMIR Mhealth Uhealth* 2018 Jul 23;6(7):e10723 [FREE Full text] [doi: [10.2196/10723](https://doi.org/10.2196/10723)] [Medline: [30037787](https://pubmed.ncbi.nlm.nih.gov/30037787/)]
62. Earle KA, Istepanian RS, Zitouni K, Sungoor A, Tang B. Mobile telemonitoring for achieving tighter targets of blood pressure control in patients with complicated diabetes: a pilot study. *Diabetes Technol Ther* 2010 Jul;12(7):575-579. [doi: [10.1089/dia.2009.0090](https://doi.org/10.1089/dia.2009.0090)] [Medline: [20597833](https://pubmed.ncbi.nlm.nih.gov/20597833/)]
63. Arulnathan A, Vaaheesan S, Denecke K. A mobile application for self-monitoring for patients with heart failure. *Stud Health Technol Inform* 2019;259:113-116. [Medline: [30923286](https://pubmed.ncbi.nlm.nih.gov/30923286/)]
64. Cajita M, Gleason KT, Han HR. A systematic review of mHealth-based heart failure interventions. *J Cardiovasc Nurs* 2016;31(3):E10-E22 [FREE Full text] [doi: [10.1097/JCN.0000000000000305](https://doi.org/10.1097/JCN.0000000000000305)] [Medline: [26544175](https://pubmed.ncbi.nlm.nih.gov/26544175/)]
65. Lee J, Nguyen AL, Berg J, Amin A, Bachman M, Guo Y, et al. Attitudes and preferences on the use of mobile health technology and health games for self-management: interviews with older adults on anticoagulation therapy. *JMIR Mhealth Uhealth* 2014 Jul 23;2(3):e32 [FREE Full text] [doi: [10.2196/mhealth.3196](https://doi.org/10.2196/mhealth.3196)] [Medline: [25098413](https://pubmed.ncbi.nlm.nih.gov/25098413/)]
66. Lefler LL, Rhoads SJ, Harris M, Funderburg AE, Lubin SA, Martel ID, et al. Evaluating the use of mobile health technology in older adults with heart failure: mixed-methods study. *JMIR Aging* 2018 Dec 04;1(2):e12178 [FREE Full text] [doi: [10.2196/12178](https://doi.org/10.2196/12178)] [Medline: [31518257](https://pubmed.ncbi.nlm.nih.gov/31518257/)]
67. Park J, Saha S, Chee B, Taylor J, Beach MC. Physician use of stigmatizing language in patient medical records. *JAMA Netw Open* 2021 Jul 01;4(7):e2117052 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.17052](https://doi.org/10.1001/jamanetworkopen.2021.17052)] [Medline: [34259849](https://pubmed.ncbi.nlm.nih.gov/34259849/)]
68. P Goddu A, O'Connor KJ, Lanzkron S, Saheed MO, Saha S, Peek ME, et al. Do words matter? Stigmatizing language and the transmission of bias in the medical record. *J Gen Intern Med* 2018 May;33(5):685-691 [FREE Full text] [doi: [10.1007/s11606-017-4289-2](https://doi.org/10.1007/s11606-017-4289-2)] [Medline: [29374357](https://pubmed.ncbi.nlm.nih.gov/29374357/)]
69. Eibner CK, Krull H, Brown KM, Cefalu M, Mulcahy AW, Pollard MS, et al. Current and projected characteristics and unique health care needs of the patient population served by the Department of Veterans Affairs. RAND. URL: https://www.rand.org/pubs/research_reports/RR1165z1.html [accessed 2022-05-27]
70. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res* 2016 Nov 10;26(13):1753-1760. [doi: [10.1177/1049732315617444](https://doi.org/10.1177/1049732315617444)] [Medline: [26613970](https://pubmed.ncbi.nlm.nih.gov/26613970/)]
71. Thorne S. Untangling the misleading message around saturation in qualitative nursing studies. *Nurse Author Editor* 2020 Mar 05;30(1):1-9. [doi: [10.1111/j.1750-4910.2020.tb00005.x](https://doi.org/10.1111/j.1750-4910.2020.tb00005.x)]

Abbreviations

HF: heart failure

mHealth: mobile health

VHA: Veterans Health Administration

Edited by A Mavragani; submitted 21.07.22; peer-reviewed by M Stein, WQ Koh; comments to author 20.10.22; revised version received 09.11.22; accepted 14.11.22; published 20.12.22.

Please cite as:

Foster M, Xiong W, Quintiliani L, Hartmann CW, Gaehde S

Preferences of Older Adult Veterans With Heart Failure for Engaging With Mobile Health Technology to Support Self-care: Qualitative Interview Study Among Patients With Heart Failure and Content Analysis

JMIR Form Res 2022;6(12):e41317

URL: <https://formative.jmir.org/2022/12/e41317>

doi: [10.2196/41317](https://doi.org/10.2196/41317)

PMID: [36538348](https://pubmed.ncbi.nlm.nih.gov/36538348/)

©Marva Foster, Wei Xiong, Lisa Quintiliani, Christine W Hartmann, Stephan Gaehde. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Feasibility, Usability and Acceptability of a mHealth Intervention to Reduce Cardiovascular Risk in Rural Hispanic Adults: Descriptive Study

Sheri Rowland¹, APRN, PhD; Athena K Ramos², MBA, MS, PhD; Natalia Trinidad², MPH; Sophia Quintero², MPH; Rebecca Johnson Beller¹, BSN; Leeza Struwe¹, PhD; Bunny Pozehl³, APRN, PhD

¹College of Nursing, University of Nebraska Medical Center, Lincoln, NE, United States

²College of Public Health, University of Nebraska Medical Center, Omaha, NE, United States

³College of Nursing, University of Nebraska Medical Center, Omaha, NE, United States

Corresponding Author:

Sheri Rowland, APRN, PhD

College of Nursing

University of Nebraska Medical Center

550 N. 19th Street

Lincoln, NE, 68508-0620

United States

Phone: 1 402 472 5959

Fax: 1 402 472 7345

Email: sheri.rowland@unmc.edu

Abstract

Background: Mobile health (mHealth) technology using apps or devices to self-manage health behaviors is an effective strategy to improve lifestyle-related health problems such as hypertension, obesity, and diabetes. However, few studies have tested an mHealth intervention with Hispanic/Latino adults, and no studies were found testing mHealth with rural Hispanic/Latino adults, the fastest-growing population in rural areas.

Objective: The purpose of this study was to evaluate the feasibility, usability, and acceptability of an mHealth cardiovascular risk self-management intervention with rural Hispanic/Latino adults.

Methods: A descriptive study using quantitative and qualitative methods was used to evaluate the feasibility, usability, and acceptability of delivering a 12-week mHealth self-management intervention to reduce cardiovascular risk with rural Hispanic/Latino adults who were randomized to 1 of 2 groups. Both groups were asked to use MyFitnessPal to self-monitor daily steps, weight, and calories. The intervention group received support to download, initiate, and troubleshoot technology challenges with MyFitnessPal (Under Armour) and a smart scale, while the enhanced usual care group received only a general recommendation to use MyFitnessPal to support healthy behaviors. The usability of MyFitnessPal and the smart scale was measured using an adapted Health Information Technology Usability Evaluation Scale (Health-ITUES). Adherence data in the intervention group (daily steps, weight, and calories) were downloaded from MyFitnessPal. Acceptability was evaluated using semistructured interviews in a subsample (n=5) of intervention group participants.

Results: A sample of 70 eligible participants (enhanced usual care group n=34; intervention group n=36) were enrolled between May and December 2019. The overall attrition was 28% at 12 weeks and 54% at 24 weeks. mHealth usability in the intervention group increased at each time point (6, 12, and 24 weeks). Adherence to self-monitoring using mHealth in the intervention group after week 1 was 55% for steps, 39% for calories, and 35% for weights; at the end of the 12-week intervention, the adherence to self-monitoring was 31% for steps, 11% for weight, and 8% for calories. Spikes in adherence coincided with scheduled in-person study visits. Structured interviews identified common technology challenges including scale and steps not syncing with the app and the need for additional technology support for those with limited mHealth experience.

Conclusions: Recruitment of rural Hispanic/Latino adults into the mHealth study was feasible using provider and participant referrals. The use of MyFitnessPal, the smart scale, and SMS text messages to self-monitor daily steps, weights, and calories was acceptable and feasible if technology support was provided. Future research should evaluate and support participants' baseline technology skill level, provide training if needed, and use a phone call or SMS text message follow-ups as a strategy to minimize attrition. A wearable device, separate from the smartphone app, is recommended for activity tracking.

KEYWORDS

mHealth; health behavior; self-management; Hispanic/Latino; rural; apps; feasibility; acceptability; participation; engagement; wearable device; tracking; smartphone

Introduction

Background

Mobile health (mHealth) refers to the use of mobile and wireless technology to achieve a health objective. While there is no standard for what constitutes mHealth, common features used alone or in combination include smartphone apps, SMS text messaging, and wearable monitoring devices. mHealth interventions targeting cardiovascular health are effective in supporting self-management of glycemic control [1,2], blood pressure [3], weight [4], and physical activity behavior [5]. Despite the evidence of mHealth intervention effectiveness, few studies have included Hispanic/Latino adults [6,7], and no studies focused on rural-living Hispanic/Latino adults were found [1]. In urban-living Hispanic/Latino adults with type II diabetes, mHealth improved medication adherence, increased fruit and vegetable consumption, and increased physical activity with SMS text messages 3 times a day for 3 weeks [8]. Another study reported lowered hemoglobin A_{1c} (HbA_{1c}) levels by sending SMS text messages (informational, motivational, and prompting) over 6 months [2].

Hispanic/Latino adults, a population underrepresented in mHealth intervention studies targeting cardiovascular risk, have notable cardiovascular health disparities. In the general population, type II diabetes affects 17% of Hispanic/Latino adults compared to 8% of non-Hispanic White adults [9]. Cardiovascular morbidity and mortality outcomes are worse for the Hispanic/Latino adult population with a greater risk of developing heart failure [10] and a 50% higher likelihood of death from complications related to diabetes than non-Hispanic White adults [11]. Among Hispanic/Latino adults living in rural areas, the prevalence of obesity is higher (Hispanic/Latino adults: 36%; non-Hispanic White adults: 32%), and health status is reported as low or poor (Hispanic/Latino adults: 28%; non-Hispanic White adult: 19%) [12]. Healthy diet and physical

activity behaviors, which can be self-managed using mHealth, are less common among Hispanic/Latino adults, particularly among those with type II diabetes [13,14]. Clearly, the Hispanic/Latino adult population has increased cardiovascular risk that is not adequately managed.

An estimated 19% of the total United States population is Hispanic with the fastest growth of this population observed in rural areas [15]. In the United States, an estimated 79% of Hispanic/Latino adults and 71% of rural-living adults own a smartphone [16]. Most adults in rural areas access the internet daily (76%); however, high-speed internet access is a major concern for rural adults (24%) compared to urban adults (13%) [17]. While Hispanic/Latino adults are noted to use smartphones to text and access email [18], it remains unclear how likely they are to use smartphone apps to actually manage their health [19]. One study in rural-living Hispanic/Latino adults indicated that 81% had the willingness to use mHealth, but few (15%) were aware of how mHealth could support the self-management of chronic health conditions like hypertension and diabetes [20]. A national survey on health app use found that Hispanic/Latino adults were significantly more likely to use a health app compared to non-Hispanic adults, and the most common feature desired was the ability to communicate with a health care provider through the health app [21].

Theoretical/Conceptual Framework

The intervention in this pilot study was based on the concept of chronic disease self-management. Specific self-management skills include decision-making (what is my priority health problem?), problem-solving (what can I do to manage this problem?), action-taking (how do I begin managing this problem?), accessing resources, and partnering with a health care provider [22]. Goal setting theory [23] and self-efficacy, a concept of social cognitive theory [24], support the mHealth intervention to build self-management skills and establish new patterns of healthy behavior (Textbox 1).

Textbox 1. mHealth intervention components.**Goal setting**

- Priority health conditions selected
 - Participant and nurse practitioner (NP) determine together
- Investigator-set participant goals
 - Weigh daily using scale synced to MyFitnessPal (MFP)
 - Log all food and drink consumed in MFP
 - Achieve prescribed daily calorie target based on age, gender, and activity level [25]
 - Walk 10,000 steps per day using MFP step tracker

NP visit

- Baseline-15 min
- 6 weeks-15 min
- 12 weeks-15 min

Self-efficacy

- Physiologic feedback
 - Biometrics, lab tests, and fitness test performed
 - Results reviewed and personalized report given to the participant

NP visit

- Baseline-15 min
- 6 weeks-15 min
- 12 weeks-15 min
- Verbal persuasion

SMS text messages to prompt MFP use, provide health information on priority health conditions, and encouragement

- Twice weekly via Remind smartphone app
- Mastery experiences
 - Use digital scale to see sync to MFP
 - Find daily step count in MFP
 - Practice logging food/drink items in MFP
- Tech support visit
 - Baseline-30 min
 - 6 weeks-15 min
 - 12 weeks-15 min

Purpose

The purpose of this study was to evaluate the feasibility, usability, and acceptability of a stand-alone mHealth cardiovascular risk self-management intervention using MyFitnessPal, a compatible smart scale, and an engaged bilingual health care provider in a rural Hispanic/Latino adult population.

Given the limited number of mHealth interventions to reduce cardiovascular risk in the Hispanic/Latino adult population, it is important to understand what aspects of mHealth interventions are most engaging and usable.

Methods

Study Design

A descriptive study design using quantitative and qualitative methods was used to evaluate the feasibility, usability, and acceptability of a cardiovascular risk self-management mHealth intervention conducted with rural Hispanic/Latino adults between May 2019 and June 2020. The 12-week intervention was tested using an unblinded randomized 2-group design. Both groups were asked to use MyFitnessPal to self-monitor daily steps, weight, and calories. The intervention group received support to download, initiate, and troubleshoot technology challenges with MyFitnessPal and a smart scale, while the enhanced usual care group received only a general recommendation to use MyFitnessPal to support healthy behaviors. This report details the feasibility and acceptability of the mHealth intervention and compares mHealth usability between the 2 groups. The cardiometabolic health outcomes in both groups are reported separately [26].

Sample and Setting

A sample of 70 rural-living Hispanic/Latino adults was recruited from 2 rural communities in Nebraska. Inclusion criteria were Hispanic/Latino adults (aged 19-65 years); English or Spanish speaking; living within 50 miles of either community; self-report hypertension, obesity, type II diabetes, or dyslipidemia; having a smartphone; and using Bluetooth or Wi-Fi. Exclusion criteria were current participation in a health behavior program (eg, Weight Watchers), BMI \geq 45 unless a clinician granted permission, or high risk for a cardiac event during fitness testing. Pregnant women were excluded from the step fitness test, BMI, and waist measurement.

In this study, a minimum sample of 54 (27 per group) was determined by setting the desired margin of error for the mean metabolic equivalents (METs), the primary outcome of cardiorespiratory fitness, at no more than 0.80. Using a standard deviation for walking METs of 2.00, a minimum sample size of 27 per group yields a margin of error of 0.75. A margin of error assumed a 95% confidence interval. With an assumed attrition rate of 30%, the total sample goal was 70 (35 per group), which is appropriate for pilot work, allowing for descriptive statistics, estimation of effect sizes, and hypothesis generation [27].

Intervention Development

To support intervention development and delivery, a community advisory board was established. Members of the board included two Hispanic/Latino community residents and representatives of the local federally qualified health clinic, local public health department, Hispanic community center, and educational service unit. The board had 2 in-person meetings and communicated through email to advise on community resources, recruitment strategies, and the cultural relevancy of the SMS text messages that were developed for the intervention group.

mHealth Intervention Group

The intervention group received the MyFitnessPal premium version (US \$50), the Withings Body+ smart scale (US \$75),

consultation with a bilingual nurse practitioner and tech support person, and SMS text messages twice a week via the free Remind app in the participant preferred language (English or Spanish) (Textbox 1). MyFitnessPal premium was selected because of the availability in Spanish, the large food database (6 million items), and the capacity to track adherence through a “diary share” feature and a downloadable CSV data file. Bilingual tech support was provided to initialize all technology including assisting with downloading and setting up the MyFitnessPal app on participants’ phones. During the 30-minute tech support visit, each participant received verbal and written instructions on how to use both the app and the smart scale. The teach-back method was used to confirm understanding.

Enhanced Usual Care Group

This group received a general recommendation to use the free version of MyFitnessPal to support a healthy lifestyle, but they did not receive technical support to download or initiate the app. At the 6-week visit, the participants had a weight and blood pressure check and were again encouraged to use MyFitnessPal but were not provided technical support. At the last 24-week visit, this group received the smart scale, a consultation with the bilingual nurse practitioner, and a visit with a tech support person.

Procedures

To standardize the intervention the following were used: a training manual with topic checklist, set time allowances for study activities, a script for participant visits, and role-playing participant visits prior to beginning recruitment. Participants were recruited using bilingual flyers posted in the community, zip code–targeted Facebook ads, and within the local clinic. A study coordinator took calls from those interested in the study, discussed eligibility, and scheduled study visits at either the mobile clinic or the local community center. Following informed consent and baseline measures, the participants were randomized to either the enhanced usual care group or the intervention group. Both groups had visits at baseline, 12 weeks, and 24 weeks to complete paper questionnaires, collect lab tests (HbA_{1c} and lipids), measure biometrics (blood pressure, BMI, and waist circumference), and complete a 2-minute step test to evaluate cardiorespiratory fitness. At a 6-week booster visit, weight and blood pressure were measured. Study visits were conducted in a reserved area of a community center, a conference room in a large family practice clinic, or in a small community clinic during off-clinic hours.

Ethics Approval

The study was approved by the institutional review board at the University of Nebraska Medical Center (reference number: 26). Informed consent was completed in person and in the language preferred by the participant (English or Spanish). Personal identifying information collected was limited to name, postal address, date of birth, and telephone number, and these data were stored in a locked file box and on a secure, password-protected server. Participant data were collected using paper instruments identified with a unique identification number. Participants were given a US \$25 gift card for each completed visit.

Measures

Usability

An adapted Health Information Technology Usability Evaluation Scale (Health-ITUES) was used to evaluate the usability of the MyFitnessPal app and compatible smart scale. This 20-item tool uses a 5-point Likert scale (strongly disagree to strongly agree) across 4 subscales: impact on daily life, perceived usefulness, perceived ease of use, and user control [28]. Examples of items for each of the subscales include, “I think MyFitnessPal and smart scale would improve the quality of life for a person living with chronic cardiovascular conditions,” “Using MyFitnessPal and smart scale makes easier to self-manage my cardiovascular-related conditions,” “I find MyFitnessPal premium and smart scale easy to use,” and “Whenever I make a mistake while using MyFitnessPal and smart scale, I recover easily and quickly.” The scale has an overall mean possible score range of 20-100 with higher scores indicating higher usability. In our study, the overall scale had excellent reliability, Cronbach $\alpha=.94$. Subscales also had good reliability with Cronbach α ranging from .82 to .97. Participants completed this measure in paper form in their preferred language (Spanish or English). The adapted Health-ITUES was completed by all intervention group participants at 6, 12, and 24 weeks. In the enhanced usual care group, only those who reported downloading/using MyFitnessPal were asked to complete the adapted Health-ITUES at 6, 12, and 24 weeks.

Adherence

Adherence was only measured in the intervention group because only the premium version of MyFitnessPal allows access to usage data. An Excel file reporting on steps (daily total), weight (by date), and calories (data total) was accessed for each intervention group participant.

Steps

Participants were asked to carry their phones with them throughout the day to track steps using the MyFitnessPal step tracker and work toward 10,000 steps per day. The step tracker is an automatic feature that was turned on at the first tech support visit.

Weight

Participants were asked to weigh themselves naked each morning using the scale that was synced to MyFitnessPal at the first tech support visit. Participants were shown how to manually enter a weight into MyFitnessPal in the event an automatic sync did not occur.

Calories

Participants were asked to log all food and nonwater beverages consumed each day using the MyFitnessPal diet tracker. They were also asked to work toward achieving a daily calorie limit based on their age, gender, and activity level [25].

Acceptability

A stratified sampling approach was used to select intervention group participants for a 15-minute semistructured interview on the acceptability of mHealth upon completion of the intervention. Interviewees were selected based on gender (male/female), age (under/over 40), and level of engagement during the study (high/low). Participants were asked specifically about their experience with (1) MyFitnessPal to track steps and calories, (2) the smart scale to track daily weight, and (3) twice weekly SMS text messages. The interviews were conducted face-to-face in a private area of the data collection space (conference room and clinic room) in the language preferred by the participant.

Analysis

Quantitative data were examined for missingness, errors, normality, and equality of variances. Continuous variables were analyzed using means and standard deviations. Categorical variables were analyzed using frequencies and percentages. Only participants who responded to a measure at a specific time point were included in analyses; therefore, the number reported at each time point varied. Qualitative data were collected through audiorecorded interviews that were transcribed verbatim in Spanish and then translated to English for thematic analysis [29].

Results

Feasibility

During enrollment (May-December 2019), 77 people expressed interest in participating, but 7 were not enrolled because they lacked a qualifying health condition. Of the 70 participants who enrolled, 45 were referred by someone already in the study, and 25 were referred by their health care provider. Table 1 details the sample demographic characteristics by group. Group differences in the demographic data were assessed with independent *t* tests, chi-square tests, and Fisher exact tests. No demographic variables were significantly different except for the preferred language in the study, where 77% (26/34) of the enhanced usual care group and 94% (34/36) of the intervention group preferred to complete the study in Spanish (questionnaires, informational SMS text messages, and MyFitnessPal; Fisher exact test $P=.04$).

The overall attrition was 28% ($n=19$) at 12 weeks and 54% ($n=38$) at 24 weeks. Attrition by group was similar at 12 weeks (enhanced usual care: $n=9$, 26%; intervention: $n=11$, 31%) and 24 weeks (enhanced usual care: $n=18$, 53%; intervention: $n=21$, 57%). Although there was missing data, this was primarily due to physical limitations with biometric testing, a mechanical limitation of blood analyzer (result too high to read), unanswered questions, and suspension of in-person data collection due to COVID-19 during the last 2 months of the study. Instead, phone calls were made to 15 participants to collect responses to the questionnaires and a self-reported weight.

Table 1. Baseline participant characteristics by group.

	Enhanced usual care group (n=34)		Intervention group (n=36)	
	Responders, n	Value	Responders, n	Value
Age (years), mean (SD)	34	40.8 (9.9)	36	41.4 (9.8)
Female, n (%)	34	23 (68)	36	31 (86)
Married/cohabitating, n (%)	33	27 (82)	35	29 (83)
Education, n (%)				
Less than 12th grade	34	21 (58)	36	16 (47)
More than 12th grade	34	15 (44)	36	18 (50)
Country of origin, n (%)				
United States	34	9 (26)	35	7 (19)
Mexico	34	21 (62)	35	19 (53)
Guatemala	34	1 (3)	35	4 (11)
El Salvador	34	2 (6)	35	3 (8)
Honduras	34	1 (3)	35	2 (6)
Language, n (%)				
English proficient	33	11 (33)	35	9 (26)
Preferred Spanish during the study	34	26 (77)	36	34 (94)
Household income per year (US \$), n (%)				
<10,000	33	8 (24)	35	6 (17)
10,000-50,999	33	23 (68)	35	22 (61)
≥60,000	33	2 (6)	35	3 (8)
Unemployed	33	16 (47)	35	17 (47)
Number of people living in a household, n (%)				
1-3	34	5 (15)	33	7 (19)
4	34	15 (44)	33	13 (36)
≥5	34	14 (41)	33	13 (36)
No health insurance, n (%)	33	18 (53)	36	18 (50)

Usability

Usability was higher in the intervention group at all 3 time points and increased at each time point (6, 12, and 24 weeks). Fewer participants in the enhanced usual care group completed the usability measure compared to the intervention group at all

3 time points. This was due to only collecting the Health-ITUES from those who reported downloading the free version of MyFitnessPal (Table 2). Given the small numbers in the enhanced usual care group, descriptive statistics are reported rather than a statistical test of difference between groups.

Table 2. Perceived usability of mHealth by groups.

	Enhanced usual care ^a (n=34)				Intervention (n=36)			
	Responders, n	Min	Max	Score, mean (SD)	Responders, n	Min	Max	Score, mean (SD)
At 6 weeks								
Impact on daily life	3	6	15	9.33 (4.93)	24	5	15	12.71 (3.06)
Usefulness	3	18	45	27.33 (15.31)	24	18	45	35.92 (9.36)
Ease of use	3	14	25	21.33 (6.35)	23	11	25	20.43 (5.09)
User control	3	7	15	11.33 (4.04)	24	3	15	11.50 (3.74)
Composite score	3	45	100	69.33 (28.04)	21	56	100	82.29 (15.47)
At 12 weeks								
Impact on daily life	8	11	15	13.50 (1.69)	23	9	15	13.78 (2.087)
Usefulness	8	27	45	36.75 (6.39)	22	24	45	40.45 (7.29)
Ease of use	8	15	25	22.50 (3.54)	23	7	25	21.26 (5.21)
User control	8	8	15	12.38 (2.77)	23	7	15	12.69 (2.82)
Composite score	8	66	100	85.13 (12.30)	22	48	100	88.23 (15.64)
At 24 weeks								
Impact on daily life	6	11	15	12.83 (1.72)	22	11	15	13.91 (1.44)
Usefulness	6	31	45	36.67 (6.59)	20	31	45	42.10 (4.22)
Ease of use	6	20	25	22.67 (2.58)	22	12	25	21.59 (4.35)
User control	6	9	15	11.83 (2.04)	21	6	15	11.91 (3.22)
Composite score	6	76	100	84.00 (9.44)	19	71	100	90.95 (10.48)

^aOnly those who reported downloading MyFitnessPal were asked to complete usability measure.

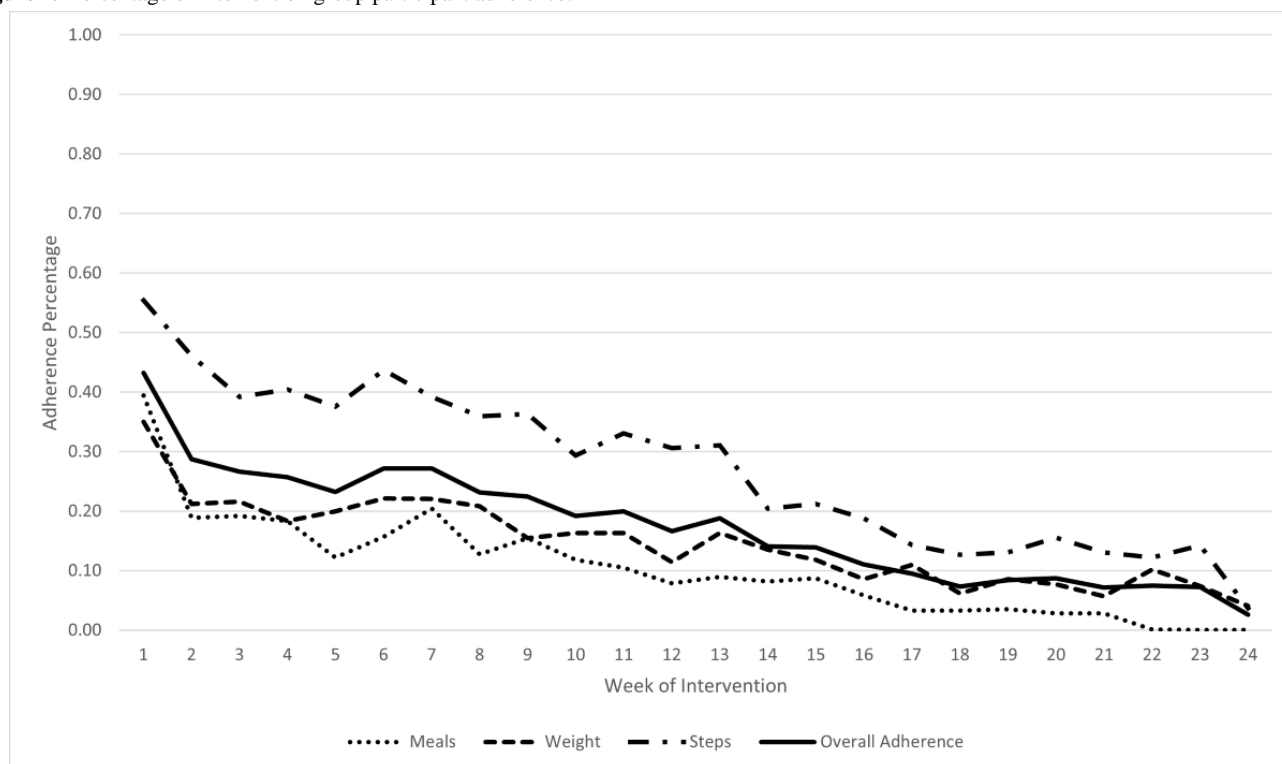
Adherence

Intervention group adherence data were obtained from the MyFitnessPal premium app, which reported a daily step total, weight by date, and a daily calorie total. A dichotomous variable was created to indicate the presence or absence of captured data for each of the self-monitoring activities: daily steps, weight, and calories/meal (breakfast, lunch, and dinner). Weekly adherence for steps, weight, and calories was determined by the number of days per week of data capture. In this study, participants were adherent to mHealth self-monitoring if there

was data captured ≥ 4 days per week for each week during the intervention (weeks 1-12) and postintervention (weeks 13-24).

Steps

In the first week of data collection, 55% (n=20) of the intervention group participants were adherent to the daily step capture using MyFitnessPal premium. Step adherence fell gradually to 31% (n=11) at the end of the intervention at week 12. Step adherence continued to gradually decline to 0% at 24 weeks. An observable spike in step adherence was noted at weeks 6 and 11 (Figure 1). The increases in step adherence coincide with in-person study visits at weeks 6 and 12.

Figure 1. Percentage of intervention group participant adherence.

Weights

In the first week of data collection, 35% (n=12) of the intervention group participants were adherent to the daily weight capture using MyFitnessPal premium and compatible smart scale. Weight adherence trended down to 11% (n=4) at the end of the 12-week intervention and to 0% at week 24. Similar to step adherence, there was a spike in weight adherence at weeks 5 and 11, which corresponds with in-person study visits at weeks 6 and 12, respectively.

Calories

In the first week of data collection, 39% (n=14) of the intervention group participants were adherent to logging food and nonwater beverages consumed each day using the MyFitnessPal diet tracker. Food logging trended down to 8% (n=3) for the intervention group participants at week 12 and to 0% at 24 weeks.

Acceptability

Five participants completed an interview about the intervention. Three interviewees were low engagers (2 female participants and 1 male participant) and 2 were high engagers (1 female participant and 1 male participant). Three themes were identified: (1) mHealth is useful, (2) mHealth has challenges, and (3) changes may be needed for success with a future study (Textbox 2). Participants described mHealth as useful by creating “self-awareness,” “personal accountability,” “empowerment,” and “motivation” to work on their health behaviors. The Spanish setting in MyFitnessPal and Spanish SMS text messages were appropriate and understandable. One participant remarked, “it [mHealth] works not only for you but for the other members of your family.” Participants also described exploring and using other mHealth features including macronutrient information, BMI calculation, and the visual weight trend on the scale.

Textbox 2. Intervention group participant quotes.

<p>Mobile health (mHealth) is useful</p> <ul style="list-style-type: none"> • “I was in control of my food, I was able to register it, and I was able to give myself feedback with regard to what I was doing. I was able to see it there, in real time or what I was eating, if I was losing [weight], if I was doing enough exercise.” (ID 6) • “It helps me control my food intake because I can see that I am close to reaching my daily goal.” (ID 2) • “They [text messages] make me feel important. And also, it motivates me to keep putting in the effort.” (ID 49) • “You notice whether or not you exercised that day...if I was not able today, I can do it tomorrow.” (ID 40) <p>mHealth has challenges</p> <ul style="list-style-type: none"> • “It [MyFitnessPal] did not have the data with typical Latino foods.... For example, there was fried Asian rice, but it is not the same as fried Latino rice.” (ID 6) • “Registering the foods...I think it is harder to do because we are Latinos and we cook in large quantities and we serve many portions.” (ID 13) • “It is a bit complicated to be able to register my foods.” (ID 49) • “I always carry my phone with me, but I don’t always have the location on. Therefore, I knew that sometimes I personally had walked, let us say, 10,000 steps, but only 7,000 or 6,000 were registered—it did not help me to stay motivated because I would say to myself “I did so much, and it only registered 6,000 steps.” (ID 6) • “I had to restart the scale five times along with my phone.” (ID 6) <p>Recommendations for a future study</p> <ul style="list-style-type: none"> • “You need to give yourself time to register or to learn because I really have not learned to do it well.” (ID 40) • “Explain to us a bit more about the portions that we need to eat during the day.” (ID 40) • “If I had had a different device, I think I would have liked it better. But to just use the phone to keep track of steps, no.” (ID 6)

Participants experienced disruptions in the automatic sync function of the app, resulting in steps and weights that were not captured. These technical issues would either be ignored, managed, or reported to the technology support person during in-person meetings. The accuracy was questioned by participants. They reported the counter was not picking up steps or they may not have had their phone with them when being active. The accuracy of calorie tracking was also questioned because portion size estimation was difficult, and the food database had limited ethnic foods. Three participants noted time as a barrier to logging food into the app.

The low mHealth engagers suggested more training or orientation for those who were not as skilled with smartphone use. Additional training was also suggested for entering recipes into MyFitnessPal so that homemade food could be accurately tracked. Most of the interviewees “liked” the SMS text messages and felt they were “good” and helpful to “remind me to do something for my health.” One participant did not find the SMS text messages “motivating” and suggested a feature to “opt-out” of receiving messages. A wearable activity monitor instead of the MyFitnessPal app was suggested to increase the accuracy of activity tracking.

Discussion

To our knowledge, this is the first study to evaluate the feasibility, usability, and acceptability of a randomized mHealth cardiovascular risk self-management intervention with rural Hispanic/Latino adults. The feasibility of recruitment or enrollment was evaluated as adequate, and participant retention was evaluated as mixed. The low number of participants in the

enhanced usual care group who downloaded and used the MyFitnessPal app as advised suggests that the mere recommendation by a health care provider to use a widely available commercial app to self-manage health behaviors is not enough to trigger and sustain mHealth use. Among those who received mHealth support in the intervention group, the usability of the MyFitnessPal app and smart scale to self-monitor daily steps, weight, and calorie intake was high and increased over time. Adherence to using the app and scale followed a typical gradual decline trajectory with upticks corresponding with in-person visits. Indications that the mHealth intervention was not acceptable were not detected in the subsample of intervention group participants who completed semistructured interviews.

Engaged health care providers and referrals from enrolled participants allowed the target sample of 70 participants to be enrolled over an 8-month period. For a larger, fully powered trial, it will be important to ensure buy-in and support of key local health care providers and consider incentivizing enrolled participants to refer others from their social network to participate. While the COVID-19 pandemic likely contributed to the high attrition rate at 24 weeks (n=38, 54%), the attrition rate at 12 weeks (n=19, 28%) was higher than the average rate of 18% observed in health behavior trials [30] but less than the 40% observed in app-based interventions targeting chronic disease management [31]. In a 24-week mHealth intervention targeting glycemic control in 126 urban-living Hispanic/Latino adults, 16% (n=11) of the intervention group and 5% (n=3) of the usual care group did not complete any follow-ups [2]. Participants presented for in-person data collection at baseline, 3 months, and 6 months; however, a study coordinator called

the patient if a blood glucose reading exceeded a set range or if a blood sugar reading was not received for 1 week [2]. Phone calls, to build *personalismo*, may be a particularly culturally relevant strategy for reaching and maintaining contact with Hispanic study participants [32]. Future studies may consider the measurement and evaluation of *personalismo* in relation to mHealth engagement by Hispanic participants [32].

Increasing usability, as measured by the adapted Health-ITUES, over 3 time points within the intervention group is an indication that with time and experience participants become more comfortable, and therefore, mHealth may grow more useful for managing one's own health. Difficulties with app use reported by low engagers during the structured interviews indicate a need to evaluate participants on their skill and confidence with mHealth technology prior to using this type of mHealth in the future. Even when participants owned a smartphone and used the SMS text messaging feature, it did not ensure competency with a commercially available health app like MyFitnessPal. Others have found that Hispanics may not have the knowledge to fully benefit from prevalent functionalities (ie, calorie tracker, macronutrient information, and sharing food logs with peers for accountability) offered by health apps [33]. An eHealth Literacy Assessment Toolkit is available, which includes reliable and valid methods of assessing not only health literacy but also technology familiarity, confidence, and incentives for engagement [34]. For future studies, assessing technology skills and confidence at the beginning of the intervention would allow for adjustment and tailoring of the mHealth training and orientation to better meet participant needs. In the enhanced usual care group, very few participants downloaded the free version of MyFitnessPal as advised. This limited the feasibility of comparing usability outcomes between groups, and future studies should consider alternative designs.

mHealth adherence in this study followed a trajectory similar to that observed in other intervention studies using a health app to self-monitor diet, weight, and activity in both rural and urban adults [35,36]. Specifically, adherence waned over the course of the study with sharper declines following the end of the intervention. In this study, adherence increases were observed around 6, 12, and 24 weeks when in-person visits occurred. It may be that the in-person visits reminded participants that they were being observed, and therefore, they altered their behavior. Group differences in the theoretical measures of self-management (self-efficacy, activation, and self-regulation) are reported in [26]. Although not significant, a medium intervention effect on self-regulation was observed at 12 and 24 weeks. In this study, the intervention group received in-person visits with a bilingual nurse practitioner and technology support person, which may have facilitated the development of self-monitoring skills (weight, calorie intake, and physical activity).

The qualitative structured interviews informed on acceptability of the mHealth intervention. While participants may have been reluctant to be overly critical of the study with the research team, indications of complete unacceptability were not detected. Low engagers acknowledged their own barriers to using the self-monitoring tools (MyFitnessPal app and smart scale) as well as provided suggestions on how to reduce barriers to mHealth usage. Acceptability may be enhanced by providing those with limited experience or low confidence with technology additional training, use of a wearable activity monitor instead of using an app to capture steps, training/troubleshooting on entering recipes into a food database, and culturally tailored education on portion size. Other methods of evaluating acceptability should be considered. For example, ecological momentary assessments (EMAs) provide a real-time data collection strategy delivered in the user's natural environment. Among low-income Spanish-speaking adults, EMAs and language content analysis of brief SMS text message responses allowed for a deeper understanding of participant responses to depression treatment [37].

A limitation of this study was that adherence data were not collected in the enhanced usual care group, as the free version of MyFitnessPal does not have data extraction features. A future study would need to have adherence data collected from all participants to examine theoretical drivers of adherence. Adherence increases may also be explained by technology fixes that occurred during in-person visits such as re-establishing the capture of steps and scale and app syncing, although the increases are noted to begin the week prior to in-person visits. We acknowledge that the broad use of "Hispanic/Latino" does not recognize the full diversity of the population in the United States. Future studies should include additional demographic variables and a sample large enough to be able to denote specific subgroup differences.

Despite the increased use of mHealth in the United States, few culturally and linguistically tailored mHealth interventions have been tested with the Hispanic/Latino adult population. We found that a mHealth intervention using MyFitnessPal, a smart scale, and SMS text messages with rural Hispanic/Latino adults to self-monitor daily steps, weights, and calories was acceptable and feasible if technology support was provided. Like most mHealth interventions with predominately White populations, mHealth engagement with our Hispanic/Latino adult sample waned over time, and culturally tailored strategies to combat attrition are recommended. Professionals who are considering using mHealth technologies in practice should tailor them according to individual skill level, experience, and confidence. App developers should consider cultural and linguistic tailoring to meet the needs of diverse populations. In conclusion, mHealth can be leveraged to promote public health and help patients self-manage cardiovascular risk factors, particularly in the absence or limited supply of health care providers in rural communities.

Acknowledgments

This project was funded by the Center for Patient, Family, and Community Engagement in Chronic Care Management Center at the University of Nebraska Medical Center. We thank the participants for their willingness to engage in research and Good Neighbor Community Health Center in Columbus, Nebraska for supporting recruitment and data collection. Special thanks to Reina Griffith, Kevin Harm, Ricardo Sanabria, and Chuck Sepers for their valuable contributions.

Data Availability

Access to raw data in .sav format is available upon request.

Conflicts of Interest

None declared.

References

1. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. *Adv Nutr* 2017 May 15;8(3):449-462 [FREE Full text] [doi: [10.3945/an.116.014100](https://doi.org/10.3945/an.116.014100)] [Medline: [28507010](https://pubmed.ncbi.nlm.nih.gov/28507010/)]
2. Fortmann AL, Gallo LC, Garcia MI, Taleb M, Euyoque J, Clark T, et al. Dulce digital: an mHealth SMS-based intervention improves glycemic control in Hispanics with type 2 diabetes. *Diabetes Care* 2017 Oct;40(10):1349-1355 [FREE Full text] [doi: [10.2337/dc17-0230](https://doi.org/10.2337/dc17-0230)] [Medline: [28600309](https://pubmed.ncbi.nlm.nih.gov/28600309/)]
3. Delva S, Waligora Mendez KJ, Cajita M, Koirala B, Shan R, Wongvibulsin S, et al. Efficacy of mobile health for self-management of cardiometabolic risk factors. *J Cardiovasc Nurs* 2020 Feb 10;36(1):34-55 [FREE Full text] [doi: [10.1097/jcn.0000000000000659](https://doi.org/10.1097/jcn.0000000000000659)]
4. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2018 Jan 17;6(1):e23. [doi: [10.2196/mhealth.8873](https://doi.org/10.2196/mhealth.8873)]
5. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. *Sports Med Open* 2018 Sep 03;4(1):42 [FREE Full text] [doi: [10.1186/s40798-018-0157-9](https://doi.org/10.1186/s40798-018-0157-9)] [Medline: [30178072](https://pubmed.ncbi.nlm.nih.gov/30178072/)]
6. Arambepola C, Ricci-Cabello I, Manikavasagam P, Roberts N, French DP, Farmer A. The impact of automated brief messages promoting lifestyle changes delivered via mobile devices to people with type 2 diabetes: a systematic literature review and meta-analysis of controlled trials. *J Med Internet Res* 2016 Apr 19;18(4):e86 [FREE Full text] [doi: [10.2196/jmir.5425](https://doi.org/10.2196/jmir.5425)] [Medline: [27095386](https://pubmed.ncbi.nlm.nih.gov/27095386/)]
7. Mateo GF, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. *J Med Internet Res* 2015 Nov 10;17(11):e253 [FREE Full text] [doi: [10.2196/jmir.4836](https://doi.org/10.2196/jmir.4836)] [Medline: [26554314](https://pubmed.ncbi.nlm.nih.gov/26554314/)]
8. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. *Diabetes Technol Ther* 2012 Jun;14(6):492-496. [doi: [10.1089/dia.2011.0252](https://doi.org/10.1089/dia.2011.0252)] [Medline: [22524591](https://pubmed.ncbi.nlm.nih.gov/22524591/)]
9. Hispanic/Latino Americans and type 2 diabetes. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/diabetes/library/features/hispanic-diabetes.html> [accessed 2022-04-05]
10. Metha H, Armstrong A, Swett K, Shah S, Allison M, Hurwitz B. Burden of systolic and diastolic left ventricular dysfunction among Hispanics in the United States: insights from the Echocardiographic Study of Latinos (ECHO-SOL). *Circ Heart Fail* 2016;9(4):e002733 [FREE Full text] [doi: [10.1161/circheartfailure.115.002733](https://doi.org/10.1161/circheartfailure.115.002733)]
11. ¡A la Buena Salud! ? To good health!. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/vitalsigns/hispanic-health/index.html> [accessed 2018-08-05]
12. James CV, Moonesinghe R, Wilson-Frederick SM, Hall JE, Penman-Aguilar A, Bouye K. Racial/ethnic health disparities among rural adults—United States, 2012–2015. *MMWR Surveill Summ* 2017 Nov 17;66(23):1-9 [FREE Full text] [doi: [10.15585/mmwr.ss6623a1](https://doi.org/10.15585/mmwr.ss6623a1)] [Medline: [29145359](https://pubmed.ncbi.nlm.nih.gov/29145359/)]
13. Nothwehr F, Stump T. Health-promoting behaviors among adults with type 2 diabetes: findings from the health and retirement study. *Prevent Med* 2000 May;30(5):407-414. [doi: [10.1006/pmed.2000.0658](https://doi.org/10.1006/pmed.2000.0658)]
14. Nwasuruba C, Khan M, Egede LE. Racial/ethnic differences in multiple self-care behaviors in adults with diabetes. *J Gen Intern Med* 2007 Jan 9;22(1):115-120 [FREE Full text] [doi: [10.1007/s11606-007-0120-9](https://doi.org/10.1007/s11606-007-0120-9)] [Medline: [17351850](https://pubmed.ncbi.nlm.nih.gov/17351850/)]
15. U.S. Census Bureau QuickFacts: United States. US Census Bureau. URL: <https://www.census.gov/quickfacts/fact/table/US/RHI725219> [accessed 2021-08-24]
16. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-01-17]

17. Digital gap between rural and nonrural America persists. Pew Research Center. 2021. URL: <https://www.pewresearch.org/fact-tank/2019/05/31/digital-gap-between-rural-and-nonrural-america-persists/> [accessed 2021-01-17]
18. Price M, Williamson D, McCandless R, Mueller M, Gregoski M, Brunner-Jackson B, et al. Hispanic migrant farm workers' attitudes toward mobile phone-based telehealth for management of chronic health conditions. *J Med Internet Res* 2013 Mar 30;15(4):e76. [doi: [10.2196/jmir.2500](https://doi.org/10.2196/jmir.2500)]
19. Racial and ethnic differences in how people use mobile technology. Pew Research Center. 2021. URL: <https://www.pewresearch.org/fact-tank/2015/04/30/racial-and-ethnic-differences-in-how-people-use-mobile-technology/> [accessed 2021-01-17]
20. Price M, Williamson D, McCandless R, Mueller M, Gregoski M, Brunner-Jackson B, et al. Hispanic migrant farm workers' attitudes toward mobile phone-based telehealth for management of chronic health conditions. *J Med Internet Res* 2013 Mar 30;15(4):e76 [FREE Full text] [doi: [10.2196/jmir.2500](https://doi.org/10.2196/jmir.2500)] [Medline: [23624105](https://pubmed.ncbi.nlm.nih.gov/23624105/)]
21. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. *JMIR Mhealth Uhealth* 2015 Nov 04;3(4):e101 [FREE Full text] [doi: [10.2196/mhealth.4924](https://doi.org/10.2196/mhealth.4924)] [Medline: [26537656](https://pubmed.ncbi.nlm.nih.gov/26537656/)]
22. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [doi: [10.1207/s15324796abm2601_01](https://doi.org/10.1207/s15324796abm2601_01)]
23. Latham GP, Locke EA. Self-regulation through goal setting. *Organ Behav Hum Decis Process* 1991 Dec;50(2):212-247. [doi: [10.1016/0749-5978\(91\)90021-K](https://doi.org/10.1016/0749-5978(91)90021-K)]
24. Bandura A. Health promotion from the perspective of social cognitive theory. *Psychol Health* 1998 Jul;13(4):623-649. [doi: [10.1080/08870449808407422](https://doi.org/10.1080/08870449808407422)]
25. 2015-2020 Dietary Guidelines for Americans. US Department of Health & Human Services. URL: <http://health.gov/dietaryguidelines/2015/guidelines/> [accessed 2022-12-03]
26. Rowland SA, Ramos AK, Trinidad N, Quintero S, Johnson-Beller R, Struwe L, et al. mHealth intervention to improve cardiometabolic health in rural Hispanic adults. *J Cardiovasc Nurs* 2022 Jan 12;37(5):439-445. [doi: [10.1097/jcn.0000000000000882](https://doi.org/10.1097/jcn.0000000000000882)]
27. Hertzog MA. Considerations in determining sample size for pilot studies. *Res Nurs Health* 2008 Apr;31(2):180-191. [doi: [10.1002/nur.20247](https://doi.org/10.1002/nur.20247)] [Medline: [18183564](https://pubmed.ncbi.nlm.nih.gov/18183564/)]
28. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. *JMIR Mhealth Uhealth* 2018 Jan 05;6(1):e4 [FREE Full text] [doi: [10.2196/mhealth.8851](https://doi.org/10.2196/mhealth.8851)] [Medline: [29305343](https://pubmed.ncbi.nlm.nih.gov/29305343/)]
29. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
30. Crutzen R, Viechtbauer W, Spigt M, Kotz D. Differential attrition in health behaviour change trials: a systematic review and meta-analysis. *Psychol Health* 2015 Jan;30(1):122-134 [FREE Full text] [doi: [10.1080/08870446.2014.953526](https://doi.org/10.1080/08870446.2014.953526)] [Medline: [25109224](https://pubmed.ncbi.nlm.nih.gov/25109224/)]
31. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. *J Med Internet Res* 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: [10.2196/20283](https://doi.org/10.2196/20283)] [Medline: [32990635](https://pubmed.ncbi.nlm.nih.gov/32990635/)]
32. Davis RE, Lee S, Johnson TP, Rothschild SK. Measuring the elusive construct of among Mexican American, Puerto Rican, and Cuban American adults. *Hisp J Behav Sci* 2019 Feb 10;41(1):103-121 [FREE Full text] [doi: [10.1177/0739986318822535](https://doi.org/10.1177/0739986318822535)] [Medline: [34135543](https://pubmed.ncbi.nlm.nih.gov/34135543/)]
33. Williams JP, Schroeder D. Popular glucose tracking apps and use of mHealth by Latinos with diabetes: review. *JMIR Mhealth Uhealth* 2015 Aug 25;3(3):e84 [FREE Full text] [doi: [10.2196/mhealth.3986](https://doi.org/10.2196/mhealth.3986)] [Medline: [26307533](https://pubmed.ncbi.nlm.nih.gov/26307533/)]
34. Karnoe A, Furstrand D, Christensen KB, Norgaard O, Kayser L. Assessing competencies needed to engage with digital health services: development of the eHealth literacy assessment toolkit. *J Med Internet Res* 2018 May 10;20(5):e178 [FREE Full text] [doi: [10.2196/jmir.8347](https://doi.org/10.2196/jmir.8347)] [Medline: [29748163](https://pubmed.ncbi.nlm.nih.gov/29748163/)]
35. Jospe MR, Roy M, Brown RC, Williams SM, Osborne HR, Meredith-Jones KA, et al. The effect of different types of monitoring strategies on weight loss: a randomized controlled trial. *Obesity (Silver Spring)* 2017 Sep 13;25(9):1490-1498. [doi: [10.1002/oby.21898](https://doi.org/10.1002/oby.21898)] [Medline: [28703448](https://pubmed.ncbi.nlm.nih.gov/28703448/)]
36. Reed JR, Struwe L, Bice MR, Yates BC. The impact of self-monitoring food intake on motivation, physical activity and weight loss in rural adults. *Appl Nurs Res* 2017 Jun;35:36-41. [doi: [10.1016/j.apnr.2017.02.008](https://doi.org/10.1016/j.apnr.2017.02.008)] [Medline: [28532724](https://pubmed.ncbi.nlm.nih.gov/28532724/)]
37. Hernandez-Ramos R, Altszyler E, Figueroa CA, Avila-Garcia P, Aguilera A. Linguistic analysis of Latinx patients' responses to a text messaging adjunct during cognitive behavioral therapy for depression. *Behav Res Ther* 2022 Mar;150:104027 [FREE Full text] [doi: [10.1016/j.brat.2021.104027](https://doi.org/10.1016/j.brat.2021.104027)] [Medline: [35026608](https://pubmed.ncbi.nlm.nih.gov/35026608/)]

Abbreviations

EMA: ecological momentary assessment

HbA_{1c}: hemoglobin A_{1c}

Health-ITUES: Health Information Technology Usability Evaluation Scale

MFP: MyFitnessPal
mHealth: mobile health

Edited by A Mavragani; submitted 17.06.22; peer-reviewed by Y Ranjit, B Kaveladze, D Baretta; comments to author 17.10.22; revised version received 07.11.22; accepted 25.11.22; published 23.12.22.

Please cite as:

Rowland S, Ramos AK, Trinidad N, Quintero S, Johnson Beller R, Struwe L, Pozehl B

Feasibility, Usability and Acceptability of a mHealth Intervention to Reduce Cardiovascular Risk in Rural Hispanic Adults: Descriptive Study

JMIR Form Res 2022;6(12):e40379

URL: <https://formative.jmir.org/2022/12/e40379>

doi: [10.2196/40379](https://doi.org/10.2196/40379)

PMID: [36563025](https://pubmed.ncbi.nlm.nih.gov/36563025/)

©Sheri Rowland, Athena K Ramos, Natalia Trinidad, Sophia Quintero, Rebecca Johnson Beller, Leeza Struwe, Bunny Pozehl. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 23.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Viewpoint

Spotlighting Disability in a Major Electronic Health Record: Michigan Medicine's Disability and Accommodations Tab

Heather Halkides^{1*}, BAS, MAS; Tyler G James^{2,3*}, MCHES, PhD; Michael M McKee^{2,3*}, MD, MPH; Michelle A Meade^{2,3,4*}, PhD; Christa Moran^{3,5*}, MEd, NIC, CoreCHI; Sophia Park^{3*}, BA

¹Health Information & Technology Services, Michigan Medicine, Ann Arbor, MI, United States

²Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

³Center for Disability Health and Wellness, Michigan Medicine, University of Michigan Medical School, Ann Arbor, MI, United States

⁴Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI, United States

⁵Interpreter Services, Office of Patient Experience, Michigan Medicine, Ann Arbor, MI, United States

* all authors contributed equally

Corresponding Author:

Tyler G James, MCHES, PhD
Department of Family Medicine
University of Michigan
1018 Fuller St
Ann Arbor, MI, 48104
United States
Phone: 1 734 998 7120
Email: jamesty@med.umich.edu

Abstract

People with disabilities represent the largest minority group in the United States and a priority population for health services research. Despite federal civil rights law, people with disabilities face inaccessible health care environments that fail to accommodate their disability. We present Michigan Medicine's Disability and Accommodations Tab. This patient-facing questionnaire and shared data field in the electronic health record enables the collection and reporting of patient disability-related accommodations. The Disability Tab seeks to address provider- and clinic staff-reported barriers to providing accommodations and fosters an opportunity to redesign health care to meet the needs of people with disabilities.

(*JMIR Form Res* 2022;6(12):e38003) doi:[10.2196/38003](https://doi.org/10.2196/38003)

KEYWORDS

patients with disabilities; disability accommodations; electronic health records; patient-centered care; Affordable Care Act; Americans with Disabilities Act; disability; disabilities; affordable care; EHR; accommodation; minority; equity; accessibility; accessible; inclusive; inclusivity; health care; health service; environment; accommodate; reporting; data collection; barrier

Introduction

People with disabilities represent a considerable proportion of the US population, with 27% of adults (in 2019) and 4% of children (in 2019) having a disability [1,2]. This population experiences widespread health inequities, largely due to stigma in society and inaccessibility of health care services. Due to these inequities, national public health and health care organizations (eg, the Agency for Health care Research and Quality [3]) deem people with disabilities as a priority population in health services research and detail several national health objectives (eg, Healthy People 2030 [4]) to improve the health of this population. To achieve these health objectives, however, we must address the inaccessibility of health care

services and tailor these services to meet the needs of people with disabilities [5].

Federal civil rights laws in the United States—including Section 504 of the Rehabilitation Act of 1973, the Americans with Disabilities Act of 1990, and Section 1557 of the Patient Protection and Affordable Care Act—delineate the responsibility of health care organizations to be accessible to people with disabilities. Despite these mandates, many health care environments remain inaccessible to people with disabilities [6-13]. The failure to provide disability accommodations appears, at least in part, due to a lack of knowledge on disability accommodations among care team members. A recent study by Iezzoni et al [14] found that 71% of US physicians provided incorrect answers on who makes decisions about reasonable

accommodations for people with disabilities, and 68% believed they were at risk for an Americans with Disabilities Act lawsuit because of accommodation issues. Misunderstanding accommodation needs and responsibilities may be due to the fact that disability accommodations are not systematically addressed in most health care systems. In many health systems, people with disabilities who require disability-related accommodations have to request the accommodation to care team staff prior to or during every medical encounter [7,15]. A lack of advanced knowledge of accommodation needs is a primary barrier to providing accommodations that can be attributed to a lack of widescale, systematic accommodation reporting [15]. This lack of centralized reporting of patient disability status and requested accommodations has been a source of inaccessibility for people with disabilities and an increasing risk for litigation for health systems [15-17].

To address barriers in communicating accommodations, the Centers for Medicare and Medicaid Services recommends collecting disability-related information at the point of care [18]. The use of health informatics, specifically electronic health records (EHRs), can improve the provision of accommodations for people with disabilities, and therefore, improve the delivery of patient care and promote health equity [15,17,19,20]. However, the description of tools to systematically collect patient disability-related information is limited.

Michigan Medicine's Disability and Accommodations Tab

Development

Michigan Medicine, the health care system owned and operated by the University of Michigan Medical School, is one of the largest health care systems in Michigan, serving over 2.7 million patient encounters per year at 3 hospitals and 40 outpatient clinics. Michigan Medicine has already developed strong commitments to improving the health of people with disabilities, including the establishment of model clinics, such as the Deaf Health Clinic in Family Medicine, and recognition and financial support for the Center for Disability Health and Wellness [21,22].

In 2019, Michigan Medicine's faculty and staff (including CM and MMM), with the support of Michigan Medicine's Disability Resource Group, met with MiChart Ambulatory Team (including HH) to create the Disability and Accommodations Tab (or "Disability Tab"). The Disability Tab is a shared data field based on a questionnaire (ie, SmartForm) within Michigan Medicine's version of Epic, called MiChart, which collects discrete data from both the patient-facing portal (Figure 1) and care team members. Patients wishing to report needed disability-related accommodations through the patient portal can complete the optional questionnaire in the same location as the Gender Identity, Sexual Orientation, and Poke Plan questionnaires. These questionnaires have the functionality to be routinely pushed to patients prior to ambulatory care encounters for intake. For care team members, the Disability Tab questionnaire can be completed through MiChart. The

description of the questionnaire in the patient portal is as follows:

Michigan Medicine is working to improve accessibility for patients with disabilities. This form is to identify accommodations that patients with disabilities may need when accessing Michigan Medicine clinics and hospitals. Completing this form does not guarantee that your Michigan Medicine clinic or facility has the accommodation available. If a specific accommodation is not available, Michigan Medicine is committed to working with you to find an effective alternative. Please directly inform your care team for any specific and urgent accessibility requests.

Initial disability classifications and accommodation options listed in the Disability Tab were created by subject matter experts at Michigan Medicine in collaboration with people with disabilities, disability advocacy groups and service centers, and care team staff who work routinely with people with disabilities. In June of 2021, the University of Michigan's Center for Disability Health and Wellness established a work group to manage the development, pilot testing, and implementation of the Disability Tab. This group further refined the disability and accommodation options to meet federal, state, and local regulatory guidance, in addition to common accommodations for different disabilities (Table 1). For example, initially we implemented a general accommodation to indicate the need for visitor or mask exemptions related to the COVID-19 pandemic. Based on guidance from the Centers for Disease Control and Prevention [23] and Michigan Medicine's Patient Civil Rights Coordinator, we refined the accommodations to be disability-specific (eg, a patient with blindness or low vision may have an indicated need for a visitor to be present to assist but not a mask mandate exemption).

Of note, the disability options listed on the Disability Tab differ from the Washington Group questions and the American Community Survey questions, as those specific questions measure only functional or activity limitations (eg, difficulty seeing and difficulty concentrating) [24,25]. For instance, one question on the American Community Survey asks, "Because of a physical, mental, or emotional problem, do you have difficulty doing errands alone such as visiting a doctor's office or shopping?" This question combines multiple functionally and qualitatively different disability categories that have different indicated accommodation needs. To assist in streamlining the provision of accommodations, we opted to be specific with respect to the type of disability a patient presents.

Data from this questionnaire are displayed on patient's Storyboard in MiChart (Figure 2), notifying care team members of the patient's disability and accommodation needs. On the Storyboard, in the patient's demographic section, there is a section for disability accommodations displayed next to medical interpreter needs and gender identity. The brief section in the demographic column identifies only the patient's indicated disability or disabilities. MiChart users can then click this section to open a section of the medical record listing the indicated accommodation needs.

The Disability Tab became active in the EHR and patient portal in October 2020. As of December 2021, however, the Disability Tab had not been promoted by Michigan Medicine nor had systematic data collection been incorporated into the clinic workflows due to the COVID-19 pandemic. In addition, no patient portal (MyChart or MyUofMHealth) reminder had prompted patients to complete the questionnaire. Despite this lack of promotion, as of December 13, 2021, almost 3000 patients (n=2941) had completed questionnaires (for reference, Michigan Medicine has over 240,000 active primary care patients). Among these patients, 1 in 4 (n=738, 25.1%) report mobility disabilities, followed by patients reporting mental health disabilities (n=441, 15%); patients who are deaf, hard of hearing, or deafblind (n=426, 14.5%); patients with cognitive disabilities (n=388, 13.2%); speech or other communication disabilities (n=209, 7.1%); blindness or low vision (n=185, 6.3%); and upper body or fine motor skill impairment (n=161, 5.5%).

Several health system wide initiatives have recognized the potential benefits of the Disability Tab to collect and display information about accommodation needs. One of these initiatives was the May 2022 rollout of MyChart Bedside, a tablet-based inpatient portal tool to improve the patient experience. The use of MyChart Bedside enables patients to complete the Disability Tab questionnaire during their inpatient stays directly from their MyChart Bedside tablets. These patient responses populate the Disability Tab field in the Storyboard for care team members to see. As of September 1, 2022, a total of 4732 patients have Disability Tab information in their medical records. Among these patients, mobility disabilities and wheelchair use were the most common (n=1134, 24%), followed by no disabilities (n=939, 19.8%); ‘other’ disabilities (n=800, 16.9%); hard of hearing, deafness, or deafblindness (n=793, 16.8%); mental health disabilities (n=676, 14.3%); cognitive disabilities (n=653, 13.8%); speech disabilities (n=356, 7.5%); blindness (n=330, 7%), upper body and fine motor skill impairment (n=256, 5.3%); other sensory disabilities (n=223, 4.7%); and respiratory disabilities (n=43, 0.9%).

Figure 1. Example of the Disability and Accommodations Tab questionnaire in the patient-facing portal. As the tool is specific to requesting disability-related accommodations, the requested accommodation is required after selecting a disability classification. © 2022 Epic Systems Corporation. The MyUofMHealth App is powered by MyChart® licensed from Epic Systems Corporation, © 1999-2022.

Disability Accommodations

Michigan Medicine is working to improve accessibility for patients with disabilities. This form is to identify accommodations that patients with disabilities may need when accessing Michigan Medicine clinics and hospitals. Completing this form does not guarantee that your Michigan Medicine clinic/facility has the accommodation available. If a specific accommodation is not available, Michigan Medicine is committed to working with you to find an effective alternative. Please directly inform your care team for any specific and urgent accessibility requests.

This questionnaire was recently updated. Please review your answers. You may need to re-answer some questions.

* Indicates a required field.

What TYPE OF DISABILITY(S) do you have?
Select all that apply.

Blind/Low vision Cognitive, Intellectual, or Developmental **Hard of Hearing, Deaf, Deaf-Blind**
Mental Health Mobility Disability and/or Wheelchair Use Respiratory Speech/communication
Other Sensory Upper Body and Fine Motor Skill Impairment None

* If you are Hard of Hearing, Deaf, Deaf-Blind, what accommodation(s) do you require?
Select all that apply.

Assistive listening devices (ALDS) ASL interpreter **Closed captioning during video visits**
Provider(s) and staff wear a clear mask Provide directions/follow-up in writing
Quiet space for communication Real time captioning Use written communication/information
Modifications to the COVID-19 visitor policy Other, please specify:

Continue Finish later Cancel

Table 1. Disability and accommodation options in Michigan Medicine's Disability and Accommodations Tab.

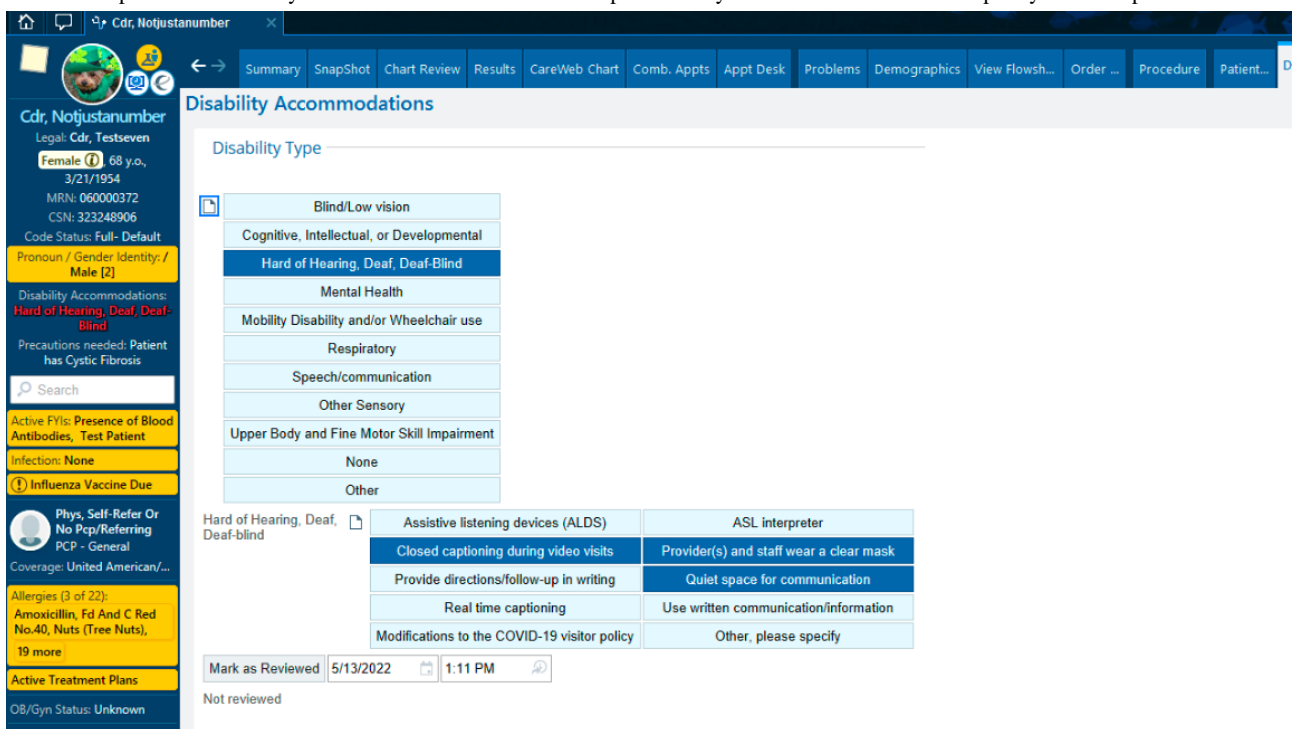
Disability classification	Available accommodation options
Blind/low vision	<ul style="list-style-type: none"> • Audio descriptors • Braille • Provide documents in large print • Screen readers • Human guide • Exceptions to the COVID-19 visitor policy • Other, please specify
Cognitive, intellectual, or developmental	<ul style="list-style-type: none"> • Assistance with completing surveys/patient intake • Check for understanding • Closed captioning during video visits • I want to give people information in advance, before going to the clinic (see 'other') • I have a support person, please involve them in my medical discussions • Provide directions/follow-up in writing • Use visuals or pictures to explain concepts • Modifications to the COVID-19 mask policy • Modifications to the COVID-19 visitor policy • Need for reduced sensory input • Other, please specify
Hard of hearing, deaf, or deafblind	<ul style="list-style-type: none"> • Assistive listening devices • ASL^a interpreter • Closed captioning during video visits • Provider(s) and staff wear a clear mask • Provide directions/follow-up in writing • Quiet space for communication • Real-time captioning • Use written communication or information • Modifications to the COVID-19 visitor policy • Other, please specify
Mental health	<ul style="list-style-type: none"> • Additional structure and assistance regulating emotions • Clear protocols to help me prepare for care • Need for reduced sensory input • Provide directions/follow-up in writing • Modifications to the COVID-19 mask policy • Modifications to the COVID-19 visitor policy • Other, please specify
Mobility disability or wheelchair use	<ul style="list-style-type: none"> • Adjustable tables • Assistance with transfers and walking • Availability of transfer equipment (eg. a lift, a transfer board) • Human assistance with transfers • Larger exam rooms • Wheelchair scales • Modifications to the COVID-19 mask policy • Modifications to the COVID-19 visitor policy • Other, please specify
Respiratory	<ul style="list-style-type: none"> • Modifications to the COVID-19 mask policy • Need for oxygen tank • Plug outlet for oxygen concentrator • Other, please specify
Speech/communication	<ul style="list-style-type: none"> • Closed captioning during video visits • Confirm that I understand • Give me additional time to speak • Understanding prompts from the provider • Whiteboards for communication • Other, please specify
Other sensory	<ul style="list-style-type: none"> • Fragrance-free environment • Limit touch • Placement in room early • Other, please specify

Disability classification	Available accommodation options
Upper body and fine motor skill impairment	<ul style="list-style-type: none"> Assistance with clothing management Assistance with completing surveys/patient intake Assistance with transfers Modifications to the COVID-19 mask policy Modifications to the COVID-19 visitor policy Other, please specify
Other (please specify)	<ul style="list-style-type: none"> Other, please specify
None	__b

^aASL: American Sign Language.

^b__: Not applicable.

Figure 2. Example of the Disability and Accommodations Tab on the patient Storyboard in MiChart. © 2022 Epic Systems Corporation.



Updating the Disability Tab

Modifications to the SmartForm used for the Disability Tab requires substantial time for revising the code and testing for quality assurance before revisions are ‘live’ in the EHR. Further, ‘retiring’ disability or accommodation categories can lead to missing data if not properly coded. For this reason, we have implemented a Change Review Board process to review requested changes (eg, adding an accommodation field) with respect to the financial, reporting, and human resource impacts. The Disability Tab Change Review Board solicits regulatory feedback from both the Americans with Disabilities Act Coordinator and Patient Civil Rights Coordinator, and other interested parties (eg, the Office of Patient Experience).

With the implementation of the Disability Tab in MyChart Bedside (in May 2022), we recognized the need for a ‘mark as reviewed’ function on the questionnaire. This is particularly important when patient disability-related accommodation needs change due to onset or progression of disability. To address this

need, we developed a button on the SmartForm that records the last user and the date or time the questionnaire was reviewed.

Identifying Opportunities for Ambulatory Care Workflow Integration

To identify opportunities to integrate disability data collection within the workflow in outpatient clinics, the Disability Tab started being actively used (eg, patients’ disability-related needs being documented by the care team on the Disability Tab) at the Dexter Health Center, operated by Michigan Medicine’s Department of Family Medicine, in March 2022. At Dexter Health Center, patients are asked to complete a paper version of the Disability Tab questionnaire, which then gets added to the electronic Disability Tab by clerical or medical assistants. Once entered, this information becomes available across all clinical encounters across Michigan Medicine Health Systems. Staff at this clinic are using the Disability Tab to prepare accommodations in advance of future clinic appointments. To understand modifications to their workflow and acceptability among care team staff, we are collecting data in an ongoing

quality improvement study. Results of this study, focused on operation workflow of the Disability Tab (for collecting information and providing accommodations), will be disseminated at a later date.

Discussion

Principal Findings

With the widespread lack of centralized disability services in health care systems, health care providers and clinic staff are often responsible for determining and providing disability accommodation needs [16]. This process, however, presents challenges, particularly with staff communication across patient encounters regarding disability accommodations. Studies, including an unpublished quality improvement study at Michigan Medicine [26], find that being unaware of the need for an accommodation in advance is a common barrier to providing accommodations to people with disabilities [15]. This is concerning since clinic staff are integral to arranging accommodations for upcoming appointments for people with disabilities.

The Disability Tab provides an opportunity to specifically address provider and clinic staff barriers in providing accommodations to people with disabilities. Patients can report their specific disabilities and accommodation needs by filling out the questionnaire within the patient portal or being prompted by care team staff at the point of care; this information appears for health care providers and clinic staff in the Storyboard, as part of the patient's demographic information. Moreover, this tool provides the opportunity to identify and articulate the role of caregivers as well as any alterations related to patient autonomy, addressing barriers often experienced by patients who have cognitive, intellectual, or developmental disabilities.

By design, the Disability Tab reports disability-related accommodation needs information systemwide, making this available to all care team members across all clinical encounters, not just the patients' primary care providers, or hidden within free-text clinical notes, as is common in health care [15]. This tool is critical in minimizing the information gap during patient transitions and handoffs. Moreover, the development of the Disability Tab in Epic, a major EHR system, allows for significant scalability to other health care systems, as this field could be requested and implemented by other systems that use Epic. As such, the design and implementation of the Disability Tab has the potential to facilitate the identification of individuals with disability and the standardization of disability-related information in health care as outlined under the Section 4302 of the Affordable Care Act [18].

Limitations and Future Work

The Disability Tab is still relatively early in the implementation phase and there are several opportunities for future work, particularly to respond to potential limitations. First, the Disability Tab questionnaire uses disability categories that are not standardized to the American Community Survey or Washington Group disability items. Therefore, information from the Disability Tab will not facilitate population health comparisons. Although this may be a limitation, we determined early in the Disability Tab's development that the goal of this tool was to facilitate accommodation access, not solely collect disability prevalence information [17,20].

Given the early stage of implementation, there are several outstanding questions. One question is whether the presence of disability identity impacts the process of care. For example, during the initial impact of the COVID-19 pandemic, there was considerable discussion of rationing (ie, declining) care to people with disabilities [27,28]. How care team members use the information for care, outside of simply providing accommodations, is undetermined. An additional concern is how to clearly establish a workflow that enables outpatient clinics and inpatient services to provide requested accommodations. The Disability Tab meets the need of providing information to care team staff; it does not, however, ensure that staff are interacting with or acting on this information. Lastly, as with all clinical informatics interventions, there are concerns that the Disability Tab may further widen the gap among people with disabilities who are multiply marginalized through intervention-generated inequality [29]. The Disability Tab is currently only available in English, and it is possible that patients with limited English proficiency are not asked their disability and accommodation needs at the bedside. Further, patients who do not have reliable internet access or patient portal access are unable to complete the questionnaire online. Future efforts should focus on further equitizing this intervention.

Conclusions

In its infancy, the Disability Tab demonstrates the opportunity to leverage EHR systems and health informatics to systematically collect disability-related accommodation needs to improve the quality of care delivered to people with disabilities and improve accessibility to health care environments. We encourage other health care systems to adopt similar approaches to address the health care needs of people with disabilities.

Acknowledgments

The authors would like to thank the Michigan Medicine Disability Resource Group for their support and advice on early conceptualizations of the Disability and Accommodations Tab.

HH, TGJ, MMM, and MAM were supported by the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), part of the Administration on Community Living (ACL), under award 90DPHF0011-01-00. Additional support for this project has been provided by the Center for Disability Health and Wellness, funded by Michigan Medicine. The contents of this paper do not necessarily represent the policy of NIDILRR, ACL, or Human and Health Services (HHS).

Authors' Contributions

All authors contributed equally.

Conflicts of Interest

None declared.

References

1. Young NAE. Childhood disability in the United States. United States Census Bureau. 2021 Mar 25. URL: <https://www.census.gov/library/publications/2021/acs/acsbr-006.html> [accessed 2022-11-30]
2. Varadaraj V, Deal JA, Campanile J, Reed NS, Swenor BK. National prevalence of disability and disability types among adults in the US, 2019. *JAMA Netw Open* 2021 Oct 01;4(10):e2130358 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.30358](https://doi.org/10.1001/jamanetworkopen.2021.30358)] [Medline: [34673966](https://pubmed.ncbi.nlm.nih.gov/34673966/)]
3. Healthcare research and quality act of 1999. Agency for Healthcare research and Quality. URL: <https://www.ahrq.gov/policymakers/hrqa99a.html> [accessed 2022-11-30]
4. Healthy people 2030 objectives. U.S. Department of Health and Human Services. URL: <https://health.gov/healthypeople/objectives-and-data/browse-objectives> [accessed 2020-09-08]
5. Meade MA, Mahmoudi E, Lee S. The intersection of disability and healthcare disparities: a conceptual framework. *Disabil Rehabil* 2015;37(7):632-641. [doi: [10.3109/09638288.2014.938176](https://doi.org/10.3109/09638288.2014.938176)] [Medline: [25060038](https://pubmed.ncbi.nlm.nih.gov/25060038/)]
6. James TG, McKee MM, Sullivan MK, Ashton G, Hardy SJ, Santiago Y, et al. Community-engaged needs assessment of deaf American Sign Language users in Florida, 2018. *Public Health Rep* 2022;137(4):730-738. [doi: [10.1177/00333549211026782](https://doi.org/10.1177/00333549211026782)] [Medline: [34161191](https://pubmed.ncbi.nlm.nih.gov/34161191/)]
7. James TG, Coady KA, Stacciarini JR, McKee MM, Phillips DG, Maruca D, et al. "They're not willing to accommodate deaf patients": communication experiences of deaf American Sign Language users in the emergency department. *Qual Health Res* 2022 Jan;32(1):48-63. [doi: [10.1177/10497323211046238](https://doi.org/10.1177/10497323211046238)] [Medline: [34823402](https://pubmed.ncbi.nlm.nih.gov/34823402/)]
8. Bauer SE, Schumacher JR, Hall A, Marlow NM, Friedel C, Scheer D, et al. Disability and physical and communication-related barriers to health care related services among Florida residents: A brief report. *Disabil Health J* 2016 Jul;9(3):552-556. [doi: [10.1016/j.dhjo.2016.03.001](https://doi.org/10.1016/j.dhjo.2016.03.001)] [Medline: [27101882](https://pubmed.ncbi.nlm.nih.gov/27101882/)]
9. Stillman MD, Bertocci G, Smalley C, Williams S, Frost KL. Healthcare utilization and associated barriers experienced by wheelchair users: a pilot study. *Disabil Health J* 2017 Oct;10(4):502-508. [doi: [10.1016/j.dhjo.2017.02.003](https://doi.org/10.1016/j.dhjo.2017.02.003)] [Medline: [28245968](https://pubmed.ncbi.nlm.nih.gov/28245968/)]
10. Iezzoni LI, Killeen MB, O'Day BL. Rural residents with disabilities confront substantial barriers to obtaining primary care. *Health Serv Res* 2006 Aug;41(4 Pt 1):1258-1275 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00534.x](https://doi.org/10.1111/j.1475-6773.2006.00534.x)] [Medline: [16899006](https://pubmed.ncbi.nlm.nih.gov/16899006/)]
11. Mudrick NR, Breslin ML, Liang M, Yee S. Physical accessibility in primary health care settings: results from California on-site reviews. *Disabil Health J* 2012 Jul;5(3):159-167. [doi: [10.1016/j.dhjo.2012.02.002](https://doi.org/10.1016/j.dhjo.2012.02.002)] [Medline: [22726856](https://pubmed.ncbi.nlm.nih.gov/22726856/)]
12. McKee MM, Lin FR, Zazove P. State of research and program development for adults with hearing loss. *Disabil Health J* 2018 Oct;11(4):519-524. [doi: [10.1016/j.dhjo.2018.07.010](https://doi.org/10.1016/j.dhjo.2018.07.010)] [Medline: [30087082](https://pubmed.ncbi.nlm.nih.gov/30087082/)]
13. James TG, Argyeni MS, Guardino DL, McKee MM, Wilson JAB, Sullivan MK, et al. Communication access in mental health and substance use treatment facilities for deaf American Sign Language users. *Health Aff (Millwood)* 2022 Oct 01;41(10):1413-1422. [doi: [10.1377/hlthaff.2022.00408](https://doi.org/10.1377/hlthaff.2022.00408)] [Medline: [36190883](https://pubmed.ncbi.nlm.nih.gov/36190883/)]
14. Iezzoni LI, Rao SR, Resselam J, Bolcic-Jankovic D, Agaronnik ND, Lagu T, et al. US physicians' knowledge about the Americans with disabilities act and accommodation of patients with disability. *Health Aff (Millwood)* 2022 Jan;41(1):96-104. [doi: [10.1377/hlthaff.2021.01136](https://doi.org/10.1377/hlthaff.2021.01136)] [Medline: [34982624](https://pubmed.ncbi.nlm.nih.gov/34982624/)]
15. Mudrick NR, Breslin ML, Nielsen KA, Swager LC. Can disability accommodation needs stored in electronic health records help providers prepare for patient visits? A qualitative study. *BMC Health Serv Res* 2020 Oct 16;20(1):958 [FREE Full text] [doi: [10.1186/s12913-020-05808-z](https://doi.org/10.1186/s12913-020-05808-z)] [Medline: [33066788](https://pubmed.ncbi.nlm.nih.gov/33066788/)]
16. Sunderland v. Bethesda Health, Inc. Casetext. URL: <https://casetext.com/case/sunderland-v-bethesda-health-inc-2> [accessed 2022-11-30]
17. Morris MA, Hasnain-Wynia R. A research agenda for documenting disability status within healthcare organizations to address disparities in care. *J Healthc Qual* 2014;36(2):7-12; quiz 12. [doi: [10.1111/jhq.12059](https://doi.org/10.1111/jhq.12059)] [Medline: [24350672](https://pubmed.ncbi.nlm.nih.gov/24350672/)]
18. Improving health care for adults with disabilities: an overview of federal data sources. Centers for Medicare and Medicaid Services. Baltimore, MD URL: <https://www.cms.gov/files/document/federaldatadisability508.pdf> [accessed 2022-01-24]
19. James TG, Sullivan MK, Butler JD, McKee MM. Promoting health equity for deaf patients through the electronic health record. *J Am Med Inform Assoc* 2021 Dec 28;29(1):213-216 [FREE Full text] [doi: [10.1093/jamia/ocab239](https://doi.org/10.1093/jamia/ocab239)] [Medline: [34741507](https://pubmed.ncbi.nlm.nih.gov/34741507/)]
20. Morris MA, Hamer MK, Eberle K, Jensen KM, Wong AA. Implementation of collection of patients' disability status by centralized scheduling. *Jt Comm J Qual Patient Saf* 2021 Oct;47(10):627-636. [doi: [10.1016/j.jcjq.2021.05.007](https://doi.org/10.1016/j.jcjq.2021.05.007)] [Medline: [34144917](https://pubmed.ncbi.nlm.nih.gov/34144917/)]

21. Pertz L, Plegue M, Diehl K, Zazove P, McKee M. Addressing mental health needs for deaf patients through an integrated health care model. *J Deaf Stud Deaf Educ* 2018 Jul 01;23(3):240-248. [doi: [10.1093/deafed/eny002](https://doi.org/10.1093/deafed/eny002)] [Medline: [29562357](https://pubmed.ncbi.nlm.nih.gov/29562357/)]
22. Panzer K, Park J, Pertz L, McKee M. Teaming together to care for our deaf patients: insights from the Deaf Health Clinic. *JADARA* 2020 Apr;53(2):58-77 [[FREE Full text](#)]
23. Order: wearing of face masks while on conveyances and at transportation hubs. Centers for Disease Control and Prevention. URL: <https://web.archive.org/web/20221027135351/https://www.cdc.gov/quarantine/masks/mask-travel-guidance.html> [accessed 2022-11-29]
24. WG short set on functioning (WG-SS). Washington Group on Disability Statistics. London: University College URL: <https://www.washingtongroup-disability.com/question-sets/wg-short-set-on-functioning-wg-ss/> [accessed 2022-11-30]
25. How disability data are collected on the American Community Survey. United States Census Bureau. URL: <https://www.census.gov/topics/health/disability/guidance/data-collection-acs.html> [accessed 2022-03-10]
26. MiChart DTCG. MiChart disability and accommodations tab. Michigan Medicine. URL: <https://disabilityhealth.medicine.umich.edu/clinical/michart-disability-accommodations-tab> [accessed 2022-11-30]
27. Andrews EE, Ayers KB, Brown KS, Dunn DS, Pilarski CR. No body is expendable: medical rationing and disability justice during the COVID-19 pandemic. *Am Psychol* 2021 Apr;76(3):451-461. [doi: [10.1037/amp0000709](https://doi.org/10.1037/amp0000709)] [Medline: [32700936](https://pubmed.ncbi.nlm.nih.gov/32700936/)]
28. Lund EM, Ayers KB. Raising awareness of disabled lives and health care rationing during the COVID-19 pandemic. *Psychol Trauma* 2020 Aug;12(S1):S210-S211. [doi: [10.1037/tra0000673](https://doi.org/10.1037/tra0000673)] [Medline: [32463287](https://pubmed.ncbi.nlm.nih.gov/32463287/)]
29. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018 Aug 01;25(8):1080-1088. [doi: [10.1093/jamia/ocy052](https://doi.org/10.1093/jamia/ocy052)] [Medline: [29788380](https://pubmed.ncbi.nlm.nih.gov/29788380/)]

Abbreviations

EHR: electronic health record

Edited by A Mavragani; submitted 15.03.22; peer-reviewed by K Ayers, M Morris; comments to author 03.10.22; revised version received 12.10.22; accepted 28.10.22; published 02.12.22.

Please cite as:

Halkides H, James TG, McKee MM, Meade MA, Moran C, Park S

Spotlighting Disability in a Major Electronic Health Record: Michigan Medicine's Disability and Accommodations Tab

JMIR Form Res 2022;6(12):e38003

URL: <https://formative.jmir.org/2022/12/e38003>

doi: [10.2196/38003](https://doi.org/10.2196/38003)

PMID: [36459406](https://pubmed.ncbi.nlm.nih.gov/36459406/)

©Heather Halkides, Tyler G James, Michael M McKee, Michelle A Meade, Christa Moran, Sophia Park. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Influences, Barriers, and Facilitators to COVID-19 Vaccination: Cross-sectional Survey on Vaccine Hesitancy in 2 Rural States

Elaine Nguyen¹, MPH, PharmD; Melanie Wright¹, PhD; John Holmes², MPH, PharmD; Kevin Cleveland¹, PharmD; Catherine Oliphant¹, PharmD; Mary Nies³, PhD; Renee Robinson⁴, MPH, MSPharm, PharmD

¹Department of Pharmacy Practice and Administration, College of Pharmacy, Idaho State University, Meridian, ID, United States

²Department of Pharmacy Practice and Administration, College of Pharmacy, Idaho State University, Pocatello, ID, United States

³School of Nursing, College of Health, Pocatello, ID, United States

⁴Department of Pharmacy Practice and Administration, College of Pharmacy, University of Alaska/Idaho State University, Anchorage, AK, United States

Corresponding Author:

Renee Robinson, MPH, MSPharm, PharmD
Department of Pharmacy Practice and Administration
College of Pharmacy
University of Alaska/Idaho State University
2533 Providence Drive PSB111
Anchorage, AK, 99508
United States
Phone: 1 907 786 6233
Email: robiren2@isu.edu

Abstract

Background: Vaccination remains one of the most effective ways to limit the spread of infectious diseases such as that caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for COVID-19. Unfortunately, vaccination hesitancy continues to be a threat to national and global health. Further research is necessary to determine the modifiable and nonmodifiable factors contributing to COVID-19 vaccine hesitancy in under-resourced, underserved, and at-risk rural and urban communities.

Objective: This study aimed to identify, understand, and address modifiable barriers and factors contributing to COVID-19 vaccine hesitancy among vaccine-eligible individuals with access to the vaccine in Alaska and Idaho.

Methods: An electronic survey based on the World Health Organization (WHO) Strategic Advisory Group on Experts (SAGE) on Immunization survey tool and investigators' previous work was created and distributed in June 2021 and July 2021. To be eligible to participate in the survey, individuals had to be ≥18 years of age and reside in Alaska or Idaho. Responses were grouped into 4 mutually exclusive cohorts for data analysis and reporting based on intentions to be vaccinated. Respondent characteristics and vaccine influences between cohorts were compared using Chi-square tests and ANOVA. Descriptive statistics were also used.

Results: There were data from 736 usable surveys with 40 respondents who did not intend to be vaccinated, 27 unsure of their intentions, 8 who intended to be fully vaccinated with no doses received, and 661 fully vaccinated or who intended to be vaccinated with 1 dose received. There were significant differences in characteristics and influences between those who were COVID-19 vaccine-hesitant and those who had been vaccinated. Concerns related to possible side effects, enough information on long-term side effects, and enough information that is specific to the respondent's health conditions were seen in those who did not intend to be fully vaccinated and unsure about vaccination. In all cohorts except those who did not intend to be fully vaccinated, more information about how well the vaccine works was a likely facilitator to vaccination.

Conclusions: These survey results from 2 rural states indicate that recognition of individual characteristics may influence vaccine choices. However, these individual characteristics represent only a starting point to delivering tailored messages that should come from trusted sources to address vaccination barriers.

(*JMIR Form Res* 2022;6(12):e39109) doi:[10.2196/39109](https://doi.org/10.2196/39109)

KEYWORDS

COVID-19; COVID-19 vaccines; vaccine hesitancy; cross-sectional studies; rural populations

Introduction

Immunization is the greatest public health achievement of all time, saving over 3 million lives worldwide each year [1,2]. State and national immunization programs have been so successful that many Americans view the risks of vaccine-preventable diseases such as measles, pertussis, and polio as minimal [1]. Vaccination remains one of the most effective ways to limit the spread of infectious diseases such as that caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus responsible for COVID-19 [3]. However, waning public confidence in vaccines, especially the COVID-19 vaccine, remains a cause for concern [4-9]. In addition, believed threshold requirements for vaccination to achieve herd immunity have shifted and increased with new variants [10,11]. As of August 2022, ~67% of the total population had been fully vaccinated (ie, primary series completed) [12], but there continues to be a need for both primary series completion and booster doses [13].

Vaccine hesitancy, one of the top 10 threats to global health, is the delay or refusal to receive a vaccine, despite access, availability, and perceived effectiveness of the vaccine [5]. Preliminary research suggests that health decisions, such as to receive or not receive a vaccine, are highly influenced by social and cultural factors (eg, political ideology, past experiences with health services, family histories, the moral dilemma between individual autonomy and the greater public health) [6,13,14]. In addition, several other complex factors may be contributing to the increased hesitancy that has been noted with the COVID-19 vaccine [5,8,15].

Given the importance of this topic, researchers have been seeking to better understand acceptability of COVID-19 vaccination and drivers of hesitancy. Global surveys conducted in different countries have shown concerns for vaccine safety and effectiveness [16-21]. At the beginning of the pandemic in May 2020, an online survey of Americans found that 69% of respondents were willing to receive a COVID-19 vaccine [16]. There were statistically significant differences in those willing and not willing to get vaccinated based on how well the vaccine works and the number of people infected with COVID-19. Since this initial survey was conducted, the COVID-19 vaccination landscape has continued to shift in the United States and globally.

With complex factors impacting vaccination decisions, further research is necessary to determine the modifiable and nonmodifiable factors contributing to COVID-19 vaccine hesitancy in under-resourced, underserved, and at-risk rural and urban communities. The goal of this project was to identify, understand, and address modifiable barriers and factors contributing to COVID-19 vaccine hesitancy among vaccine-eligible individuals with access to the vaccine in Alaska and Idaho. The primary goal of this paper was to present the results from a vaccine hesitancy survey.

Methods**Survey Details**

The survey was based on the World Health Organization (WHO) Strategic Advisory Group on Experts (SAGE) on Immunization survey tool and investigators' previous qualitative work with residents of Alaska and Idaho who remain hesitant to receive the COVID-19 vaccine [1,22,23]. Some survey questions from the WHO SAGE were previously validated, and some were from field experts. The draft survey underwent several revisions and was reviewed multiple times by project investigators (n=7) and the project advisory board. The advisory board (n=11) was composed of community members, health care providers, and public health organization representatives. Although the final survey was extensively reviewed, it did not undergo any formal validation studies.

The survey (Multimedia Appendix 1) was estimated to take 10 minutes and included 31 questions divided into 4 parts: introduction, barriers and facilitators, influences, and demographics. All questions, except those in the introduction (eligibility screening and vaccination status), were optional. To be eligible to participate in the survey, individuals had to be ≥18 years of age and reside in Alaska or Idaho. The focus on this population was to capture adult perspectives from primarily rural states.

The survey was created and made available via Qualtrics online survey software. It was distributed through project investigators and advisory board member contacts. It was also promoted in Facebook advertisements. A broad defined audience was used for the Facebook advertisements, with advertisements targeting only location (Alaska or Idaho) and age (≥18 years). The survey was available for approximately 1 month, with responses collected from June 11, 2021, through July 16, 2021.

Data Analysis

Given the broad survey distribution, responses were reviewed for validity, and responses deemed potentially invalid were removed from analysis. Responses were removed if at least one of the following criteria were met: The survey was not finished, completed multiple times from the same IP address with no unique free-text responses, or not completed in Alaska and Idaho (as determined by GPS coordinates). IP address and GPS data are automatically collected in the Qualtrics survey platform.

Eligible and valid responses were included in the data analysis. Respondents were grouped into 4 mutually exclusive cohorts: (1) did not intend to be fully vaccinated, (2) were unsure of their vaccination intentions, (3) intended to be fully vaccinated but had not yet received their first dose, and (4) were fully vaccinated or intended to be fully vaccinated and had received at least 1 dose. When comparing respondent characteristics and vaccine influences between cohorts, Chi-square tests (for nominal data) and ANOVA (for continuous data) were used. *P* values <.05 were considered statistically significant when comparing respondent characteristics. When comparing vaccine

influences, a Bonferroni correction with a P value $<.007$ was considered statistically significant. Descriptive statistics (counts and percentages) were used to describe barriers, facilitators, and trust in sources for vaccine information. Cronbach alpha was also used to measure reliability for barriers, facilitators, and trust in sources for vaccine information.

Ethical Considerations

This work underwent an expedited review and was approved by the Idaho State University Institutional Review Board (IRB-FY2021-256). Respondents indicated their consent to participate after reading survey background information (eg, purpose, estimated completion time) and by continuing to the next survey page.

After determining if a response was valid and categorizing the response into a mutually exclusive cohort, any individual response data not associated with survey responses were removed for analysis to protect respondent privacy. To incentivize participation, a raffle to be entered to win one of 20 US \$100 Amazon electronic gift cards was offered. To maintain respondent privacy, raffle information was collected in a separate form to keep survey responses anonymous.

Results

After removal of invalid responses, there were data from 736 usable surveys: 40 respondents did not intend to be vaccinated, and 27 were unsure of their intentions. Although 8 respondents had not yet received any COVID-19 vaccine, they had intended to be fully vaccinated. Lastly, 661 respondents were fully vaccinated ($n=654$) or intended to be fully vaccinated with 1 dose received ($n=7$). Characteristics of survey respondents are presented in [Table 1](#). There were statistically significant differences across cohorts among all characteristics evaluated (see P values in [Table 1](#)).

The intended to be vaccinated cohort with no doses received had the lowest mean age (33.3 years), but this sample size was small. Those who were fully vaccinated or intended to be fully vaccinated with at least 1 dose received had the oldest mean age (59.1 years). Of those who did not intend to be vaccinated, the lowest age of a respondent was 35 years, whereas the other cohorts had younger respondents. The distribution of respondents' ages by cohorts is available in [Multimedia Appendix 2](#).

Nearly 90% (7/8, 88%) of respondents in the intended to be fully vaccinated but had not yet received their first dose cohort were men versus only 18.0% (119/661) in the fully vaccinated/intended to be fully vaccinated with 1 dose received cohort. One-half (20/40, 50%) of those who did not intend to be fully vaccinated identified as Christian, whereas only 34.5% (231/669) of those who intended to be fully vaccinated or were already fully vaccinated were Christian. Conversely, a reverse

pattern was seen across these cohorts with those who were agnostic, atheist, or believed in nothing in particular. In those who did not intend to be vaccinated, the largest political preference was Republican (17/40, 43%). In those fully vaccinated/intended to be fully vaccinated with 1 dose received, the largest political preference was Democrat (314/661, 47.5%). Across cohorts, there were statistical differences in race and ethnicity, but overall, the groups were predominantly White and not Hispanic nor Latino. The majority of respondents had health insurance (702/736, 95.4%), with only 22 (22/736, 3.0%) reporting no insurance; 12 (12/736, 1.6%) were unsure or chose not to share their health insurance status.

When assessing vaccine influences, there continued to be differences across cohorts ([Table 2](#)). Only 25% (10/40) of those who did not intend to be vaccinated had a medium-high perceived risk of getting COVID-19 versus 43.0% (284/661) of those fully vaccinated/intended to be fully vaccinated. Interestingly, the percentage of those who had prior COVID-19 and had been really sick was twice as high in those with no plans of vaccination versus those fully vaccinated/intended to be fully vaccinated (5/40, 13% vs 39/661, 5.9%). Of the respondents, 90.2% (664/736) knew somebody who had COVID-19, and nearly one-third (239/736, 32.5%) knew somebody who had died from the disease. The proportion of those knowing somebody who died from COVID-19 was lower in those who did not intend to be vaccinated than in the other cohorts. Significant differences were seen in beliefs that vaccines work to prevent diseases and, related to this belief, the typical receipt of the influenza vaccine.

Data represented in [Tables 3-5](#) represent descriptive statistics only. From these data, concerns related to possible side effects, enough information on long-term side effects, and enough information that is specific to respondents' health conditions were seen in those who did not intend to be fully vaccinated and unsure about vaccination ([Table 3](#)). Practical factors for vaccination (ie, scheduling, time away from daily responsibilities for vaccination and side effects, child supervision) were not seen as barriers. The Cronbach alpha for barriers was 0.8184. In all cohorts except those who did not intend to be fully vaccinated, more information about how well the vaccine works is a likely facilitator to vaccination ([Table 4](#)). Factors such as payment to get the vaccine or requirements for work or travel were not likely facilitators or motivators to vaccination across all cohorts. The Cronbach alpha for facilitators was 0.7279.

In those not planning to receive the vaccine, there was low trust from most sources of information ([Table 5](#)). In those unsure, a primary care provider or doctor and pharmacist were the most trusted sources of information. This trend was also seen in those who intended to be fully vaccinated or those already fully vaccinated. The Cronbach alpha for trust in sources of vaccine information was 0.8239.

Table 1. Respondent characteristics.

Characteristic	Did NOT intend to be fully vaccinated (n=40)	Unsure of full vaccination intentions (n=27)	Intended to be fully vaccinated with no doses received (n=8)	Fully vaccinated OR intended to be with 1 dose received (n=661)	All (N=736)	P value
Age (years), mean (SD)	56.6 (11.0)	54.9 (15.6)	33.3 (12.0)	59.1 (14.5)	58.4 (14.7)	<.001
Age (years), median (range)	55.0 (35-92)	55.0 (25-80)	29.5 (23-61)	63.0 (18-85)	62.0 (18-92)	— ^a
Gender identity, n (%)						
Man	12 (30.0)	4 (14.8)	7 (87.5)	119 (18.0)	142 (19.3)	.001
Woman	26 (65.0)	22 (81.5)	1 (12.5)	531 (80.3)	580 (78.8)	
Other	1 (2.5)	0 (0)	0 (0)	5 (0.8)	6 (0.8)	
Prefer not to say/no response	1 (2.5)	1 (3.7)	0 (0)	6 (1.0)	8 (1.1)	
Race^b, n (%)						
Alaska Native	0 (0)	1 (3.4)	1 (12.5)	1 (0.2)	3 (0.4)	<.001
American Indian/Native American	2 (5.0)	1 (3.4)	0 (0)	7 (1.1)	10 (1.4)	
Asian	0 (0)	0 (0)	0 (0)	9 (1.4)	9 (1.2)	
Black/African American	0 (0)	0 (0)	2 (25.0)	2 (0.3)	4 (0.5)	
Native Hawaiian/Pacific Islander	0 (0)	0 (0)	0 (0)	2 (0.3)	2 (0.3)	
White	34 (85.0)	26 (89.7)	5 (62.5)	633 (95.8)	698 (94.8)	
Other	2 (5.0)	0 (0)	0 (0)	7 (1.1)	9 (1.2)	
Prefer not to say/no response	2 (5.0)	1 (3.4)	1 (12.5)	13 (2.0)	17 (2.3)	
Ethnicity, n (%)						
Hispanic/Latino	1 (2.5)	3 (11.1)	1 (12.5)	12 (1.8)	17 (2.3)	.009
Not Hispanic/Latino	29 (72.5)	19 (70.4)	6 (75.0)	547 (82.8)	601 (81.7)	
Other	1 (2.5)	1 (3.7)	0 (0)	43 (6.5)	45 (5.1)	
Prefer not to say/no response	9 (22.5)	4 (14.8)	1 (12.5)	59 (8.9)	73 (9.9)	
State, n (%)						
Alaska	0 (0)	5 (18.5)	5 (62.5)	15 (2.3)	25 (3.4)	<.001
Idaho	40 (100)	22 (81.5)	3 (37.5)	646 (97.7)	711 (96.6)	
Religion^b, n (%)						
Agnostic/atheist/nothing in particular	3 (7.5)	5 (18.5)	2 (25.0)	227 (34.3)	237 (32.2)	<.001
Christian	20 (50.0)	11 (40.7)	5 (62.5)	226 (34.2)	262 (35.6)	
Jewish	1 (2.5)	0 (0)	0 (0)	6 (0.9)	7 (1.0)	
Mormon	5 (12.5)	6 (22.2)	0 (0)	109 (16.5)	120 (16.3)	
Muslim	1 (2.5)	0 (0)	1 (12.5)	0 (0)	2 (0.3)	
Roman Catholic	3 (7.5)	3 (11.1)	0 (0)	56 (8.5)	62 (8.4)	
Other	2 (5.0)	2 (7.4)	0 (0)	37 (5.6)	41 (5.6)	
Prefer not to say/no response	6 (15.0)	3 (11.1)	0 (0)	38 (5.7)	47 (6.4)	
Political preference^b, n (%)						

Characteristic	Did NOT intend to be fully vaccinated (n=40)	Unsure of full vaccination intentions (n=27)	Intended to be fully vaccinated with no doses received (n=8)	Fully vaccinated OR intended to be with 1 dose received (n=661)	All (N=736)	P value
Democrat	2 (5.0)	5 (18.5)	1 (12.5)	314 (47.5)	322 (42.8)	<.001
Republican	17 (42.5)	7 (25.9)	4 (50.0)	110 (16.6)	138 (18.8)	
Independent	9 (22.5)	7 (25.9)	3 (37.5)	193 (29.2)	212 (28.8)	
Other	2 (5.0)	3 (11.1)	0 (0)	51 (7.7)	56 (7.6)	
Prefer not to say/no response	10 (25.0)	7 (25.9)	0 (0)	55 (8.3)	72 (9.8)	
Highest grade finished/degree received, n (%)						
High school graduate/GED or less	2 (5.0)	1 (3.7)	1 (12.5)	26 (3.9)	30 (4.1)	.003
Some college, no degree/associate degree	16 (40.0)	12 (44.4)	2 (25.0)	150 (22.7)	180 (24.5)	
Bachelor degree	11 (27.5)	4 (14.8)	2 (25.0)	226 (34.2)	243 (33.0)	
Postbachelor degree	9 (22.5)	6 (22.2)	3 (37.5)	242 (36.6)	260 (35.3)	
Other	2 (5.0)	4 (14.8)	0 (0)	14 (2.1)	20 (2.7)	
Prefer not to say/no response	0 (0)	0 (0)	0 (0)	3 (0.5)	3 (0.4)	
Employment status^b, n (%)						
Employed by government (local, state, and federal)	4 (10.0)	3 (11.1)	3 (37.5)	106 (16.0)	116 (15.8)	<.001
Employed by a private company (for-profit and non-profit)	12 (30.0)	6 (25.9)	4 (50.0)	130 (19.7)	152 (20.7)	
Other	18 (45.0)	17 (63.0)	1 (12.5)	441 (66.7)	477 (64.8)	
Prefer not to say/no response	8 (20.0)	2 (7.4)	0 (0)	11 (1.7)	21 (2.9)	

^aNot calculated.

^bRespondents could select all that apply; the sum of the percentages may be >100.

Table 2. Factors influencing the decision to be vaccinated.

Influence	Did NOT intend to be fully vaccinated (n=40), n (%)	Unsure of full vaccination intentions (n=27), n (%)	Intended to be fully vaccinated with no doses received (n=8), n (%)	Fully vaccinated OR intended to be with 1 dose received (n=661), n (%)	All (N=736), n (%)	P value
Perceived risk of getting COVID-19						
None	9 (22.5)	0 (0)	1 (12.5)	37 (5.6)	47 (6.4)	.004
Low	21 (52.5)	15 (55.6)	5 (62.5)	340 (51.4)	381 (51.8)	
Medium	9 (22.5)	9 (33.3)	1 (12.5)	198 (30.0)	217 (29.5)	
High	1 (2.5)	3 (11.1)	1 (12.5)	86 (13.0)	91 (12.4)	
Perceived risk of getting really sick from COVID-19						
None	11 (27.5)	4 (14.8)	0 (0)	53 (8.0)	68 (9.2)	__a
Low	20 (50.0)	10 (37.0)	4 (50.0)	283 (42.8)	317 (43.1)	
Medium	7 (17.5)	6 (22.2)	1 (12.5)	199 (30.1)	213 (28.9)	
High	2 (5.0)	7 (25.9)	3 (37.5)	125 (18.9)	137 (18.6)	
No response	0 (0)	0 (0)	0 (0)	1 (0.2)	1 (0.1)	
Prior COVID-19						
No	20 (50.0)	10 (37.0)	6 (75.0)	498 (75.3)	534 (72.6)	<.001
Unsure	10 (25.0)	9 (33.3)	0 (0)	76 (11.5)	95 (12.9)	
Yes, minor/no symptoms	4 (10.0)	5 (18.5)	2 (25.0)	46 (7.0)	57 (7.7)	
Yes, really sick	5 (12.5)	3 (11.1)	0 (0)	39 (5.9)	47 (6.4)	
No response	1 (2.5)	0 (0)	0 (0)	2 (0.3)	3 (0.4)	
Known somebody who had COVID-19, worst outcome						
No	5 (12.5)	1 (3.7)	3 (37.5)	41 (6.2)	50 (6.8)	<.001
Unsure	0 (0)	2 (7.4)	0 (0)	18 (2.7)	20 (2.7)	
Yes, only minor/no symptoms	18 (45.0)	7 (25.9)	3 (37.5)	121 (18.3)	149 (20.2)	
Yes, really sick	12 (30.0)	7 (25.9)	0 (0.0)	257 (38.9)	276 (37.5)	
Yes, died	4 (10.0)	10 (37.0)	2 (25.0)	223 (33.7)	239 (32.5)	
No response	1 (2.5)	0	0	1 (0.2)	2 (0.3)	
Belief that vaccines work to prevent diseases						
Not at all	4 (10.0)	0 (0)	0 (0)	0 (0)	4 (0.5)	<.001
A little	8 (20.0)	6 (22.2)	0 (0)	5 (0.8)	19 (2.6)	
A moderate amount	14 (35.0)	5 (18.5)	4 (50.0)	47 (7.1)	70 (9.5)	
A lot	13 (32.5)	16 (59.3)	4 (50.0)	608 (92.0)	641 (87.1)	
No response	1 (2.5)	0 (0)	0 (0)	1 (0.2)	2 (0.3)	
Typical receipt of flu vaccine						
No	31 (77.5)	13 (48.1)	1 (12.5)	92 (13.9)	137 (18.6)	<.001
Unsure	0 (0)	1 (3.7)	2 (25.0)	7 (1.1)	10 (1.4)	
Yes	8 (20.0)	12 (44.4)	5 (62.5)	557 (84.3)	582 (79.1)	
Prefer not to say/no response	1 (2.5)	1 (3.7)	0 (0)	5 (0.8)	7 (1.0)	
Worrisome allergies with COVID-19 vaccine						

Influence	Did NOT intend to be fully vaccinated (n=40), n (%)	Unsure of full vaccination intentions (n=27), n (%)	Intended to be fully vaccinated with no doses received (n=8), n (%)	Fully vaccinated OR intended to be with 1 dose received (n=661), n (%)	All (N=736), n (%)	P value
No	26 (65.0)	12 (44.4)	3 (37.5)	588 (89.0)	629 (85.5)	<.001
Unsure	6 (15.0)	3 (11.1)	2 (25.0)	27 (4.1)	38 (5.2)	
Yes	5 (12.5)	12 (44.4)	3 (37.5)	42 (6.4)	62 (8.4)	
Prefer not to say	3 (7.5)	0 (0)	0 (0)	4 (0.6)	7 (1.0)	

^aNot performed due to low cell counts.

Table 3. Barriers to vaccination.

Barrier ^a	Did NOT intend to be fully vaccinated (n=40), n (%)		Unsure of full vaccination intentions (n=27), n (%)		Intended to be fully vaccinated with no doses received (n=8), n (%)		Fully vaccinated OR intended to be with 1 dose received (n=661), n (%)	
	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot
Enough trusted information about the vaccine	18 (45.0)	22 (55.0)	8 (29.6)	18 (66.7)	1 (12.5)	7 (87.5)	573 (86.7)	87 (13.2)
Enough information about the vaccine in respondent language	30 (75.0)	10 (25.0)	19 (70.4)	7 (25.9)	3 (37.5)	5 (62.5)	611 (92.4)	43 (6.5)
Enough information on short-term vaccine side effects	22 (55.0)	18 (45.0)	8 (29.6)	18 (66.7)	3 (37.5)	5 (62.5)	560 (84.7)	95 (14.4)
Enough information on long-term vaccine side effects	11 (27.5)	29 (72.5)	4 (14.8)	22 (81.5)	2 (25.0)	6 (75.0)	519 (78.5)	134 (20.3)
Enough information about the vaccine that is specific to respondent's health conditions	13 (32.5)	27 (67.5)	6 (22.2)	21 (77.8)	2 (25.0)	6 (75.0)	552 (83.5)	96 (14.5)
Process of scheduling a vaccine appointment	40 (100)	0 (0)	25 (92.6)	1 (3.7)	5 (62.5)	3 (37.5)	518 (78.4)	142 (21.5)
Possible side effects from the vaccine	8 (20.0)	32 (80.0)	5 (18.5)	22 (81.5)	0 (0)	8 (100)	596 (90.2)	61 (9.2)
Time it takes to get the vaccine	39 (97.5)	1 (2.5)	23 (85.2)	3 (11.1)	3 (37.5)	5 (62.5)	603 (91.2)	52 (7.9)
Time off needed from daily responsibilities if side effects were experienced	26 (65.0)	14 (35.0)	17 (63.0)	9 (33.3)	3 (37.5)	5 (62.5)	587 (88.8)	68 (10.3)
Child supervision ^b	30 (75.0)	1 (2.5)	21 (77.8)	2 (7.4)	7 (87.5)	1 (12.5)	434 (65.7)	14 (2.1)

^aThere are no responses that are not shown in the table.

^bNot applicable for all respondents

Table 4. Facilitators to vaccination.

Facilitator ^a	Did NOT intend to be fully vaccinated (n=40), n (%)		Unsure of full vaccination intentions (n=27), n (%)		Intended to be fully vaccinated with no doses received (n=8), n (%)		Fully vaccinated OR intended to be with 1 dose received (n=661), n (%)	
	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot
Somebody trusted tells to get the vaccine	39 (97.5)	1 (2.5)	21 (77.8)	6 (22.2)	1 (12.5)	7 (87.5)	427 (64.6)	227 (34.3)
Interaction with other people who are at high risk of getting really sick from COVID-19	36 (90.0)	4 (10.0)	15 (55.6)	12 (44.4)	2 (25.0)	6 (75.0)	294 (44.5)	361 (54.6)
People around the respondent get the vaccine	40 (100)	0 (0)	23 (85.2)	4 (14.8)	3 (37.5)	5 (62.5)	428 (64.8)	224 (33.9)
More information about how well the vaccine works	35 (87.5)	5 (12.5)	13 (48.1)	14 (51.9)	0 (0)	8 (100)	207 (31.3)	450 (68.1)
Getting the vaccine at primary care provider's office	37 (92.5)	1 (2.5)	18 (66.7)	9 (33.3)	3 (37.5)	5 (62.5)	568 (85.9)	85 (12.9)
Getting the vaccine close to respondent	37 (92.5)	1 (2.5)	19 (70.4)	8 (29.6)	3 (37.5)	5 (62.5)	258 (39.0)	395 (59.8)
Paid to get the vaccine	38 (95.0)	1 (2.5)	24 (88.9)	3 (11.1)	3 (37.5)	5 (62.5)	645 (97.6)	11 (1.7)
Required for work	37 (92.5)	2 (5.0)	18 (66.7)	9 (33.3)	0 (0)	8 (100)	634 (95.9)	20 (3.0)
Required for travel	39 (97.5)	0 (0)	19 (70.4)	7 (25.9)	1 (12.5)	7 (87.5)	559 (84.6)	93 (14.1)
No longer have to wear a mask	38 (95.0)	1 (2.5)	19 (70.4)	7 (25.9)	0 (0)	8 (100)	462 (69.9)	193 (29.2)

^aThere are no responses that are not shown in the table.

Table 5. Trust in sources for vaccine information

Source ^a	Did NOT intend to be fully vaccinated (n=40), n (%)		Unsure of full vaccination intentions (n=27), n (%)		Intended to be fully vaccinated with no doses received (n=8), n (%)		Fully vaccinated OR intended to be with 1 dose received (n=661), n (%)	
	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot	Not at all/little	Moderately/a lot
Family	31 (77.5)	8 (20.0)	21 (77.8)	6 (22.2)	2 (25.0)	6 (75.0)	386 (58.4)	269 (40.7)
Friends	32 (80.0)	6 (15.0)	22 (81.5)	4 (14.8)	4 (50.0)	4 (50.0)	445 (67.3)	206 (31.2)
Primary care provider/doctor	33 (82.5)	7 (17.5)	11 (40.7)	16 (59.3)	1 (12.5)	7 (87.5)	51 (7.7)	606 (91.7)
Pharmacist	33 (82.5)	7 (17.5)	14 (51.9)	13 (48.1)	1 (12.5)	7 (87.5)	86 (13.0)	568 (85.9)
Community leaders	35 (87.5)	0 (0)	22 (81.5)	3 (11.1)	3 (37.5)	4 (50.0)	409 (61.9)	194 (29.3)
Local news	34 (85.0)	0 (0)	24 (88.9)	2 (7.4)	4 (50.0)	3 (37.5)	402 (60.8)	204 (30.9)
National news	35 (87.5)	0 (0)	22 (81.5)	3 (11.1)	2 (25.0)	5 (62.5)	310 (46.9)	300 (45.4)
Social media	36 (90.0)	0 (0)	24 (88.9)	2 (7.4)	2 (25.0)	5 (62.5)	539 (81.5)	73 (11.0)
Celebrities	34 (85.0)	0 (0)	24 (88.9)	2 (7.4)	3 (37.5)	4 (50.0)	576 (87.1)	36 (5.4)

^aThere are no responses that are not shown in the table.

The survey also included open-ended responses related to reasons for vaccination intentions, barriers, and facilitators. Most of these responses corroborated trends seen in the quantitative data. In those who did not intend to be vaccinated, other notable reasons for not getting vaccinated included the lack of Food and Drug Administration (FDA) approval and the politics and related political pressures surrounding vaccination. Many people in this cohort also noted that “nothing” would make them choose to get fully vaccinated. In those who intended

to be fully vaccinated or were fully vaccinated, a major reason noted for their choice was to prevent the spread of COVID-19 and protecting themselves and others. Lastly, in those who had been fully vaccinated, some noted no concerns with vaccination, while others noted the new vaccine, speed of development, effectiveness, and allergic reactions as concerns.

Discussion

Principal Findings

The project survey results showed a significant difference in characteristics and influences between those who were COVID-19 vaccine-hesitant (refusing or delaying) and those who had been vaccinated. Similar to national surveys [24,25], there were differences across gender and political preferences, with a greater percentage of men and Republicans in the did not intend to be fully vaccinated cohort versus the vaccinated cohort. Likewise, COVID-19 risk perceptions among those not planning vaccination are lower [24]. Such characteristics, as well as others identified in Table 1, are especially relevant given the demographic characteristics of Alaska and Idaho. For example, more Alaskans and Idahoans politically identify as Republican than Democrat or Independent [26,27].

Although the sample size for respondents in the unsure or intended to be vaccinated cohorts were smaller, these individuals may be especially important to target in vaccination efforts. Addressing vaccine safety, transparency, and sources of information may encourage some individuals to get vaccinated. More information about how well the vaccine works was also seen as an important facilitator. Conveying this information to the lay public can be difficult given the technical and scientific details related to vaccine mechanism of action and how vaccine effectiveness data are calculated, reported, and interpreted [28].

The safety of vaccines (particularly long-term side effects) was noted as a barrier in 80% (28/35) of respondents in the unsure and intended to be vaccinated cohorts. Although case reports have revealed legitimate safety concerns with vaccination (eg, myocarditis and pericarditis with messenger RNA vaccines), risks of severe adverse events are still low [29]. Furthermore, risks of severe adverse events are even higher during and after SAR-CoV-2 infection [30]. A frequent message from the Centers for Disease Control and Prevention has been that COVID-19 vaccines are safe and effective; however, this messaging is not tailored and does not address individual-specific concerns [31]. At least 75% (27/35, 77%) of survey respondents unsure of or delaying vaccination (but intended to be vaccinated) indicated that having enough information about the vaccine that is specific to their health conditions was also a barrier. Therefore, communication needs to also be personalized. It may also be worthwhile to explore ways that local influencers can share their personal experiences or even health systems sharing local data on demographics or characteristics of those vaccinated and their outcomes.

Comparison With Prior Work

Although much previous related work has been done on this topic, the results presented here are unique given the focus on 2 rural states with continued lower vaccination rates. The identified barriers can be utilized with other resources to facilitate vaccination. Along with addressing individual-specific concerns related to vaccine safety and side effects, national organizations have also made recommendations on word choices to improve vaccine acceptance [32]. A 2020 survey by the de Beaumont Foundation (n=1400) found that family was an especially important motivator for vaccination [32]. Therefore,

when discussing the benefits of vaccination, focusing on family may be more helpful than focusing on the country, community, or friends. Although the de Beaumont Foundation data suggest that some individuals may be motivated to receive vaccination for their family, results from the VACCINE project survey indicated that being told by a trusted source to get the COVID-19 vaccine was not a facilitator in most cohorts. Furthermore, family members were only seen as a trusted source of vaccine information by ~20% (14/67, 21%) of respondents in the did not intend to be vaccinated and unsure of vaccination cohorts. However, 75% (6/8) of respondents in the intending to be vaccinated cohorts stated that family was a trusted source, but these results may be skewed due to the lower sample size.

For those unsure or intended to be vaccinated, health care personnel (primary care provider or doctor and pharmacist) were the most trusted source for vaccine information. Health care providers can leverage their position to provide a strong recommendation for vaccination, which has been shown in the past to increase likelihood of vaccination against influenza [31,33-35]. Professionals play a key role in influencing the decision to receive the influenza vaccine. Information about influenza and its vaccine needs to be combined with improvements in service provision if overall target uptake rates of 70% (65% in those aged 65 years and over) are to be achieved [33,34]. Of concern is vaccine misinformation (and disinformation) especially from health care providers [36,37], which has been discussed in recent news outlets. Misinformation can impact intention to vaccinate [38] and has been identified by the US Surgeon General as an urgent public threat [39].

Limitations

The landscape of the COVID-19 pandemic is rapidly changing, which may impact some of the findings from this cross-sectional work. Since the VACCINE survey was disseminated in mid-June, the proportion of the circulating virus as the Omicron variant increased from 6.3% to 78.4% (as of March 2022), and new variants have arisen. In August 2021, the FDA approved the first COVID-19 vaccine [40]. Since this time period, there have also been increases in vaccination requirements (eg, as seen in President Biden's COVID-19 Action Plan). Although these national changes have been significant, vaccination rates in Alaska and Idaho are still dismal and below the national average. There continues to be a need to address barriers contributing to vaccine hesitancy in these rural states. Because this work focused specifically on respondents from these 2 states, the broader generalizability to other populations may be limited. Other limitations of this work include a small sample size, especially for those who intended to be fully vaccinated with no doses received, and the use of Facebook advertisements to recruit participants such as this may introduce response bias.

Conclusions

Efforts to counter vaccine misinformation, address hesitancy, and increase confidence continue to be underway to increase COVID-19 vaccination rates [41]. It is important that targeted approaches are taken in diverse communities (eg, rural areas). The project survey results from 2 rural states indicate that recognition of individual characteristics may influence vaccine choices. However, these individual characteristics represent

only a starting point in delivering tailored messages that should come from trusted sources to address vaccination barriers.

Acknowledgments

Thank you to the respondents for their time in completing our survey and to Lee Ann Waldron for her assistance with survey promotion. We appreciate the support of our advisory board members and our colleagues who helped in survey refinement.

This work was funded by a Merck Investigators Studies Program Grant: VAccination Challenges through Community-Identified NEEds (VACCINE). The funder was not directly involved in any of the study work.

Conflicts of Interest

KC works as a consultant for Seqirus, Inc and Idaho Immunization Coalition. All other authors declare that they have no conflicts of interest related to the manuscript.

Multimedia Appendix 1

Study survey.

[[DOCX File, 42 KB - formative_v6i12e39109_app1.docx](#)]

Multimedia Appendix 2

Age distribution.

[[PDF File \(Adobe PDF File\), 40 KB - formative_v6i12e39109_app2.pdf](#)]

References

1. Andre F, Booy R, Bock H, Clemens J, Datta S, John T, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull World Health Organ* 2008 Feb 1;86(2):140-146 [FREE Full text] [doi: [10.2471/blt.07.040089](#)] [Medline: [18297169](#)]
2. Immunization. UNICEF. URL: <https://www.unicef.org/immunization> [accessed 2022-09-09]
3. Myths and Facts about COVID-19 Vaccines. Centers for Disease Control and Prevention. 2022 Jul 20. URL: <https://tinyurl.com/bdcp4weh> [accessed 2022-09-09]
4. Eibensteiner F, Ritschl V, Nawaz FA, Fazel SS, Tsagkaris C, Kulnik ST, et al. People's willingness to vaccinate against COVID-19 despite their safety concerns: Twitter poll analysis. *J Med Internet Res* 2021 Apr 29;23(4):e28973 [FREE Full text] [doi: [10.2196/28973](#)] [Medline: [33872185](#)]
5. Siddiqui M, Salmon DA, Omer SB. Epidemiology of vaccine hesitancy in the United States. *Hum Vaccin Immunother* 2013 Dec 01;9(12):2643-2648 [FREE Full text] [doi: [10.4161/hv.27243](#)] [Medline: [24247148](#)]
6. Building Confidence in COVID-19 Vaccines. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/vaccines/covid-19/vaccinate-with-confidence.html> [accessed 2022-09-09]
7. World Health Organization. Report of the Sage Working Group on Vaccine Hesitancy. *MedBox*. 2014. URL: <https://www.medbox.org/document/report-of-the-sage-working-group-on-vaccine-hesitancy#GO> [accessed 2022-09-09]
8. Kumar D, Chandra R, Mathur M, Samdariya S, Kapoor N. Vaccine hesitancy: understanding better to address better. *Isr J Health Policy Res* 2016 Feb 1;5(1):2 [FREE Full text] [doi: [10.1186/s13584-016-0062-y](#)] [Medline: [26839681](#)]
9. Lewis JR. What is driving the decline in people's willingness to take the COVID-19 vaccine in the United States? *JAMA Health Forum* 2020 Nov 18;1(11):e201393. [doi: [10.1001/jamahealthforum.2020.1393](#)]
10. Herd immunity and COVID-19: What you need to know. *Mayo Clinic*. URL: <https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/herd-immunity-and-coronavirus/art-20486808> [accessed 2022-09-09]
11. Bolotin S, Wilson S, Murti M. Achieving and sustaining herd immunity to SARS-CoV-2. *CMAJ* 2021 Jul 19;193(28):E1089-E1089 [FREE Full text] [doi: [10.1503/cmaj.210892](#)] [Medline: [34281966](#)]
12. COVID Data Tracker. Centers for Disease Control and Prevention. URL: <https://covid.cdc.gov/covid-data-tracker/#datatracker-home> [accessed 2022-09-09]
13. Karafillakis E, Larson H. The paradox of vaccine hesitancy among healthcare professionals. *Clin Microbiol Infect* 2018 Aug;24(8):799-800 [FREE Full text] [doi: [10.1016/j.cmi.2018.04.001](#)] [Medline: [29649595](#)]
14. Lyu H, Wang J, Wu W, Duong V, Zhang X, Dye TD, et al. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. *Intell Med* 2022 Feb;2(1):1-12 [FREE Full text] [doi: [10.1016/j.imed.2021.08.001](#)] [Medline: [34457371](#)]
15. Savoia E, Piltch-Loeb R, Goldberg B, Miller-Idriss C, Hughes B, Montrond A, et al. Predictors of COVID-19 vaccine hesitancy: socio-demographics, co-morbidity, and past experience of racial discrimination. *Vaccines (Basel)* 2021 Jul 09;9(7):767 [FREE Full text] [doi: [10.3390/vaccines9070767](#)] [Medline: [34358184](#)]

16. Reiter PL, Pennell ML, Katz ML. Acceptability of a COVID-19 vaccine among adults in the United States: How many people would get vaccinated? *Vaccine* 2020 Sep 29;38(42):6500-6507 [FREE Full text] [doi: [10.1016/j.vaccine.2020.08.043](https://doi.org/10.1016/j.vaccine.2020.08.043)] [Medline: [32863069](https://pubmed.ncbi.nlm.nih.gov/32863069/)]
17. Lin Y, Hu Z, Zhao Q, Alias H, Danaee M, Wong LP. Understanding COVID-19 vaccine demand and hesitancy: A nationwide online survey in China. *PLoS Negl Trop Dis* 2020 Dec 17;14(12):e0008961 [FREE Full text] [doi: [10.1371/journal.pntd.0008961](https://doi.org/10.1371/journal.pntd.0008961)] [Medline: [33332359](https://pubmed.ncbi.nlm.nih.gov/33332359/)]
18. Zhou Y, Zhang J, Wu W, Liang M, Wu Q. Willingness to receive future COVID-19 vaccines following the COVID-19 epidemic in Shanghai, China. *BMC Public Health* 2021 Jun 09;21(1):1103 [FREE Full text] [doi: [10.1186/s12889-021-11174-0](https://doi.org/10.1186/s12889-021-11174-0)] [Medline: [34107930](https://pubmed.ncbi.nlm.nih.gov/34107930/)]
19. Machida M, Nakamura I, Kojima T, Saito R, Nakaya T, Hanibuchi T, et al. Acceptance of a COVID-19 vaccine in Japan during the COVID-19 pandemic. *Vaccines (Basel)* 2021 Mar 03;9(3):210 [FREE Full text] [doi: [10.3390/vaccines9030210](https://doi.org/10.3390/vaccines9030210)] [Medline: [33802285](https://pubmed.ncbi.nlm.nih.gov/33802285/)]
20. Bell S, Clarke R, Mounier-Jack S, Walker JL, Paterson P. Parents' and guardians' views on the acceptability of a future COVID-19 vaccine: A multi-methods study in England. *Vaccine* 2020 Nov 17;38(49):7789-7798 [FREE Full text] [doi: [10.1016/j.vaccine.2020.10.027](https://doi.org/10.1016/j.vaccine.2020.10.027)] [Medline: [33109389](https://pubmed.ncbi.nlm.nih.gov/33109389/)]
21. Kumari A, Ranjan P, Chopra S, Kaur D, Kaur T, Upadhyay AD, et al. Knowledge, barriers and facilitators regarding COVID-19 vaccine and vaccination programme among the general population: A cross-sectional survey from one thousand two hundred and forty-nine participants. *Diabetes Metab Syndr* 2021 May;15(3):987-992 [FREE Full text] [doi: [10.1016/j.dsx.2021.04.015](https://doi.org/10.1016/j.dsx.2021.04.015)] [Medline: [33984818](https://pubmed.ncbi.nlm.nih.gov/33984818/)]
22. Larson H, Schulz W, Tucker J, Smith D. Measuring vaccine confidence: introducing a global vaccine confidence index. *PLoS Curr* 2015 Feb 25;7:1 [FREE Full text] [doi: [10.1371/currents.outbreaks.ce0f6177bc97332602a8e3fe7d7f7cc4](https://doi.org/10.1371/currents.outbreaks.ce0f6177bc97332602a8e3fe7d7f7cc4)] [Medline: [25789200](https://pubmed.ncbi.nlm.nih.gov/25789200/)]
23. Larson HJ, Jarrett C, Schulz WS, Chaudhuri M, Zhou Y, Dube E, SAGE Working Group on Vaccine Hesitancy. Measuring vaccine hesitancy: The development of a survey tool. *Vaccine* 2015 Aug 14;33(34):4165-4175 [FREE Full text] [doi: [10.1016/j.vaccine.2015.04.037](https://doi.org/10.1016/j.vaccine.2015.04.037)] [Medline: [25896384](https://pubmed.ncbi.nlm.nih.gov/25896384/)]
24. KFF COVID-19 Vaccine Monitor: July 2021. Kaiser Family Foundation. 2021 Aug 04. URL: <https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-july-2021/> [accessed 2022-09-09]
25. Mentzer R. GOP Men Are Most Likely To Say They'll Refuse COVID-19 Vaccine. Wisconsin Public Radio. 2021 Mar 17. URL: <https://www.wpr.org/gop-men-are-most-likely-say-theyll-refuse-covid-19-vaccine> [accessed 2022-09-09]
26. Party affiliation among adults in Alaska. Pew Research Center. URL: <https://www.pewforum.org/religious-landscape-study/state/alaska/party-affiliation/> [accessed 2022-09-09]
27. Party affiliation among adults in Idaho. Pew Research Center. URL: <https://www.pewforum.org/religious-landscape-study/state/idaho/party-affiliation/> [accessed 2022-09-09]
28. Okuhara T, Ishikawa H, Ueno H, Okada H, Kato M, Kiuchi T. Readability assessment of vaccine information: A systematic review for addressing vaccine hesitancy. *Patient Educ Couns* 2022 Feb;105(2):331-338. [doi: [10.1016/j.pec.2021.05.039](https://doi.org/10.1016/j.pec.2021.05.039)] [Medline: [34090717](https://pubmed.ncbi.nlm.nih.gov/34090717/)]
29. Myocarditis and Pericarditis After mRNA COVID-19 Vaccination. Centers for Disease Control and Prevention. 2022 Jul 14. URL: <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/myocarditis.html> [accessed 2022-09-09]
30. Barda N, Dagan N, Ben-Shlomo Y, Kepten E, Waxman J, Ohana R, et al. Safety of the BNT162b2 mRNA Covid-19 vaccine in a nationwide setting. *N Engl J Med* 2021 Sep 16;385(12):1078-1090. [doi: [10.1056/nejmoa2110475](https://doi.org/10.1056/nejmoa2110475)]
31. Vaccine Safety. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/vaccinesafety/index.html> [accessed 2022-09-09]
32. New Poll Reveals Most Effective Language to Improve COVID-19 Vaccine Acceptance. de Beaumont. 2020 Dec. URL: <https://debeaumont.org/changing-the-covid-conversation/vaccineacceptance/> [accessed 2022-09-09]
33. Lewis-Parmer H, McCann R. Achieving national influenza vaccine targets--an investigation of the factors affecting influenza vaccine uptake in older people and people with diabetes. *Commun Dis Public Health* 2002 Jun;5(2):119-126. [Medline: [12166297](https://pubmed.ncbi.nlm.nih.gov/12166297/)]
34. Lin CJ, Nowalk MP, Zimmerman RK, Ko F, Zoffel L, Hoberman A, et al. Beliefs and attitudes about influenza immunization among parents of children with chronic medical conditions over a two-year period. *J Urban Health* 2006 Sep 13;83(5):874-883 [FREE Full text] [doi: [10.1007/s11524-006-9084-z](https://doi.org/10.1007/s11524-006-9084-z)] [Medline: [16770701](https://pubmed.ncbi.nlm.nih.gov/16770701/)]
35. Kahn KE, Santibanez TA, Zhai Y, Bridges CB. Association between provider recommendation and influenza vaccination status among children. *Vaccine* 2018 Jun 07;36(24):3486-3497 [FREE Full text] [doi: [10.1016/j.vaccine.2018.04.077](https://doi.org/10.1016/j.vaccine.2018.04.077)] [Medline: [29764679](https://pubmed.ncbi.nlm.nih.gov/29764679/)]
36. Levy A. The social network for doctors is full of vaccine disinformation. CNBC. 2021 Aug 06. URL: <https://www.cnbc.com/2021/08/06/doximity-social-network-for-doctors-full-of-antivax-disinformation.html> [accessed 2022-09-09]
37. Brumfiel G. This Doctor Spread False Information About COVID. She Still Kept Her Medical License. NPR. 2021 Sep 14. URL: <https://www.npr.org/sections/health-shots/2021/09/14/1035915598/doctors-covid-misinformation-medical-license> [accessed 2022-09-09]

38. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat Hum Behav* 2021 Mar 05;5(3):337-348. [doi: [10.1038/s41562-021-01056-1](https://doi.org/10.1038/s41562-021-01056-1)] [Medline: [33547453](https://pubmed.ncbi.nlm.nih.gov/33547453/)]
39. Health Misinformation. Office of the U.S. Surgeon General. URL: <https://www.hhs.gov/surgeongeneral/reports-and-publications/health-misinformation/index.html> [accessed 2022-09-09]
40. FDA Approves First COVID-19 Vaccine. Food & Drug Administration. 2021 Aug 23. URL: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine> [accessed 2022-09-09]
41. What Is Vaccine Confidence? Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/vaccines/covid-19/vaccinate-with-confidence/building-trust.html> [accessed 2022-09-07]

Abbreviations

FDA: Food and Drug Administration

SAGE: Strategic Advisory Group on Experts

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

WHO: World Health Organization

Edited by A Mavragani; submitted 06.05.22; peer-reviewed by CH Lai, PCI Pang, J Luo; comments to author 29.07.22; revised version received 19.08.22; accepted 02.09.22; published 01.12.22.

Please cite as:

Nguyen E, Wright M, Holmes J, Cleveland K, Oliphant C, Nies M, Robinson R

Influences, Barriers, and Facilitators to COVID-19 Vaccination: Cross-sectional Survey on Vaccine Hesitancy in 2 Rural States
JMIR Form Res 2022;6(12):e39109

URL: <https://formative.jmir.org/2022/12/e39109>

doi: [10.2196/39109](https://doi.org/10.2196/39109)

PMID: [36067411](https://pubmed.ncbi.nlm.nih.gov/36067411/)

©Elaine Nguyen, Melanie Wright, John Holmes, Kevin Cleveland, Catherine Oliphant, Mary Nies, Renee Robinson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 01.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Impact of Telehealth on the Delivery of Prenatal Care During the COVID-19 Pandemic: Mixed Methods Study of the Barriers and Opportunities to Improve Health Care Communication in Discussions About Pregnancy and Prenatal Genetic Testing

Caitlin G Craighead¹, MPH; Christina Collart¹, MEd; Richard Frankel², PhD; Susannah Rose³, PhD; Anita D Misra-Hebert⁴, MD, MPH; Brownsyne Tucker Edmonds⁵, MD, MPH; Marsha Michie⁶, PhD; Edward Chien¹, MD, MBA; Marissa Coleridge^{1,7}, MS; Oluwatosin Goje¹, MD; Angela C Ranzini⁸, MD; Ruth M Farrell^{1,7,9}, MD, MA

¹Obstetrics and Gynecology and Women's Health Institute, Cleveland Clinic, Cleveland, OH, United States

²Research Innovation and Education, Cleveland Clinic Lerner College of Medicine, Cleveland Clinic, Cleveland, OH, United States

³Center for Patient Experience, Cleveland Clinic, Cleveland, OH, United States

⁴Department of Internal Medicine, Cleveland Clinic Community Care, Cleveland Clinic, Cleveland, OH, United States

⁵Department of Obstetrics and Gynecology, Indiana University, Indianapolis, IN, United States

⁶Department of Bioethics, Case Western Reserve University, Cleveland, OH, United States

⁷Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH, United States

⁸Department of Obstetrics and Gynecology, MetroHealth Medical Center, Cleveland, OH, United States

⁹Center for Bioethics, Cleveland Clinic, Cleveland, OH, United States

Corresponding Author:

Ruth M Farrell, MD, MA

Obstetrics and Gynecology and Women's Health Institute

Cleveland Clinic

9500 Euclid Ave

Cleveland, OH, 44195

United States

Phone: 1 216 445 7085

Email: farrelr@ccf.org

Abstract

Background: The COVID-19 pandemic brought significant changes in health care, specifically the accelerated use of telehealth. Given the unique aspects of prenatal care, it is important to understand the impact of telehealth on health care communication and quality, and patient satisfaction. This mixed methods study examined the challenges associated with the rapid and broad implementation of telehealth for prenatal care delivery during the pandemic.

Objective: In this study, we examined patients' perspectives, preferences, and experiences during the COVID-19 pandemic, with the aim of supporting the development of successful models to serve the needs of pregnant patients, obstetric providers, and health care systems during this time.

Methods: Pregnant patients who received outpatient prenatal care in Cleveland, Ohio participated in in-depth interviews and completed the Coronavirus Perinatal Experiences-Impact Survey (COPE-IS) between January and December 2021. Transcripts were coded using NVivo 12, and qualitative analysis was used, an approach consistent with the grounded theory. Quantitative data were summarized and integrated during analysis.

Results: Thematic saturation was achieved with 60 interviews. We learned that 58% (35/60) of women had telehealth experience prior to their current pregnancy. However, only 8% (5/60) of women had used both in-person and virtual visits during this pregnancy, while the majority (54/60, 90%) of women participated in only in-person visits. Among 59 women who responded to the COPE-IS, 59 (100%) felt very well supported by their provider, 31 (53%) were moderately to highly concerned about their child's health, and 17 (29%) reported that the single greatest stress of COVID-19 was its impact on their child. Lead themes focused on establishing patient-provider relationships that supported shared decision-making, accessing the information needed for shared decision-making, and using technology effectively to foster discussions during the COVID-19 pandemic. Key findings

indicated that participants felt in-person visits were more personal, established greater rapport, and built better trust in the patient-provider relationship as compared to telehealth visits. Further, participants felt they could achieve a greater dialogue and ask more questions regarding time-sensitive information, including prenatal genetic testing information, through an in-person visit. Finally, privacy concerns arose if prenatal genetic testing or general pregnancy conversations were to take place outside of the health care facility.

Conclusions: While telehealth was recognized as an option to ensure timely access to prenatal care during the COVID-19 pandemic, it also came with multiple challenges for the patient-provider relationship. These findings highlighted the barriers and opportunities to achieve effective and patient-centered communication with the continued integration of telehealth in prenatal care delivery. It is important to address the unique needs of this population during the pandemic and as health care increasingly adopts a telehealth model.

(*JMIR Form Res* 2022;6(12):e38821) doi:[10.2196/38821](https://doi.org/10.2196/38821)

KEYWORDS

prenatal health care delivery; health care communication; telehealth; access to health care; COVID-19; pregnancy

Introduction

Health care delivery has changed dramatically as a result of the COVID-19 pandemic, with telehealth as a major resource to maintain health care access during this time [1,2]. Telehealth had been developing into an accepted modality to deliver health care prior to the COVID-19 pandemic, with data emerging about health care quality, patient-provider communication, and patient satisfaction with this alternative approach to in-person care [3-5]. Although the application of telehealth in obstetric care is not new, the integration of telehealth within prenatal care had not been broadly implemented in many practices across the United States prior to the pandemic [6]. The pandemic accelerated its implementation and, in doing so, shed light on some of the most significant benefits, drawbacks, and challenges of its rapid implementation across diverse patient populations, in addition to the need to study outcomes using this modality [6-9].

Telehealth was a particularly important prenatal care strategy to maintain health care access while helping to prevent viral exposure to pregnant patients and health care providers, as well as communities [6-8]. Early on in the pandemic, it was established that pregnant patients infected with SARS-CoV-2 had increased risks of intensive care unit admission and death [10]. Thus, it was critical to use telehealth to help avoid exposure, since there was limited data about the best approaches for infection prevention and management among the pregnant population. The need for regular and timely access to care was additionally significant; prenatal care delivery involved multiple time-sensitive and potentially complex health care decisions. Prenatal genetic screens and diagnostic tests are examples, which are time-sensitive with respect to the gestational age at which the testing windows open and close [11]. A delay in access to information about these tests, or use of these tests, can have major implications for the outcome of the pregnancy [12]. Because of factors such as these, telehealth was rapidly implemented across average-risk and high-risk patients, and at different gestational ages [13-15]. Yet, there was limited opportunity to understand the impact of uptake on patient experience and key markers of health care quality during the pandemic, a time in which patients had a new and additional set of informational priorities about the prevention and

management of SARS-CoV-2 infection in the health care discussion.

Prenatal care is complex, even at the best of times, and clear communication from clinicians and comprehension by patients can be challenging. The increased use of non-face-to-face communication modalities (eg, telephone and virtual video visits) during the pandemic has introduced greater complexity as well as opportunities and risks to interpersonal communication. Little is currently known about the impact of virtual visits in novel contexts such as a pandemic. To address this gap, we conducted a mixed methods study to better understand patients' lived experiences of virtual prenatal care.

Methods

Study Design

This study was developed as a mixed methods study to explore emerging concepts and themes as they relate to obstetric health care delivery and patient experience during the pandemic.

Recruitment

Participants were 18 years of age or older, were English speaking, had a viable intrauterine pregnancy, and received outpatient obstetric care. We recruited pregnant women at outpatient centers within the Cleveland Clinic and MetroHealth health care systems between January and December 2021.

Participants were contacted by means of a recruitment letter. The letter invited patients, who met the inclusion criteria and were interested in sharing their knowledge and opinions of decision-making surrounding prenatal testing in light of the COVID-19 pandemic, to contact the research team. Recruitment was structured to seek input from 2 groups of patients who represented patients at different significant time points in pregnancy. One group included patients in the first trimester of pregnancy to capture prenatal care needs, preferences, and experiences at the onset of pregnancy and prenatal care delivery (Group 1). A second group included patients in the second trimester, who had already considered or undergone prenatal genetic screening or diagnostic testing at the time of the interview (Group 2). Recruitment continued until thematic saturation in interviews was reached.

Data Collection

After an informed consent process, each participant participated in a telephone interview to maintain consistency with the health care systems' recommendations for social distancing and patient contact for research purposes at the onset of the pandemic. Interviews were conducted by a member of the research team using a semistructured interview guide, which contained questions about knowledge and opinions on COVID-19, prenatal care delivery during the pandemic, accessing information about prenatal genetic testing during the pandemic, accessing prenatal genetic testing during the pandemic, health care system resources to support patients, and demographic and reproductive history. This guide was developed in conjunction with content experts in obstetrics, clinical genetics, medical decision - making, patient experience, and maternal - fetal medicine. With the participants' permission, the interviews were audio recorded and then transcribed verbatim for analysis.

Data were also collected using the self-administered Coronavirus Perinatal Experiences-Impact Survey (COPE-IS) [16]. The COPE-IS is a newly developed survey to understand the experiences of pregnant women during the COVID-19 pandemic. It has not been psychometrically tested at this time [17-19]. The survey was administered to participants after the telephone interview was completed to assess both the events and circumstances of women's lives as new or expectant mothers during the time of the global pandemic. The survey was administered via REDCap Survey accessed on a computer or mobile device, or if the patient preferred, a hard copy was mailed with a stamped envelope to send back to the study team once completed.

Statistical Analysis

Qualitative analysis was approached as an iterative and progressive process of data immersion, coding, memoing, and theme identification, which is an inductive process consistent with the grounded theory [20,21]. We identified content domains

and categories in transcripts to create a coding tree used to organize the data. A companion codebook was created to serve as a reference for the analysis. The coding and analysis processes were led by 2 members of the study team (RMF and CGC) using NVivo (version 12; QSR International). The research team held weekly meetings to review data coding and memoing, and identify themes. Themes identified were contextualized with information about the trimester of pregnancy, gravity/parity, and previous pregnancies. Data from the COPE-IS and demographic information were summarized as frequency and mean. Quantitative data were summarized and integrated during analysis.

Ethical Considerations

This study was reviewed in advance, approved, and monitored by the Cleveland Clinic Institutional Review Board (IRB number 20-1333). The Institutional Review Board approved a waiver of remote written consent from participants via DocuSign, a 21 Code of Federal Regulations Part 11-compliant electronic signature platform. All identifying information inadvertently disclosed by the study participants during the interviews were deleted from the original data file, and all study participant data were deidentified. Additionally, all participants received a US \$50 gift card after the completion of the interview and COPE-IS.

Results

Participant Demographics

Thematic saturation was achieved with 60 interviews. Of the 60 patients, 30 were in their first trimester (Group 1) and 30 were in their second trimester (Group 2). The average age of the participants was 31 (SD 4.28) years. Moreover, this was the first pregnancy for 22 (37%) women, and 17 (28%) were considered to have an advanced maternal age (Table 1). Of the 60 women, 35 (58%) had telehealth experience prior to their pregnancy; however, 54 (90%) participated in only in-person visits during this pregnancy.

Table 1. Participant demographics.

Demographic	Value (N=60)
Age (years), mean (SD)	31.1 (4.28)
AMA^a status, n (%)	
Non-AMA (<35 years)	43 (72)
AMA (≥35 years)	17 (28)
Race, n (%)	
White	49 (82)
Black	3 (5)
Asian	3 (5)
Multiracial	2 (3)
Declined to answer	3 (5)
Reproductive history, n (%)	
Primigravida	22 (37)
Multigravida	38 (63)
Trimester of pregnancy	
1st trimester	30 (50)
2nd trimester	30 (50)
Prior telehealth experience	
Yes	35 (58)
No	20 (33)
Unsure	5 (8)
Visit type during this pregnancy	
In-person visit only	54 (90)
Virtual visit only	1 (2)
Hybrid, used both in-person and virtual visits	5 (8)

^aAMA: advanced maternal age.

COPE-IS

Almost all of the participants completed the COPE-IS (59/60, 98%), and among these, 59 (100%) indicated feeling very well supported by their primary care provider, 31 (53%) reported

feeling moderately to highly concerned about the impact of COVID-19 on their child's health, and 17 (29%) reported the single greatest stress due to COVID-19 was its impact on their child (Table 2). All data from the COPE-IS are presented in Multimedia Appendix 1.

Table 2. Coronavirus Perinatal Experiences-Impact Survey results.

Question and response	Value (N=59), n (%)
How well are you currently being supported by your primary prenatal care provider(s)?	
Very well supported	59 (100)
Somewhat well supported	0 (0)
Not very well supported	0 (0)
Has the support you receive from your prenatal care changed due to the COVID-19 outbreak?	
Significantly worsened	0 (0)
Somewhat worsened	1 (2)
No change	49 (83)
Somewhat improved	6 (10)
Significantly improved	3 (5)
Do you have any concerns about your child's health as a result of the COVID-19 outbreak?	
No	24 (41)
Yes (score)	35 (59)
1, no concern	0 (0)
2	0 (0)
3	4 (11)
4	6 (17)
5	10 (29)
6	7 (20)
7, highly concerned	8 (23)
In general, how distressed are you about your own COVID-19–related symptoms or potential illness? (score)	
1, no distress	18 (31)
2	6 (10)
3	10 (17)
4	6 (10)
5	13 (22)
6	5 (9)
7, highly distressed	1 (2)
How has the COVID-19 outbreak changed your stress levels or mental health?	
Worsened them significantly	6 (10)
Worsened them moderately	30 (51)
No change	20 (34)
Improved them moderately	3 (5)
Improved them significantly	0 (0)
Overall level of stress related to the COVID-19 outbreak (score)	
1, nothing	6 (10)
2	9 (15)
3	15 (25)
4	13 (22)
5	11 (19)
6	2 (3)
7, extreme	3 (5)

Question and response	Value (N=59), n (%)
What is the single greatest source of stress due to the COVID-19 outbreak right now? (check only one)	
Impact on your child	17 (29)
Health concerns	13 (22)
Impact on family members (eg, elderly parents)	9 (15)
Financial concerns	5 (9)
General well-being due to social distancing and/or quarantine	5 (9)
Impact on society	4 (7)
I am not stressed	3 (5)
Impact on your partner	2 (3)
Access to baby supplies (eg, formula, diapers, and wipes)	1 (2)
Impact on your community	0 (0)
Impact on close friends	0 (0)
Access to food	0 (0)
Access to mental health care	0 (0)
Stress about other aspects (open field)	0 (0)

Themes

Qualitative analysis identified the following primary themes: (1) establishing patient-provider relationships that supported

shared decision-making, (2) accessing the information needed for a shared decision-making process, and (3) using technology effectively to foster discussions during the COVID-19 pandemic. These themes and example quotes are presented in [Table 3](#).

Table 3. Themes and example quotes.

Theme	Quotes
Establishing patient-provider relationships that supported shared decision-making	<ul style="list-style-type: none"> • “Although like if she’s giving...if you’re giving somebody really terrible news, I think you would want them...I mean, it’s probably [easier] to be in-person just to make sure the person understands everything you know?” [Group 1, Participant #5] • “I would always put the face-to-face above because of the fluidity and the conversation that happens when you’re there in-person.” [Group 1, Participant #21] • “For me, it’s not like she spelled everything out for me and went over every little thing. She just said, ‘This was available to you. Do you know what you want to do?’ and, ‘Do you have any questions about it?’ And, I really didn’t at the time. I know they talked about the book they provided, and I looked up some of my own stuff online and heard stuff from other people. So, I just came to my own conclusions. But, for some people who might be really worried about prenatal testing or are on the fence about it, an in-person might be better to talk about it. They do feel a little different [in-person vs VV conversations]. It almost feels awkward in a way. That’s just what I got. It feels wrong. It felt a little more rushed, and I don’t know if that was because she was running a few minutes late so then she was running late to another one afterwards, or, she really didn’t have a lot of questions for me and I didn’t have a lot of questions for her. But yeah, it did have a little of a rushed feeling.” [Group 2, Participant #2] • “I was grateful to be able to have a pretty long in-person conversation with my OB/GYN about genetic testing on several occasions prior to the actual conversation with the genetic counselor. That made me feel at ease that I had the space and the time to have a pretty long conversation and ask a lot of questions on a couple of different occasions. So I’m grateful that have that the first time, be in-person and have the space to ask a lot of questions.” [Group 2, Participant #26] • “I think the benefits... I just... like that one-on-one interaction. Especially now being a stay at home mom with very little outings ... if I had a virtual appointment he’s [her son] all over me. So, it’s a little distracting, where I feel like I get...I have more of a clear head for a one-on-one interaction without him around.” [Group 2, Participant #14] • “For me personally like someone who’s gonna be best involved in intimate care, I just feel more comfortable meeting with the person.” [Group 1, Participant #14]
Accessing the information needed for a shared decision-making process	<ul style="list-style-type: none"> • “More of a personal touch...I think...that’s one of the hardest things right now is if you have to separate from people...so that’s kinda nice to have actual interaction face-to face...and then to me it just feels like...it is more secure, you are able to ask questions...I don’t have to worry about a screaming child in the back so I don’t have any distractions...I just like in-person visits better.” [Group 1, Participant #2] • “I think there could be benefits for the right people who are comfortable enough and confident enough and asking the right questions over, over video and things like that.” [Group 1, Participant #11] • “The downside, I keep saying, I think like it’s those questions that are asked I feel like are different then when I’ve done a virtual visit. I went to one virtual visit and it was very quick and then I ended up seeing my doctor in-person a few months later and the way questions were asked and things were discussed were completely different feeling when I did a virtual visit.” [Group 2, Participant #18] • “I forget to ask questions cause I find the virtual visits a little bit awkward and instead of being an in-person thing...” [Group 2, Participant #22]
Using technology effectively to foster discussions during the COVID-19 pandemic	<ul style="list-style-type: none"> • “For some reason, it didn’t connect for a while. So, I didn’t think that ... I thought I lost her. Like, I thought I almost missed the appointment because it wouldn’t connect or it was being goofy.” [Group 1, Participant #8] • “I like to talk to people in-person. I don’t like to be over the phone or people in my house. So, going inside the hospital is much better for me.” [Group 2, Participant #25] • “I do as long as I can find a space to have that conversation. Usually those conversations happen while I am at work. So I’ve had to remove myself and find a private...go sit in my car. But I am not concerned about my information being, you know the security of it.” [Group 2, Participant #26] • “Hardest thing for me was finding a quite space while the kids were home to have them [the appointments].” [Group 2, Participant #18]

Establishing Patient-Provider Relationships as the Basis for Shared Decision-Making During the COVID-19 Pandemic

The strength of the patient-provider relationship and trust were important themes. Participants expressed different opinions about the quality of the patient-provider relationship when engaging in in-person or telehealth visits. Some participants characterized telehealth visits as a more convenient and safer way to receive prenatal care during the pandemic. One participant who experienced telehealth during pregnancy stated:

It’s a minimizing exposure thing. I think it’s very empowering to the client to have to, or the patient to have to take their own blood pressure, it helps with their accountability and their engagement in their care to have to be able to, Doppler, their fetal heart rate and all that stuff. So, I think that would be a benefit. Also, being in the comfort of your own home is kind of nice. [Group 1, Participant #19]

For some, the level of preference regarding telehealth was a function of prior experience with pregnancy.

I guess because of all of that, because of the convenience, because of COVID, I now feel comfortable with my second one [pregnancy] to do some visits virtually and don't really have concerns that I wouldn't see her as often. [Group 2, Participant #13]

In addition, some did not see a significant difference in how health care communication may unfold in this setting.

I'm sure they would give you the same information either way, I mean the information is the same, how it is being delivered is different. [Group 2, Participant #27]

However, most of the participants indicated that in-person visits were “more personal” and had a greater potential for “establishing a rapport” with the provider.

The drawback [to virtual visits] is definitely the personal touch, just talking to somebody through a screen. So you're not actually there and maybe feeling like, if you were worried about something you wouldn't feel the same emotions I guess coming from your doctor. [Group 2, Participant #2]

For some, this modality hindered the development of trust with their provider.

I don't feel the same trust that you can build with a doctor if you are face to face with them versus on a tele-call. [Group 2, Participant #27]

This trust was particularly important during pregnancy and with respect to the nature of the issues and decisions that are often made in this setting. Compared to other clinical settings, prenatal care was described as a “personal” visit with a unique “level of intimacy” needed when discussing issues of reproduction and pregnancy.

I do think it's easier to build a relationship with your provider in-person, which I think for this kind of thing, for pregnancy, is important, at least for me. [Group 1, Participant #11]

Trust and relationship building were important in themselves, but also contributed to the quality of patient-provider communication. Participants reflected that they were accustomed to in-person visits and the kinds of patient-provider interactions that take place in the consultation or exam room. The transition to a virtual visit modality was unfamiliar to many, and for this reason, they reported that this was a very different health care delivery experience for them. This difference was due, in part, to their prior experience of in-person visits and having a level of comfort with discussions in this format. One participant who reflected on her preference for in-person visits stated:

I feel like in-person I'd definitely feel more at ease talking about it then over the phone or wherever virtually... [Group 1, Participant #16]

This was also due to an overall lower familiarity with using a video conferencing platform for prenatal care prior to the pandemic, particularly at the beginning of a new patient-provider relationship. One participant stated:

During a virtual visit I think it will be harder for me I guess ... just like you breaking out of the shell like, you know, on a Zoom call to bring something up. [Group 2, Participant #5]

In-person discussions were described as interactions in which “conversation is more natural than like through Zoom or whatever...FaceTime” [Group 2, Participant #18]. This was the case even if there was baseline familiarity with using these platforms for nonmedical reasons.

I think it's just easier to communicate. You know we've all had one million Zoom meetings at this point. So, you know, I think we've all gotten better at communicating through the virtual methods. But I also think that there's just something that can't be replaced about, you know, questions that come up in the moment. And it's sort of easier to talk in-person... [Group 1, Participant #4]

Overall, participants reflected on the need to learn about how best to use telehealth to obtain the same level of experience they expected and were accustomed to with in-person visits.

The need for in-person interaction was even more important in situations that called for an accurate, time-sensitive, and patient-centered discussion for care planning. One participant who reflected on a specific connection with her health care provider now that her pregnancy was at increased risk because of maternal age stated:

I like to look my doctor in the eye and I feel like I'm older in pregnancy now. So, I feel like I just...I need that one-on-one. [Group 1, Participant #6]

This type of interaction was even more important when there was a chance of a difficult conversation, including in cases when the patient received information about a potential problem with the pregnancy.

It's probably easier to be in-person just to make sure the person understands everything you know? [Group 1, Participant #5]

An in-person dynamic also played a role in the setting of conversations regarding care decisions that may be complex (eg, multiple different options or steps in a testing algorithm), entrenched in patient values, or potentially sensitive. One participant described the nature of prenatal genetic testing decisions and the need to have accurate information from their health care provider about the choices as follows:

I think it's [prenatal genetic testing] a touchy subject and there's a lot of different, you know, advice out there about having that knowledge and the benefits and also cons of it. So, I think having that conversation in-person it made me feel more comfortable. [Group 1, Participant #10]

This level of communication was also seen as a benefit to help mitigate the uncertainty that can be associated with these decisions. One participant stated:

For me personally, I knew I wanted to do the genetic testing. So, I don't think it would have made a huge difference but I can see, for some people again, who

might be unsure of what the right choice is for them, I think again in-person, it just makes that communication easier to talk about different options of why one may or may not choose genetic testing. [Group 1, Participant #18]

Some participants noted that nonverbal communication was a key component of successful conversations and that in-person interactions were an important opportunity for the provider to assess how to approach a conversation or the level of the patient's understanding and for the patient to know how to interpret the provider's information.

I think it's important for them [the healthcare provider] to just see what you look like and, you know, how you're doing, how you're reacting to things, like I think there's a lot that goes into nonverbal communication and that's harder to do when you're not in person. [Group 1, Participant #18]

Participants discussed potential opportunities for issues to be minimized. One such opportunity was for health care provider education in order to be certain that a telehealth visit would provide an equivalent level of patient experience and health care communication that fostered a shared decision-making process. One participant stated:

I also had conversations with the genetic counselors and she was very good at the skill of virtual visits so she allowed that kind of space. But I don't think that is characteristic of every healthcare professional. She did settle in and have a long conversation and a long phone conversation as well and sat through my repetitive questions. Because she was good at the skill, a virtual appointment was okay. But I don't think that style of conversation is indicative of every clinical provider. [Group 2, Participant #26]

Another suggested opportunity was a hybrid model that alternated telehealth visits with in-person visits. This approach would allow for the patient to establish a relationship and obtain additional information in a conventional health care format as needed for prenatal care decision-making.

So, maybe at this point in the game, I would be okay with doing a virtual visit as long as I knew maybe next month we'd meet in person. I wouldn't want to do virtual visits my entire pregnancy but I would be okay with doing it every now and then. [Group 2, Participant #4]

The success of such a hybrid model would be related to an individualized approach that would be dependent upon the estimated gestational age of the pregnancy with associated prenatal care milestones and the needs of the patient.

Accessing the Information Needed for a Shared Decision-Making Process During the COVID-19 Pandemic

Underlying these comments was a concern that key prenatal care discussions may be limited and that telehealth could impede access to information about care choices. Many participants suggested that they might obtain more information about their

obstetric care during an in-person visit. The preference for in-person discussion primarily pertained to the ability to have a "more natural" and dynamic discussion in which there was a shared decision-making conversation with the provider.

When you're face to face there's a little more openness and ability to think on your feet as to what other questions or areas of concern that you can bring up with the doctor. [Group 1, Participant #21]

Participants noted that in-person visits allowed them to ask questions to obtain the information they needed, while attending to their reproductive and medical history needs as well.

One participant for whom this was her first pregnancy stated:

This is my first pregnancy. There is a lot that I don't know. So, I find the flow of the conversation moves a million times better in an in-person conversation and there is more space to stop and think when someone asks, 'Do you have any questions?' I find it that the full conversation becomes truncated over the phone and there is less time for thinking and working through questions or letting information arrive in a conversation ... I find that more information gets volunteered, details get talked through in an in-person conversation versus a phone conversation ... The drawbacks [of a virtual visit], I see, are a less comprehensive visit, less comprehensive care, less opportunity for more information to come to light in the conversation. [Group 2, Participant #26]

For this participant, her absence of prior experience with pregnancy made her more uncertain about what questions she should ask, with concerns that if she did not initiate the question, she may not receive information that was important to her. In addition, some participants expressed preferences for modalities in which they learned best, including how they received the information and were able to retain and integrate it into health care decisions. One participant stated:

I feel like it helps to be in-person because of the fact that it's easier to retain the information than over a video call. [Group 1, Participant #30]

Thus, there were questions about whether and how the telehealth visit may affect their ability to obtain the information they desired and process that information in a way that would support informed decision-making.

For several participants, in-person visits were preferred due to the chance to mentally prepare for the visit and think through questions and goals for the discussion.

If I am running around the house and then I log on I'm probably not thinking about the visit as much or like the questions that I want to ask her... And sometimes I think- maybe I think of more questions cause I'm sitting in the waiting room and I'm actually think about it, rather than like running around the house and I'm logging on real quick for an appointment. [Group 2, Participant #8]

Participants also spoke of concerns about "distractions" in the home environment that may negatively impact the ability to

obtain information about their health care and potentially decrease the level of communication that allows for a shared decision-making process.

There is so much going on with being at home during the pandemic, that it is difficult to find the focus to concentrate on the discussion at hand. [Group 1, Participant #25]

Using Technology Effectively to Foster Discussions During the COVID-19 Pandemic

Participants also discussed concerns about technology-related factors associated with virtual visits, something they did not have to worry about during an in-person visit. One participant discussed her difficulties with obtaining a clear connection with her health care provider as follows:

There's just technical difficulties and usually it's on like an internet connection type level, sometimes the calls a little choppy or sometimes patients have a difficult time getting on, you know, just some basic, technical issues. [Group 2, Participant #17]

Participants also expressed concerns about how technological or internet issues could distract them from focusing on what was important to them during the conversation with their provider. One participant stated:

Maybe someone who doesn't know how to use the platform to do the appointment might get a little confused and frustrated. I will say, when I did my first appointment virtually, it wasn't a problem, but my doctor was late. And, I was like, 'Oh shoot. Wait...do I pick up the phone and call the front desk and ask if we need to reschedule or is she just running behind?' Usually, when you're in the waiting room or in the room waiting for the doctor to come in, which happens all the time, you know they are eventually going to come in. So, I feel like when you're waiting on a call like that and it's just a blank screen you're like, 'Shoot. Is something wrong with my computer or is it like she's running late?' I feel like there is just more question with that. [Group 2, Participant #2]

For this participant, the uncertainty about missing the appointment or having a failed connection caused her to be more distracted during the visit, with less time to prepare questions that she aimed to ask during the visit.

Participants identified concerns about privacy and the level of privacy that could be acquired during the visit. One level of concern pertained to safety and security associated with using a mobile or other personal device for health care, particularly applicable in the context of discussing topics relevant to reproductive history. One participant spoke of her reluctance to use internet-enabled devices for private discussions as follows:

I feel like personally we all, everybody, knows that the government kind of watches us and tracks us through our phones and everything. [Group 1, Participant #30]

Issues of privacy also pertained to a participant's ability to access a location for the telehealth visit when it took place outside of a consultation or exam room in the clinic, especially for those participants who worked outside of the home during the pandemic.

I don't have a private office. So, sometimes even calling to make a doctor's appointment is like a little... you know... I go out to my car or I try to get into a conference room to make that call just because that's a personal thing. You don't want your coworkers hearing, especially if you're calling to make your eight week appointment saying, 'Hey, I just took a pregnancy test. When can I come in?' So yeah, if I wasn't working from home, I would definitely want to go into the doctor's office in-person. [Group 2, Participant #3]

While some participants could find a private space at work or in their car, others struggled with whether and where they would have the resources for the kinds of discussions they needed with their health care provider. Privacy was also an issue for those participants who remained at home during the pandemic, several of whom raised concerns about access to a private space in the home for conversations about potentially highly sensitive and personal topics related to the pregnancy and their reproductive health. Some participants commented as follows:

I can see where someone would feel uncomfortable depending on where their setting is. [Group 2, Participant #27]

I don't have other kids in the house. I don't have other family. It's just my husband and I in our home. So, I have privacy in our home to carry on those conversations. I don't feel like I need to be in a doctor's office to have a safe conversation with somebody or private conversation with someone. So, I can see how for other people that that might not be their situation. They may not have as much privacy at home. [Group 2, Participant #4]

Discussion

Principal Findings

Prenatal care delivery is uniquely complex given the complexity and time-sensitive nature of the decisions that need to be made regarding maternal and fetal health. The rapid and robust introduction of telehealth has added an additional layer of complexity, with numerous variables that could interfere with effective patient-provider discussions. Our study demonstrates that there is a spectrum of opinions regarding pregnant patients' perceptions of the effects of telehealth on health care quality and satisfaction during the COVID-19 pandemic. Although telehealth has been available for many years, increase in its use during the pandemic had a wide-scale impact on obstetric health care providers and patients. Studies conducted both prior to and during the pandemic demonstrated that telehealth visits may improve access to health care, decrease childcare needs, eliminate transportation and parking costs, reduce office wait times, and, most importantly, minimize exposure to COVID-19 [22-25]. In response, a series of authors have established

protocols for integrating telehealth and hybrid models into prenatal care episodes [26-28]. Yet, a parallel discussion has suggested that increased satisfaction and convenience observed with telehealth visits may not equate to the same levels of health care quality and patient-centered care as observed during in-person visits [29]. Data are emerging that some patients may prefer lower technology visits, such as via the telephone, over those that involve a video component, according to factors related to patient characteristics [30]. As health care increasingly adopts telehealth models, it is important to determine how systems will adapt these methods to diverse patient and provider populations with different knowledge, resources, and skills to utilize telehealth overall and in the unique setting of prenatal care, in which often complex and time-sensitive decisions with significant implications for obstetric outcomes must be made. In addition, it is important to understand how factors, such as demographics and the presence of stressors (eg, the impact of the COVID-19 pandemic on individuals and pregnant patients), may play a role in telehealth implementation and utilization.

The findings of this study are significant as we identified several additional issues, apart from those related to technology. One important theme we identified was a concern about the barriers to health care communication resulting from the conversion of in-person visits to telehealth encounters. These barriers have both clinical and ethical implications as access to accurate and patient-centered information is a component of health care quality. Technological issues, such as the ability to access and use mobile devices with the appropriate level of broadband internet, in addition to familiarity with telehealth platforms, are factors contributing to the digital divide [31]. Issues with technology may lead to a cascade of downstream implications for patient-provider communication. This may begin with how individuals in the health care discussion express informational priorities in addition to how they exchange and receive relevant information in health care discussions. Ultimately, the interaction by and among individuals in the clinical encounter can have ramifications for the medical options presented by providers and the health care decisions that patients make during pregnancy.

We identified several other issues in addition to those associated with using technology. These included the degree of effective communication, trust in the therapeutic relationship, and patient-centered care that they had expected for their prenatal care or had experienced in prior pregnancies. These factors all related to the ability to seek and acquire information in a way that supported an informed decision-making process about their prenatal care. These are factors that may also relate to the degree of impact experienced from the COVID-19 pandemic, as demonstrated by participant responses to the COPE-IS. In part, these barriers were attributed to patients' unfamiliarity with differences in communication styles between telehealth and in-person visits, which patients were not aware of prior to the visits, and thus, they may not have had an opportunity to adapt their behaviors or actions accordingly. These findings echo observations of other researchers. For example, studies demonstrated that communication in telehealth visits is different from that in in-person visits in significant ways. Telehealth visits may be more physician-centered than patient-centered,

characterized by a communication style driven by provider-centered behaviors that make assumptions about the patients' interests and needs [29,32]. Telehealth visits have also been associated with less discussion about and orientation to agenda setting or additional patient concerns that come up during the visit, which represent aspects that are of key importance in the delivery of prenatal care [12]. In addition, there may be limited opportunities for patients to ask questions and relay their understanding of the key concepts of the conversation with the provider, which is a key aspect of patient-centered care [32]. These issues must be addressed for all patients, particularly those who face existing health care disparities and may face additional challenges that interfere with shared decision-making [31]. Notably, previous studies involved general internal medicine visits, raising the question of how the unique aspects of prenatal care delivery may be affected by variations in communication and patient-centeredness.

Our findings also bring to light novel issues of privacy with the integration of telehealth into reproductive health care. Prenatal care visits may address what patients may consider private, personal, and sensitive topics relating to parenthood and reproduction. Those discussions can be additionally salient when discussing prenatal genetic screening and diagnostic testing, in which issues related to heritable genetic risk factors, family history (eg, issues of paternity for the current pregnancy), disability, and pregnancy termination are discussed [12]. Participants raised concerns about prenatal telehealth relating to not only internet security, but also the ability to find a space in their home or workplace away from family, friends, or coworkers to have those conversations. This may be a particularly important factor among patient populations of lower socioeconomic status, where household crowding and housing instability are more common. In addition, home internet access may be unaffordable, requiring patients to access care in public libraries or other public venues. In turn, these concerns about privacy may have limited their ability to ask questions, provide responses, build trust, and engage in shared decision-making.

As telehealth continues to be integrated into prenatal care, it will be critical to discuss ways to prepare patients for some of the differences they may encounter between telehealth and in-person visits. These may include establishing resources, such as health care extenders, who can educate patients about the telehealth visit prior to their appointment, providing an orientation to the telehealth interface, and educating patients about differences they can expect from a virtual versus in-person visit. Though there was little opportunity to develop such strategies during the pandemic due to the urgent need to protect patients, especially pregnant patients, from exposure to SARS-CoV-2 infection, emerging data highlight the need for additional efforts to improve telehealth visits in future circumstances. Existing theoretical frameworks, such as interpersonal communication theory and symbol interaction theory, provided a basis for developing effective approaches to health care communication in telehealth applications [33-37]. It is important to contextualize these foundational theories with the perspectives of pregnant patients who can inform the best practices moving forward.

At the same time, it is also critical to educate obstetric providers about how to facilitate the visit in a way that is most supportive of patient-centered communication. It is important for health care providers to recognize that they may need to adjust how they conduct a telehealth visit compared to an in-person visit in response to patients' familiarity, receptivity, and resources in order to optimize this format. This includes an awareness of how both verbal and nonverbal cues and information may differ in a telehealth visit compared to an in-person visit [33,37,38] and, in turn, the effect of those different modes on patients' access to information and medical decision-making. In establishing these practices, it may be of benefit to utilize one or more of the developed approaches to improve communication in telehealth visits [38,39]. For example, the health literacy universal precautions approach "assumes that all patients are at risk for miscommunication and misunderstanding" [39-41]. Using this approach and reflecting, providers can prepare for telehealth visits with communication techniques that will support high-quality prenatal care and the skill set to rapidly transition between in-person and telehealth modalities quickly in a busy clinic. In developing the best practices moving forward, it is important to consider the diversity of pregnant patients' needs and preferences. Existing protocols focus on prioritizing telehealth visits for encounters that do not involve a procedure that requires an in-person visit or for average-risk patients [28]. Yet, patients in this study suggested that visits in which important, complex, and time-sensitive discussions must be made (eg, discussions about prenatal genetic screening and diagnostic testing) are also significant events, for which some of the dynamics of an in-person visit would be of benefit. These findings call for additional research to understand how best to individualize a plan of in-person and telehealth visits for patients based on their resources, needs, and preferences, independent of reproductive history.

Limitations

While our study provides insights into the clinical and ethical challenges with implementing telehealth, the findings should be contextualized with the limitations of this study. The study was based on patients from health care systems in Ohio that

adopted telehealth protocols in similar ways during the pandemic. Nonetheless, it is possible that there were subtle differences in the ways in which the practices associated with them were implemented. Although we sought a broad demographic representation in our recruitment efforts, most participants were <35 years of age (72%), self-described White (82%), and from the same geographic area. As a result, our results, by design, are not meant to be generalizable. We acknowledge that other health care systems and geographic areas of the United States may have had other experiences or practices with respect to telehealth delivery. In our population, more than half of the participants had a telehealth experience prior to pregnancy in addition to having at least one telehealth visit during the current pregnancy. While our sample represented patients with different reproductive histories, our sample was limited in racial and ethnic representation. Despite these limitations, the study brings to light important findings for which further research is needed to elucidate about larger and more diverse patient populations.

Conclusion

The variables that affect health care communication are complex factors that may differ based on in-person versus telehealth interactions. While telehealth was utilized as a mechanism to ensure timely access to prenatal care during the COVID-19 pandemic, it also comes with multiple challenges and opportunities to develop best practices around its continued integration into health care delivery. Our study speaks to the variability in patient perceptions of the utility and usability of telehealth for prenatal care delivery and the need to identify evidence-based approaches to individualize care. This includes education and strategies to support effective patient-centered communication so that patients can access the information and decision support needed to make the often complex, time-sensitive, and critical decisions that characterize prenatal health care. As health care communication is a key component of health care quality and patient safety, it is essential that we understand how to develop best practices around telehealth as its role in the delivery of prenatal care grows.

Acknowledgments

We would like to thank Bailey Diaz, Melissa Kinan, and Toni Tritto for their recruitment efforts at the MetroHealth health care system.

Conflicts of Interest

CGC, CC, RF, ADMH, BTE, MM, EC, MC, OG, ACR, and RMF do not have any relevant conflicts of interest to declare. SR received speaking honorariums and travel funding within the past 3 years from Siemens Healthineers, Panagora Pharma, Healthcare Information and Management Systems Society, Inc (HIMSS), Next Generation Patient Experience (NGPX), and health care systems in Sweden and Saudi Arabia for topics related to public health, bioethics, and health policy.

Multimedia Appendix 1

Overall and individual Coronavirus Perinatal Experience-Impact Survey data.
[DOCX File, 29 KB - [formative_v6i12e38821_app1.docx](#)]

References

1. Litchfield I, Shukla D, Greenfield S. Impact of COVID-19 on the digital divide: a rapid review. *BMJ Open* 2021 Oct 12;11(10):e053440 [FREE Full text] [doi: [10.1136/bmjopen-2021-053440](https://doi.org/10.1136/bmjopen-2021-053440)] [Medline: [34642200](https://pubmed.ncbi.nlm.nih.gov/34642200/)]
2. Eruchalu CN, Pichardo MS, Bharadwaj M, Rodriguez CB, Rodriguez JA, Bergmark RW, et al. The expanding digital divide: Digital health access inequities during the COVID-19 pandemic in New York City. *J Urban Health* 2021 Apr;98(2):183-186 [FREE Full text] [doi: [10.1007/s11524-020-00508-9](https://doi.org/10.1007/s11524-020-00508-9)] [Medline: [33471281](https://pubmed.ncbi.nlm.nih.gov/33471281/)]
3. DeNicola N, Grossman D, Marko K, Sonalkar S, Butler Tobah YS, Ganju N, et al. Telehealth interventions to improve obstetric and gynecologic health outcomes: A systematic review. *Obstet Gynecol* 2020 Feb;135(2):371-382 [FREE Full text] [doi: [10.1097/AOG.0000000000003646](https://doi.org/10.1097/AOG.0000000000003646)] [Medline: [31977782](https://pubmed.ncbi.nlm.nih.gov/31977782/)]
4. Lowery C. Telehealth: A new frontier in ob/gyn. *Contemporary OB/GYN*. 2018 Aug 15. URL: <https://www.contemporaryobgyn.net/view/telehealth-new-frontier-obgyn> [accessed 2022-11-19]
5. van den Heuvel JFM, Teunis CJ, Franx A, Crombag NMTH, Bekker MN. Home-based telemonitoring versus hospital admission in high risk pregnancies: a qualitative study on women's experiences. *BMC Pregnancy Childbirth* 2020 Feb 04;20(1):77 [FREE Full text] [doi: [10.1186/s12884-020-2779-4](https://doi.org/10.1186/s12884-020-2779-4)] [Medline: [32019499](https://pubmed.ncbi.nlm.nih.gov/32019499/)]
6. Spiegelman J, Krenitsky N, Syeda S, Sutton D, Moroz L. Rapid development and implementation of a Covid-19 telehealth clinic for obstetric patients. *NEJM Catalyst* 2020:1-13 [FREE Full text]
7. No authors listed. Implementing telehealth in practice. *Obstet Gynecol* 2020 Feb;135(2):e73-e79. [doi: [10.1097/AOG.0000000000003671](https://doi.org/10.1097/AOG.0000000000003671)] [Medline: [31977796](https://pubmed.ncbi.nlm.nih.gov/31977796/)]
8. Zork NM, Aubey J, Yates H. Conversion and optimization of telehealth in obstetric care during the COVID-19 pandemic. *Semin Perinatol* 2020 Oct;44(6):151300 [FREE Full text] [doi: [10.1016/j.semperi.2020.151300](https://doi.org/10.1016/j.semperi.2020.151300)] [Medline: [32928561](https://pubmed.ncbi.nlm.nih.gov/32928561/)]
9. Danylchuk N, Cook L, Shane-Carson K, Cacioppo C, Hardy M, Nusbaum R, et al. Telehealth for genetic counseling: A systematic evidence review. *J Genet Couns* 2021 Oct;30(5):1361-1378. [doi: [10.1002/jgc4.1481](https://doi.org/10.1002/jgc4.1481)] [Medline: [34355839](https://pubmed.ncbi.nlm.nih.gov/34355839/)]
10. Pregnant and Recently Pregnant People: At Increased Risk for Severe Illness from COVID-19. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/pregnant-people.html> [accessed 2022-11-19]
11. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Obstetrics, Committee on Genetics, Society for Maternal-Fetal Medicine. Screening for fetal chromosomal abnormalities: ACOG Practice Bulletin, Number 226. *Obstet Gynecol* 2020 Oct;136(4):e48-e69. [doi: [10.1097/AOG.0000000000004084](https://doi.org/10.1097/AOG.0000000000004084)] [Medline: [32804883](https://pubmed.ncbi.nlm.nih.gov/32804883/)]
12. Farrell RM, Pierce M, Collart C, Tucker Edmonds B, Chien E, Coleridge M, et al. Making the most of the first prenatal visit: The challenge of expanding prenatal genetic testing options and limited clinical encounter time. *Prenat Diagn* 2020 Sep 19;40(10):1265-1271. [doi: [10.1002/pd.5752](https://doi.org/10.1002/pd.5752)] [Medline: [32441820](https://pubmed.ncbi.nlm.nih.gov/32441820/)]
13. Aziz A, Zork N, Aubey JJ, Baptiste CD, D'Alton ME, Emeruwa UN, et al. Telehealth for high-risk pregnancies in the setting of the COVID-19 pandemic. *Am J Perinatol* 2020 Jun;37(8):800-808 [FREE Full text] [doi: [10.1055/s-0040-1712121](https://doi.org/10.1055/s-0040-1712121)] [Medline: [32396948](https://pubmed.ncbi.nlm.nih.gov/32396948/)]
14. Galle A, Semaan A, Huysmans E, Audet C, Asefa A, Delvaux T, et al. A double-edged sword-telemedicine for maternal care during COVID-19: findings from a global mixed-methods study of healthcare providers. *BMJ Glob Health* 2021 Feb 25;6(2):e004575 [FREE Full text] [doi: [10.1136/bmjgh-2020-004575](https://doi.org/10.1136/bmjgh-2020-004575)] [Medline: [33632772](https://pubmed.ncbi.nlm.nih.gov/33632772/)]
15. Peahl AF, Powell A, Berlin H, Smith RD, Krans E, Waljee J, et al. Patient and provider perspectives of a new prenatal care model introduced in response to the coronavirus disease 2019 pandemic. *Am J Obstet Gynecol* 2021 Apr;224(4):384.e1-384.e11 [FREE Full text] [doi: [10.1016/j.ajog.2020.10.008](https://doi.org/10.1016/j.ajog.2020.10.008)] [Medline: [33039393](https://pubmed.ncbi.nlm.nih.gov/33039393/)]
16. Thomason M, Graham A, VanTieghem M. COPE: Coronavirus Perinatal Experiences - Impact Survey (COPE-IS). PhenX Toolkit. 2020. URL: https://www.phenxtoolkit.org/toolkit_content/PDF/COS_COPE_IS_Perinatal.pdf [accessed 2022-11-19]
17. Kinser PA, Jallo N, Amstadter AB, Thacker LR, Jones E, Moyer S, et al. Depression, anxiety, resilience, and coping: The experience of pregnant and new mothers during the first few months of the COVID-19 pandemic. *J Womens Health (Larchmt)* 2021 May;30(5):654-664 [FREE Full text] [doi: [10.1089/jwh.2020.8866](https://doi.org/10.1089/jwh.2020.8866)] [Medline: [33844945](https://pubmed.ncbi.nlm.nih.gov/33844945/)]
18. Werchan DM, Hendrix CL, Ablow JC, Amstadter AB, Austin AC, Babineau V, et al. Behavioral coping phenotypes and associated psychosocial outcomes of pregnant and postpartum women during the COVID-19 pandemic. *Sci Rep* 2022 Jan 24;12(1):1209 [FREE Full text] [doi: [10.1038/s41598-022-05299-4](https://doi.org/10.1038/s41598-022-05299-4)] [Medline: [35075202](https://pubmed.ncbi.nlm.nih.gov/35075202/)]
19. Motrico E, Mateus V, Bina R, Felice E, Bramante A, Kalcev G, et al. Good practices in perinatal mental health during the COVID-19 pandemic: a report from task-force RISEUP-PPD COVID-19. *Clínica y Salud* 2020;31(3):155-160 [FREE Full text] [doi: [10.5093/clysa2020a26](https://doi.org/10.5093/clysa2020a26)]
20. Strauss A, Corbin J. Grounded theory methodology: An overview. In: Denzin NK, Lincoln YS, editors. *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, Inc; 1994:273-285.
21. Ryan GW, Bernard H. Techniques to identify themes. *Field Methods* 2003 Feb;15(1):85-109 [FREE Full text] [doi: [10.1177/1525822x02239569](https://doi.org/10.1177/1525822x02239569)]
22. Madden N, Emeruwa UN, Friedman AM, Aubey JJ, Aziz A, Baptiste CD, et al. Telehealth uptake into prenatal care and provider attitudes during the COVID-19 pandemic in New York City: A quantitative and qualitative analysis. *Am J Perinatol* 2020 Aug;37(10):1005-1014 [FREE Full text] [doi: [10.1055/s-0040-1712939](https://doi.org/10.1055/s-0040-1712939)] [Medline: [32516816](https://pubmed.ncbi.nlm.nih.gov/32516816/)]

23. Holcomb D, Faucher MA, Bouzid J, Quint-Bouzid M, Nelson DB, Duryea E. Patient perspectives on audio-only virtual prenatal visits amidst the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic. *Obstet Gynecol* 2020 Aug;136(2):317-322. [doi: [10.1097/AOG.0000000000004026](https://doi.org/10.1097/AOG.0000000000004026)] [Medline: [32544144](https://pubmed.ncbi.nlm.nih.gov/32544144/)]
24. Almuslim H, AlDossary S. Models of incorporating telehealth into obstetric care during the COVID-19 pandemic, its benefits and barriers: A scoping review. *Telemed J E Health* 2022 Jan 01;28(1):24-38. [doi: [10.1089/tmj.2020.0553](https://doi.org/10.1089/tmj.2020.0553)] [Medline: [33819434](https://pubmed.ncbi.nlm.nih.gov/33819434/)]
25. Wu H, Sun W, Huang X, Yu S, Wang H, Bi X, et al. Online antenatal care during the COVID-19 pandemic: Opportunities and challenges. *J Med Internet Res* 2020 Jul 22;22(7):e19916 [FREE Full text] [doi: [10.2196/19916](https://doi.org/10.2196/19916)] [Medline: [32658860](https://pubmed.ncbi.nlm.nih.gov/32658860/)]
26. Peahl AF, Smith RD, Moniz MH. Prenatal care redesign: creating flexible maternity care models through virtual care. *Am J Obstet Gynecol* 2020 Sep;223(3):389.e1-389.e10 [FREE Full text] [doi: [10.1016/j.ajog.2020.05.029](https://doi.org/10.1016/j.ajog.2020.05.029)] [Medline: [32425200](https://pubmed.ncbi.nlm.nih.gov/32425200/)]
27. Peahl AF, Zahn CM, Turrentine M, Barfield W, Blackwell SD, Roberts SJ, et al. The Michigan plan for appropriate tailored healthcare in pregnancy prenatal care recommendations. *Obstet Gynecol* 2021 Oct 01;138(4):593-602. [doi: [10.1097/AOG.0000000000004531](https://doi.org/10.1097/AOG.0000000000004531)] [Medline: [34352810](https://pubmed.ncbi.nlm.nih.gov/34352810/)]
28. Kern-Goldberger AR, Srinivas SK. Obstetrical telehealth and virtual care practices during the COVID-19 pandemic. *Clin Obstet Gynecol* 2022 Mar 01;65(1):148-160 [FREE Full text] [doi: [10.1097/GRF.0000000000000671](https://doi.org/10.1097/GRF.0000000000000671)] [Medline: [35045037](https://pubmed.ncbi.nlm.nih.gov/35045037/)]
29. Agha Z, Roter DL, Schapira RM. An evaluation of patient-physician communication style during telemedicine consultations. *J Med Internet Res* 2009 Sep 30;11(3):e36 [FREE Full text] [doi: [10.2196/jmir.1193](https://doi.org/10.2196/jmir.1193)] [Medline: [19793720](https://pubmed.ncbi.nlm.nih.gov/19793720/)]
30. Alkureishi MA, Choo Z, Rahman A, Ho K, Benning-Shorb J, Lenti G, et al. Digitally disconnected: Qualitative study of patient perspectives on the digital divide and potential solutions. *JMIR Hum Factors* 2021 Dec 15;8(4):e33364 [FREE Full text] [doi: [10.2196/33364](https://doi.org/10.2196/33364)] [Medline: [34705664](https://pubmed.ncbi.nlm.nih.gov/34705664/)]
31. Paige SR, Bunnell BE, Bylund CL. Disparities in patient-centered communication via telemedicine. *Telemed J E Health* 2022 Feb 01;28(2):212-218. [doi: [10.1089/tmj.2021.0001](https://doi.org/10.1089/tmj.2021.0001)] [Medline: [33913764](https://pubmed.ncbi.nlm.nih.gov/33913764/)]
32. Gordon HS, Solanki P, Bokhour BG, Gopal RK. "I'm not feeling like I'm part of the conversation" patients' perspectives on communicating in clinical video telehealth visits. *J Gen Intern Med* 2020 Jun;35(6):1751-1758 [FREE Full text] [doi: [10.1007/s11606-020-05673-w](https://doi.org/10.1007/s11606-020-05673-w)] [Medline: [32016705](https://pubmed.ncbi.nlm.nih.gov/32016705/)]
33. Bylund CL, Peterson EB, Cameron KA. A practitioner's guide to interpersonal communication theory: an overview and exploration of selected theories. *Patient Educ Couns* 2012 Jun;87(3):261-267 [FREE Full text] [doi: [10.1016/j.pec.2011.10.006](https://doi.org/10.1016/j.pec.2011.10.006)] [Medline: [22112396](https://pubmed.ncbi.nlm.nih.gov/22112396/)]
34. Elias N. *The Symbol Theory*. Thousand Oaks, CA: SAGE Publications Ltd; 1991.
35. Carter MJ, Fuller C. Symbols, meaning, and action: The past, present, and future of symbolic interactionism. *Current Sociology* 2016 Jul 09;64(6):931-961 [FREE Full text] [doi: [10.1177/0011392116638396](https://doi.org/10.1177/0011392116638396)]
36. Berger CR, Roloff ME. *Interpersonal Communication*. In: Stacks DW, Salwen MB, Eichhorn KC, editors. *An Integrated Approach to Communication Theory and Research*. New York, NY: Routledge; 2019.
37. Charmaz K, Belgrave LL. *Modern Symbolic Interaction Theory and Health*. In: Cockerham W, editor. *Medical Sociology on the Move*. Dordrecht, Netherlands: Springer; 2013:11-39.
38. Miller EA. Telemedicine and doctor-patient communication: a theoretical framework for evaluation. *J Telemed Telecare* 2002;8(6):311-318. [doi: [10.1258/135763302320939185](https://doi.org/10.1258/135763302320939185)] [Medline: [12537917](https://pubmed.ncbi.nlm.nih.gov/12537917/)]
39. Coleman C. Health literacy and clear communication best practices for telemedicine. *Health Lit Res Pract* 2020 Nov 06;4(4):e224-e229 [FREE Full text] [doi: [10.3928/24748307-20200924-01](https://doi.org/10.3928/24748307-20200924-01)] [Medline: [33170288](https://pubmed.ncbi.nlm.nih.gov/33170288/)]
40. Health Literacy Universal Precautions Toolkit. AHRQ. URL: <https://www.ahrq.gov/health-literacy/improve/precautions/index.html> [accessed 2022-11-19]
41. Coleman C, Hudson S, Pederson B. Prioritized health literacy and clear communication practices for health care professionals. *Health Lit Res Pract* 2017 Jul;1(3):e91-e99 [FREE Full text] [doi: [10.3928/24748307-20170503-01](https://doi.org/10.3928/24748307-20170503-01)] [Medline: [31294254](https://pubmed.ncbi.nlm.nih.gov/31294254/)]

Abbreviations

COPE-IS: Coronavirus Perinatal Experience-Impact Survey

Edited by A Mavragani; submitted 03.05.22; peer-reviewed by N Hu, F Petrazzuoli; comments to author 23.09.22; revised version received 14.10.22; accepted 19.10.22; published 05.12.22.

Please cite as:

Craighead CG, Collart C, Frankel R, Rose S, Misra-Hebert AD, Tucker Edmonds B, Michie M, Chien E, Coleridge M, Goje O, Ranzini AC, Farrell RM

Impact of Telehealth on the Delivery of Prenatal Care During the COVID-19 Pandemic: Mixed Methods Study of the Barriers and Opportunities to Improve Health Care Communication in Discussions About Pregnancy and Prenatal Genetic Testing

JMIR Form Res 2022;6(12):e38821

URL: <https://formative.jmir.org/2022/12/e38821>

doi: [10.2196/38821](https://doi.org/10.2196/38821)

PMID: [36383634](https://pubmed.ncbi.nlm.nih.gov/36383634/)

©Caitlin G Craighead, Christina Collart, Richard Frankel, Susannah Rose, Anita D Misra-Hebert, Brownsyne Tucker Edmonds, Marsha Michie, Edward Chien, Marissa Coleridge, Oluwatosin Goje, Angela C Ranzini, Ruth M Farrell. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Assessing Associations Between COVID-19 Symptomology and Adverse Outcomes After Piloting Crowdsourced Data Collection: Cross-sectional Survey Study

Natalie Flaks-Manov¹, PhD; Jiawei Bai², PhD; Cindy Zhang³; Anand Malpani³, PhD; Stuart C Ray¹, MD; Casey Overby Taylor¹, PhD

¹Johns Hopkins University School of Medicine, Baltimore, MD, United States

²Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

³Johns Hopkins Whiting School of Engineering, Baltimore, MD, United States

Corresponding Author:

Casey Overby Taylor, PhD

Johns Hopkins University School of Medicine

3101 Wyman Park Dr.

Baltimore, MD, 21218

United States

Phone: 1 443 287 6657

Email: cot@jhu.edu

Abstract

Background: Crowdsourcing is a useful way to rapidly collect information on COVID-19 symptoms. However, there are potential biases and data quality issues given the population that chooses to participate in crowdsourcing activities and the common strategies used to screen participants based on their previous experience.

Objective: The study aimed to (1) build a pipeline to enable data quality and population representation checks in a pilot setting prior to deploying a final survey to a crowdsourcing platform, (2) assess COVID-19 symptomology among survey respondents who report a previous positive COVID-19 result, and (3) assess associations of symptomology groups and underlying chronic conditions with adverse outcomes due to COVID-19.

Methods: We developed a web-based survey and hosted it on the Amazon Mechanical Turk (MTurk) crowdsourcing platform. We conducted a pilot study from August 5, 2020, to August 14, 2020, to refine the filtering criteria according to our needs before finalizing the pipeline. The final survey was posted from late August to December 31, 2020. Hierarchical cluster analyses were performed to identify COVID-19 symptomology groups, and logistic regression analyses were performed for hospitalization and mechanical ventilation outcomes. Finally, we performed a validation of study outcomes by comparing our findings to those reported in previous systematic reviews.

Results: The crowdsourcing pipeline facilitated piloting our survey study and revising the filtering criteria to target specific MTurk experience levels and to include a second attention check. We collected data from 1254 COVID-19-positive survey participants and identified the following 6 symptomology groups: abdominal and bladder pain (Group 1); flu-like symptoms (loss of smell/taste/appetite; Group 2); hoarseness and sputum production (Group 3); joint aches and stomach cramps (Group 4); eye or skin dryness and vomiting (Group 5); and no symptoms (Group 6). The risk factors for adverse COVID-19 outcomes differed for different symptomology groups. The only risk factor that remained significant across 4 symptomology groups was influenza vaccine in the previous year (Group 1: odds ratio [OR] 6.22, 95% CI 2.32-17.92; Group 2: OR 2.35, 95% CI 1.74-3.18; Group 3: OR 3.7, 95% CI 1.32-10.98; Group 4: OR 4.44, 95% CI 1.53-14.49). Our findings regarding the symptoms of abdominal pain, cough, fever, fatigue, shortness of breath, and vomiting as risk factors for COVID-19 adverse outcomes were concordant with the findings of other researchers. Some high-risk symptoms found in our study, including bladder pain, dry eyes or skin, and loss of appetite, were reported less frequently by other researchers and were not considered previously in relation to COVID-19 adverse outcomes.

Conclusions: We demonstrated that a crowdsourced approach was effective for collecting data to assess symptomology associated with COVID-19. Such a strategy may facilitate efficient assessments in a dynamic intersection between emerging infectious diseases, and societal and environmental changes.

KEYWORDS

COVID-19; coronavirus; symptoms; symptomology; crowdsourcing; adverse outcomes; data quality

Introduction

COVID-19 represents a global public health concern [1-3]. While extensive measures are being implemented to control the outbreak, the high speed of transmission makes collecting data needed to inform clinical management and public health planning a challenge. Efficiently collecting high-quality data to characterize disease severity enables accurate information to be disseminated in a timely manner for such planning.

To understand and predict the adverse health outcomes in patients affected by COVID-19, many scientific efforts studying sociodemographic, clinical, and symptomatic risk factors are underway. Findings from those efforts, however, are not all consistent, with conflicting evidence on the risk factors associated with adverse COVID-19 outcomes [4-6]. Furthermore, infected people have reported a wide range of symptoms, from asymptomatic to severe illness [2-12]. Common symptoms include fever, cough, fatigue, shortness of breath, and loss of the sense of smell or taste, and less frequent symptoms are gastrointestinal and neurological symptoms [4-6,13-16]. Although there has been a concerted effort to describe patients' symptoms [7,17,18], there is no evidence yet as to whether symptoms differ between people with different characteristics, such as chronic diseases and demographic backgrounds [19,20]. As individual symptoms cannot predict COVID-19 adverse outcomes [21], knowledge of a patient's profile of symptoms (ie, symptomology) holds promise to improve estimations of the risk of adverse outcomes [22].

A crowdsourcing model is a useful way to rapidly collect information in the context of the COVID-19 pandemic [23,24]. Recent work to classify different types of crowdsourcing used to tackle the COVID-19 crisis [23] found that the most common configuration to deal with information and knowledge management problems was open crowdsourcing (described as a one-to-many configuration with potentially unlimited contributors, and without any form of preselection). Most initiatives falling under this category, however, demonstrated a desire to locate and assemble information. The *COVID Near You* website [25], for example, uses crowdsourced data to visualize maps to identify current and potential pandemic hotspots. An important emphasis for crowdsourced data, however, is to collect high-quality data. Indeed, the risk of bias can be great when building COVID-19 diagnosis and prognosis prediction models trained on small or low-quality data sets. The majority of COVID-19 prediction models to date, for example, show a high risk of bias (n=226, 97%) [26].

To eliminate substandard crowd data submissions, we used a "crowdsourcing via a broker" strategy with broker services that allowed for filtering participants and their responses, and testing data quality before finalizing the crowdsourcing data collection strategy. We chose to use the Amazon Mechanical Turk (MTurk) crowdsourcing platform that provides filtering

mechanisms via setting qualifications [27-30]. Through the MTurk platform, entities known as "requesters" can hire independent contractors, known as "workers," to perform a wide variety of remote jobs, known as "human intelligence tasks" (HITs). A worker's reputation is indicated by their HIT acceptance rate [30]. The emphasis on obtaining good-quality data through setting qualifications, however, has created some concern about "superworkers." These are experienced and very active MTurk workers to whom researchers often target survey distribution. This oversampling from experienced workers can lead to an issue of worker nonnaivete as workers are frequently exposed to common methods in research studies. Recent research shows that nonsuperworkers can also produce high-quality data [31], and our strategy thus incorporated a pilot phase with broad inclusion criteria according to experience qualifications. Rather than defaulting to experienced workers, the pilot data collection allowed us to determine what filtering criteria were best suited to our needs.

In this paper, we describe (1) a pipeline to enable data quality and population representation checks in a pilot setting prior to deploying the final survey to MTurk workers, (2) an assessment of COVID-19 symptomology among MTurk worker survey respondents who reported a previous positive COVID-19 result, and (3) an assessment of the associations of symptomology groups and underlying chronic conditions with adverse outcomes due to COVID-19.

Methods

Study Design and Instrument

This was a cross-sectional study. We developed 2 web-based surveys using Qualtrics. One survey was for individuals (ie, individual survey) who indicated a self-reported positive test for COVID-19, and another survey was for individuals whose relatives (ie, family survey), living in the same house, tested positive for COVID-19. We hosted both surveys on MTurk between August and December 2020. To improve the quality of data collection through MTurk and to make our study sample more representative of the target population, we followed the best practices suggested by Young et al [30].

A few restrictions were implemented to exclude certain survey responses from the final data analysis. First, only those participants who provided an existing COVID-19 test type (nasal/throat/blood/sputum) answer in response to our screening question could continue with the survey. Second, participants could fill the survey only once for themselves (individual survey) and for 1 family member (family survey). Third, a quality control question was included during the questionnaire, which stated, "Do not answer this question (Please click NEXT to go to the next question)." If the question was answered, the survey responses were excluded from the data analyses. Fourth, in the family survey, we asked the participants about their confidence level in their responses regarding their family

member. Responses with low reported confidence were excluded from the final analyses.

Ethical Considerations

This study was judged as imposing only minimal risks on participants and was determined to be exempt research by the Johns Hopkins University (IRB00248053) Institutional Review Board.

Variables and Definitions

The web-based surveys had 5 blocks as follows ([Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)):

1. *Introduction and screening*: Introduction to the aim of the study, including the estimated amount of time needed to complete the survey, the compensation amount, information about the voluntary nature of the survey, and instructions on not filling out the questionnaire more than once.
2. *Symptoms of COVID-19*: We asked participants to select all the symptoms that they experienced following a COVID-19 infection.
3. *Adverse outcomes of COVID-19*: We asked questions about hospitalizations related to COVID-19 and connection to a mechanical ventilator.
4. *Medical history*: We asked questions about background medical conditions, smoking status, and influenza vaccine status in a previous season.
5. *Demographic characteristics*: Participants reported on their age, sex they were assigned at birth, race, ethnicity, last year's income (monthly and yearly), and the highest level of education.

Survey measures were from the Johns Hopkins University COVID-19 community response survey guidance toolkit that draws from multiple sources [32]. An additional data source we used beyond the toolkit to compile COVID-19 symptoms was Twitter [18,32,33].

Recruitment

Population

The inclusion criteria for this study were individuals living in the United States, adults (aged 18 years or older), and MTurk workers with a self-reported positive COVID-19 result. For the family survey, the participants could complete the survey for 1 family member, even if the family member, who lived in the same household, was below 18 years old. Thus, the target population of this study was COVID-19 patients living in the United States and having sufficient skills to use the MTurk platform. The participants were compensated according to a standard minimum wage and our estimate of completion time (about 5-10 min).

Crowdsourcing Pipeline

Before posting the final survey to MTurk, we conducted a pilot from August 5, 2020, to August 14, 2020, to assess the quality of responses among workers with different levels of experience. The pilot analysis stratified the worker sample into the following 3 experience groups: those who previously completed 100-499 HITs, 500-999 HITs, and 1000+ HITs. First, a worker would complete a qualification test asking them to verify that they or

a family member tested positive for COVID-19. If qualified, the worker could then start the MTurk HIT that included a link to the 26-question web-based Qualtrics survey. The first part of the survey was a screening question (age ≥ 18 years) and comprehension check. In response to the comprehension check, if a worker selected an invalid COVID-19 test type (eg, urine test), they could not continue the survey.

For those passing the screening test, responses were labeled as "high quality" according to the following criteria: sufficient time taken (threshold of more than 60 s); matching codes and IDs between Qualtrics and MTurk; each code being associated with only 1 worker; and worker had not taken the survey previously (ie, nonduplicate response). A worker's response was included in the "high-quality" group if they passed all of these criteria.

Separately, we assessed "nonduplicate responses." A nonduplicate response indicates that the respondent completed the survey only once. This criterion was considered under the assumption that workers who attempted to complete the survey multiple times to receive more compensation did not read through survey instructions carefully, and thus, they may provide lower quality responses than those who attempted to complete the survey once.

The general characteristics of age, sex, race, education, and income were extracted and compared among experience groups. Chi-square analysis was conducted to evaluate if there was a significant difference between experience groups in the number of high-quality and nonduplicate responses. Findings from this analysis were used to refine our filtering criteria in the final crowdsourcing pipeline.

Statistical Analysis

Outcomes

We assessed the following 2 primary adverse outcomes related to COVID-19: hospital admission due to COVID-19 and use of mechanical ventilation during admission.

Statistical Analyses

We used descriptive statistics to characterize the total cohort of participants. Bivariate analyses, using Pearson χ^2 tests, were performed to assess differences in participant characteristics between those hospitalized and those not hospitalized, and between those who needed mechanical ventilation during admission and those who did not need mechanical ventilation. We then fitted multivariate logistic regression models to identify the association of COVID-19 symptoms with hospitalization and mechanical ventilation due to COVID-19, adjusted for sociodemographic characteristics and comorbid conditions. Thereafter, hierarchical cluster analysis was conducted to search for patterns based on COVID-19 symptoms. The similarity measure was cosine similarity, and the linkage method was Ward minimum variance. To describe clusters, we calculated frequencies of the risk factors for each cluster of symptoms. We then developed logistic regression models for hospitalization and mechanical ventilation as outcomes, using symptomology groups as risk factors. Finally, we developed a logistic regression model for each symptomology group to identify the significant

risk factors for hospitalization among individuals with different symptomatology. All analyses were performed using R version 3.6.2 (R Foundation for Statistical Computing).

Validation Assessment

To validate our findings, we performed a comparison with existing systematic review or meta-analysis papers that assessed symptoms as risk factors for COVID-19 adverse outcomes. Articles for which the analyses occurred prior to our data collection were selected for comparison.

For each article and this study, individual symptoms were checked for being reported as (1) a significant risk factor for an adverse outcome (“yes”) and (2) a nonsignificant risk factor for an adverse outcome (“no”). We also noted if a symptom was not assessed (“NA”). When synthesizing findings across studies, if we found a statistically significant association between an adverse outcome and a symptom that was not studied by others, we labeled it “New.” If there was agreement between this study and at least one other study in identifying a symptom as a risk factor (significant or nonsignificant), we labeled it “1.” Symptoms we did not assess were labeled “NA.”

Results

Pilot Findings

Pilot survey data were collected from 259 respondents who passed both the qualification test and the screening questions, and of these, 147 (56.8%) were considered to have “high quality” responses. For the experience groups 100-499, 500-999, and 1000+ HITs, the proportions of high-quality responses were 58% (48/83), 43% (41/95), and 72% (58/81), respectively (Table 1). There was no significant difference between the experience groups for obtaining high-quality responses ($P=.14$). There was, however, a significant difference between the groups for nonduplicate responses ($P<.001$). Comparisons of demographic characteristics across all experience groups among MTurk workers are shown in Multimedia Appendix 3.

Two modifications were made to our crowdsourcing pipeline following the pilot. First, we included only workers with 500+ prior HITs in our final filtering criteria. Given the differences in nonduplicate responses between groups, we reasoned that for tasks requiring a higher cognitive ability, workers with 500+ HITs may provide more high-quality responses than those with 100-499 HITs. Second, we added an attention check question to the Qualtrics survey (ie, “don’t answer this question”).

Table 1. Comparison of the approval rates and percentage of quality responses out of approved responses between the different experience groups.

Variable	Experience group			P value
	100-499 HITs ^a , n/N (%)	500-999 HITs, n/N (%)	1000+ HITs, n/N (%)	
High-quality responses	48/83 (58)	41/95 (43)	58/81 (72)	0.135
High-quality and nonduplicate responses	10/48 (21)	41/41 (100)	49/58 (85)	<0.001

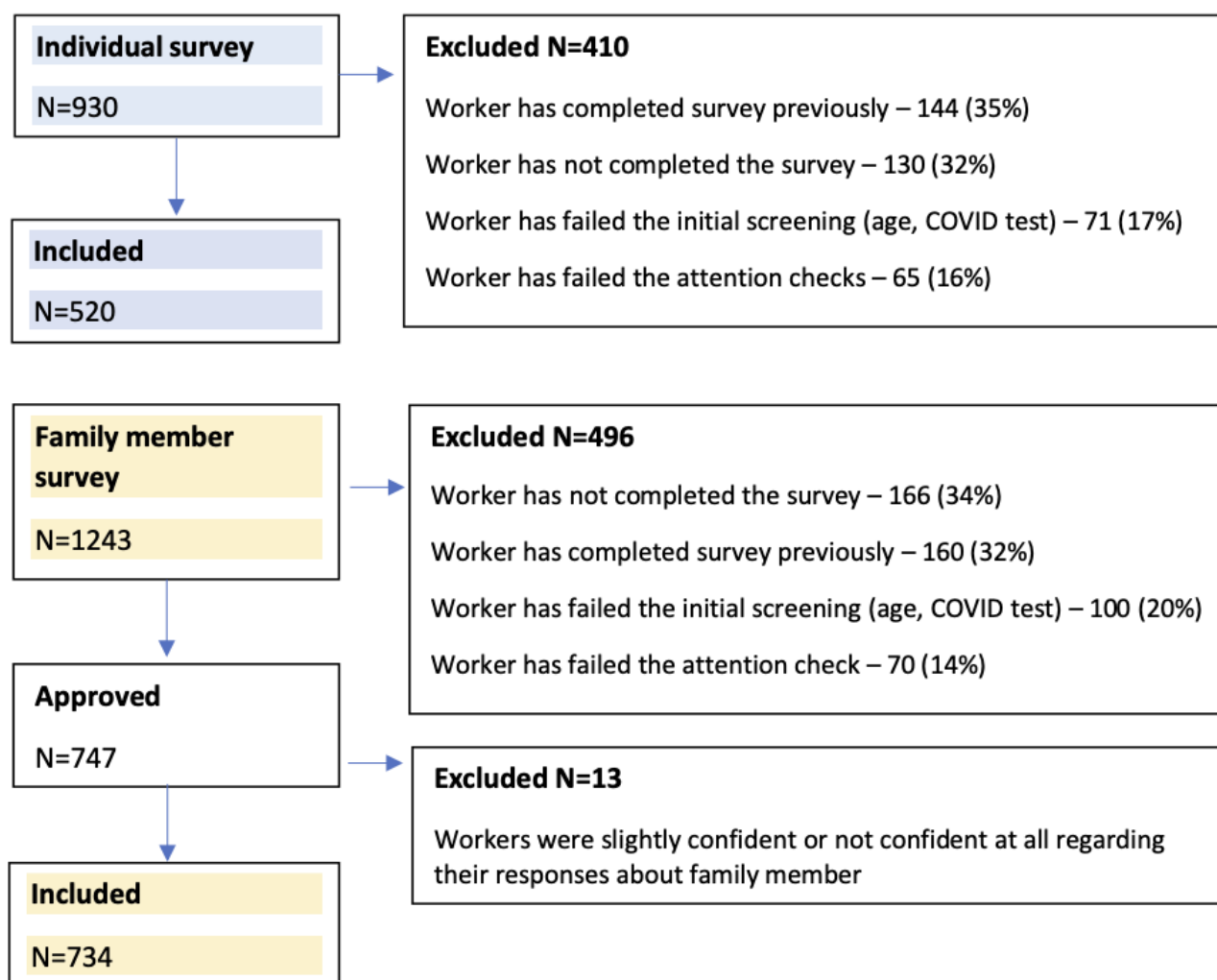
^aHIT: human intelligence task.

Survey Responses

After implementing our final crowdsourcing pipeline, we collected data from 930 individual surveys and 1243 family surveys; however, data from 410 individual surveys and 496 family surveys were excluded (late August to December 31, 2020). The reasons for exclusion were completion of the survey previously, noncompletion of the survey, initial screening failure for age or comprehension check, and attention check failure (Figure 1). Thus, we finally collected data from 1267 eligible COVID-19-positive participants, and of these, 520 were from individual surveys and 747 were from family surveys. Thirteen participants were further excluded as they were either only

slightly confident (n=12) or not confident at all (n=1) regarding their responses in the family survey. Thus, data from 1254 surveys were analyzed. The average time required to complete the general survey was 5.5 minutes.

Regarding family survey respondents, 68.3% (501/734) provided answers about a first-degree family member, 25.7% (189/734) provided answers about a second-degree family member, and only 6.0% (44/734) provided answers about a third-degree relative. There were no statistically significant differences in characteristics or outcomes between the individual respondents and the persons the respondents completed the family survey for, except for age (Multimedia Appendix 4). Therefore, the analysis presented here combined data from both surveys.

Figure 1. Study Inclusion and Exclusion of Amazon Mechanical Turk Worker Responses.

Demographic Characteristics

Over 90% (1159/1254, 92.4%) of the participants were up to 65 years old, and only 1.2% (15/1254) were less than 18 years old. Moreover, 52.0% (652/1254) were male, 81.2% (1018/1254) were white, 79.5% (997/1254) were not Hispanic or Latino, 68.4% (858/1254) had a bachelor's degree or any postgraduate degree, 14.4% (180/1254) had yearly income of US \$75,000 or more, 39.6% (496/1254) were smokers, and 46.8% (587/1254) had an influenza vaccine in the last season (Multimedia Appendix 5). Eight responders mentioned "passed away" in the family survey. As reflected in our sample, the MTurk worker population tended to be younger than the overall US population, with household incomes below that of the average US population [27].

Findings From Assessing Individual Symptoms Associated With Adverse COVID-19 Outcomes

Hospitalization

Overall, 47.6% (597/1254) of participants were hospitalized due to COVID-19. Bivariate analysis showed statistically significant differences between hospitalized and nonhospitalized COVID-19 participants for most demographic factors, except gender (Multimedia Appendix 5). Chronic conditions, including depression, hypertension, asthma, alcohol disorder, anemia,

weight loss, ulcer, lung/respiratory disease, bladder problems, bowel disease, and angina, were associated with more COVID-19 hospitalizations (Multimedia Appendix 6). COVID-19 symptoms associated with higher risk for hospitalization were cough with sputum, sneezing, abdominal pain, vomiting, confusion, bladder pain, dry eyes, dry skin, skin rash, and seizure (Multimedia Appendix 7).

From the logistic regression analysis of the total study population (Multimedia Appendix 8), we found statistically significant associations between the following participant characteristics and COVID-19 hospitalization compared with baseline: being in any age group over 24 years; having a bachelor's degree (odds ratio [OR] 2.57, 95% CI 1.43-4.66); smoking every day, smoking some days, or past smoking with quitting less than a year ago (OR 2.06, 95% CI 1.18-3.61; OR 3.41, 95% CI 2.2-5.33; and OR 3.39, 95% CI 1.88-6.24, respectively); and influenza vaccine in the last season (OR 3.09, 95% CI 2.18-4.41). Chronic conditions associated with higher risk for hospitalization were depression (OR 1.77, 95% CI 1.18-2.67), asthma (OR 3.83, 95% CI 2.22-6.78), diabetes (OR 2.66, 95% CI 1.46-4.96), and bladder problems (OR 5.51, 95% CI 1.28-27.13). COVID-19 symptoms associated with higher risk for hospitalization were abdominal pain (OR 2.02, 95% CI 1.15-3.59), bladder pain (OR 3.20, 95% CI 1.25-9.3), cough

with sputum (OR 2.60, 95% CI 1.77-3.86), fever with a temperature over 100.4°F (OR 1.50, 95% CI 1.04-2.17), and shortness of breath (OR 2.75, 95% CI 1.8-4.23).

Mechanical Ventilation

Overall, 66.8% (399/597) of hospitalized participants were connected to a mechanical ventilator (31.8% of all participants). There were 11 hospitalized participants from the family survey whose mechanical ventilation use was unknown to the survey respondents, and these participants were not included in the subsequent mechanical ventilation analysis. Smoking every day (OR 3.51, 95% CI 1.45-9.1), influenza vaccine in the last season (OR 3.65, 95% CI 2.29-5.89), loss of appetite (OR 2.07, 95% CI 1.09-4.02), tiredness and fatigue (OR 2.36, 95% CI 1.04-5.44), and vomiting (OR 2.68, 95% CI 1.3-5.71) were significantly associated with higher risk for mechanical ventilation (Multimedia Appendix 8).

Findings From Assessing COVID-19 Symptomology

We identified the following 6 symptomology groups using hierarchical cluster analysis (Figure 2): Group 1, abdominal and bladder pain; Group 2, flu-like symptoms (loss of smell/taste/appetite); Group 3, hoarseness and sputum production; Group 4, joint aches and stomach cramps; Group 5, skin or eye dryness and vomiting; and Group 6, no symptoms. We found sociodemographic and clinical differences between the symptomology groups (Table 2).

The flu-like symptoms group (Group 2) mostly represented the general study population. The abdominal and bladder pain group (Group 1) and the skin or eye dryness group (Group 5) had the highest hospitalization frequencies (153/227, 67.4% and 134/196, 68.4%, respectively). Both groups were characterized by a lower chance of high income (19/227, 8.4% and 21/196, 10.7%, respectively), more smoking (121/227, 53.3% and 102/196, 52.0%, respectively), and more influenza vaccinations (144/227, 63.4% and 102/196, 52.0%, respectively). The group with abdominal and bladder pain symptoms (Group 1) had higher proportions of Hispanic participants (82/227, 36.1%), asthma patients (64/227, 28.2%), alcohol disorder patients (64/227, 28.2%), and anemia patients (54/227, 23.8%). The group with skin or eye dryness (Group 5) had higher proportions of patients with depression (64/196, 32.7%), diabetes (28/196, 14.3%), weight loss (27/196, 13.8%), and ulcers (25/196, 12.8%). The group with joint aches and stomach cramps (Group 4) had lower proportions of hospitalization (65/158, 41.1%) and mechanical ventilation (31/158, 47.7%). Compared with the general study population, the asymptomatic group (Group 6) was younger (age 18-44 years; 70/85, 82.4%), had more males (54/85, 63.5%), had less white participants (60/85, 70.6%), had less Hispanic participants (7/85, 8.2%), had more participants with a high income (18/85, 21.2%), had less smokers (26/85, 30.6%), had less influenza vaccinations reported (30/85, 35.3%), had a higher proportion of participants with no chronic conditions (45/85, 52.9%), and had a very low risk for hospitalization (12/85, 14.1%).

Figure 2. Dendrogram for COVID-19 symptom clusters.

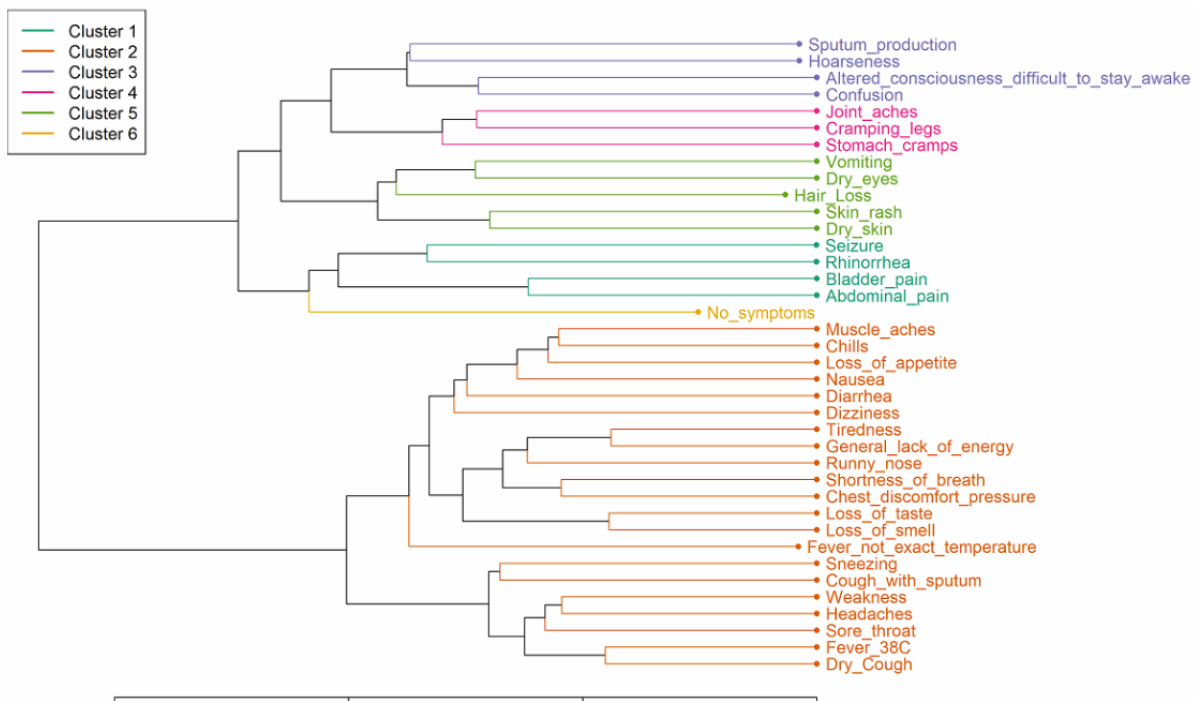


Table 2. Descriptive characteristics of COVID-19 symptomology groups.

Characteristic	Group 1 (abdominal and bladder pain) (N=227), n (%)	Group 2 (flu-like symptoms) (N=1139), n (%)	Group 3 (hoarseness, sputum production) (N=144), n (%)	Group 4 (joint aches, stomach cramps) (N=158), n (%)	Group 5 (skin or eye dryness) (N=196), n (%)	Group 6 (no symptoms) (N=85), n (%)
Hospitalization	153 (67.4)	621 (54.5)	79 (54.9)	65 (41.1)	134 (68.4)	12 (14.1)
Mechanical ventilation (among hospitalized patients)	112 (73.2)	366 (58.9)	46 (58.2)	31 (47.7)	86 (64.2)	10 (83.3)
Demographic characteristics						
Male gender	107 (47.1)	588 (51.6)	64 (44.4)	87 (55.1)	91 (46.4)	54 (63.5)
Age 18-44 years	148 (65.2)	727 (63.8)	82 (56.9)	94 (59.5)	133 (67.9)	70 (82.4)
Age ≥45 years	76 (33.5)	379 (33.3)	54 (37.5)	58 (36.7)	57 (29.1)	15 (17.6)
White race	186 (81.9)	929 (81.6)	118 (81.9)	134 (84.8)	168 (85.7)	60 (70.6)
Hispanic or Latino ethnicity	82 (36.1)	230 (20.2)	23 (16.0)	27 (17.1)	43 (21.9)	7 (8.2)
US \$75,000 or more yearly income	19 (8.4)	162 (14.2)	20 (13.9)	28 (17.7)	21 (10.7)	18 (21.2)
Smoking	121 (53.3)	450 (39.5)	53 (36.8)	45 (28.5)	102 (52.0)	26 (30.6)
Flu vaccination	144 (63.4)	533 (46.8)	62 (43.1)	69 (43.7)	102 (52.0)	30 (35.3)
Chronic conditions						
Depression	38 (16.7)	285 (25.0)	35 (24.3)	37 (23.4)	64 (32.7)	12 (14.1)
Obesity	32 (14.1)	165 (14.5)	35 (24.3)	37 (23.4)	29 (14.8)	6 (7.1)
Asthma	64 (28.2)	156 (13.7)	29 (20.1)	19 (12.0)	29 (14.8)	6 (7.1)
Alcohol or substance use disorder	64 (28.2)	128 (11.2)	19 (13.2)	17 (10.8)	25 (12.8)	7 (8.2)
Diabetes, uncomplicated	12 (5.3)	116 (10.2)	15 (10.4)	19 (12.0)	28 (14.3)	1 (1.2)
Mental illness	36 (15.9)	98 (8.6)	23 (16.0)	22 (13.9)	17 (8.7)	4 (4.7)
Migraines	23 (10.1)	81 (7.1)	15 (10.4)	23 (14.6)	15 (7.7)	4 (4.7)
Weight loss	20 (8.8)	76 (6.7)	14 (9.7)	15 (9.5)	27 (13.8)	3 (3.5)
Anemia	54 (23.8)	70 (6.1)	12 (8.3)	13 (8.2)	22 (11.2)	3 (3.5)
High cholesterol	7 (3.1)	69 (6.1)	14 (9.7)	20 (12.7)	9 (4.6)	3 (3.5)
Ulcer	10 (4.4)	68 (6.0)	5 (3.5)	11 (7.0)	25 (12.8)	3 (3.5)
No chronic condition	34 (15.0)	264 (23.2)	23 (16.0)	31 (19.6)	29 (14.8)	45 (52.9)

Symptomology Groups Associated With Adverse COVID-19 Outcomes

Our findings from the logistic regression models, using symptomology groups as risk factors for adverse COVID-19 outcomes and adjusted for all sociodemographic characteristics and comorbid conditions, showed the following 3 groups

associated with hospitalization: abdominal and bladder pain group (Group 1; OR 1.5, 95% CI 1.01-2.34); flu-like symptoms group (Group 2; OR 3.33, 95% CI 1.97-5.79); and skin or eye dryness group (Group 5; OR 1.63, 95% CI 1.07-2.52). No symptomology group was associated with a high risk for mechanical ventilation (Table 3).

Table 3. Associations between COVID-19 symptomology groups and adverse COVID-19 outcomes.

Symptomology group ^a	Hospitalization ^b			Mechanical ventilation ^b		
	OR ^c	95% CI	P value	OR	95% CI	P value
Abdominal and bladder pain group (Group 1)	1.54	1.01-2.34	.04	1.12	0.66-1.92	.68
Flu-like symptoms group (Group 2)	3.33	1.97-5.79	<.001	0.17	0.04-0.54	.01
Hoarseness and sputum production group (Group 3)	1.51	0.92-2.48	.10	1.09	0.57-2.11	.80
Joint aches and stomach cramps group (Group 4)	0.56	0.35-0.88	.01	0.54	0.27-1.07	.08
Skin or eye dryness group (Group 5)	1.63	1.07-2.52	.02	1.25	0.76-2.05	.39

^aGroup 6 (no symptoms) is excluded.

^bMultivariate logistic models adjusted for sociodemographic characteristics and comorbid conditions.

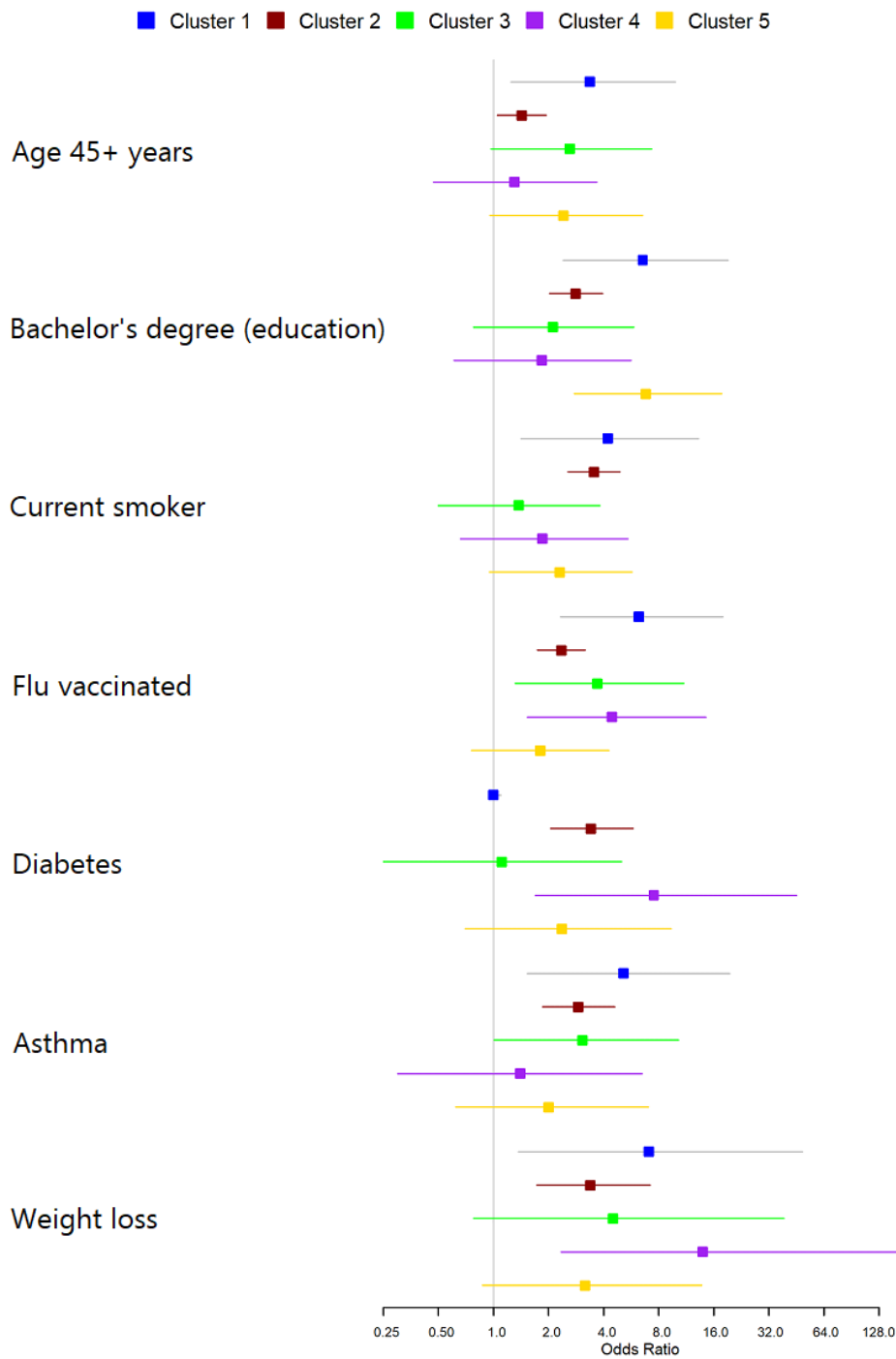
^cOR: odds ratio.

Risk Factors for COVID-19 Hospitalization Among Symptomology Groups

Finally, we developed 5 logistic regression models for symptomology groups to compare the risk factors for COVID-19 hospitalization among those groups (asymptomatic participants were excluded from this analysis). The results of those models are presented as a forest plot of significant variables in at least one symptomology group (Figure 3). The risk factors differed between participants from different symptomology groups. The only risk factor that was significant for 4 out of 5 groups was

influenza vaccine in the last season (Group 1: OR 6.22, 95% CI 2.32-17.92; Group 2: OR 2.35, 95% CI 1.74-3.18; Group 3: OR 3.7, 95% CI 1.32-10.98; Group 4: OR 4.44, 95% CI 1.53-14.49). Smoking (OR 4.22, 95% CI 1.42-13.26) and asthma (OR 5.14, 95% CI 1.53-19.56) were significant risk factors for hospitalization in the abdominal and bladder pain group (Group 1). Weight loss was a risk factor in the joint aches and stomach cramps group (Group 4; OR 13.9, 95% CI 2.34-161.64) and in the abdominal and bladder pain group (Group 1; OR 7.05, 95% CI 1.37-49.01). Diabetes was a risk factor in the joint aches and stomach cramps group (Group 4; OR 7.5, 95% CI 1.69-45.28).

Figure 3. Risk factors for hospitalization among individuals in different symptomology groups.



Findings From the Validation Assessment

A comparison of our findings with those of other studies can be found in [Multimedia Appendix 9](#). At the time of our analysis, we found 3 systematic review or meta-analysis studies mapping the association of symptoms with the risk of adverse outcomes of COVID-19 [19-21].

We found agreement between this study and previous studies for 18 symptoms, 6 of which were associated with adverse outcomes (abdominal pain, cough, dyspnea/shortness of breath, fever, fatigue, and vomiting). In addition, we assessed 14 symptoms that were not previously studied by others, 6 of which

were associated with adverse outcomes (bladder pain, dry eyes, dry skin, loss of appetite, seizure, and skin rash).

Discussion

Principal Findings

Our results identified individual symptoms and behaviors associated with COVID-19 adverse outcomes. Among these, some were well-known and some were new. We also identified 6 symptomology groups, with 3 groups showing statistically significant associations with COVID-19 outcomes. Furthermore, the findings of this work increase our understanding of the

MTurk population and show that with precautionary measures, high-quality data can be obtained.

Well-known single COVID-19 symptoms identified (ie, abdominal pain, cough, fever, and shortness of breath) were associated with hospitalization [5,6]. Less common symptoms identified, such as bladder pain, eye dryness, and skin dryness were also associated with adverse COVID-19 outcomes. We provided additional validation of our findings by comparing the results with the findings of systematic review and meta-analysis studies. The individual symptoms we identified as being associated with adverse COVID-19 outcomes were consistent with the symptoms in those studies.

Our analysis of chronic conditions and associations with COVID-19 adverse outcomes showed that patients with preexisting asthma, diabetes, depression, and bladder problems were at high risk for hospitalization, similar to the findings in previous studies. Although previous studies have shown an increased risk of severe COVID-19 among people with obesity [34], our study did not find a significant increase in the risk of hospitalization among obese people. This result may be due to the participants in our sample being younger than those in other studies, resulting in a weaker link between obesity and chronic diseases that are the actual drivers of COVID-19 severity.

When studying behaviors influencing adverse COVID-19 outcomes, like previous studies, we found that smoking increased the risk of severe COVID-19 outcomes [35-37]. Current smokers and past smokers who quit less than a year ago had a higher risk of hospitalization, and every day smokers also had a higher risk for mechanical ventilation. Our finding showing an effect of influenza vaccination on adverse outcomes contradicts the findings in some other studies. For example, it has been previously reported that influenza vaccination could be considered a protective factor against severe cases of COVID-19 infection [38,39]. Our data, however, suggested that COVID-19-positive respondents who were vaccinated against influenza in Autumn 2019 had higher odds of hospitalization and mechanical ventilation after adjusting for demographic factors, chronic conditions, and COVID-19 symptoms, as the influenza vaccination status might be associated with preexisting comorbidities and a person's demographics. This is not an isolated finding as others have reported that there is a positive association between influenza vaccination rates and COVID-19 death rates [40], that influenza vaccination coverage in a country is a risk factor associated with higher infection rates of COVID-19 [41], and that there is a need to investigate the potential impact of influenza vaccination on COVID-19 risk and severity [42].

In addition to studying individual symptoms and behaviors, this study identified 6 COVID-19 symptomology groups by cluster analysis and assessed their associations with adverse outcomes of the disease. Three symptomology groups (flu-like symptoms, abdominal and bladder pain symptoms, and eye and skin dryness symptoms) were highly associated with a high risk for hospitalization. While the characteristics of respondents in the flu-like symptoms group were similar to the characteristics of the general population, the abdominal and bladder pain group included survey respondents who had lower income, and were

more likely to have smoked and to be influenza vaccinated. They also tended to have chronic conditions, such as asthma and anemia, and alcohol disorder. The survey respondents in the eye and skin dryness group were generally older and had a greater possibility of being white. They were also more likely to have smoked and to be influenza vaccinated. This group also had a very high percentage of survey respondents with depression, diabetes, and ulcers.

Characterizing patients according to clusters using artificial intelligence devices and machine learning is a pioneering method in a variety of infectious and noninfectious diseases. The use of scientific methods to identify clusters of patients with similar characteristics and specific disease risks might improve awareness of heterogeneity in symptomology, and may enable targeted interventions to reduce disease severity. Other studies of COVID-19 disease trajectories have been able to identify vulnerable population clusters that could benefit from specific health resources, and have provided insights for public health targets for managing the pandemic [43,44]. One previous study identified 3 symptomatic groups and 1 asymptomatic group among COVID-19 patients [43]. However, that study did not analyze the associations between the symptomology groups and COVID-19 outcomes. Our analysis of 6 symptomatic groups found that the risk factors for COVID-19 adverse outcomes differed between participants from the different symptomology groups. For the asymptomatic COVID-19 group, other studies have shown that asymptomatic carriers account for 15% to 60% of the infected population and play a key role in disease transmission [45]. Adding to our understanding of asymptomatic carriers, our findings indicated that the asymptomatic symptomology group had a low percentage of hospitalization; a high percentage of young non-Hispanic men with high income; and a low percentage of people with chronic conditions, smoking, and influenza vaccination. These characteristics add to those described in a review study of asymptomatic COVID-19 carriers' characteristics that found young age alone to be a significant factor for having no symptoms [46-48]. Another study of Mexican outpatients found a lower frequency of smokers and influenza vaccination among asymptomatic responders [43].

The percentage of those connected to a mechanical ventilator among hospitalized patients may seem high in our study (61.8%); however, the management of patients hospitalized with COVID-19 has changed considerably over the course of the pandemic. More than half of the study population had been hospitalized, and two-thirds of them were on ventilators. Since the survey was conducted in the first months of the COVID-19 pandemic, many people who got sick with COVID-19 were hospitalized and then connected to a mechanical ventilator. Over time, fewer people with COVID-19 were hospitalized, and among those who were hospitalized, only patients with more severe disease were put on ventilators. Other studies have also shown a high percentage (68%) of ventilator use among hospitalized COVID patients [49].

This work also showed that with precautionary measures to ensure high-quality data collection, a crowdsourcing model can be used to collect data to characterize symptomology for COVID-19 diagnosis and prognosis. There are many studies

assessing health data on MTurk as a source of high-quality and rapidly collected data, and it has demonstrated good reliability [30,50,51]. However, to improve data quality on MTurk, there are recommendations to include workers with an “approval rate” above 95% and keep the “number of HITs approved” to at least 100 [30,31,51]. Prior studies have not investigated data quality from workers by comparing survey responses of 3 experience levels (according to the number of HITs approved) in a pilot study. By launching a pilot study, we found no difference in the approval rate of workers from different experience groups; thus, all could provide adequate data to satisfy the basic approval criterion. For specific tasks requiring higher cognitive ability, however, workers with more experience may provide higher quality data. In our case, we found that those with 500+ HITs submitted fewer duplicate responses than those with 100-499 HITs. While this may exacerbate the superworker issue, the tradeoff of quality data for the use of more experienced workers may be necessary depending on the task. To provide additional validation of our findings, we compared the findings of individual symptoms associated with COVID-19 to the findings of other researchers and identified many concordant findings.

Limitations

A major limitation of this study was the self-reported data, which can be less reliable than physiological assessments. Our crowdsourced approach, however, allowed for reaching many participants, which helped mitigate the noise, and the fast data collection process was helpful during this pandemic. In addition, during this pandemic, many risk factors of COVID-19 were discovered through social media and other self-reported surveys [52-56]. To use those data sources, crowdsourced practices are emerging in research fields such as *infodemiology* (defined as

collecting and analyzing data in real time through an electronic medium with the aim to inform public health decision makers) [57-59]. Another growing field is *digital epidemiology*, in which researchers are using internet data for epidemiological purposes [60-62]. The techniques of capturing relevant real-world data are promising but need to be further developed to meet the possible public health challenges in the future. Second, some of our findings warrant further validation. The risk factors first reported in our study, such as bladder pain symptoms and eye or skin dryness symptoms, need to be more extensively studied so that they can be used in clinical assessments. Furthermore, the influence of influenza vaccination on COVID-19 adverse outcomes should be further investigated as it appears now that humans will have to co-exist with both diseases for a long time even after this pandemic.

Conclusions

Our work demonstrated that a crowdsourced approach was effective for collecting data to assess the symptomology associated with COVID-19. Conducting a pilot study to assess data quality and population representation facilitated refining the filtering criteria for our final data collection strategy. We validated our approach by comparing the findings from assessing individual symptoms associated with COVID-19 to those identified by others and found highly concordant results. In our assessment of symptomology groups, we discovered that the bladder pain and skin or eye dryness groups had a high risk of COVID-19 hospitalization. Given these findings, we believe that a crowdsourcing strategy, such as the one proposed here, should be considered by others for quick and cost-effective assessments in a rapidly changing spectrum of infectious diseases, and societal and environmental factors.

Acknowledgments

We would like to thank Dr Dina Demner-Fushman (National Library of Medicine, National Institutes of Health) for her helpful feedback. We would also like to thank the Johns Hopkins COVID-19 Research Response Program Community Research Working Group for providing access to the toolkit used by this research team to create a survey that includes harmonized data elements. This group is led by Dr Shruti Mehta (Public Health), Dr Jason Farley (Nursing), and Dr Jacky Jennings (Medicine). This work was supported in part by a Microsoft Investigator Fellowship awarded to COT.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Individual survey of COVID-19 symptoms.

[[DOCX File, 27 KB - formative_v6i12e37507_app1.docx](#)]

Multimedia Appendix 2

Family survey of COVID-19 symptoms.

[[DOCX File, 31 KB - formative_v6i12e37507_app2.docx](#)]

Multimedia Appendix 3

Comparison of demographic characteristics across all experience groups among approved Amazon Mechanical Turk workers.

[[DOCX File, 27 KB - formative_v6i12e37507_app3.docx](#)]

Multimedia Appendix 4

Descriptive characteristics of participants in the individual and family surveys.

[[DOCX File , 32 KB - formative_v6i12e37507_app4.docx](#)]

Multimedia Appendix 5

Demographic characteristics of the study participants.

[[DOCX File , 27 KB - formative_v6i12e37507_app5.docx](#)]

Multimedia Appendix 6

Chronic condition characteristics of the study participants.

[[DOCX File , 31 KB - formative_v6i12e37507_app6.docx](#)]

Multimedia Appendix 7

COVID-19 symptom characteristics of the study participants.

[[DOCX File , 30 KB - formative_v6i12e37507_app7.docx](#)]

Multimedia Appendix 8

Associations between symptoms and adverse COVID-19 outcomes adjusted for sociodemographic factors and chronic conditions (multivariate logistic regression).

[[DOCX File , 41 KB - formative_v6i12e37507_app8.docx](#)]

Multimedia Appendix 9

Comparison of our findings with those of systematic review and meta-analysis studies regarding the association between COVID-19 symptoms and adverse outcomes.

[[DOCX File , 18 KB - formative_v6i12e37507_app9.docx](#)]

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2022-11-10]
2. Symptoms of COVID-19. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html> [accessed 2022-11-10]
3. CDC COVID-19 Response Team. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Mar 27;69(12):343-346 [FREE Full text] [doi: [10.15585/mmwr.mm6912e2](https://doi.org/10.15585/mmwr.mm6912e2)] [Medline: [32214079](https://pubmed.ncbi.nlm.nih.gov/32214079/)]
4. Booth A, Reed AB, Ponzo S, Yassaee A, Aral M, Plans D, et al. Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. *PLoS One* 2021 Mar 4;16(3):e0247461 [FREE Full text] [doi: [10.1371/journal.pone.0247461](https://doi.org/10.1371/journal.pone.0247461)] [Medline: [33661992](https://pubmed.ncbi.nlm.nih.gov/33661992/)]
5. Berek MA, Aziz MA, Islam MS. Impact of age, sex, comorbidities and clinical symptoms on the severity of COVID-19 cases: A meta-analysis with 55 studies and 10014 cases. *Heliyon* 2020 Dec;6(12):e05684 [FREE Full text] [doi: [10.1016/j.heliyon.2020.e05684](https://doi.org/10.1016/j.heliyon.2020.e05684)] [Medline: [33344791](https://pubmed.ncbi.nlm.nih.gov/33344791/)]
6. Li J, Huang DQ, Zou B, Yang H, Hui WZ, Rui F, et al. Epidemiology of COVID-19: A systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. *J Med Virol* 2021 Mar 25;93(3):1449-1458 [FREE Full text] [doi: [10.1002/jmv.26424](https://doi.org/10.1002/jmv.26424)] [Medline: [32790106](https://pubmed.ncbi.nlm.nih.gov/32790106/)]
7. Jain V, Yuan J. Predictive symptoms and comorbidities for severe COVID-19 and intensive care unit admission: a systematic review and meta-analysis. *Int J Public Health* 2020 Jun 25;65(5):533-546 [FREE Full text] [doi: [10.1007/s00038-020-01390-7](https://doi.org/10.1007/s00038-020-01390-7)] [Medline: [32451563](https://pubmed.ncbi.nlm.nih.gov/32451563/)]
8. Rodriguez-Morales AJ, Cardona-Ospina JA, Gutiérrez-Ocampo E, Villamizar-Peña R, Holguin-Rivera Y, Escalera-Antezana JP, Latin American Network of Coronavirus Disease 2019-COVID-19 Research (LANCOVID-19). Electronic address: <https://www.lancovid.org>. Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis. *Travel Med Infect Dis* 2020 Mar;34:101623 [FREE Full text] [doi: [10.1016/j.tmaid.2020.101623](https://doi.org/10.1016/j.tmaid.2020.101623)] [Medline: [32179124](https://pubmed.ncbi.nlm.nih.gov/32179124/)]
9. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. *J Infect* 2020 Aug;81(2):e16-e25 [FREE Full text] [doi: [10.1016/j.jinf.2020.04.021](https://doi.org/10.1016/j.jinf.2020.04.021)] [Medline: [32335169](https://pubmed.ncbi.nlm.nih.gov/32335169/)]
10. Ahmed A, Ali A, Hasan S. Comparison of Epidemiological Variations in COVID-19 Patients Inside and Outside of China-A Meta-Analysis. *Front Public Health* 2020;8:193 [FREE Full text] [doi: [10.3389/fpubh.2020.00193](https://doi.org/10.3389/fpubh.2020.00193)] [Medline: [32574293](https://pubmed.ncbi.nlm.nih.gov/32574293/)]
11. Li L, Huang T, Wang Y, Wang Z, Liang Y, Huang T, et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *J Med Virol* 2020 Jun 23;92(6):577-583 [FREE Full text] [doi: [10.1002/jmv.25757](https://doi.org/10.1002/jmv.25757)] [Medline: [32162702](https://pubmed.ncbi.nlm.nih.gov/32162702/)]
12. Sun P, Qie S, Liu Z, Ren J, Li K, Xi J. Clinical characteristics of hospitalized patients with SARS-CoV-2 infection: A single arm meta-analysis. *J Med Virol* 2020 Jun;92(6):612-617 [FREE Full text] [doi: [10.1002/jmv.25735](https://doi.org/10.1002/jmv.25735)] [Medline: [32108351](https://pubmed.ncbi.nlm.nih.gov/32108351/)]

13. Ioannou GN, Locke E, Green P, Berry K, O'Hare AM, Shah JA, et al. Risk Factors for Hospitalization, Mechanical Ventilation, or Death Among 10 131 US Veterans With SARS-CoV-2 Infection. *JAMA Netw Open* 2020 Sep 01;3(9):e2022310 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.22310](https://doi.org/10.1001/jamanetworkopen.2020.22310)] [Medline: [32965502](https://pubmed.ncbi.nlm.nih.gov/32965502/)]
14. Nobel YR, Phipps M, Zucker J, Lebowitz B, Wang TC, Sobieszczyk ME, et al. Gastrointestinal Symptoms and Coronavirus Disease 2019: A Case-Control Study From the United States. *Gastroenterology* 2020 Jul;159(1):373-375.e2 [FREE Full text] [doi: [10.1053/j.gastro.2020.04.017](https://doi.org/10.1053/j.gastro.2020.04.017)] [Medline: [32294477](https://pubmed.ncbi.nlm.nih.gov/32294477/)]
15. Gupta A, Madhavan MV, Sehgal K, Nair N, Mahajan S, Sehrawat TS, et al. Extrapulmonary manifestations of COVID-19. *Nat Med* 2020 Jul 10;26(7):1017-1032. [doi: [10.1038/s41591-020-0968-3](https://doi.org/10.1038/s41591-020-0968-3)]
16. Wang H, Li X, Yan Z, Sun X, Han J, Zhang B. Potential neurological symptoms of COVID-19. *Ther Adv Neurol Disord* 2020 Mar 28;13:1756286420917830 [FREE Full text] [doi: [10.1177/1756286420917830](https://doi.org/10.1177/1756286420917830)] [Medline: [32284735](https://pubmed.ncbi.nlm.nih.gov/32284735/)]
17. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The Prevalence of Symptoms in 24,410 Adults Infected by the Novel Coronavirus (SARS-CoV-2; COVID-19): A Systematic Review and Meta-Analysis of 148 Studies from 9 Countries. *SSRN Journal* 2020:3582819. [doi: [10.2139/ssrn.3582819](https://doi.org/10.2139/ssrn.3582819)]
18. Santosh R, Schwartz HA, Eichstaedt J, Ungar L, Guntuku SC. Detecting Emerging Symptoms of COVID-19 using Context-based Twitter Embeddings. In: *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. 2020 Presented at: 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020; December 2020; Online. [doi: [10.18653/v1/2020.nlpccovid19-2.35](https://doi.org/10.18653/v1/2020.nlpccovid19-2.35)]
19. Li J, Chen Z, Nie Y, Ma Y, Guo Q, Dai X. Identification of Symptoms Prognostic of COVID-19 Severity: Multivariate Data Analysis of a Case Series in Henan Province. *J Med Internet Res* 2020 Jun 30;22(6):e19636 [FREE Full text] [doi: [10.2196/19636](https://doi.org/10.2196/19636)] [Medline: [32544071](https://pubmed.ncbi.nlm.nih.gov/32544071/)]
20. Talukder A, Razu SR, Alif SM, Rahman MA, Islam SMS. Association Between Symptoms and Severity of Disease in Hospitalised Novel Coronavirus (COVID-19) Patients: A Systematic Review and Meta-Analysis. *J Multidiscip Healthc* 2022;15:1101-1110 [FREE Full text] [doi: [10.2147/JMDH.S357867](https://doi.org/10.2147/JMDH.S357867)] [Medline: [35592814](https://pubmed.ncbi.nlm.nih.gov/35592814/)]
21. Cui L, Wang C, Wu Z, Peng D, Huang J, Zhang C, et al. Symptomatology differences of major depression in psychiatric versus general hospitals: A machine learning approach. *J Affect Disord* 2020 Jan 01;260:349-360. [doi: [10.1016/j.jad.2019.09.030](https://doi.org/10.1016/j.jad.2019.09.030)] [Medline: [31521873](https://pubmed.ncbi.nlm.nih.gov/31521873/)]
22. Sudre CH, Lee KA, Lochlainn MN, Varsavsky T, Murray B, Graham MS, et al. Symptom clusters in COVID-19: A potential clinical prediction tool from the COVID Symptom Study app. *Sci Adv* 2021 Mar 19;7(12):eabd4177 [FREE Full text] [doi: [10.1126/sciadv.abd4177](https://doi.org/10.1126/sciadv.abd4177)] [Medline: [33741586](https://pubmed.ncbi.nlm.nih.gov/33741586/)]
23. Vermicelli S, Cricelli L, Grimaldi M. How can crowdsourcing help tackle the COVID - 19 pandemic? An explorative overview of innovative collaborative practices. *R&D Management* 2020 Nov 15;51(2):183-194. [doi: [10.1111/radm.12443](https://doi.org/10.1111/radm.12443)]
24. Desai A, Warner J, Kuderer N, Thompson M, Painter C, Lyman G, et al. Crowdsourcing a crisis response for COVID-19 in oncology. *Nat Cancer* 2020 May 21;1(5):473-476 [FREE Full text] [doi: [10.1038/s43018-020-0065-z](https://doi.org/10.1038/s43018-020-0065-z)] [Medline: [35121981](https://pubmed.ncbi.nlm.nih.gov/35121981/)]
25. Crowdsourcing COVID-19. Harvard College. URL: <https://hms.harvard.edu/news/crowdsourcing-covid-19> [accessed 2022-11-10]
26. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 2020 Apr 07;369:m1328 [FREE Full text] [doi: [10.1136/bmj.m1328](https://doi.org/10.1136/bmj.m1328)] [Medline: [32265220](https://pubmed.ncbi.nlm.nih.gov/32265220/)]
27. Difallah D, Filatova E, Ipeirotis P. Demographics and Dynamics of Mechanical Turk Workers. In: *WSDM '18: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 2018 Presented at: Eleventh ACM International Conference on Web Search and Data Mining; February 5-9, 2018; Marina Del Rey, CA, USA p. 135-143. [doi: [10.1145/3159652.3159661](https://doi.org/10.1145/3159652.3159661)]
28. Moss AJ, Rosenzweig C, Robinson J, Litman L. Demographic Stability on Mechanical Turk Despite COVID-19. *Trends Cogn Sci* 2020 Sep;24(9):678-680 [FREE Full text] [doi: [10.1016/j.tics.2020.05.014](https://doi.org/10.1016/j.tics.2020.05.014)] [Medline: [32553445](https://pubmed.ncbi.nlm.nih.gov/32553445/)]
29. Mortensen K, Hughes TL. Comparing Amazon's Mechanical Turk Platform to Conventional Data Collection Methods in the Health and Medical Research Literature. *J Gen Intern Med* 2018 Apr 4;33(4):533-538 [FREE Full text] [doi: [10.1007/s11606-017-4246-0](https://doi.org/10.1007/s11606-017-4246-0)] [Medline: [29302882](https://pubmed.ncbi.nlm.nih.gov/29302882/)]
30. Young J, Young KM. Don't Get Lost in the Crowd: Best Practices for Using Amazon's Mechanical Turk in Behavioral Research. *Journal of the Midwest Association for Information Systems* 2019;2019(2):Article 2 [FREE Full text] [doi: [10.17705/3jmwa.000050](https://doi.org/10.17705/3jmwa.000050)]
31. Robinson J, Rosenzweig C, Moss AJ, Litman L. Tapped out or barely tapped? Recommendations for how to harness the vast and largely unused potential of the Mechanical Turk participant pool. *PLoS One* 2019 Dec 16;14(12):e0226394 [FREE Full text] [doi: [10.1371/journal.pone.0226394](https://doi.org/10.1371/journal.pone.0226394)] [Medline: [31841534](https://pubmed.ncbi.nlm.nih.gov/31841534/)]
32. COVID-19 Community Response Survey. PhenX Toolkit. URL: <https://www.phenxtoolkit.org/protocols/view/940301> [accessed 2022-11-14]
33. Sarker A, Lakamana S, Hogg-Bremer W, Xie A, Al-Garadi MA, Yang YC. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. *J Am Med Inform Assoc* 2020 Aug 01;27(8):1310-1315 [FREE Full text] [doi: [10.1093/jamia/ocaa116](https://doi.org/10.1093/jamia/ocaa116)] [Medline: [32620975](https://pubmed.ncbi.nlm.nih.gov/32620975/)]

34. Steenblock C, Hassanein M, Khan EG, Yaman M, Kamel M, Barbir M, et al. Obesity and COVID-19: What are the Consequences? *Horm Metab Res* 2022 Aug 09;54(8):496-502 [[FREE Full text](#)] [doi: [10.1055/a-1878-9757](https://doi.org/10.1055/a-1878-9757)] [Medline: [35724688](https://pubmed.ncbi.nlm.nih.gov/35724688/)]
35. Reddy RK, Charles WN, Sklavounos A, Dutt A, Seed PT, Khajuria A. The effect of smoking on COVID-19 severity: A systematic review and meta-analysis. *J Med Virol* 2021 Feb 13;93(2):1045-1056 [[FREE Full text](#)] [doi: [10.1002/jmv.26389](https://doi.org/10.1002/jmv.26389)] [Medline: [32749705](https://pubmed.ncbi.nlm.nih.gov/32749705/)]
36. Patanavanich R, Glantz S. Smoking Is Associated With COVID-19 Progression: A Meta-analysis. *Nicotine Tob Res* 2020 Aug 24;22(9):1653-1656 [[FREE Full text](#)] [doi: [10.1093/ntr/ntaa082](https://doi.org/10.1093/ntr/ntaa082)] [Medline: [32399563](https://pubmed.ncbi.nlm.nih.gov/32399563/)]
37. van Zyl-Smit RN, Richards G, Leone FT. Tobacco smoking and COVID-19 infection. *The Lancet Respiratory Medicine* 2020 Jul;8(7):664-665. [doi: [10.1016/s2213-2600\(20\)30239-3](https://doi.org/10.1016/s2213-2600(20)30239-3)]
38. Yang M, Rooks BJ, Le TT, Santiago IO, Diamond J, Dorsey NL, et al. Influenza Vaccination and Hospitalizations Among COVID-19 Infected Adults. *J Am Board Fam Med* 2021 Feb 23;34(Supplement):S179-S182. [doi: [10.3122/jabfm.2021.s1.200528](https://doi.org/10.3122/jabfm.2021.s1.200528)]
39. Zein JG, Whelan G, Erzurum SC. Safety of influenza vaccine during COVID-19. *J. Clin. Trans. Sci* 2020 Sep 17;5(1):e49. [doi: [10.1017/cts.2020.543](https://doi.org/10.1017/cts.2020.543)]
40. Walach H, Klement R. Influenza Vaccination Rates Predict 30% of the Variance in Covid-19 Related Deaths in Europe – A Modeling Approach. Zenodo. URL: <https://zenodo.org/record/4067724#.Y2y0auRBzDc> [accessed 2022-11-10]
41. Lisewski AM. Association between Influenza Vaccination Rates and SARS-CoV-2 Outbreak Infection Rates in OECD Countries. SSRN Journal 2020:3558270. [doi: [10.2139/ssrn.3558270](https://doi.org/10.2139/ssrn.3558270)]
42. De Wals P, Divangahi M. Could seasonal influenza vaccination influence COVID-19 risk? medRxiv. URL: <https://www.medrxiv.org/content/10.1101/2020.09.02.20186734v1.full> [accessed 2022-11-10]
43. Fernández-Rojas MA, Luna-Ruiz Esparza MA, Campos-Romero A, Calva-Espinosa DY, Moreno-Camacho JL, Langle-Martínez AP, et al. Epidemiology of COVID-19 in Mexico: Symptomatic profiles and presymptomatic people. *Int J Infect Dis* 2021 Mar;104:572-579 [[FREE Full text](#)] [doi: [10.1016/j.ijid.2020.12.086](https://doi.org/10.1016/j.ijid.2020.12.086)] [Medline: [33434668](https://pubmed.ncbi.nlm.nih.gov/33434668/)]
44. San-Cristobal R, Martín-Hernández R, Ramos-Lopez O, Martinez-Urbistondo D, Micó V, Colmenarejo G, et al. Longwise Cluster Analysis for the Prediction of COVID-19 Severity within 72 h of Admission: COVID-DATA-SAVE-LIFES Cohort. *J Clin Med* 2022 Jun 10;11(12):3327 [[FREE Full text](#)] [doi: [10.3390/jcm11123327](https://doi.org/10.3390/jcm11123327)] [Medline: [35743398](https://pubmed.ncbi.nlm.nih.gov/35743398/)]
45. Nikolai LA, Meyer CG, Kreamsner PG, Velavan TP. Asymptomatic SARS Coronavirus 2 infection: Invisible yet invincible. *Int J Infect Dis* 2020 Nov;100:112-116 [[FREE Full text](#)] [doi: [10.1016/j.ijid.2020.08.076](https://doi.org/10.1016/j.ijid.2020.08.076)] [Medline: [32891737](https://pubmed.ncbi.nlm.nih.gov/32891737/)]
46. Zhang W, Wu X, Zhou H, Xu F. Clinical characteristics and infectivity of asymptomatic carriers of SARS-CoV-2 (Review). *Exp Ther Med* 2021 Feb;21(2):115 [[FREE Full text](#)] [doi: [10.3892/etm.2020.9547](https://doi.org/10.3892/etm.2020.9547)] [Medline: [33335578](https://pubmed.ncbi.nlm.nih.gov/33335578/)]
47. Hu Z, Song C, Xu C, Jin G, Chen Y, Xu X, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Sci China Life Sci* 2020 May 04;63(5):706-711 [[FREE Full text](#)] [doi: [10.1007/s11427-020-1661-4](https://doi.org/10.1007/s11427-020-1661-4)] [Medline: [32146694](https://pubmed.ncbi.nlm.nih.gov/32146694/)]
48. Wang Y, Liu Y, Liu L, Wang X, Luo N, Li L. Clinical Outcomes in 55 Patients With Severe Acute Respiratory Syndrome Coronavirus 2 Who Were Asymptomatic at Hospital Admission in Shenzhen, China. *J Infect Dis* 2020 May 11;221(11):1770-1774 [[FREE Full text](#)] [doi: [10.1093/infdis/jiaa119](https://doi.org/10.1093/infdis/jiaa119)] [Medline: [32179910](https://pubmed.ncbi.nlm.nih.gov/32179910/)]
49. Anesi G, Jablonski J, Harhay MO, Atkins JH, Bajaj J, Baston C, et al. Characteristics, Outcomes, and Trends of Patients With COVID-19-Related Critical Illness at a Learning Health System in the United States. *Ann Intern Med* 2021 May;174(5):613-621 [[FREE Full text](#)] [doi: [10.7326/M20-5327](https://doi.org/10.7326/M20-5327)] [Medline: [33460330](https://pubmed.ncbi.nlm.nih.gov/33460330/)]
50. Hauser DJ, Schwarz N. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav Res Methods* 2016 Mar 12;48(1):400-407. [doi: [10.3758/s13428-015-0578-z](https://doi.org/10.3758/s13428-015-0578-z)] [Medline: [25761395](https://pubmed.ncbi.nlm.nih.gov/25761395/)]
51. Moss A, Litman L. After the Bot Scare: Understanding What's Been Happening With Data Collection on MTurk and How to Stop It. CloudResearch. URL: <https://www.cloudresearch.com/resources/blog/after-the-bot-scare-understanding-whats-been-happening-with-data-collection-on-mturk-and-how-to-stop-it/> [accessed 2022-11-10]
52. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med* 2020 Jul 11;26(7):1037-1040 [[FREE Full text](#)] [doi: [10.1038/s41591-020-0916-2](https://doi.org/10.1038/s41591-020-0916-2)] [Medline: [32393804](https://pubmed.ncbi.nlm.nih.gov/32393804/)]
53. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat Med* 2021 Jan 29;27(1):73-77. [doi: [10.1038/s41591-020-1123-x](https://doi.org/10.1038/s41591-020-1123-x)] [Medline: [33122860](https://pubmed.ncbi.nlm.nih.gov/33122860/)]
54. Zens M, Brammertz A, Herpich J, Südkamp N, Hinterseer M. App-Based Tracking of Self-Reported COVID-19 Symptoms: Analysis of Questionnaire Data. *J Med Internet Res* 2020 Sep 09;22(9):e21956 [[FREE Full text](#)] [doi: [10.2196/21956](https://doi.org/10.2196/21956)] [Medline: [32791493](https://pubmed.ncbi.nlm.nih.gov/32791493/)]
55. Peters A, Rospleszcz S, Greiser KH, Dallavalle M, Berger K. Complete list of authors available under: Collaborators. The Impact of the COVID-19 Pandemic on Self-Reported Health. *Dtsch Arztebl Int* 2020 Dec 11;117(50):861-867 [[FREE Full text](#)] [doi: [10.3238/arztebl.2020.0861](https://doi.org/10.3238/arztebl.2020.0861)] [Medline: [33295275](https://pubmed.ncbi.nlm.nih.gov/33295275/)]

56. Dreyer NA, Reynolds M, DeFilippo Mack C, Brinkley E, Petruski-Ivleva N, Hawaldar K, et al. Self-reported symptoms from exposure to Covid-19 provide support to clinical diagnosis, triage and prognosis: An exploratory analysis. *Travel Med Infect Dis* 2020 Nov;38:101909 [FREE Full text] [doi: [10.1016/j.tmaid.2020.101909](https://doi.org/10.1016/j.tmaid.2020.101909)] [Medline: [33152512](https://pubmed.ncbi.nlm.nih.gov/33152512/)]
57. Eysenbach G. Infodemiology: the epidemiology of (mis)information. *The American Journal of Medicine* 2002 Dec;113(9):763-765. [doi: [10.1016/s0002-9343\(02\)01473-0](https://doi.org/10.1016/s0002-9343(02)01473-0)]
58. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](https://pubmed.ncbi.nlm.nih.gov/21521589/)]
59. Mavragani A. Tracking COVID-19 in Europe: Infodemiology Approach. *JMIR Public Health Surveill* 2020 Apr 20;6(2):e18941 [FREE Full text] [doi: [10.2196/18941](https://doi.org/10.2196/18941)] [Medline: [32250957](https://pubmed.ncbi.nlm.nih.gov/32250957/)]
60. Salathé M. Digital epidemiology: what is it, and where is it going? *Life Sci Soc Policy* 2018 Jan 04;14(1):1 [FREE Full text] [doi: [10.1186/s40504-017-0065-7](https://doi.org/10.1186/s40504-017-0065-7)] [Medline: [29302758](https://pubmed.ncbi.nlm.nih.gov/29302758/)]
61. He Z, Zhang CJP, Huang J, Zhai J, Zhou S, Chiu JW, et al. A New Era of Epidemiology: Digital Epidemiology for Investigating the COVID-19 Outbreak in China. *J Med Internet Res* 2020 Sep 17;22(9):e21685 [FREE Full text] [doi: [10.2196/21685](https://doi.org/10.2196/21685)] [Medline: [32805703](https://pubmed.ncbi.nlm.nih.gov/32805703/)]
62. Tarkoma S, Alghnam S, Howell MD. Fighting pandemics with digital epidemiology. *EClinicalMedicine* 2020 Sep;26:100512 [FREE Full text] [doi: [10.1016/j.eclinm.2020.100512](https://doi.org/10.1016/j.eclinm.2020.100512)] [Medline: [32864592](https://pubmed.ncbi.nlm.nih.gov/32864592/)]

Abbreviations

HIT: human intelligence task

MTurk: Amazon Mechanical Turk

OR: odds ratio

Edited by A Mavragani; submitted 23.02.22; peer-reviewed by H Matsui, T Karen; comments to author 04.07.22; revised version received 21.09.22; accepted 02.11.22; published 06.12.22.

Please cite as:

Flaks-Manov N, Bai J, Zhang C, Malpani A, Ray SC, Taylor CO

Assessing Associations Between COVID-19 Symptomology and Adverse Outcomes After Piloting Crowdsourced Data Collection: Cross-sectional Survey Study

JMIR Form Res 2022;6(12):e37507

URL: <https://formative.jmir.org/2022/12/e37507>

doi: [10.2196/37507](https://doi.org/10.2196/37507)

PMID: [36343205](https://pubmed.ncbi.nlm.nih.gov/36343205/)

©Natalie Flaks-Manov, Jiawei Bai, Cindy Zhang, Anand Malpani, Stuart C Ray, Casey Overby Taylor. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Engagement, Use, and Impact of Digital Mental Health Resources for Diverse Populations in COVID-19: Community-Partnered Evaluation

Kenneth Wells^{1,2,3*}, MPH, MD; April Denise Thames^{2*}, PhD; Alexander S Young^{1,2*}, MSHS, MD; Lily Zhang^{1*}, MS; MarySue V Heilemann^{4*}, PhD; Daniela Flores Romero^{1*}, BA; Adrian Oliva^{5*}, BS; Felica Jones^{5*}; Lingqi Tang^{1*}, PhD; Melissa Brymer^{2*}, PsyD, PhD; Thomas Elliott^{6*}, MD; Armen Arevian^{7*}, MD, PhD; Together for Wellness/Juntos Collaborators and Writing Group^{8*}

¹Research Center for Health Services and Society, Jane and Terry Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, United States

²Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, United States

³Department of Health Policy and Management, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA, United States

⁴School of Nursing, University of California, Los Angeles, Los Angeles, CA, United States

⁵Healthy African American Families II, Los Angeles, CA, United States

⁶National Clinician Scholars Program, Division of General Internal Medicine and Health Services Research, Department of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

⁷Chorus Innovations, Inc, Long Beach, CA, United States

⁸See Acknowledgements

*all authors contributed equally

Corresponding Author:

Kenneth Wells, MPH, MD

Research Center for Health Services and Society

Jane and Terry Semel Institute for Neuroscience and Human Behavior

University of California, Los Angeles

10920 Wilshire Blvd Suite 300

Los Angeles, CA, 90024

United States

Phone: 1 310 794 3728

Fax: 1 310 794 3724

Email: KWells@mednet.ucla.edu

Abstract

Background: The COVID-19 pandemic increased disparities for communities burdened by structural barriers such as reduced affordable housing, with mental health consequences. Limited data are available on digital resources for public mental health prevention during the COVID-19 pandemic.

Objective: The study aim was to evaluate engagement in and impact of free digital resources on the Together for Wellness/Juntos por Nuestro Bienestar (T4W/Juntos) website during COVID-19 in California.

Methods: A pilot evaluation of T4W/Juntos was performed, with partner agencies inviting providers, clients, and partners to visit the website and complete surveys at baseline (September 20, 2021, to April 4, 2022) and at 4–6-week follow-up (October 22, 2021, to May 17, 2022). Website use was assessed by three engagement items (ease of use, satisfaction, relevance), comfort in use, and use of six resource categories. Primary outcomes at follow-up were depression and anxiety (scores ≥ 3 on Patient Health Questionnaire-2 item [PHQ2] and Generalized Anxiety Disorder-2 item [GAD2] scales). Secondary outcomes were post-pre differences in PHQ2 and GAD2 scores, and use of behavioral health hotlines and services the month before follow-up.

Results: Of 366 eligible participants, 315 (86.1%) completed baseline and 193 (61.3%) completed follow-up surveys. Of baseline participants, 72.6% identified as female, and 21.3% identified as lesbian, gay, bisexual, transgender, queer/questioning, and others (LGBTQ+). In terms of ethnicity, 44.0% identified as Hispanic, 17.8% as African American, 26.9% as non-Hispanic white, and

11.4% as other ethnicity. Overall, 32.7% had moderate anxiety or depression (GAD2/PHQ2 \geq 3) at baseline. Predictors of baseline website engagement included being Hispanic versus other race/ethnicity ($\beta=.27$, 95% CI .10-.44; $P=.002$) and number of COVID-19–related behavior changes ($\beta=.09$, 95% CI .05-.13; $P<.001$). Predictors of comfort using the website were preferring English for website use (odds ratio [OR] 5.57, 95% CI 2.22-13.96; $P<.001$) and COVID-19–related behavior changes (OR 1.37, 95% CI 1.12-1.66; $P=.002$); receiving overnight behavioral health treatment in the prior 6 months (OR 0.15, 95% CI 0.03-0.69, $P=.015$) was associated with less comfort in website use. The main predictor of depression at follow-up (PHQ2 \geq 3) was baseline depression (OR 6.24, 95% CI 2.77-14.09; $P<.001$). Engagement in T4W/Juntos was associated with lower likelihood of depression (OR 0.54, 95% CI 0.34-0.86; $P=.01$). Website use the month before follow-up was associated with a post-pre reduction in PHQ2 score ($\beta=-.62$, 95% CI -1.04 to -0.20 ; $P=.004$). The main predictor of GAD2 \geq 3 at follow-up was baseline GAD2 \geq 3 (OR 13.65, 95% CI 6.06-30.72; $P<.001$). Greater baseline website engagement predicted reduced hotline use (OR 0.36, 95% CI 0.18-0.71; $P=.004$).

Conclusions: Ethnicity/language and COVID-19–related behavior changes were associated with website engagement; engagement and use predicted reduced follow-up depression and behavioral hotline use. Findings are based on participants recommended by community agencies with moderate follow-up rates; however, significance was similar when weighting for nonresponse. This study may inform research and policy on digital mental health prevention resources.

(*JMIR Form Res* 2022;6(12):e42031) doi:[10.2196/42031](https://doi.org/10.2196/42031)

KEYWORDS

digital mental health; prevention; COVID-19; depression; hotline use; health disparity; community health; public health; health resource; mental well-being; ethnic; website engagement; minority population; digital resource

Introduction

Background

The COVID-19 pandemic increased disparities for communities burdened by structural barriers such as shortage of affordable housing or work opportunities, with mental health consequences [1]. Stressors related to COVID-19 include loss/grief, self-quarantining, physical distancing, social isolation, business and school closures, financial impacts, lack of access to basic needs, and stress, including that related to racial discrimination, with mental health consequences [2]. There is also documentation of provider stress such as burnout, fear of infection, and loss [3]. There are racial/ethnic disparities in prevalence and economic impacts, and disparities exist for other marginalized groups such as lesbian, gay, bisexual, transgender, queer/questioning and others (LGBTQ+), rural farm workers, and persons with mental health challenges [1,3-6]. These impacts are noted across age groups, with concerns about school/work-related distress, substance use, suicidality, abuse/violence, and overall mental health consequences [7,8], highlighting the importance of public-facing interventions.

With advances in social media and digital technology, the importance of digital interventions for public health is clear; some countries use digital resources to promote mental health prevention, including in youth portals [9]. Such strategies may reduce the need for services, while increasing access to services for those in need [10]. There are limited data on the effectiveness of digital interventions for public mental health and in disasters [10]. Over the past decade, mental health care has been supplemented by mobile mental health apps and internet resources, ranging from meditation and mindfulness apps to symptom diaries and self-management tools such as cognitive behavioral therapy (CBT) [11]. While there is evidence for the efficacy of internet-based self-directed CBT and behavioral activation, evidence for many apps is limited, with inconsistent documentation of how apps are informed by evidence, as well

as how they are accessible and tailored for underresourced groups [11-13]. Commentaries on next directions emphasize reframing tools as service enhancements and designing tools with end-user input, attending to the community and system contexts [14]. Some studies used consumer advisor input to tailor apps, such as sexual minority men to manage generalized anxiety and major depressive disorder [15] or consumer advisors' partnered digital resources for youth or underresourced communities, but with limited evidence on impacts [16,17]. A recent review suggested greater uptake and effectiveness of digital resources with human contact in some studies, including a few rigorous randomized trials, although some showed no significant added effects [18]. Commentaries emphasize “solution-focused” approaches to digital mental health, with attention to user experience and an intervention target of engagement (usefulness, usability, satisfaction, use, contacts with others), consistent with community-engaged approaches [19-21].

Digital Mental Health Tools and Together for Wellness Development

Given a history of partnered interventions, including those delivered during disasters [22,23], California Health Care Services Division of Behavioral Health included in its Federal Emergency Management Agency (FEMA)/Substance Abuse and Mental Health Services Administration (SAMHSA) a COVID-19 crisis counseling program, supporting free digital mental health resources (Together for Wellness/Juntos por Nuestro Bienestar [T4W/Juntos]) with partner input. T4W/Juntos features evidence-informed or evidence-based resources reviewed with lead agencies (Latinx, Black, Asian American, LGBTQ+, parent support for youths, older adults, persons with mental illness). The public website was developed iteratively, initially focusing on mindfulness, and adding resources vetted by community partners on coping with stress; grief; connecting to others; social justice issues such as racial discrimination; information on COVID-19; and resources for families with

children, teachers, and older adults. Given the input, toolkits included visuals and videos, resources in multiple languages, and video orientations by provider and community partners; the initial partnered development process was previously described in a commentary [24].

This article describes further development of T4W/Juntos and results of a partnered evaluation, informed by principles of community-partnered participatory research (CPPR), including trust, respect, and two-way input [21]. The website and evaluation are informed by the technology acceptance model (TAM) and Behavioral Model for Vulnerable Populations, with COVID-19-related behavior changes [25-27]. The TAM explains use of technologies by professionals and patients, emphasizing reported engagement/satisfaction and comfort in use, contributing to actual use [25]. The Vulnerable Populations model emphasizes factors affecting underresourced groups [26], coupled here with COVID-19-related changes [27], to inform mechanisms of action (engagement, comfort in use) for promoting mental well-being, attending to the individual and system contexts [19,20].

The research questions for this study were as follows:

- (1) What are specific individual and social factors that promote engagement and use of the T4W/Juntos website?
- (2) Does reported engagement in and use of T4W/Juntos reduce symptoms of depression and anxiety (primary), and use of crisis hotlines and behavioral health services (secondary)?

Based on input from community and policy partners, we hypothesized that individuals with higher need (greater depression, anxiety symptoms) and more COVID-19-related

behavior changes (such as social distancing) and COVID-19-related stressors (such as loss or financial stress) would engage more in website use, and that engagement and use between baseline and follow-up would be associated with reduced depression, anxiety, and use of hotlines at follow-up, as evaluated by participant surveys.

Methods

T4W/Juntos Resources

The website was iteratively reviewed by community partners in group meetings (largely by Zoom) in English and Spanish. Meetings included representatives of 11 agencies, youth and older adult advisory groups, and involved interaction with investigators and technology design leaders. Resources were reviewed and updated. Suggestions led to modifications in design, addition of other languages, and others. Based on this feedback, 133 resources were selected for T4W/Juntos: websites (n=62), videos (n=18), YouTube links (n=2), PDFs (n=32), apps (n=10), and hotlines (n=9). The 6 main categories with examples are illustrated in Table 1. The website is available in 13 languages (videos in Mixteco were published in the summer of 2022). Once resources were made available on the website, we used an iterative process through focus group sessions (N=4) that consisted of youth (n=8) and older adult (n=10) advisors. Community partners were asked to provide feedback about their experience using the website and their review of resources. Each focus group was led by a moderator and notes were taken by members of the T4W/Juntos team. Feedback was communicated to the technology team and the website was revised/updated accordingly.

Table 1. Examples of digital mental health resources selected for the Together for Wellness/Juntos por Nuestro Bienestar website.

Resource	Description
Learn About COVID-19	
Multilingual Resource Hub: Translate-COVID.org ^a	COVID-19 information in over 40 languages on the disease, impacts, vaccines, tips to prevent the spread, and more. Created by the UCLA ^b Asian American Studies Center
COVID-19 Guide for Trans people	COVID-19 guide for the transgender community and their families. Developed by the National Center for Transgender Equality
Covid-19 in Pregnancy ^c	Video featuring Yalda Afshar, MD, PhD, and Rashmi R Rao, MD, at UCLA, who explain COVID-19 and answer questions regarding pregnancy during this pandemic. Video created by UCLA Health
Covid Coach ^d	App created for everyone, including veterans and service members, to support self-care and overall mental health during the COVID-19 pandemic. Developed by the mobile mental health team of the National Center for PTSD ^c , Dissemination & Training Division
Be Safe and Healthy with Potter the Otter	A kid-friendly guide on staying safe during the COVID-19 pandemic starring First 5's Potter the Otter. Created by First 5 LA
Soothe Anxiety and Stress	
Deep Breathing for Beginners ^d	YouTube video demonstrating deep breathing, and giving tips for improving oxygen uptake and reducing stress and fatigue. YouTube channel led by Michelle Kenway
UCLA Mindfulness App ^d	App on mindfulness in meditations that help reduce stress and anxieties, created by the UCLA Mindful Awareness Research Center
Mindshift App ^{a,d}	App based on cognitive behavioral therapy, developed to help you learn to relax and be mindful, develop more effective ways of thinking, and use active steps to take charge of your anxiety. Developed by Anxiety Canada
The Safe Place ^d	Minority mental health app geared toward the Black community, with meditations, breathing and exercise tips, coping with police brutality, and more. Developed by Jasmin Pierre
Stress Relief for Caregivers and Kids during COVID-19 ^d	Guide with tips for parents and caregivers on how to take care of themselves and their families. Developed by the California Surgeon General's office
Support Resilience	
123 Sesame Street: Resilience ^d	Webpage with videos, articles, and other resources for parents and caregivers, with tips for helping children manage anxiety during this time. Produced by Sesame Workshop
First Aid for Feelings ^d	A book of activities for children to understand the changes that are occurring in their lives due to COVID-19. The activities will provide emotional support and resilience skills. Developed by Denise Daniels, RN, MS; Scholastic; and the Yale Child Study Center
Dav Pilkey at Home	Webpage that offers children activities for reading, drawing, creating, as well as videos, featuring well-known cartoon characters. Developed by Scholastic
Cope with a Recent Loss	
Apart of Me ^c	Game designed by grief experts to help users come to terms with the loss they have suffered. Developed by Apart of Me
Healing After Death	App with guided meditations designed to assist you in supporting the spirit of someone you love who has passed away. Developed by Vanessa Callison-Burch
VA: Deal with Loss ^{c,d}	Guide for taking care of yourself after a loss. Developed by the National Center for PTSD
Connect with People & Support Social Justice	
TEACH. PLAY. LOVE. - Important Conversations About Social Justice for Kids ^{c,d}	Podcast led by child development experts who share strategies to help you teach your child about diversity, equity, and inclusion. Developed by Bright Horizons
Covid-19 Resources to Stand Against Racism	Webpage with a list of resources and helpful tips, primarily for AAPI ^f communities, in light of the increased hate crimes that occurred during the pandemic. Developed by Asian Americans Advancing Justice
Explore Justice	A video series designed to help us unpack and examine the current and historical perspectives that shape social justice. Developed by 211LA

^aEvidence-based.^bUCLA: University of California, Los Angeles.^cExpert opinion.

^dEvidence-informed.

^ePTSD: posttraumatic stress disorder.

^fAAPI: Asian American and Pacific Islander.

Evidence Basis

As shown in [Table 1](#), most resources are evidence-informed or evidence-based for specific populations, rather than for public prevention. For example, there are studies of effectiveness of mindfulness on attentional and emotional self-regulation [28]; one study of community health promotion used a pre-post design [29] and another was a randomized trial for depressed English and Spanish adults, which was found to reduce depression [28]. Mindfulness is thus considered to be “evidence-informed” for public prevention. Some resources are based on CBT and mindfulness (such as Mindshift App and Sanvello), whereas others provide information (eg, on depression) with resources for subgroups (eg, children).

Design

The website evaluation was conducted in collaboration with 11 California community-based agencies representing diverse populations. Each agency gave input in planning meetings on survey measures and design, signed an agreement, and received US \$1000 for their work to invite participants for the survey, including clients, partners, and staff, as groups impacted by COVID-19. For agencies with large lists of potential participants, random selection tools were provided. Assuming a 50% response rate, we encouraged the agencies to invite 80-100 individuals to ultimately enroll 30-40 per agency; that number was later expanded to 70 per agency to increase enrollment. The goal was to have 300-350 participants review the website and complete the baseline survey online; of these, 10% would be invited for a qualitative telephone interview 2 weeks later and all would be invited for a 4-6-week online follow-up survey. The goal was to apply the TAM/Vulnerable Population Model to describe engagement in website use [30], and predictors of engagement and anxiety/depression (Generalized Anxiety Disorder scale-2 [GAD2]/Patient Health Questionnaire-2 [PHQ2]), including demographics, services use, and COVID-19-related behavior changes and stressors [22,27,31-33]. For the evaluation, we created a duplicate version of the entire public-access website that allowed us to track enrollment by the inviting agency. Potential participants were emailed invitations (with 1-2 reminders) and a link to the eligibility screener (participants needed to have online access, speak English or Spanish, be aged 18+ years). Eligible participants who consented online were registered, given a link to the duplicate website that they were asked to visit, and asked to complete a baseline survey after visiting the website. Those who completed the baseline survey and consented to follow-up were invited to participate online surveys 4-6 weeks later, and (if selected) to be contacted for a qualitative interview 2 weeks after baseline (10% of total sample). Surveys were provided in English or Spanish. The survey period was September 20, 2021, to April 4, 2022, for the baseline, and October 22, 2021, to May 17, 2022, for follow-up.

Ethics Approval

The study was reviewed and approved by the University of California, Los Angeles Institutional Review Board for Human Subjects (20-002163-AM-00008). All participants provided informed consent online and were advised that participation was voluntary, data would be deidentified, their contact information would be used for follow-up, and linking data (an ID code) would be stored on a secure server separately from study data. Informed consent was affirmed at follow-up assessments. For each participation event (survey, interview), participants received a US \$25 electronic gift card. Inviting agencies were not informed of the participation of individuals. The study team used the separately stored contact information to invite participants to follow-up. Participation in each activity was voluntary.

Measures

Baseline

Demographics included age in years, gender (female, male, genderqueer, questioning, trans man, trans woman, other/unknown, not stated), sexual orientation (gay/lesbian, bisexual/pansexual, queer, not sexual/none, questioning, other/unknown, not stated), race/ethnicity (American Indian/Native American/Alaskan; Black/African American/African; East Asian; South East Asian; Hispanic, Latino/Spanish Origin; Middle Eastern; Pacific Islander; white/Caucasian; unknown/not stated), website language preference (English, Spanish, Cantonese, Vietnamese, Tagalog, Mandarin, Korean, Japanese, Russian, Farsi, Armenian, Arabic, Mixteco, other/unknown/not stated), education (some high school or less, high school graduate/equivalent, vocational/certificate, some college, college graduate, graduate school), and zip code [22,34-36]. Two mental health stigma items were also included at baseline (scored on 5-point Likert agreement scales plus “don’t know”) [22].

Baseline and Follow-up

Depression and anxiety screeners included responses to the PHQ2 and GAD2 (range 0-6; score ≥ 3 considered moderate depression/anxiety) [32,33]. Services use was assessed for the prior 6 months in the baseline survey and for the prior 1 month in the follow-up survey, including any hotline use for behavioral health (any and number of calls) and any behavioral health services (emergency room, primary care, mental health, substance use visits; hospitalization, residential care) [22]. Pandemic-related stress at baseline was assessed according to an adapted version of the COVID-19 Stress Scale [37], including COVID-19-related behavior changes (nine items: no changes, social distancing, isolation, caring for someone at home, working from home, not working, change in health care services, following media coverage on COVID-19, and changing travel plans). The overall impact of the pandemic was assessed over five categories (plus decline to answer). COVID-19 stressors [19] included having COVID-19; fear of acquiring or spreading COVID-19; worrying about others; stigma or discrimination;

financial, food, and home insecurity; frustration, depression, anxiety, alcohol, sleep, and sexual activity change; confusion about COVID-19; other difficulties (not contributing to greater good, having social/emotional support, having financial support from others). Additional items included were vaccination status (yes/no), vaccine acceptability (5-point Likert agreement plus don't know), date of survey completion, and racial/ethnic discrimination during the COVID-19 pandemic (assessed at baseline only: 6 items, 4-point Likert scale) [25,35].

The use of and engagement in the website at baseline were assessed according to website satisfaction, relevance of content, and ease of use, along with comfort in using the website (each with a 5-point Likert agreement scale plus "don't know"). At follow-up, use and engagement were assessed according to actual website use in the prior month, including any use (yes/no); frequency/download and use of any of the six main categories, each with five response options (did not use, used some, used and valuable, used and very valuable, used and would recommend to anyone); and if recommended website to others (yes/no), number of times and which resource categories were recommended [22,25,31].

Statistical Analyses

We performed descriptive analyses to describe the sample using means (SDs) for continuous variables, and counts and percentages for categorical variables. To describe predictors of engagement/use, we fit linear regression models for continuous variables (eg, mean engagement score) and logistic regression models for dichotomous variables (eg, used website). For preliminary bivariate analyses, predictors included demographics, need, stigma, COVID-19-related behaviors and stressors, and behavioral health services use (including predictors of follow-up nonresponse). The final regression model included predictors that were significant in preliminary models at $P < .05$.

For website use in the month before follow-up, we used logistic regression for any use and linear regression for total score of use across the six main categories. We examined predictors at

follow-up of having $\text{PHQ2} \geq 3$ and $\text{GAD2} \geq 3$, and as sensitivity analyses, the post minus pre difference in PHQ2 and GAD2 scores. For secondary analyses, we examined predictors of hotline use and any behavioral health services use (outpatient, inpatient, rehabilitation, hotlines) in the month before follow-up; an imputed version of inpatient/rehabilitation was used for sensitivity analysis, assigning 0 if the items were skipped but other use items answered. We initially examined all measures as predictors in bivariate analyses, including predictors of follow-up nonresponse. We then fit regression models with significant predictors to inform final models. For final follow-up models, we conducted sensitivity analyses using inverse propensity weighting of data for predictors of nonresponse at baseline (age) and for follow-up after baseline [38,39].

Results

Participant Characteristics

By May 31, 2022, 495 individuals had completed the eligibility screener. Of the 446 that were eligible, 367 (82.3%, 6-69 per agency) consented, with no significant difference in mean age ($P = .56$) between those that did not consent (38.7, SD 12.8 years) and did consent (39.7, SD 13.93 years). Of the 367 who consented, 315 (85.8%) completed baseline surveys, with a significantly lower mean age ($P = .002$) for completers (38.8, SD 13.5 years) than noncompleters (45.3, SD 15.1 years).

Of the baseline participants, 72.6% were female, 21.3% self-reported as LGBTQ+, 44.0% were Hispanic, 17.8% African American, 26.9% non-Hispanic white, and 11.4% endorsed another ethnicity. The sample's mean age was 38.8 years with 110 (34.9%) aged 18-30 years and 18.5% had a high school education or less. Of the baseline participants, 32.7% screened positive for moderate anxiety or depression ($\text{GAD2}/\text{PHQ2} \geq 3$) and 61.2% had used behavioral health services in the prior 6 months. For COVID-19-related behavior changes, the mean number selected of the 9 listed was 3.9 (SD 2.0). For COVID-19 stressors, the mean selected of the 19 listed was 6.8 (SD 4.1) (Table 2).

Table 2. Characteristics at baseline, overall, and by follow-up status.

Characteristics	Baseline (n=315)	Had follow-up (n=193)	No follow-up (n=122)	<i>t</i> or χ^2 (df) ^a	<i>P</i> value
Age (years), n responses, mean (SD)	315, 38.8 (13.5)	193, 38.1 (13.0)	122, 39.9 (14.3)	-1.18 (313)	.24
Gender, n (%)				10.13 (2)	.006
Responses, n	310	189	121		
Female	225 (72.6)	130 (68.8)	95 (78.5)		
Male	70 (22.6)	53 (28.0)	17 (14.1)		
Other	15 (4.8)	6 (3.2)	9 (7.4)		
Sex minority, n (%)				0.51 (1)	.48
Responses, n	300	185	115		
Yes	64 (21.3)	37 (20.0)	27 (23.5)		
No	236 (78.7)	148 (80.0)	88 (76.5)		
Race, n (%)				13.04 (3)	.005
Responses, n	298	184	114		
Hispanic	131 (44.0)	68 (37.0)	63 (55.3)		
Black/African American	53 (17.8)	32 (17.4)	21 (18.4)		
White/Caucasian	80 (26.8)	61 (33.2)	19 (16.7)		
Other	34 (11.4)	23 (12.5)	11 (9.6)		
Education, n (%)				5.15 (4)	.27
Responses, n	313	192	121		
Less than high school	23 (7.3)	10 (5.2)	13 (10.7)		
High school graduate	35 (11.2)	21 (10.9)	14 (11.6)		
Some college	87 (27.8)	56 (29.2)	31 (25.6)		
College	119 (38.0)	78 (40.6)	41 (33.9)		
Graduate school	49 (15.7)	27 (14.1)	22 (18.2)		
Language prefer to use on the website, n (%)				0.02 (2)	.99
Responses, n	313	191	122		
English	260 (83.1)	159 (83.2)	101 (82.8)		
Spanish	40 (12.8)	24 (12.6)	16 (13.1)		
Other	13 (4.2)	8 (4.2)	5 (4.1)		
PHQ2^b or GAD2^c≥3, n (%)				0.92 (1)	.34
Responses, n	312	190	122		
Yes	102 (32.7)	66 (34.7)	36 (29.5)		
No	210 (67.3)	124 (65.3)	86 (70.5)		
PHQ2 score, N responses, mean (SD)	311, 1.6(1.5)	190, 1.7(1.4)	121, 1.4(1.5)	1.87 (309)	.06
PHQ2≥3, n (%)				2.82 (1)	.09
Responses, n	311	190	121		
Yes	61 (19.6)	43 (22.6)	18 (14.9)		
No	250 (80.4)	147 (77.4)	103 (85.1)		
GAD2 score, N responses, mean (SD)	312, 1.8(1.7)	190, 1.9(1.7)	122, 1.6 (1.5)	1.81 (310)	.07
GAD2≥3, n (%)				0.37 (1)	.54
Responses, n	312	190	122		
Yes	80 (25.6)	51 (26.8)	29 (23.8)		

Characteristics	Baseline (n=315)	Had follow-up (n=193)	No follow-up (n=122)	<i>t</i> or χ^2 (df) ^a	<i>P</i> value
No	232 (74.4)	139 (73.2)	93 (76.2)		
Stigma score, N responses, mean (SD)	290, 2.8 (1.0)	181, 2.9 (1.0)	109, 2.7 (1.0)	0.96 (288)	.34
Engagement score (3 items) ^d , N responses, mean (SD)	304, 4.0 (0.7)	188, 4.0 (0.8)	116, 4.1 (0.7)	-2.03 (302)	.04
Do not feel comfortable using this website, n (%)				1.14 (4)	.89
Responses, n	305	188	117		
Strongly disagree	106 (34.8)	66 (35.1)	40 (34.2)		
Disagree	108 (35.4)	65 (34.6)	43 (36.8)		
Neither agree nor disagree	22 (7.2)	12 (6.4)	10 (8.5)		
Agree	37 (12.1)	25 (13.3)	12 (10.3)		
Strongly agree	32 (10.5)	20 (10.6)	12 (10.3)		
Any service use for emotional, mental health, alcohol, or drug problems, n (%)				0.44 (1)	.51
Responses, n	197	130	67		
Yes	121 (61.4)	82 (63.1)	39 (58.2)		
No	76 (38.6)	48 (36.9)	28 (41.8)		
Any service use for emotional, mental health, alcohol, or drug problems, imputed, n (%)				2.07 (1)	.15
Responses, n	240	152	88		
Yes	121 (50.4)	82 (53.9)	39 (44.3)		
No	119 (49.6)	70 (46.1)	49 (55.6)		
Total number of COVID-19-related changes, N responses, mean (SD)	315, 3.9 (2.0)	193, 4.0 (1.9)	122, 3.6 (2.1)	1.75 (313)	.08
Total number of COVID-19 stressors experienced, N responses, mean (SD)	315, 6.8 (4.1)	193, 7.3 (4.3)	122, 6.0 (3.8)	2.81 (313)	.005

^a χ^2 tests were used for categorical variables and *t* tests were used for continuous variables to compare groups with and without a follow-up response.

^bPHQ2: Patient Health Questionnaire-2 item.

^cGAD2: Generalized Anxiety Disorder-2 item scale.

^dItems (ease of use, relevance of topics, satisfaction) were averaged as mean engagement based on 5-point Likert scales.

Website Engagement

Of all participants, 87.1% agreed the website was easy to use, 80.5% found the topics to be relevant, 85.9% were satisfied with the website, and 70.2% were comfortable using the website. Three items (ease, relevance, satisfaction) were averaged to obtain the mean engagement score with standard Cronbach $\alpha=.736$; comfort using the website was considered a separate measure. In final regression analyses, predictors of higher baseline engagement included Hispanic versus other race/ethnicity ($\beta=.27$, 95% CI .10-.44; $P=.002$) and COVID-19-related behavior changes ($\beta=.09$, 95% CI .05-.13; $P<.001$). At baseline, 83.1% (260/313) of participants reported English as their preferred language for website use versus 12.8% (40/313) Spanish or 4.2% other (13/313) (Table 2). Predictors for comfort using the website in final regression analyses were preferring English compared to other languages (odds ratio [OR] 5.57, 95% CI 2.22-13.96; $P<.001$) and number of COVID-19-related behavior changes (OR 1.37, 95% CI 1.12-1.66; $P=.002$). In addition, those having overnight treatment for behavioral health within 6 months before baseline ($n=8$ inpatients, $n=10$ rehabilitation patients) had less comfort using the website (OR 0.15, 95% CI 0.03-0.69; $P=.02$).

Follow-up Response and Predictors

Of the 315 participants completing the baseline survey, 193 (61.3%) completed follow-up surveys. Table 2 shows baseline factors in bivariate analyses that predicted follow-up survey completion (significance of differences between groups with and without follow-up response were evaluated with χ^2 tests for categorical variables and with *t* tests for continuous variables). In final logistic regression models, baseline predictors of follow-up survey response were: (1) non-Hispanic white versus other race/ethnicity (OR 2.04, 95% CI 1.11-3.76; $P=.02$), (2) lower mean website engagement (OR 0.66, 95% CI 0.45-0.95; $P=.03$), and (3) greater number of COVID-19-related stressors (OR 1.07, 95% CI 1.01-1.14; $P=.03$). These variables were included as predictors in follow-up models and were excluded if not significant.

Follow-up Website Use

Among the follow-up participants, 119/193 (61.7%) visited the website or used resources in the month prior to follow-up surveys. By category, the rate of resource use was 50.0% (96/192) for (1) Learn about COVID, 53.9% (103/191) for (2) Soothe Anxiety and Stress, 40.0% (76/190) for (3) Support Resilience in Kids and Families, 35.6% (68/191) for (4) Cope

with a Recent Loss, 42.2% (81/192) for (5) Build Community and Connect with People, and 31.3% (60/192) for (6) Need to Talk to Someone? (Helplines). For any use in the month prior to follow-up, unique baseline predictors in the regressions were: PHQ2 score (higher=more use), called hotline for behavioral health before baseline (higher=more use), and greater alcohol use during COVID-19; these predictors were also significant when weighted for nonresponse (Table 3). The mean score of category use (1-5, from "no use" to "recommend to everyone") varied from a high of 2.30 (SD 1.45) for Soothe Anxiety and

Stress to a low of 1.75 (SD 1.26) for Need to Talk to Someone; mean scores were significantly higher for Categories 1-3 and 5 than for 6; categories 1 and 2 than for 4 and 5; and category 2 than category 3 (all $P \leq .002$). Predictors of higher total use summed across the six categories were: Hispanic ethnicity, baseline PHQ2 or GAD2 score ≥ 3 , and caring for someone at home during COVID-19 (Table 3). COVID-19-related change in sexual activity was associated with reduced total use; these predictors were also significant with nonresponse weighting (Table 3).

Table 3. Final models for website use in the month before follow-up.

Variables	Main analysis (unweighted)			Sensitivity analysis (IPW ^a , nonresponse)		
	Statistic ^b	95% CI	P value	Statistic	95% CI	P value
Any website use^c						
Female	0.52	0.25 to 1.08	.08	0.61	0.27 to 1.35	.22
PHQ2 ^d ≥ 3	1.33	1.04 to 1.69	.02	1.44	1.12 to 1.85	.004
Called hotline for behavioral health before baseline	6.50	1.68 to 25.24	.007	6.48	1.25 to 33.58	.03
COVID-19 stressor, increased alcohol/substance use	0.26	0.11 to 0.60	.002	0.21	0.09 to 0.52	<.001
Use of 6 categories^{e,f}						
Hispanic or Latino (vs other race/ethnicity)	2.74	0.75 to 4.73	.007	2.56	0.41 to 4.72	.02
PHQ-2 or GAD-2 ^g ≥ 3	3.39	1.39 to 5.40	.001	4.51	2.18 to 6.83	<.001
COVID-related change, caring for someone at home	2.86	0.77 to 4.95	.008	3.24	0.97 to 5.51	.006
COVID stressor, change in sexual activity	-3.58	-5.88 to -1.28	.002	-3.97	-6.04 to -1.91	<.001

^aIPW: inverse probability weighting for nonresponse predictors at baseline and follow-up.

^bThe effect is presented as the odds ratio for any website use and as β for use of 6 categories.

^cAnalytical N=186 for main analysis, N=181 for sensitivity analysis.

^dPHQ2: Patient Health Questionnaire-2 item.

^eAnalytical N=175 for main analysis, N=170 for sensitivity analysis.

^fTotal score across 6 categories (Learn about COVID, Soothe Anxiety and Stress, Supporting Resilience in Kids and Families, Cope with a Recent Loss, Build Community and Connect With People, and Need to Talk to Someone) each of which has 5 responses (not use to recommend to anyone).

^gGAD2: Generalized Anxiety Disorder scale-2 item.

Associations With Primary and Secondary Follow-up Outcomes

The main predictor of follow-up depression (PHQ2 ≥ 3) was baseline depression (Table 4). In addition, greater reported engagement in T4W/Juntos was associated with a lower likelihood of depression at follow-up (Table 4). PHQ2 scores can range from 0 to 6; for our sample, the baseline mean PHQ2 score was 1.6 (SD 1.5). Predictors of a post-pre change in the PHQ2 score included: (1) Caucasian/white versus other race/ethnicity, which was associated with an increase in PHQ2; (2) using the website resources in the month before follow-up, which was associated with a reduced post-pre PHQ2 score; and (3) having more COVID-19-related stressors at baseline, which was associated with a reduced post-pre PHQ2 score (Table 4). The findings were similar when weighted for nonresponse. For GAD2 ≥ 3 at follow-up, the main predictor was baseline GAD2. For a post-pre change in the GAD2 score, the main predictors were non-Hispanic white versus other race/ethnicity, age, and

COVID-19-related stressors; however, only COVID-19-related stressors and age were significant in the analysis weighting for nonresponse (Table 4).

Predictors of using hotlines for behavioral health in the month before follow-up included: baseline depression or anxiety (PHQ2 or GAD2 ≥ 3) ($P=.006$) and use of hotlines prior to baseline assessment ($P=.002$). Greater baseline mean website engagement, greater comfort using the website at baseline, and greater mean total website use in the month before follow-up each predicted reduced hotline use at follow-up, which remained significant with nonresponse weighting (Table 5). For use of any behavioral health services (outpatient, inpatient, rehabilitation, hotlines) prior to follow-up, the main predictor in logistic regression was use of such services prior to baseline. Increased reported mean engagement with the website at baseline was associated with a borderline trend toward reduced probability of behavioral services use, which was significant when weighting for nonresponse (Table 5).

Table 4. Final models for follow-up Patient Health Questionnaire (PHQ2) and Generalized Anxiety Disorder (GAD2) associated with website use.

Variables	Main analysis (unweighted)			Sensitivity analysis (IPW ^a , nonresponse)		
	Statistic ^b	95% CI	P value	Statistic	95% CI	P value
Follow-up PHQ2^c≥3						
Engagement mean score, 3 items ^d (baseline)	0.54	0.34 to 0.86	.01	0.55	0.33 to 0.91	.02
PHQ2 ^c ≥3 (baseline)	6.34	2.77 to 14.09	.04	6.53	2.78 to 15.84	<.001
Follow-up GAD2^c≥3						
GAD2 ^c ≥3 (baseline)	13.65	6.06 to 30.72	<.001	11.45	4.81 to 22.27	<.001
PHQ2 mean score post-pre change^f						
Non-Hispanic white (vs other race/ethnicity)	.46	.03 to .90	.04	.44	.00 to .87	.049
Visited T4W/Juntos ^g or used resources month before follow-up	-.62	-1.04 to -.20	.004	-.52	-.93 to -.19	.02
Total number of COVID-19 stressors (baseline)	-.07	-.12 to -.02	.004	-.08	-.13 to -.02	.005
GAD2 mean score post-pre change^h						
Non-Hispanic white (vs other race/ethnicity)	.44	.00 to .87	.048	.30	-.14 to .74	.19
Age	.02	.00 to .03	.02	.02	.00 to .03	.049
Total number of COVID-19 stressors (baseline)	-.06	-.11 to -.01	.02	-.08	-.13 to -.03	.002

^aIPW: inverse probability weighting for nonresponse predictors at baseline and follow-up.

^bThe effect is presented as the odds ratio for follow-up scores ≥ 3 and as β for mean post-pre changes in scores.

^cAnalytical N=187 for main analysis, N=185 for sensitivity analysis.

^dItems (ease of use, relevance of topics, satisfaction) were averaged as mean engagement based on 5-point Likert scales.

^eAnalytical N=187 for main analysis, N=187 for sensitivity analysis.

^fAnalytical N=181 for main analysis, N=176 for sensitivity analysis.

^gT4W/Juntos: Together for Wellness/Juntos por Nuestro Bienestar.

^hAnalytical N=178 for main analysis, N=173 for sensitivity analysis.

Table 5. Final models for follow-up hotline and behavioral health service use.

Variables	Main analysis (unweighted)			Sensitivity analysis (IPW ^a , nonresponse)		
	OR ^b	95% CI	P value	OR	95% CI	P value
Use of hotlines for behavioral health in month before follow-up^c						
PHQ2 ^d or GAD2 ^e ≥3 (baseline)	7.89	1.79-34.77	.006	9.58	1.68-54.83	.01
Using hotline for behavioral health prior to baseline	10.44	2.34-46.47	.002	10.75	1.98-58.38	.006
Engagement mean score, 3 items ^f (baseline)	0.36	0.17-0.76	.007	0.36	0.15-0.84	.02
Comfort using website (baseline)	0.23	0.06-0.91	.04	0.18	0.04-0.91	.04
Total score of website category use in past month before follow-up ^g	1.15	1.03-1.28	.01	1.15	1.07-1.23	<.001
Use of any behavioral health services in past month before follow-up^h						
Visited any behavioral health provider in 6 months prior to baseline	14.78	5.57-39.25	<.001	15.02	5.06-44.61	<.001
Engagement mean score, 3 item (baseline) ^f	0.61	0.35-1.07	.08	0.59	0.35-1.00	.05

^aIPW: inverse probability weighting for nonresponse predictors at baseline and follow-up.

^bOR: odds ratio.

^cAnalytical N=174 for main analysis, N=172 for sensitivity analysis.

^dPHQ2: Patient Health Questionnaire-2 item.

^eGAD2: Generalized Anxiety Disorder-2 item.

^f3 items (ease of use, relevance of topics, satisfaction) averaged as mean engagement; 5-point Likert scales for agreement.

^gTotal score across 6 categories (Learn about COVID, Soothe Anxiety and Stress, Supporting Resilience in Kids and Families, Cope with a Recent Loss, Build Community and Connect, and Need to Talk to Someone) each of which has 5 responses (not use to recommend to anyone).

^hAnalytical N=137 for main analysis, N=136 for sensitivity analysis.

Discussion

Principal Findings

This article presents an evaluation of engagement in and impact of free digital mental health resources developed with community advisor input to support well-being during COVID-19 in California. Informed by the TAM, Behavioral Health Model for Vulnerable Populations, COVID-19 stressors, and CPPR principles [21,25-27], we hypothesized that individuals with higher need (depression, anxiety) and more COVID-19-related behavior changes and stressors would engage more in website use, and that higher engagement and use would be associated with reduced depression, anxiety, and hotline use at follow-up. Findings on predictors of website use were somewhat consistent with these hypotheses, but the details differed for baseline and follow-up. For main outcomes, findings were consistent with hypotheses for depression and hotline use, in that greater engagement or use before follow-up was associated with lower follow-up depression and greater reduction in depression from baseline to follow-up, but not reduced anxiety, for the overall sample. The main findings were consistent in sensitivity analyses weighting for nonresponse, and a borderline significant trend for baseline website engagement was significantly associated with reduction in any behavioral health service use when weighting for nonresponse.

These findings suggest that recruiting participants through community agencies can generate a diverse sample in race/ethnicity, age, gender identity, and sexual orientation [1,6-8]. Baseline completers were somewhat younger than noncompleters, suggesting that more support for participation

may be needed for older adults. Baseline website engagement was higher for those of Hispanic ethnicity, whereas comfort in website use was greater for people preferring English, suggesting that website modification or human support may be important for non-English speakers [18]. In addition, less comfort in use was associated with having a prior overnight behavioral health stay, which could reflect disability or need for support. Website use before follow-up was lower for those with greater alcohol use during the COVID-19 pandemic. Total website use across categories was higher for those caring for someone at home and was lower for those with a change in sexual activity during the pandemic. These findings highlight potential differences in website use patterns for specific COVID-19 changes, which is an area for future research. The main findings for depression and hotline use are largely consistent with policymaker goals for developing the website for prevention to reduce mental health needs and crisis calls.

Limitations

Limitations included that, as a nonrandomized study, the findings could reflect selection effects or reverse causality, although the main finding of website use associated with reduced depression was robust in analyses of end status and post-pre change with and without nonresponse weighting. Higher use of this website could be associated with use of other websites that could have had an unmeasured (confounding) effect on outcomes. Further, the study used a convenience sample recruited by partnering agencies in one state representing diverse populations rather than a general population sample during a specific period of the pandemic (September 2021-May 2022). While effects of COVID-19 evolved over time, during this

period, the date of survey completion was not significant and not included in final models. Future research may include experimental designs to clarify the causation and mechanisms of action. Further, the study had a moderate response at follow-up (61.3%). Predictors of response (non-Hispanic white, lower baseline engagement, greater number of COVID-19 stressors) were included as predictors in final models if significant, with weighting for nonresponse in sensitivity analyses. Recruited participants included clients, partners, and providers of partnering agencies, and we did not track participant roles; however, baseline education data suggest the sample included some nonproviders. In future research, it may be important to include demographic or cultural differences in stigma of mental health that may affect website use to inform the tailoring of resources, and it may be important to evaluate the added value of human support to enhance website use and survey response, particularly for non-English speakers or older adults [18].

Public Health Implications

Given the consistency of main outcome findings with hypotheses and policy consumer advisor goals for developing digital

resources, the findings were considered by policy partners as encouraging for further website development and evaluation as next steps, including expansion to youth/young adults. Descriptive findings reinforce the importance of community agency partnerships to achieve a diverse sample. These findings also raise the issue of potential impacts for diverse populations of prevention-oriented, free digital mental health resources during COVID-19 to reduce depression and crisis hotline use, an important issue for future randomized trials or quality improvement initiatives with public health implications. The consumer advisor engagement in this study reinforces the public health value of a partnered participatory effort in intervention development and evaluation. The findings suggest next steps for research in public health, such as exploring human support for digital resources, especially for non-English speakers, even for a website available in 13 languages. Although this represents only a preliminary evaluation, given limited data on digital mental health for preventive public health goals [10], the findings may inform the next steps for development of resources and evaluation efforts to inform public mental health prevention.

Acknowledgments

The members of the Together for Wellness/Juntos Collaborators and Writing Group are: Sergio Aguilar-Gaxiola MD PhD, (UC Davis Center for Reducing Health Disparities) lead for outreach to Latinx and farmworker communities; Patricia Lester MD, Director, UCLA Nathanson Family Resilience Center, Co-investigator for development of children and youth resources; Minhxuan Tran (Chorus Innovations), staff coordinator for technology and stakeholder input in website development. Community partner leads for website and evaluation design input, participant recruitment and input on analysis findings: Melissa Hannah MS (United Parents), Senait Admassu (African Communities Public Health Coalition), Leslie Grace Sulite Xu (Boat People SOS), Tiffany Carter MS (Cal Voices), Dannie Cesena MPH (CA LGBTQ Health and Human Services Network), Guadalupe Rodriguez (Health Education Council), Ragini Lai (at the time of collaboration, with NAMI California), Benita Ramsey MA, JD (Rainbow Pride Youth Alliance), Gerry Balcazar and Maria Lemus (Vision Y Compromiso), Elizabeth Moore MD (initially, UCLA Clinical Scholars Program; currently University of Pennsylvania Department of Psychiatry/VA), informing evaluation survey measures. This study received funding from the G-31202 California Health Care Foundation 9/1/2020-8/31/202, Role: PI (Wells) Evaluation of Digital Mental Health (for COVID-19 Response).

Authors' Contributions

KW wrote the manuscript. DFR provided the data for Table 1. LZ performed the statistical analyses. LZ and LT provided the data for Tables 2-5. ADT, ASY, and MSVH provided edits to the manuscript. DFR, AO, and FJ facilitated community input into the design, measures, and manuscript. TE and MB led youth stakeholder engagement and developed hypotheses. AA led technology services for website resources and data collection for surveys, along with the associated write-up. All authors reviewed the final manuscript.

Conflicts of Interest

AA is founder and CEO of Chorus Innovations and Arevian Technologies. The other authors have no conflicts to declare.

References

1. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA* 2020 Jun 23;323(24):2466-2467 [FREE Full text] [doi: [10.1001/jama.2020.8598](https://doi.org/10.1001/jama.2020.8598)] [Medline: [32391864](https://pubmed.ncbi.nlm.nih.gov/32391864/)]
2. Galea S, Merchant RM, Lurie N. The mental health consequences of COVID-19 and physical distancing: the need for prevention and early intervention. *JAMA Intern Med* 2020 Jun 01;180(6):817-818. [doi: [10.1001/jamainternmed.2020.1562](https://doi.org/10.1001/jamainternmed.2020.1562)] [Medline: [32275292](https://pubmed.ncbi.nlm.nih.gov/32275292/)]
3. Feist JB, Feist JC, Cipriano P. Stigma compounds the consequences of clinician burnout during COVID-19: a call to action to break the culture of silence. *NAM Perspect* 2020 Aug 6;2020:10.31478/202008b [FREE Full text] [doi: [10.31478/202008b](https://doi.org/10.31478/202008b)] [Medline: [35291748](https://pubmed.ncbi.nlm.nih.gov/35291748/)]

4. Salerno JP, Williams ND, Gattamorta KA. LGBTQ populations: psychologically vulnerable communities in the COVID-19 pandemic. *Psychol Trauma* 2020 Aug;12(S1):S239-S242 [[FREE Full text](#)] [doi: [10.1037/tra0000837](https://doi.org/10.1037/tra0000837)] [Medline: [32551761](https://pubmed.ncbi.nlm.nih.gov/32551761/)]
5. Logan RI, Castañeda H. Addressing health disparities in the rural United States: advocacy as caregiving among community health workers and Promotores de Salud. *Int J Environ Res Public Health* 2020 Dec 10;17(24):9223 [[FREE Full text](#)] [doi: [10.3390/ijerph17249223](https://doi.org/10.3390/ijerph17249223)] [Medline: [33321718](https://pubmed.ncbi.nlm.nih.gov/33321718/)]
6. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 14;69(32):1049-1057. [doi: [10.15585/mmwr.mm6932a1](https://doi.org/10.15585/mmwr.mm6932a1)] [Medline: [32790653](https://pubmed.ncbi.nlm.nih.gov/32790653/)]
7. Mayne SL, Hannan C, Davis M, Young JF, Kelly MK, Powell M, et al. COVID-19 and adolescent depression and suicide risk screening outcomes. *Pediatrics* 2021 Sep;148(3):e2021051507. [doi: [10.1542/peds.2021-051507](https://doi.org/10.1542/peds.2021-051507)] [Medline: [34140393](https://pubmed.ncbi.nlm.nih.gov/34140393/)]
8. Krause KH, Verlenden JV, Szucs LE, Swedo EA, Merlo CL, Niolon PH, et al. Disruptions to school and home life among high School Students During the COVID-19 pandemic - Adolescent Behaviors and Experiences Survey, United States, January-June 2021. *MMWR Suppl* 2022 Apr 01;71(3):28-34 [[FREE Full text](#)] [doi: [10.15585/mmwr.su7103a5](https://doi.org/10.15585/mmwr.su7103a5)] [Medline: [35358164](https://pubmed.ncbi.nlm.nih.gov/35358164/)]
9. Piper S, Davenport TA, LaMonica H, Ottavio A, Iorfino F, Cheng VWS, et al. Implementing a digital health model of care in Australian youth mental health services: protocol for impact evaluation. *BMC Health Serv Res* 2021 May 12;21(1):452 [[FREE Full text](#)] [doi: [10.1186/s12913-021-06394-4](https://doi.org/10.1186/s12913-021-06394-4)] [Medline: [33980229](https://pubmed.ncbi.nlm.nih.gov/33980229/)]
10. Taylor CB, Ruzek JI, Fitzsimmons-Craft EE, Sadeh-Sharvit S, Topooco N, Weissman RS, et al. Using digital technology to reduce the prevalence of mental health disorders in populations: time for a new approach. *J Med Internet Res* 2020 Jul 24;22(7):e17493 [[FREE Full text](#)] [doi: [10.2196/17493](https://doi.org/10.2196/17493)] [Medline: [32706665](https://pubmed.ncbi.nlm.nih.gov/32706665/)]
11. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLoS One* 2016;11(5):e0154248 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0154248](https://doi.org/10.1371/journal.pone.0154248)] [Medline: [27135410](https://pubmed.ncbi.nlm.nih.gov/27135410/)]
12. Figueroa CA, Aguilera A. The need for a mental health technology revolution in the COVID-19 pandemic. *Front Psychiatry* 2020 Jun 3;11:523. [doi: [10.3389/fpsy.2020.00523](https://doi.org/10.3389/fpsy.2020.00523)] [Medline: [32581891](https://pubmed.ncbi.nlm.nih.gov/32581891/)]
13. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. *Curr Treat Options Psych* 2019 Jul 5;6(3):243-255. [doi: [10.1007/s40501-019-00181-z](https://doi.org/10.1007/s40501-019-00181-z)]
14. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. *J Med Internet Res* 2017 Jan 05;19(1):e10 [[FREE Full text](#)] [doi: [10.2196/jmir.6645](https://doi.org/10.2196/jmir.6645)] [Medline: [28057609](https://pubmed.ncbi.nlm.nih.gov/28057609/)]
15. Burns MN, Montague E, Mohr DC. Initial design of culturally informed behavioral intervention technologies: developing an mHealth intervention for young sexual minority men with generalized anxiety disorder and major depression. *J Med Internet Res* 2013 Dec 05;15(12):e271 [[FREE Full text](#)] [doi: [10.2196/jmir.2826](https://doi.org/10.2196/jmir.2826)] [Medline: [24311444](https://pubmed.ncbi.nlm.nih.gov/24311444/)]
16. Arevian AC, O'Hora J, Jones F, Mango J, Jones L, Williams PG, et al. Participatory technology development to enhance community resilience. *Ethn Dis* 2018;28(Suppl 2):493-502 [[FREE Full text](#)] [doi: [10.18865/ed.28.S2.493](https://doi.org/10.18865/ed.28.S2.493)] [Medline: [30202203](https://pubmed.ncbi.nlm.nih.gov/30202203/)]
17. Fleming T, Merry S, Stasiak K, Hopkins S, Patolo T, Ruru S, et al. The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. *JMIR Ment Health* 2019 May 08;6(5):e12656 [[FREE Full text](#)] [doi: [10.2196/12656](https://doi.org/10.2196/12656)] [Medline: [31066705](https://pubmed.ncbi.nlm.nih.gov/31066705/)]
18. Renfrew ME, Morton DP, Morton JK, Przybylko G. The influence of human support on the effectiveness of digital mental health promotion interventions for the general population. *Front Psychol* 2021 Aug 19;12:716106. [doi: [10.3389/fpsyg.2021.716106](https://doi.org/10.3389/fpsyg.2021.716106)] [Medline: [34489818](https://pubmed.ncbi.nlm.nih.gov/34489818/)]
19. Mohr DC, Riper H, Schueller SM. A solution-focused research approach to achieve an implementable revolution in digital mental health. *JAMA Psychiatry* 2018 Feb 01;75(2):113-114. [doi: [10.1001/jamapsychiatry.2017.3838](https://doi.org/10.1001/jamapsychiatry.2017.3838)] [Medline: [29238805](https://pubmed.ncbi.nlm.nih.gov/29238805/)]
20. Graham AK, Lattie EG, Mohr DC. Experimental therapeutics for digital mental health. *JAMA Psychiatry* 2019 Dec 01;76(12):1223-1224 [[FREE Full text](#)] [doi: [10.1001/jamapsychiatry.2019.2075](https://doi.org/10.1001/jamapsychiatry.2019.2075)] [Medline: [31433448](https://pubmed.ncbi.nlm.nih.gov/31433448/)]
21. Jones L, Wells K. Strategies for academic and clinician engagement in community-participatory partnered research. *JAMA* 2007 Jan 24;297(4):407-410. [doi: [10.1001/jama.297.4.407](https://doi.org/10.1001/jama.297.4.407)] [Medline: [17244838](https://pubmed.ncbi.nlm.nih.gov/17244838/)]
22. Arevian AC, Jones F, Tang L, Sherbourne CD, Jones L, Miranda J, Community Partners in Care Writing Group. Depression remission from community coalitions versus individual program support for services: findings from Community Partners in Care, Los Angeles, California, 2010-2016. *Am J Public Health* 2019 Jun;109(S3):S205-S213. [doi: [10.2105/AJPH.2019.305082](https://doi.org/10.2105/AJPH.2019.305082)] [Medline: [31242001](https://pubmed.ncbi.nlm.nih.gov/31242001/)]
23. Palinkas LA, Springgate BF, Sugarman OK, Hancock J, Wennerstrom A, Haywood C, et al. A rapid assessment of disaster preparedness needs and resources during the COVID-19 pandemic. *Int J Environ Res Public Health* 2021 Jan 07;18(2):425 [[FREE Full text](#)] [doi: [10.3390/ijerph18020425](https://doi.org/10.3390/ijerph18020425)] [Medline: [33430355](https://pubmed.ncbi.nlm.nih.gov/33430355/)]
24. Goodsmith N, Moore EM, Siddiq H, Barceló SE, Ulloa-Flores E, Loera G, et al. Community-partnered development of a digital mental health resource website to support diverse communities during the COVID-19 pandemic. *J Health Care Poor Underserved* 2022;33(1):506-516. [doi: [10.1353/hpu.2022.0039](https://doi.org/10.1353/hpu.2022.0039)] [Medline: [35153238](https://pubmed.ncbi.nlm.nih.gov/35153238/)]

25. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag Sci* 2000 Feb;46(2):186-204. [doi: [10.1287/mnsc.46.2.186.11926](https://doi.org/10.1287/mnsc.46.2.186.11926)]
26. Gelberg L, Andersen RM, Leake BD. The Behavioral Model for Vulnerable Populations: application to medical care use and outcomes for homeless people. *Health Serv Res* 2000 Feb;34(6):1273-1302 [[FREE Full text](#)] [Medline: [10654830](https://pubmed.ncbi.nlm.nih.gov/10654830/)]
27. Harkness A, Behar-Zusman V, Safren SA. Understanding the impact of COVID-19 on Latino sexual minority men in a US HIV hot spot. *AIDS Behav* 2020 Jul 16;24(7):2017-2023 [[FREE Full text](#)] [doi: [10.1007/s10461-020-02862-w](https://doi.org/10.1007/s10461-020-02862-w)] [Medline: [32300989](https://pubmed.ncbi.nlm.nih.gov/32300989/)]
28. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: a pilot randomised controlled trial. *PLoS One* 2018 Dec 31;13(12):e0209482 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0209482](https://doi.org/10.1371/journal.pone.0209482)] [Medline: [30596696](https://pubmed.ncbi.nlm.nih.gov/30596696/)]
29. Galla BM, O'Reilly GA, Kitil MJ, Smalley SL, Black DS. Community-based mindfulness program for disease prevention and health promotion: targeting stress reduction. *Am J Health Promot* 2015;30(1):36-41. [doi: [10.4278/ajhp.131107-QUAN-567](https://doi.org/10.4278/ajhp.131107-QUAN-567)] [Medline: [25162319](https://pubmed.ncbi.nlm.nih.gov/25162319/)]
30. Lopez-Maya E, Olmstead R, Irwin MR. Mindfulness meditation and improvement in depressive symptoms among Spanish- and English speaking adults: A randomized, controlled, comparative efficacy trial. *PLoS One* 2019 Jul 5;14(7):e0219425 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0219425](https://doi.org/10.1371/journal.pone.0219425)] [Medline: [31276540](https://pubmed.ncbi.nlm.nih.gov/31276540/)]
31. Arevian AC, O'Houra J, Jones F, Mango J, Jones L, Williams PG, et al. Participatory technology development to enhance community resilience. *Ethn Dis* 2018;28(Suppl 2):493-502 [[FREE Full text](#)] [doi: [10.18865/ed.28.S2.493](https://doi.org/10.18865/ed.28.S2.493)] [Medline: [30202203](https://pubmed.ncbi.nlm.nih.gov/30202203/)]
32. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007 Mar 06;146(5):317-325. [doi: [10.7326/0003-4819-146-5-200703060-00004](https://doi.org/10.7326/0003-4819-146-5-200703060-00004)] [Medline: [17339617](https://pubmed.ncbi.nlm.nih.gov/17339617/)]
33. Kroenke K, Spitzer RL, Williams JBW. 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](https://doi.org/10.1097/01.MLR.0000093487.78664.3C)] [Medline: [14583691](https://pubmed.ncbi.nlm.nih.gov/14583691/)]
34. National Alliance on Mental Illness California. 2020. URL: <https://namica.org/blog/2019-2020-stakeholder-surveys-for-diverse-communities-families-%EF%BB%BF/> [accessed 2021-06-02]
35. Elwood WN, Irvin VL, Liu B, Lee R, Breen N. Health-related influences of extending marital benefits to same-sex couples: results from the California Health Interview Survey. *Fam Relat* 2020 Oct 10;69(5):934-943. [doi: [10.1111/fare.12507](https://doi.org/10.1111/fare.12507)]
36. Jans M, Viana J, Grant D, Cochran SD, Lee AC, Ponce NA. Trends in sexual orientation missing data over a decade of the California Health Interview Survey. *Am J Public Health* 2015 May;105(5):e43-e50. [doi: [10.2105/AJPH.2014.302514](https://doi.org/10.2105/AJPH.2014.302514)] [Medline: [25790399](https://pubmed.ncbi.nlm.nih.gov/25790399/)]
37. Tambling RR, Russell BS, Park CL, Fendrich M, Hutchinson M, Horton AL, et al. Measuring cumulative stressfulness: psychometric properties of the COVID-19 Stressors Scale. *Health Educ Behav* 2021 Feb 11;48(1):20-28. [doi: [10.1177/1090198120979912](https://doi.org/10.1177/1090198120979912)] [Medline: [33307818](https://pubmed.ncbi.nlm.nih.gov/33307818/)]
38. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res* 2013 Jun;22(3):278-295. [doi: [10.1177/0962280210395740](https://doi.org/10.1177/0962280210395740)] [Medline: [21220355](https://pubmed.ncbi.nlm.nih.gov/21220355/)]
39. Li L, Shen C, Li X, Robins JM. On weighting approaches for missing data. *Stat Methods Med Res* 2013 Feb;22(1):14-30 [[FREE Full text](#)] [doi: [10.1177/0962280211403597](https://doi.org/10.1177/0962280211403597)] [Medline: [21705435](https://pubmed.ncbi.nlm.nih.gov/21705435/)]

Abbreviations

- CBT:** cognitive behavioral therapy
- CPPR:** community-partnered participatory research
- FEMA:** Federal Emergency Management Agency
- GAD2:** Generalized Anxiety Disorder-2 item scale.
- LGBTQ+:** lesbian, gay, bisexual, transgender, queer/questioning and others
- OR:** odds ratio
- PHQ2:** Patient Health Questionnaire-2 item
- SAMHSA:** Substance Abuse and Mental Health Services Administration (
- T4W/Juntos:** Together for Wellness/Juntos por Nuestro Bienestar
- TAM:** technology acceptance model

Edited by A Mavragani; submitted 22.08.22; peer-reviewed by M Stoner, M Kapsetaki; comments to author 23.09.22; revised version received 04.11.22; accepted 04.11.22; published 07.12.22.

Please cite as:

Wells K, Thames AD, Young AS, Zhang L, Heilemann MV, Romero DF, Oliva A, Jones F, Tang L, Brymer M, Elliott T, Arevian A, Together for Wellness/Juntos Collaborators and Writing Group

Engagement, Use, and Impact of Digital Mental Health Resources for Diverse Populations in COVID-19: Community-Partnered Evaluation

JMIR Form Res 2022;6(12):e42031

URL: <https://formative.jmir.org/2022/12/e42031>

doi: [10.2196/42031](https://doi.org/10.2196/42031)

PMID: [36346902](https://pubmed.ncbi.nlm.nih.gov/36346902/)

©Kenneth Wells, April Denise Thames, Alexander S Young, Lily Zhang, MarySue V Heilemann, Daniela Flores Romero, Adrian Oliva, Felica Jones, Lingqi Tang, Melissa Brymer, Thomas Elliott, Armen Arevian, Together for Wellness/Juntos Collaborators and Writing Group. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Continued Use of Contact-Tracing Apps in the United States and the United Kingdom: Insights From a Comparative Study Through the Lens of the Health Belief Model

Zhan Zhang^{1*}, PhD; Isaac Vaghefi^{2*}, PhD

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

²Zicklin School of Business, Baruch College, City University of New York, New York, NY, United States

* all authors contributed equally

Corresponding Author:

Isaac Vaghefi, PhD

Zicklin School of Business

Baruch College

City University of New York

55 Lexington Ave,

New York, NY, 10010

United States

Phone: 1 (646) 312 3409

Email: isaac.vaghefi@baruch.cuny.edu

Abstract

Background: To contain the spread of SARS-CoV-2, contact-tracing (CT) mobile apps were developed and deployed to identify and notify individuals who have exposure to the virus. However, the effectiveness of these apps depends not only on their adoption by the general population but also on their continued use in the long term. Limited research has investigated the facilitators of and barriers to the continued use of CT apps.

Objective: In this study, we aimed to examine factors influencing the continued use intentions of CT apps based on the health belief model. In addition, we investigated the differences between users and nonusers and between the US and UK populations.

Methods: We administered a survey in the United States and the United Kingdom. Respondents included individuals who had previously used CT technologies and those without experience. We used the structural equation modeling technique to validate the proposed research model and hypotheses.

Results: Analysis of data collected from 362 individuals showed that perceived benefits, self-efficacy, perceived severity, perceived susceptibility, and cues to action positively predicted the continued use intentions of CT apps, while perceived barriers could reduce them. We observed few differences between the US and UK groups; the only exception was the effect of COVID-19 threat susceptibility, which was significant for the UK group but not for the US group. Finally, we found that the only significant difference between users and nonusers was related to perceived barriers, which may not influence nonusers' continued use intentions but significantly reduce experienced users' intentions.

Conclusions: Our findings have implications for technological design and policy. These insights can potentially help governments, technology companies, and media outlets to create strategies and policies to promote app adoption for new users and sustain continued use for existing users in the long run.

(*JMIR Form Res* 2022;6(12):e40302) doi:[10.2196/40302](https://doi.org/10.2196/40302)

KEYWORDS

contact tracing; app adoption; app continued use; public attitudes; health belief model; COVID-19

Introduction

Overview

The COVID-19 outbreak, a consequence of SARS-CoV-2, has had a major, long-lasting effect on our society. The COVID-19 pandemic has been considered the most significant public health threat the world has experienced in the last 100 years [1]. In the United States alone, there have been over 80 million reported cases of the disease, with over a million fatal cases [2]. The COVID-19 pandemic has lasted for over 2 years and several new, high-transmissibility variants (eg, delta and omicron) have emerged during this time. Various efforts (eg, the distribution of multiple vaccines) and containment measures have been taken worldwide, including social distancing, travel restrictions, testing, and contact tracing (CT). Among these measures, digital CT apps have been considered a particularly important strategy for managing the pandemic and reducing COVID-19 cases and deaths [3,4]. These apps can rapidly identify and notify exposures to the disease [5,6].

Although traditionally performed manually [7], a set of CT apps have been developed and deployed by different countries (eg, China, the United States, the United Kingdom, Japan, Israel, and Singapore), health providers (eg, Mayo Clinic), and technology companies (eg, Apple and Google). Regardless of the differences in app design and architecture, prior research has indicated that the effectiveness of these apps is proportional to the number of people who use them [8]; that is, to suppress the epidemic, 80% of all smartphone users or 56% of the population overall need to use the CT app [3,9]. A survey conducted early in the pandemic reported that $\geq 60\%$ of the population indicated a willingness to install a hypothetical CT app [10]. Yet, the actual adoption of these apps has been considerably lower than anticipated [11], with installation rates of $\leq 10\%$ of the population in some countries where apps have been deployed [12]. This fact highlights the intention-behavior gap [13]—even though many people may have displayed the intention to use a CT app—they do not take action by installing or using it. Such low adoption rates are problematic and prevent the actual value of CT apps from being realized.

According to the World Health Organization, several other diseases could likely cause pandemics in the future, which signifies the importance of preparing for such occasions, even after the current pandemic is over [14]. Therefore, as an effective approach to containing pandemics, there is a need to understand why CT apps have not been as popular as anticipated and what factors facilitate or hinder their continued use as pandemics go on. Investigating this problem is critical not only for addressing the COVID-19 pandemic but also for understanding how to design CT technologies and apps for future health crises.

Ample research has been conducted in the early phase of the COVID-19 pandemic to investigate factors that could influence the adoption and uptake of CT apps [15]. For instance, prior work pointed to the influence of demographics (eg, age and sex) [16-18], individual beliefs and attitudes (eg, trust, privacy concerns, or access to technologies) [10,19,20], situational factors (eg, COVID-19 cases and deaths or lockdown measures) [21], and contextual factors (eg, cultural, regional, and national

differences) [15]. Despite the important contributions made by these studies, as Jamieson et al [22] stated, “the collective utility of contact tracing technology to suppress the spread of viruses depends not only on the adoption of contact tracing apps but also on their continued use.” In addition, compared with the beginning of the pandemic, our society has gained more knowledge about COVID-19 and a significant portion of the population has been vaccinated, all of which could affect people’s willingness to continue the use of CT apps. More importantly, recent evidence in health informatics shows that sustained and continued use of CT apps may have different motivators than the initial adoption of these technologies [23,24]. To our knowledge, *there is limited research that pays scholarly attention to the continued use of CT apps* [25].

Building upon prior work in this domain, we aim to address this research gap by proposing and validating a research model of the predictors of continued use of CT apps, defined as extending the use of CT apps beyond the initial stages of adoption and the first few uses. Our specific research questions include the following: (1) What are the predictors of an individual’s continued intention to use CT apps? (2) How do user behaviors and predictors differ among users who had experience versus those who did not and between individuals in the United States versus the United Kingdom? To answer these research questions, we surveyed respondents in the United States and the United Kingdom from among those who had previously used CT technologies and those without any experience. Our findings highlighted several factors (eg, perceived benefits, self-efficacy, perceived severity, perceived susceptibility, and cues to action) that have positive impacts on the continued use of CT apps. In contrast, perceived barriers could reduce people’s continued use intentions. We also observed some differences between the United States and the United Kingdom and between users and nonusers. Our study could provide important insights for governments, technology companies, and media outlets to determine how to promote CT apps better and sustain continued use in the long run.

Background

Since the outbreak of the COVID-19 pandemic, seminal research has studied user acceptance and use intention of CT apps based on several theoretical models, such as the health belief model (HBM) [26], the protection motivation theory [27], health behavior change [28], cognitive appraisal theory [29], procedural fairness theory and cultural dimension theory [19], and the unified theory of acceptance and use of technology [30-32]. These studies primarily focused on evaluating the public’s attitude; for example, individuals’ willingness or intention to install and adopt a hypothetical CT app.

This body of prior work highlighted a set of factors that influence the adoption and use intention of CT apps, ranging from individual characteristics (eg, age, sex, and experience with technology) to situational (eg, lockdown measures) and contextual factors (eg, cultural and national differences). For example, younger ages [10,16,17], higher education level [18,26,33], and experience using smartphone apps [10,34] are associated with positive intentions to download a CT app. In addition, the perceptions, trust, and acceptance of CT

technologies seem to vary in different countries and cultural backgrounds. For example, countries with collectivist values, such as those in Asian regions, usually see broad acceptance from their citizens [10,16]; in contrast, the US and European respondents generally report lower acceptance [10,17,18,35]. Situational factors, such as the total number of COVID-19 cases in a region and lockdown measures, could also affect people's willingness to download the CT app [21,25]. For example, those living in places with lockdown measures or restrictions on mobility are less likely to download and use the app [25].

In summary, prior work has primarily investigated the acceptance and intention to adopt and use CT technology. This is because most of the reviewed studies were conducted before or shortly after the introduction of CT apps; therefore, they were only able to measure people's willingness or intention to use a newly developed CT app. In addition, although some studies examined the actual use of CT apps with users [17,20,36,37], they only explored adoption at an early stage. *Thus, continued use remains to be investigated* [22,25]. To that end, in this study, we examined the continued use of CT apps beyond the initial stages of adoption and the first few interactions. We included both prior app users and nonusers in our study to investigate the factors that impact a person's continuous intention to use CT apps and how the influence of key predictors may differ between the user and nonuser groups.

Theoretical Framework and Hypotheses Development

In this work, we examine which factors influence an individual's *continued use* of a CT app by adopting the HBM [38,39]. We chose this theoretical model because the HBM is a widely recognized model in the context of health behavior change [38] and has been used by many prior studies to explain why people follow healthy choices of action in the presence of a threat. This is comparable with our study's context, where individuals may continue to use a CT app as a behavior to help counteract the threat of contracting and spreading COVID-19.

The HBM consists of several constructs including perceived susceptibility, perceived benefits, perceived barriers, perceived severity, and cues to action. The model assumes that people who anticipate a health threat are more willing to perform a protective health behavior because they believe that such an action will reduce a severe illness [40,41]. The 2 constructs of the HBM—perceived susceptibility and perceived severity—are highly related to this cognitive presumption. *Perceived susceptibility* represents an individual's perceived risk or likelihood of catching a disease due to a particular behavior. *Perceived severity* refers to individuals' beliefs about the impact of the harm of pursuing a particular behavior. Several studies have revealed that perceived susceptibility and perceived severity can cause people to take protective actions, such as using and adopting mobile health technologies [42,43]. In relation to the current COVID-19 pandemic, if people perceive themselves as liable to COVID-19 infection and related health complications, they are more inclined to continue using the app to reduce the infection risk of COVID-19. In addition, someone who perceives high health threats and severity tends to continue using the CT app. Therefore, we hypothesized the following:

Hypothesis 1: The perceived susceptibility to COVID-19 is positively associated with the continued use intention of CT apps.

Hypothesis 2: The perceived threat of COVID-19 is positively associated with the continued use intention of CT apps.

Perceived benefits refer to a person's belief that recommended health behaviors will be beneficial in preventing the disease or reducing its effect. A high perception of benefits increases the likelihood of adopting such behavior. *Perceived barriers*, on the contrary, represent the costs of or obstacles to performing the recommended health behavior, including tangible costs (eg, time, money, and knowledge acquisition) and psychological costs (eg, feeling anxious, pessimistic, and embarrassed) [44]. Low perception of barriers increases a person's willingness to adopt a particular behavior. The model assumes that the more benefits the individual believes there are from a new behavior, and the lower the obstacles to performing such behavior, the greater the chance of adopting it [45]. In our study context, perceived benefits could include personal benefits (eg, being formed of a potential infection to protect a person's health) and social benefits (eg, helping the community contain the spread of the coronavirus). Perceived barriers, as pointed out in prior works [46,47], include privacy issues and security concerns raised by the app. These concerns could present barriers to using the app in the long term. Therefore, we hypothesized as follows:

Hypothesis 3: The perceived benefits of CT apps are positively associated with the continued use intention of CT apps.

Hypothesis 4: The perceived barriers of CT apps are negatively associated with the continued use intention of CT apps.

The HBM is often complemented by constructs and factors related to health and protective behavior [48]. Therefore, we added perceived self-efficacy to the model, as we believe that the continued use of CT apps is also influenced by individuals' beliefs in their competence to use the app. *Perceived self-efficacy* is an individual's belief that he or she can successfully perform a particular health behavior. A few studies have demonstrated that self-efficacy is a significant factor in predicting health behaviors [49,50]. In addition, research has addressed the role of self-efficacy in predicting users' intention to continue using information technology systems [51]. In the context of COVID-19, if individuals have found mastery of the app, they may be more inclined to adopt the app [19,26,27]. However, whether having the ability to use CT apps could predict continued intention is understudied. Therefore, we hypothesized the following:

Hypothesis 5: Perceived self-efficacy is positively associated with the continued use intention of CT apps.

Cues to action are the circumstances that inspire the readiness to act. This construct can influence individuals' decisions on whether to engage in protective behaviors. Concerning COVID-19, cues to action include exposure to media content and the infection experience of close friends and family members. It has been 2 years since the outbreak of the pandemic

when this study was performed. People, especially existing users of CT apps, have had a variety of ways to get to know about COVID-19 and experience CT technology. Therefore, they may be more inclined to continue using the app. Our hypothesis was as follows:

Hypothesis 6: Cues to action are positively associated with the continued use intention of CT apps.

Textbox 1. Overview of contextualized constructs according to the health belief model.

- Perceived benefits of contact-tracing (CT) app use: perceptions about the positive outcomes of using CT apps.
- Perceived barriers of CT app use: perceptions about the negative outcomes or obstacles of using CT apps.
- COVID-19 threat severity: the extent to which one's health might be negatively affected by the COVID-19 pandemic.
- COVID-19 threat susceptibility: the extent to which one feels vulnerable to contracting COVID-19.
- Self-efficacy to use CT apps: belief in having the resources, skills, and ability to continue using CT apps.
- Cue to action (using CT apps): cues that trigger the use of CT apps.
- (CT apps) Continued use intention: willingness to continue the use of a mobile health app [52].

Methods

Data Collection

Pursuant to our research questions, we conducted a survey study among individuals in the United States and the United Kingdom to evaluate their intentions to continue using CT apps. We conducted the surveys in fall 2021 (October to December). To be open in our data collection and insights, we surveyed individuals who previously used CT technologies, in addition to those who have had no prior experience. This helped us gauge responses from both users and nonusers to examine what would generally determine decisions for long-term engagement with CT apps.

Individuals were recruited with the help of 2 survey companies: Amazon Mechanical Turk in the United States and Prolific in the United Kingdom. We used a web-based survey (Qualtrics) to collect the responses. No identifying information was collected to ensure anonymity. Nonetheless, we collected demographic information as shown in [Table 1](#). To calculate the minimum sample size necessary for variance-based structural equation model analysis, we followed the study by Hair et al [53], which suggests a sufficient sample that is a minimum of 10 times the number of items of the formative indicators in the model. Given that we had 24 items formatively represented in the model, we concluded that a minimum sample size of 240 was required.

[Textbox 1](#) provides an overview of the model constructs and their definitions in this study. In addition to these constructs, we included several factors that may influence the behavioral decision to use the app, such as individual characteristics (eg, age, sex, or educational level), contextual differences (United States vs United Kingdom), and COVID-19 experiences (whether they previously contracted COVID-19).

Of the 532 individuals who initiated the surveys, 363 (68.2%; [Table 1](#)) completed them (171 US and 203 UK respondents). The response rate is acceptable, given the sensitive nature of some questions [54]. Yet, we checked for nonresponse bias by comparing the demographic characteristics of respondents and nonrespondents; the results showed no significant differences, suggesting that nonresponse bias was not an issue in our study.

The measurement items were selected from previously validated measures [26,55,56] and adapted to fit the study context (see [Table 2](#) for measurement items). Although the measurements were validated in prior research, we conducted a pilot study and asked a convenience sample of 12 CT app users to review and reflect on the survey and provide qualitative feedback regarding the questionnaire guide measurement instruments. We also asked them to provide written feedback on the appropriateness, readability, and meaningfulness of the measures and their fit to the context of the study. On the basis of the qualitative feedback provided, we revised the survey guidelines to avoid possible confusion in responding and introducing any bias. Furthermore, we have slightly reworded a few items for clarity. For instance, item 1 in perceived barriers was changed from “The contact tracing app will violate my rights” to “The contact tracing app will violate my privacy.” Overall, the pilot study helped us improve the quality and accuracy of the data collection instrument.

Table 1. Demographic characteristics.

	Values (N=363), n (%)
Age (years)	
18-25	50 (14)
25-35	119 (33)
35-45	99 (27)
45-55	66 (18)
55-78	29 (8)
Sex	
Male	160 (44)
Female	200 (55)
Nonbinary	3 (1)
Education	
Less than high school	6 (2)
High school graduate	45 (12)
Some college	85 (23)
2-year degree	44 (12)
4-year degree	138 (38)
Professional degree	40 (11)
Doctorate	5 (1)
Contact-tracing app experience	
Currently using	125 (56)
Used in the past	114 (16)
Never used	124 (22)

Table 2. Measurement items and loadings.

Variable and items	Loadings
Perceived benefits of contact-tracing apps [26]	
Thanks to the contact-tracing app, I will be more on my guard when I have face-to-face contact.	0.785
Thanks to the contact-tracing app, I will take more precautions not to spread the Coronavirus myself (eg, wash my hands, maintain distance from others, and limit my outside movements).	0.719
By using the contact-tracing app, I will help public authorities combat the Coronavirus.	0.732
The contact tracing app will allow me to protect myself from the Coronavirus.	0.752
Perceived barriers of contact-tracing app use [26]	
The contact-tracing app will violate my privacy.	0.855
The contact-tracing app will create tensions between individuals who are infected by the Coronavirus and those who are not.	0.91
COVID-19 threat severity [26]	
If I get infected by the Coronavirus, it will have important health consequences for me.	0.843
If I get infected by the Coronavirus, my health will be severely affected.	0.895
If I get infected by the Coronavirus, my health will be significantly reduced.	0.914
COVID-19 threat susceptibility [26]	
I am at risk of being infected by the Coronavirus.	0.826
It is likely that I would suffer from the Coronavirus.	0.649
It is possible that I could be infected by the Coronavirus.	0.738
Self-efficacy to use contact-tracing apps [57]	
I have the knowledge needed to use the contact-tracing app.	0.914
I have the necessary resources to use the contact-tracing app.	0.915
I can get help from others if I experience difficulties using the contact-tracing app.	0.724
Cue to action (using contact-tracing apps) [26]	
To what extent do the following cues prompt the use of a contact-tracing app?	
Hearing someone near you contracted COVID-19	0.712
Website of a newspaper, TV or radio station, or magazine	0.938
App of a newspaper, TV or radio station, or magazine	0.984
News shared on social media (Facebook, YouTube, Twitter, Instagram, WhatsApp, etc)	0.901
Alerts through email and newsletters	0.908
(Contact-tracing apps) Continued use intention [52]	
I would be willing to continue using contact-tracing app.	0.912
I plan to continue using contact-tracing app.	0.920
I want to continue using contact-tracing app in the future.	0.901
Control factors	
Demographic factors: age, sex, education	— ^a
COVID-19 experience: Have you or a person close to you (ie, a close friend or family) been affected by COVID-19?	—
Contact-tracing app experience	—

^aNot available.

Data Analysis

Before testing the model, we assessed reliability and validity (ie, interconstruct correlations, Cronbach α values, composite reliability, and average variance extracted), as well as descriptive

statistics (Table 3). These variables showed good reliability. The α and composite reliability scores were all above the 0.7 acceptable thresholds [58]. The average variance extracted scores were also above 0.5, indicating good convergent validity

(ibid). Furthermore, the square roots of average variance extracted scores (on the matrix's diagonal) were higher than the correlations with the other constructs. In addition, the item loadings (see Table 2, last column) were all above 0.7 and loaded primarily onto their factor, which confirms a good discriminant validity [58]. Altogether, the results supported the reliability and validity of the constructs [59].

Subsequently, we used the structural equation modeling technique to validate the proposed research model and hypotheses. Finally, we ran additional analyses to measure the differences between users and nonusers and between-country differences. More details about these analyses are provided in the *Results* section.

Table 3. Descriptive statistics, correlations, reliability, and validity.

Constructs	Value, mean (SD)	Cronbach α	CR ^a	AVE ^b	1	2	3	4	5	6	7
1. Perceived benefits of CT ^c app use	4.40 (1.60)	.91	0.94	0.79	0.89	— ^d	—	—	—	—	—
2. Perceived barriers of CT app use	4.17 (1.58)	.71	0.87	0.77	-0.44	0.88	—	—	—	—	—
3. COVID-19 threat severity	4.22 (1.20)	.85	0.92	0.90	0.22	-0.10	0.90	—	—	—	—
4. COVID-19 threat susceptibility	4.72 (1.44)	.85	0.91	0.77	0.22	-0.21	0.49	0.88	—	—	—
5. Self-efficacy to use CT apps	4.72 (1.12)	.79	0.88	0.71	0.24	-0.08	-0.02	0.03	0.84	—	—
6. Cue to action (using CT apps)	2.61 (1.11)	.92	0.94	0.76	0.70	-0.44	0.22	0.31	0.20	0.87	—
7. (CT apps) Continued use intention	4.28 (2.03)	.98	0.99	0.96	0.72	-0.55	0.13	0.29	0.25	0.74	0.98

^aCR: composite reliability.

^bAVE: average variance extracted.

^cCT: contact tracing.

^dNot applicable.

Ethics Approval

This study was approved by the institutional review board of Pace University (IRB number 172765) before conducting any data collection.

Results

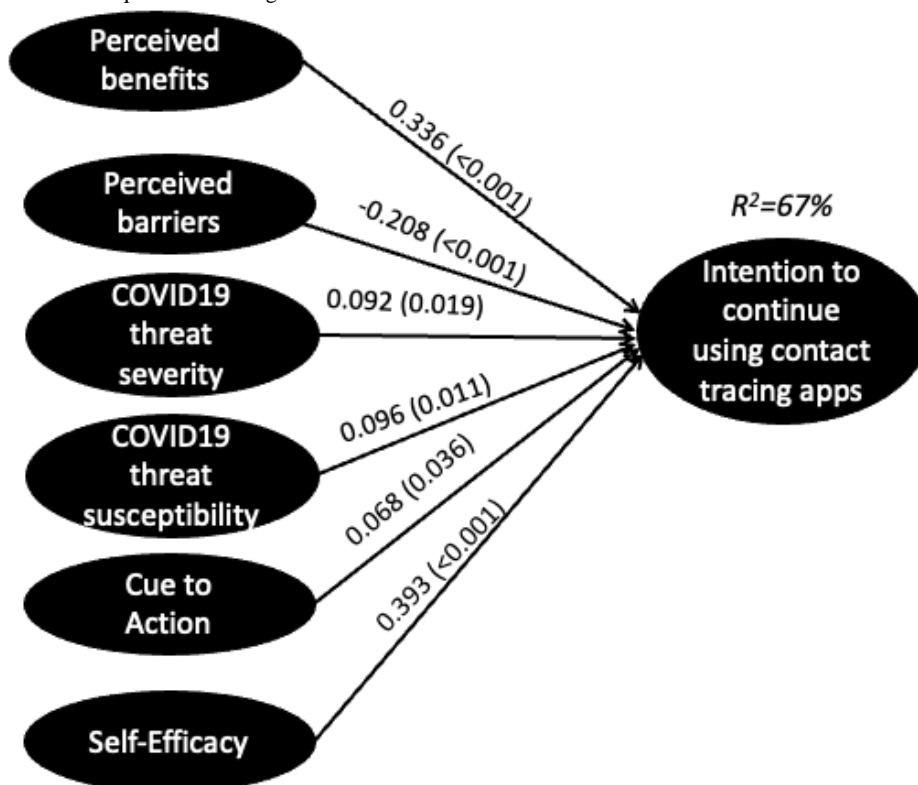
Structural Model Testing

We used a variance-based structural equation model analysis with SmartPLS (version 3; SmartPLS) to test the proposed research model. The results (Figure 1) supported the proposed hypotheses. We found that the perceived benefits of CT apps contribute to higher continued use intention ($\beta=.336$; $P<.001$), while the perceived barriers can reduce individuals' willingness

to do so ($\beta=-0.208$; $P<.001$). We also found that the perceptions of severity ($\beta=.092$; $P=.02$) and susceptibility ($\beta=.096$; $P=.01$) to COVID-19 as a significant health threat positively predict one's continued use intention. Next, we found that having self-efficacy to use CT apps, that is, being confident about having the knowledge and skills required to work with the app, can positively explain continued use intentions ($\beta=.393$; $P<.001$). Finally, cues from news, media, and peers can increase continued use intentions ($\beta=.068$; $P=.03$). Together, these factors explained 67% of the variance of the intention to continue using a CT app.

Although the direct hypothesized relationships were significant, the control variables (eg, age, sex, or education) posed no significant effect on the model and hypothesized relationships.

Figure 1. Results of the structural equation modeling.



Post Hoc Analysis 1: Testing Country-Based Differences

We performed additional analyses to check for heterogeneity in individuals’ CT app use. We tested the possible differences between app users’ intentions and their predictors (1) among individuals in the United States (n=160) versus the United Kingdom (n=203) and (2) among those who had previously used CT apps (n=239) versus those who did not (n=124). Accordingly, we followed the multigroup analysis (MGA) procedure [60] in PLS, which allows for direct nonparametric testing of the path estimates in the structural model for each bootstrap sample.

One potential concern when running an MGA is the measurement invariance [61]; hence, it is important to assess whether the construct measures are invariant between the

samples [60]. To establish measurement invariance, we checked the difference in item loadings across the 2 samples and whether the difference is statistically significant. Following the procedure by Henseler et al [62], all *P* values were above .05, except for 1 item (Benefits_4), which we dropped from further analysis (leaving 3 other reflective items for that construct). Next, we tested the differences in the complete structural model (Table 4). The results showed that the influence of predictors is largely similar in both the US and UK groups. Yet, we found a significant difference regarding the effect of COVID-19 threat susceptibility (*P*=.01), as it significantly contributed to continued use intention for the UK group but not for the US group. In addition, when considering each group individually, we found nonsignificant effects for COVID-19 threat severity and susceptibility among the US group and self-efficacy to use CT apps for the UK group.

Table 4. Results of post hoc analysis 1.

Predictor: (CT ^a apps) continued use intention	Path coefficients (United States)	<i>P</i> value (United States)	Path coefficients (United Kingdom)	<i>P</i> value (United Kingdom)	Path coefficients (difference)	<i>P</i> value (difference)
Perceived benefits of CT app use	0.32	<.001	0.36	<.001	-0.04	.74 ^b
Perceived barriers of CT app use	-0.21	<.001	-0.19	<.001	-0.02	.77 ^b
COVID-19 threat severity	0.04	.51 ^b	0.12	.01	-0.08	.29 ^b
COVID-19 threat susceptibility	-0.02	.71 ^b	0.19	<.001	-0.21	.01
Self-efficacy to use CT apps	0.14	.01	0.01	.85 ^b	0.13	.06 ^b
Cue to action (using CT apps)	0.41	<.001	0.38	<.001	0.04	.76 ^b

^aCT: contact tracing.

^bNot significant.

Post Hoc Analysis 2: Testing Experience-Based Differences

In the second post hoc analysis, we checked for potential differences in findings based on individuals' prior experience with CT apps. Following the same procedure, we assessed the measurement item invariance using the MGA procedure in PLS. We found that 2 items (Benefit_4 and CuestoAction_2) were not invariant across samples and hence were dropped before the MGA. Next, we tested for significant differences in the effects of predictors between the 2 groups (Table 5). We found

that the only significant difference is related to the effect of perceived barriers. The perceived barriers cannot influence nonusers' continued use intentions, while they can significantly reduce the experience group's intentions. Finally, when the model was tested in each subsample, we found that for experienced users, the effects of perceived threat severity and self-efficacy in using CT apps were nonsignificant. For the group with no prior experience, in addition to the nonsignificant effect of self-efficacy, the effect of perceived barriers was found to be nonsignificant.

Table 5. Results of post hoc analysis 2.

Predictors of intention	Path coefficients (exp ^a)	P value (no exp ^b)	Path coefficients (exp)	P value (no exp)	Path coefficients (difference)	P value (difference)
Perceived benefits of CT ^c app use	0.42	<.001	0.42	<.001	<0.001	.99 ^d
Perceived barriers of CT app use	-0.28	<.001	-0.05	.49 ^d	-0.23	.01
COVID-19 threat severity	0.07	.17 ^d	0.15	.03	-0.09	.33 ^d
COVID-19 threat susceptibility	0.09	.05	0.17	.03	-0.08	.34 ^d
Self-efficacy to use CT apps	0.08	.09 ^d	0.01	.90 ^d	0.07	.43 ^d
Cue to action (using CT apps)	0.26	<.001	0.43	<.001	-0.17	.08 ^d

^aExp: user with prior CT app experience.

^bNo exp: users who have had no prior experience with CT apps.

^cCT: contact tracing.

^dNot significant.

Discussion

Principal Findings

This study aimed to improve the current knowledge of the predictors for the continued use of CT apps. To this end, we draw upon the HBM [38,39] to propose a research model that shows the effect of 6 predictors. Analysis of data collected from 362 individuals showed that perceived benefits, self-efficacy, perceived severity, perceived susceptibility, and cues to action positively predicted the continued use intentions of CT apps, while perceived barriers could reduce them. Furthermore, we tested the possible differences among individuals in the United States versus the United Kingdom and those who previously used CT apps versus those who did not. These analyses revealed that the influence of critical predictors is similar in both the US and UK groups, with 1 exception—the effect of COVID-19 threat susceptibility is significant for the UK group but not for the US group. In addition, we noticed that the only significant difference between users and nonusers is related to the effect of perceived barriers; perceived barriers may not influence nonusers' continued use intentions, while they can significantly reduce experienced users' intentions.

Prior research has provided mixed results regarding the relationship between whether people are worried about COVID-19 and their intention to use a CT app [15]. Some studies reported that perceived health threat is positively associated with acceptance of CT app [37]. In addition, people who perceived lower health threats from COVID-19 tend to have a lower intention to embrace CT technology [16,37].

However, a few other studies showed contrasting findings—that perceived severity and perceived susceptibility were not related to the motivation for using CT apps [26,27]. In our study, when considering the entire sample, we found that perceived severity and susceptibility of COVID-19 could positively predict continued use intention. One possible explanation is that the prior studies found nonsignificance of a health threat in terms of CT app adoption intention at the early stage of the crisis when the government issued stay-at-home orders and mandated mask wearing. These measures limit people's contact with others, which could lead them to think that they are less susceptible to the virus. In contrast, at the time of our study, confinement measures and mandatory mask wearing were lifted, while a new highly transmissible variant (omicron) quickly spread in the community; such a situation could have influenced existing users' threat appraisal and intentions to continue using their CT app. Another explanation could be that individuals perceived adoption and continued use quite differently and considered different factors important in their decisions [63].

Regarding the impact of perceived benefits, prior studies found that social benefits (eg, using the app for the benefit of society) motivated CT app adoption [10,20,64,65]. However, mixed results were reported about the effect of personal benefits [19,66]; for example, as pointed out by Trang et al [66], compared with perceived social benefits, personal benefits seem to minimize the willingness to use a CT app among both critical and undecided respondents. In our study, we found that both perceived social benefits and personal benefits contributed to higher continued use intention of CT apps.

Many studies have highlighted the significant relationship between perceived barriers and CT app adoption intention [19,20,67]. A prominent perceived barrier for users is their concern about privacy issues raised by the app [26,34]. For example, the fear of data breaches or data misuse [18,27] and the fear of surveillance by the government [10,68] are significant barriers that prevent citizens from adopting CT apps. Consistent with prior findings, we found that perceived barriers can reduce individuals' willingness to continue using CT apps. This finding of the negative impact of perceived privacy on the uptake and continued use of CT apps highlighted the importance of informing users of how their privacy and data security are protected within the app. Privacy concern was also one of the motives behind developing and launching decentralized CT apps, which store and analyze personal data on users' devices, while the central server plays only a minor role in the CT process [69]. In addition, some CT methods have been proposed to use data-minimizing solutions such that they do not use user location data [70]. Future studies can investigate whether decentralized CT systems and clarifying data access can address these privacy concerns.

Individuals' beliefs in their competence to use the app (self-efficacy) have been associated with their acceptance or intention to use a CT app in multiple studies [19,26,27]. In particular, self-efficacy for app use decreases with age, as the younger population has more experience with digital technologies such as smartphone apps [26]. Similar to the results reported in prior work, our study found that having self-efficacy to use CT apps can positively explain continued use intentions. Future work could examine how to promote app uptake among the population with low technical proficiency and experience. This effort may also contribute to bridging the digital divide.

Our study found a positive relationship between cues to action (eg, exposure to information) and CT app continued use. This finding indicates that more (traditional and on the web) media coverage of CT apps can enhance their continued use. This finding aligns with prior work reporting that people's media consumption could influence their attitudes and intentions toward the app [71]. In this regard, more research can be conducted to investigate and analyze media coverage and web-based discussions on social media to gain insights into concerns and questions about the app. These insights can be used to inform app developers and governments to better communicate the usefulness and effectiveness of CT apps.

Our study also revealed between-country differences, even though the United States and the United Kingdom have followed similar COVID-19 measures. For example, perceived susceptibility has a significant effect on the UK group but not on the US group. In addition, we found nonsignificant effects for COVID-19 threat severity and susceptibility in the US group. This finding could be related to the upsurge of COVID-19 cases in the United Kingdom at the time of data collection, and United Kingdom residents may have felt more susceptible to the threat of COVID-19 during that period.

Finally, we observed differences in the effects of predictors between users and nonusers. More specifically, the perceived barriers were not associated with nonusers' continued use

intentions, but they could significantly reduce the intentions of the experienced group. One explanation for this difference could be that those who are yet to use CT technology cannot have an assessment of the potential challenges they may face, such as privacy and security, and this construct was not perceived as important in nonusers' decisions. As privacy concerns may not be the primary reason for nonadoption of CT apps among nonusers, future work needs to investigate the primary facilitators and barriers for nonusers to adopt and start using CT apps.

Overall, our results revealed that the factors contributing to the adoption of CT apps also play an essential role in existing users' intention to continue app use. More specifically, perceived benefits, self-efficacy, perceived severity, perceived susceptibility, and cues to action can motivate the continued use of CT apps, whereas perceived barriers could reduce an individual's intention to continue using a CT app. Another interesting finding is that perceived barriers could significantly reduce experienced user's continued use intentions, but this predictor had a limited influence on nonusers' intentions.

Practical Implications

Our study has several practical implications. First, as perceived social and personal benefits contributed to higher continued use intention of CT apps, these apps should be designed to continuously inform users about the potential social and personal benefits to ensure the effective use of CT apps in the long run. For example, these apps could present timely and updated informational resources (eg, advice on self-isolation, preventive, and testing options) to inform and assist users according to the COVID-19 trend (eg, the emergence of a new variant). In addition, a clear description of benefits to the user and society [66] as well as basic statistics that help users understand how the app aids people and society combat COVID-19 [34] are worth exploring. Second, more media coverage of CT apps might lead to continued use; as such, marketing and promoting such tools on both traditional media outlets and popular social media platforms (eg, Instagram or TikTok) could be helpful. This is similar to recent suggestions [72] about viable CT promotional strategies, such as using public health experts, independent privacy experts, and celebrities to endorse using these apps. In addition, over the past few years, we have witnessed much misinformation about COVID-19; the government and social media company decision makers should target misinformation about CT apps and provide proven factual and scientific information about these apps. Third, given that the perception of barriers (such as loss of privacy) could significantly reduce continued use, it might be useful to provide personalized options according to each user's preference for data sharing. For example, CT apps could offer an opt-in feature for users who are willing to contribute more location and personal data to obtain more useful features, while allowing those who are more concerned about data privacy risks to provide minimum data access yet be able to use the basic CT service [34].

Limitations and Future Work

There are several limitations. First, we had little representation of older adults and those with limited education levels. As age

and education level could be associated with the use of CT apps, data collection in the future would benefit from a booster sample of underrepresented and minority participants with a low level of education and technology proficiency to better understand inequalities across diverse populations. Second, we conducted a cross-sectional study rather than a longitudinal study. Future work should examine the facilitators of and barriers to the long-term use of CT apps. Finally, other models and constructs may provide important insights into the continued use of CT apps. Future studies can contribute to the understanding of CT apps' continued use and adoption by adopting other models.

Conclusions

The effectiveness of CT apps depends on not only their uptake by citizens but also their continued use during the COVID-19 pandemic. Our work contributes to the knowledge of facilitators and barriers in determining an individual's continued use intention of CT apps. We found that perceived benefits, self-efficacy, perceived severity, perceived susceptibility, and cues to action have significant positive impacts on the continued use intentions of CT apps, while perceived barriers can reduce such intentions. Further analyses revealed some degree of difference between users and nonusers. Those insights can be used by governments, technology companies, and media outlets to better promote the adoption and continued use of CT apps.

Conflicts of Interest

None declared.

References

1. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 contact tracing and data protection can go together. *JMIR Mhealth Uhealth* 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: [10.2196/19359](https://doi.org/10.2196/19359)] [Medline: [32294052](https://pubmed.ncbi.nlm.nih.gov/32294052/)]
2. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2022-05-02]
3. Colizza V, Grill E, Mikolajczyk R, Cattuto C, Kucharski A, Riley S, et al. Time to evaluate COVID-19 contact-tracing apps. *Nat Med* 2021 Mar;27(3):361-362. [doi: [10.1038/s41591-021-01236-6](https://doi.org/10.1038/s41591-021-01236-6)] [Medline: [33589822](https://pubmed.ncbi.nlm.nih.gov/33589822/)]
4. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 2020 May 08;368(6491):eabb6936 [FREE Full text] [doi: [10.1126/science.abb6936](https://doi.org/10.1126/science.abb6936)] [Medline: [32234805](https://pubmed.ncbi.nlm.nih.gov/32234805/)]
5. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health* 2020 Apr;8(4):e488-e496 [FREE Full text] [doi: [10.1016/S2214-109X\(20\)30074-7](https://doi.org/10.1016/S2214-109X(20)30074-7)] [Medline: [32119825](https://pubmed.ncbi.nlm.nih.gov/32119825/)]
6. Ho HJ, Zhang ZX, Huang Z, Aung AH, Lim WY, Chow A. Use of a real-time locating system for contact tracing of health care workers during the COVID-19 pandemic at an Infectious Disease Center in Singapore: validation study. *J Med Internet Res* 2020 May 26;22(5):e19437 [FREE Full text] [doi: [10.2196/19437](https://doi.org/10.2196/19437)] [Medline: [32412416](https://pubmed.ncbi.nlm.nih.gov/32412416/)]
7. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *J Epidemiol Community Health* 2020 Oct;74(10):861-866 [FREE Full text] [doi: [10.1136/jech-2020-214051](https://doi.org/10.1136/jech-2020-214051)] [Medline: [32576605](https://pubmed.ncbi.nlm.nih.gov/32576605/)]
8. Lewis D. Why many countries failed at COVID contact-tracing - but some got it right. *Nature* 2020 Dec;588(7838):384-387. [doi: [10.1038/d41586-020-03518-4](https://doi.org/10.1038/d41586-020-03518-4)] [Medline: [33318682](https://pubmed.ncbi.nlm.nih.gov/33318682/)]
9. Hinch R, Probert W, Nurtay A, Kendall M, Wymant C, Hall M, et al. Effective configurations of a digital contact tracing app: a report to NHSX. *The Conversation*. 2020. URL: https://cdn.theconversation.com/static_files/files/1009/Report_-_Effective_App_Configurations.pdf [accessed 2020-07-01]
10. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: cross-country survey study. *JMIR Mhealth Uhealth* 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: [10.2196/19857](https://doi.org/10.2196/19857)] [Medline: [32759102](https://pubmed.ncbi.nlm.nih.gov/32759102/)]
11. Farronato C, Iansiti M, Bartosiak M, Denicolai S, Ferretti L, Fontana R. How to get people to actually use contact-tracing apps. *Harvard Business Review*. 2020 Jul 15. URL: <https://hbr.org/2020/07/how-to-get-people-to-actually-use-contact-tracing-apps> [accessed 2022-11-22]
12. Chan S. COVID-19 contact tracing apps reach 9% adoption in most populous countries. *SensorTower*. 2020 Jul. URL: <https://sensortower.com/blog/contact-tracing-app-adoption> [accessed 2022-11-22]
13. Sutton S. Health behavior: psychosocial theories. In: Smelser NJ, Baltes B, editors. *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam, The Netherlands: Elsevier; 2001:6499-6506.
14. Simpson S, Kaufmann MC, Glozman V, Chakrabarti A. Disease X: accelerating the development of medical countermeasures for the next pandemic. *Lancet Infect Dis* 2020 May;20(5):e108-e115 [FREE Full text] [doi: [10.1016/S1473-3099\(20\)30123-7](https://doi.org/10.1016/S1473-3099(20)30123-7)] [Medline: [32197097](https://pubmed.ncbi.nlm.nih.gov/32197097/)]
15. Villius Zetterholm M, Lin Y, Jokela P. Digital contact tracing applications during COVID-19: a scoping review about public acceptance. *Informatics* 2021 Jul 22;8(3):48. [doi: [10.3390/informatics8030048](https://doi.org/10.3390/informatics8030048)]

16. Kostka G, Habich-Sobiegalla S. In times of crisis: public perceptions toward COVID-19 contact tracing apps in China, Germany, and the United States. *New Media Soc* (forthcoming) 2022 Apr 03;146144482210832. [doi: [10.1177/14614448221083285](https://doi.org/10.1177/14614448221083285)]
17. von Wyl V, Höglinger M, Sieber C, Kaufmann M, Moser A, Serra-Burriel M, et al. Drivers of acceptance of COVID-19 proximity tracing apps in Switzerland: panel survey analysis. *JMIR Public Health Surveill* 2021 Jan 06;7(1):e25701 [FREE Full text] [doi: [10.2196/25701](https://doi.org/10.2196/25701)] [Medline: [33326411](https://pubmed.ncbi.nlm.nih.gov/33326411/)]
18. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mülken M. COVID-19 contact tracing apps: predicted uptake in the Netherlands based on a discrete choice experiment. *JMIR Mhealth Uhealth* 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: [10.2196/20741](https://doi.org/10.2196/20741)] [Medline: [32795998](https://pubmed.ncbi.nlm.nih.gov/32795998/)]
19. Sharma S, Singh G, Sharma R, Jones P, Kraus S, Dwivedi YK. Digital health innovation: exploring adoption of COVID-19 digital contact tracing apps. *IEEE Trans Eng Manage* (forthcoming) 2020 Sep 15:1-17. [doi: [10.1109/tem.2020.3019033](https://doi.org/10.1109/tem.2020.3019033)]
20. Simko L, Chang J, Jiang M, Calo R, Roesner F, Kohno T. COVID-19 contact tracing and privacy: a longitudinal study of public opinion. *Digital Threats* 2022 Sep 30;3(3):1-36. [doi: [10.1145/3480464](https://doi.org/10.1145/3480464)]
21. Saw YE, Tan EY, Liu JS, Liu JC. Predicting public uptake of digital contact tracing during the covid-19 pandemic: results from a nationwide survey in Singapore. *J Med Internet Res* 2021 Feb 03;23(2):e24730 [FREE Full text] [doi: [10.2196/24730](https://doi.org/10.2196/24730)] [Medline: [33465034](https://pubmed.ncbi.nlm.nih.gov/33465034/)]
22. Jamieson J, Yamashita N, Epstein DA, Chen Y. Deciding if and how to use a COVID-19 contact tracing app: influences of social factors on individual use in Japan. *Proc ACM Hum Comput Interact* 2021 Oct 13;5(CSCW2):1-30. [doi: [10.1145/3479868](https://doi.org/10.1145/3479868)]
23. Clawson J, Pater JA, Miller AD, Mynatt ED, Mamykina L. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2015 Sep Presented at: UbiComp '15; September 7-11, 2015; Osaka, Japan p. 647-658. [doi: [10.1145/2750858.2807554](https://doi.org/10.1145/2750858.2807554)]
24. Epstein DA, Caraway M, Johnston C, Ping A, Fogarty J, Munson SA. Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. *Proc SIGCHI Conf Hum Factor Comput Syst* 2016 May;2016:1109-1113 [FREE Full text] [doi: [10.1145/2858036.2858045](https://doi.org/10.1145/2858036.2858045)] [Medline: [28503678](https://pubmed.ncbi.nlm.nih.gov/28503678/)]
25. Horvath L, Banducci S, Blamire J, Degnen C, James O, Jones A, et al. Adoption and continued use of mobile contact tracing technology: multilevel explanations from a three-wave panel survey and linked data. *BMJ Open* 2022 Jan 17;12(1):e053327 [FREE Full text] [doi: [10.1136/bmjopen-2021-053327](https://doi.org/10.1136/bmjopen-2021-053327)] [Medline: [35039293](https://pubmed.ncbi.nlm.nih.gov/35039293/)]
26. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. *JMIR Public Health Surveill* 2020 Sep 01;6(3):e20572 [FREE Full text] [doi: [10.2196/20572](https://doi.org/10.2196/20572)] [Medline: [32755882](https://pubmed.ncbi.nlm.nih.gov/32755882/)]
27. Kaspar K. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: quantitative survey study. *J Med Internet Res* 2020 Aug 27;22(8):e21613 [FREE Full text] [doi: [10.2196/21613](https://doi.org/10.2196/21613)] [Medline: [32759100](https://pubmed.ncbi.nlm.nih.gov/32759100/)]
28. Tomczyk S, Barth S, Schmidt S, Muehlan H. Utilizing health behavior change and technology acceptance models to predict the adoption of COVID-19 contact tracing apps: cross-sectional survey study. *J Med Internet Res* 2021 May 19;23(5):e25447 [FREE Full text] [doi: [10.2196/25447](https://doi.org/10.2196/25447)] [Medline: [33882016](https://pubmed.ncbi.nlm.nih.gov/33882016/)]
29. Suh A, Li M. Digital tracing during the COVID-19 pandemic: user appraisal, emotion, and continuance intention. *Sustainability* 2021 Jan 10;13(2):608. [doi: [10.3390/su13020608](https://doi.org/10.3390/su13020608)]
30. Kukuk L. Analyzing adoption of COVID-19 contact tracing apps using UTAUT. University of Twente. 2020. URL: http://essay.utwente.nl/81983/1/Kukuk_BA_EEMCS.pdf [accessed 2022-11-22]
31. Beaudry A, Vaghefi I, Bagayogo F, Lapointe L. Impact of IT user behavior: observations through a new lens. *Commun Assoc Inf Syst* 2020 Mar 4;46:331-364. [doi: [10.17705/1cais.04615](https://doi.org/10.17705/1cais.04615)]
32. Walrave M, Waeterloos C, Ponnet K. Ready or not for contact tracing? Investigating the adoption intention of COVID-19 contact-tracing technology using an extended unified theory of acceptance and use of technology model. *Cyberpsychol Behav Soc Netw* 2021 Jun;24(6):377-383. [doi: [10.1089/cyber.2020.0483](https://doi.org/10.1089/cyber.2020.0483)] [Medline: [33017171](https://pubmed.ncbi.nlm.nih.gov/33017171/)]
33. Vaghefi I, Lapointe L, Boudreau-Pinsonneault C. A typology of user liability to IT addiction. *Inf Systems J* 2017 Mar;27(2):125-169. [doi: [10.1111/isj.12098](https://doi.org/10.1111/isj.12098)]
34. Li T, Cobb C, Yang JJ, Baviskar S, Agarwal Y, Li B, et al. What makes people install a COVID-19 contact-tracing app? Understanding the influence of app design and individual difference on contact-tracing app adoption intention. *Pervasive Mob Comput* 2021 Aug;75:101439. [doi: [10.1016/j.pmcj.2021.101439](https://doi.org/10.1016/j.pmcj.2021.101439)]
35. Abhari K, Vaghefi I. Screen time and productivity: an extension of goal-setting theory to explain optimum smartphone use. *AIS Trans Human Comput* 2022 Sep 30;14(3):254-288. [doi: [10.17705/1thci.00169](https://doi.org/10.17705/1thci.00169)]
36. Garrett PM, White JP, Lewandowsky S, Kashima Y, Perfors A, Little DR, et al. The acceptability and uptake of smartphone tracking for COVID-19 in Australia. *PLoS One* 2021 Jan 22;16(1):e0244827 [FREE Full text] [doi: [10.1371/journal.pone.0244827](https://doi.org/10.1371/journal.pone.0244827)] [Medline: [33481841](https://pubmed.ncbi.nlm.nih.gov/33481841/)]
37. Munzert S, Selb P, Gohdes A, Stoetzer LF, Lowe W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat Hum Behav* 2021 Feb;5(2):247-255. [doi: [10.1038/s41562-020-01044-x](https://doi.org/10.1038/s41562-020-01044-x)] [Medline: [33479505](https://pubmed.ncbi.nlm.nih.gov/33479505/)]

38. Rosenstock IM. Historical origins of the Health Belief Model. *Health Educ Monogr* 1974 Dec 01;2(4):328-335. [doi: [10.1177/109019817400200403](https://doi.org/10.1177/109019817400200403)]
39. Rosenstock IM. The health belief model and preventive health behavior. *Health Educ Monogr* 1974 Dec 01;2(4):354-386. [doi: [10.1177/109019817400200405](https://doi.org/10.1177/109019817400200405)]
40. Glanz KE, Rimer BK, Viswanath K. *Health Behavior and Health Education: Theory, Research, and Practice*. 4th edition. Hoboken, NJ, USA: John Wiley & Sons; 2008.
41. Paige SR, Bonnar KK, Black DR, Coster DC. Risk factor knowledge, perceived threat, and protective health behaviors: implications for type 2 diabetes control in rural communities. *Diabetes Educ* 2018 Feb;44(1):63-71. [doi: [10.1177/0145721717747228](https://doi.org/10.1177/0145721717747228)] [Medline: [29241427](https://pubmed.ncbi.nlm.nih.gov/29241427/)]
42. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. *Int J Inf Manag* 2018 Dec;43:342-350. [doi: [10.1016/j.ijinfomgt.2017.08.006](https://doi.org/10.1016/j.ijinfomgt.2017.08.006)]
43. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. *JMIR Mhealth Uhealth* 2017 Dec 06;5(12):e177 [FREE Full text] [doi: [10.2196/mhealth.7886](https://doi.org/10.2196/mhealth.7886)] [Medline: [29212629](https://pubmed.ncbi.nlm.nih.gov/29212629/)]
44. Tobias G, Sgan-Cohen H, Spanier AB, Mann J. Perceptions and attitudes toward the use of a mobile health app for remote monitoring of gingivitis and willingness to pay for mobile health apps (part 3): mixed methods study. *JMIR Form Res* 2021 Oct 05;5(10):e26125 [FREE Full text] [doi: [10.2196/26125](https://doi.org/10.2196/26125)] [Medline: [34609320](https://pubmed.ncbi.nlm.nih.gov/34609320/)]
45. Champion VL. Instrument refinement for breast cancer screening behaviors. *Nurs Res* 1993;42(3):139-143. [doi: [10.1097/00006199-199305000-00003](https://doi.org/10.1097/00006199-199305000-00003)] [Medline: [8506161](https://pubmed.ncbi.nlm.nih.gov/8506161/)]
46. Lu XL, Reynolds T, Jo E, Hong H, Page X, Chen Y. Comparing perspectives around human and technology support for contact tracing. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021 May Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-15. [doi: [10.1145/3411764.3445669](https://doi.org/10.1145/3411764.3445669)]
47. Redmiles EM. User concerns and tradeoffs in technology-facilitated COVID-19 response. *Digit Gov Res Pract* 2021 Jan 31;2(1):1-12. [doi: [10.1145/3428093](https://doi.org/10.1145/3428093)]
48. Champion VL, Skinner CS. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. *Health Behavior and Health Education: Theory, Research, and Practice*. 4th edition. Hoboken, NJ, USA: Jossey-Bass; 2008:45-65.
49. Wu JH, Wang SC, Lin LM. Mobile computing acceptance factors in the healthcare industry: a structural equation model. *Int J Med Inform* 2007 Jan;76(1):66-77. [doi: [10.1016/j.ijmedinf.2006.06.006](https://doi.org/10.1016/j.ijmedinf.2006.06.006)] [Medline: [16901749](https://pubmed.ncbi.nlm.nih.gov/16901749/)]
50. Wei J, Vinnikova A, Lu L, Xu J. Understanding and predicting the adoption of fitness mobile apps: evidence from China. *Health Commun* 2021 Jul;36(8):950-961. [doi: [10.1080/10410236.2020.1724637](https://doi.org/10.1080/10410236.2020.1724637)] [Medline: [32041437](https://pubmed.ncbi.nlm.nih.gov/32041437/)]
51. Yi MY, Hwang Y. Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *Int J Human Comput Stud* 2003 Oct;59(4):431-449. [doi: [10.1016/s1071-5819\(03\)00114-9](https://doi.org/10.1016/s1071-5819(03)00114-9)]
52. Bhattacharjee A. Understanding information systems continuance: an expectation-confirmation model. *MIS Q* 2001 Sep;25(3):351-370. [doi: [10.2307/3250921](https://doi.org/10.2307/3250921)]
53. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. *J Market Theory Pract* 2011;19(2):139-152. [doi: [10.2753/mtp1069-6679190202](https://doi.org/10.2753/mtp1069-6679190202)]
54. Murdoch M, Simon AB, Polusny MA, Bangerter AK, Grill JP, Noorbaloochi S, et al. Impact of different privacy conditions and incentives on survey response rate, participant representativeness, and disclosure of sensitive information: a randomized controlled trial. *BMC Med Res Methodol* 2014 Jul 16;14:90 [FREE Full text] [doi: [10.1186/1471-2288-14-90](https://doi.org/10.1186/1471-2288-14-90)] [Medline: [25027174](https://pubmed.ncbi.nlm.nih.gov/25027174/)]
55. Major B, Richards C, Cooper ML, Cozzarelli C, Zubek J. Personal resilience, cognitive appraisals, and coping: an integrative model of adjustment to abortion. *J Pers Soc Psychol* 1998 Mar;74(3):735-752. [doi: [10.1037//0022-3514.74.3.735](https://doi.org/10.1037//0022-3514.74.3.735)] [Medline: [9523416](https://pubmed.ncbi.nlm.nih.gov/9523416/)]
56. Limayem M, Hirt SG, Cheung MK. How habit limits the predictive power of intention: the case of information systems continuance. *MIS Q* 2007 Dec;31(4):705-737. [doi: [10.2307/25148817](https://doi.org/10.2307/25148817)]
57. Bandura A. Guide for constructing self-efficacy scales. In: Urdan T, Pajares F, editors. *Self-Efficacy Beliefs of Adolescents*. Charlotte, NC, USA: Information Age Publishing; 2006:307-337.
58. Hair Jr JF, Hult GT, Ringle CM, Sarstedt M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA, USA: Sage Publications; 2021.
59. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Market Res* 1981 Feb;18(1):39-50. [doi: [10.1177/002224378101800104](https://doi.org/10.1177/002224378101800104)]
60. Sarstedt M, Henseler J, Ringle C. Multigroup analysis in partial least squares (PLS) path modeling: alternative methods and empirical results. In: Sarstedt M, Schwaiger M, Taylor CR, editors. *Measurement and Research Methods in International Marketing*. Bingley, UK: Emerald Group Publishing; 2011:195-218.
61. Henseler J, Ringle CM, Sinkovics RR. The use of partial least squares path modeling in international marketing. In: Sinkovics RR, Ghauri PN, editors. *New Challenges to International Marketing*. Bingley, UK: Emerald Group Publishing; 2009:277-319.
62. Henseler J, Ringle CM, Sarstedt M. Testing measurement invariance of composites using partial least squares. *Int Market Rev* 2016 May 9;33(3):405-431. [doi: [10.1108/IMR-09-2014-0304](https://doi.org/10.1108/IMR-09-2014-0304)]

63. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. *JMIR Mhealth Uhealth* 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: [10.2196/12983](https://doi.org/10.2196/12983)] [Medline: [31469081](https://pubmed.ncbi.nlm.nih.gov/31469081/)]
64. Williams SN, Armitage CJ, Tampe T, Dienes K. Public attitudes towards COVID-19 contact tracing apps: a UK-based focus group study. *Health Expect* 2021 Apr;24(2):377-385 [FREE Full text] [doi: [10.1111/hex.13179](https://doi.org/10.1111/hex.13179)] [Medline: [33434404](https://pubmed.ncbi.nlm.nih.gov/33434404/)]
65. Garrett PM, Wang Y, White JP, Hsieh S, Strong C, Lee YC, et al. Young adults view smartphone tracking technologies for COVID-19 as acceptable: the case of Taiwan. *Int J Environ Res Public Health* 2021 Feb 02;18(3):1332 [FREE Full text] [doi: [10.3390/ijerph18031332](https://doi.org/10.3390/ijerph18031332)] [Medline: [33540628](https://pubmed.ncbi.nlm.nih.gov/33540628/)]
66. Trang S, Trenz M, Weiger WH, Tarafdar M, Cheung CM. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. *Eur J Inf Syst* 2020 Jul 27;29(4):415-428. [doi: [10.1080/0960085x.2020.1784046](https://doi.org/10.1080/0960085x.2020.1784046)]
67. Zimmermann BM, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. *J Med Internet Res* 2021 Feb 08;23(2):e25525 [FREE Full text] [doi: [10.2196/25525](https://doi.org/10.2196/25525)] [Medline: [33503000](https://pubmed.ncbi.nlm.nih.gov/33503000/)]
68. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. *Ir J Med Sci* 2021 Aug;190(3):863-887 [FREE Full text] [doi: [10.1007/s11845-020-02389-y](https://doi.org/10.1007/s11845-020-02389-y)] [Medline: [33063226](https://pubmed.ncbi.nlm.nih.gov/33063226/)]
69. Sowmiya B, Abhijith VS, Sudersan S, Sakthi Jaya Sundar R, Thangavel M, Varalakshmi P. A survey on security and privacy issues in contact tracing application of COVID-19. *SN Comput Sci* 2021;2(3):136 [FREE Full text] [doi: [10.1007/s42979-021-00520-z](https://doi.org/10.1007/s42979-021-00520-z)] [Medline: [33728414](https://pubmed.ncbi.nlm.nih.gov/33728414/)]
70. Yasaka TM, Lehrich BM, Sahyouni R. Peer-to-Peer contact tracing: development of a privacy-preserving smartphone app. *JMIR Mhealth Uhealth* 2020 Apr 07;8(4):e18936 [FREE Full text] [doi: [10.2196/18936](https://doi.org/10.2196/18936)] [Medline: [32240973](https://pubmed.ncbi.nlm.nih.gov/32240973/)]
71. Amann J, Sleight J, Vayena E. Digital contact-tracing during the Covid-19 pandemic: an analysis of newspaper coverage in Germany, Austria, and Switzerland. *PLoS One* 2021 Feb 3;16(2):e0246524 [FREE Full text] [doi: [10.1371/journal.pone.0246524](https://doi.org/10.1371/journal.pone.0246524)] [Medline: [33534839](https://pubmed.ncbi.nlm.nih.gov/33534839/)]
72. Farrelly G, Trabelsi H, Cocosila M. COVID-19 contact tracing applications: an analysis of individual motivations for adoption and use. *First Monday* 2022 Jun 04;27(6). [doi: [10.5210/fm.v27i6.12324](https://doi.org/10.5210/fm.v27i6.12324)]

Abbreviations

CT: contact tracing

HBM: health belief model

MGA: multigroup analysis

Edited by A Mavragani; submitted 14.06.22; peer-reviewed by W Kok Hoe Mok, A Yazdanmehr; comments to author 18.10.22; revised version received 07.11.22; accepted 08.11.22; published 08.12.22.

Please cite as:

Zhang Z, Vaghefi I

Continued Use of Contact-Tracing Apps in the United States and the United Kingdom: Insights From a Comparative Study Through the Lens of the Health Belief Model

JMIR Form Res 2022;6(12):e40302

URL: <https://formative.jmir.org/2022/12/e40302>

doi: [10.2196/40302](https://doi.org/10.2196/40302)

PMID: [36351080](https://pubmed.ncbi.nlm.nih.gov/36351080/)

©Zhan Zhang, Isaac Vaghefi. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 08.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using a Proximity-Detection Technology to Nudge for Physical Distancing in a Swedish Workplace During the COVID-19 Pandemic: Retrospective Case Study

My Villius Zetterholm¹, MSci; Lina Nilsson², PhD; Päivi Jokela³, PhD

¹Department of Informatics, Faculty of Technology, Linnaeus University, Kalmar, Sweden

²eHealth Institute, Department of Medicine and Optometry, Faculty of Health and Life Sciences, Linnaeus University, Kalmar, Sweden

³Department of Informatics, Faculty of Technology, Linnaeus University, Växjö, Sweden

Corresponding Author:

My Villius Zetterholm, MSci

Department of Informatics

Faculty of Technology

Linnaeus University

Universitetsplatsen 1

Kalmar, 39231

Sweden

Phone: 46 480497711

Email: my.villiuszetterholm@lnu.se

Abstract

Background: The recent COVID-19 pandemic has contributed to the emergence of several technologies for infectious disease management. Although much focus has been placed on contact-tracing apps, another promising new tactic is proximity tracing, which focuses on health-related behavior and can be used for primary prevention. Underpinned by theories on behavioral design, a proximity-detection system can be devised that provides a user with immediate nudges to maintain physical distance from others. However, the practical feasibility of proximity detection during an infectious disease outbreak has not been sufficiently investigated.

Objective: We aimed to evaluate the feasibility of using a wearable device to nudge for distance and to gather important insights about how functionality and interaction are experienced by users. The results of this study can guide future research and design efforts in this emerging technology.

Methods: In this retrospective case study, a wearable proximity-detection technology was used in a workplace for 6 weeks during the production of a music competition. The purpose of the technology was to nudge users to maintain their physical distance using auditory feedback. We used a mixed methods sequential approach, including interviews (n=8) and a survey (n=30), to compile the experiences of using wearable technology in a real-life setting.

Results: We generated themes from qualitative analysis based on data from interviews and open-text survey responses. The quantitative data were subsequently integrated into these themes: *feasibility* (implementation and acceptance—establishing a shared problem; distance tags in context—strategy, environment, and activities; understanding and learning; and accomplishing the purpose) and *design aspects* (a purposefully annoying device; timing, tone, and proximity; and additional functions).

Conclusions: This empirical study reports on the feasibility of using wearable technology based on proximity detection to nudge individuals to maintain physical distance in the workplace. The technology supports attention to distance, but the usability of this approach is dependent on the context and situation. In certain situations, the audio signal is frustrating, but most users agree that it needs to be annoying to ensure sufficient behavioral adaption. We proposed a dual nudge that involves vibration followed by sound. There are indications that the technology also facilitates learning how to maintain a greater distance from others, and that this behavior can persist beyond the context of technology use. This study demonstrates that the key value of this technology is that it places the user in control and enables immediate action when the distance to others is not maintained. This study provides insights into the emerging field of personal and wearable technologies used for primary prevention during infectious disease outbreaks. Future research is needed to evaluate the preventive effect on transmission and investigate behavioral changes in detail and in relation to different forms of feedback.

KEYWORDS

case study; COVID-19; feasibility; mixed methods; nudging; physical distance; preventive behavior; preventive technologies; proximity detecting technology; wearables

Introduction

Background

Since the onset of the COVID-19 pandemic, governments across the globe have issued a range of different mitigation and control measures to reduce transmission. At the individual level, these measures have included, for example, promotion of hygiene routines, wearing a face mask, various levels of contact tracing, lockdowns, and restrictions in movement and crowding in an effort to maintain physical distance between individuals [1]. Consequently, there has been a worldwide interest in using wearable and mobile information technologies for preventive work [2]. Gasser et al [3] provided a typology of digital tools against COVID-19; these include 4 categories of technologies, namely *flow modeling*, *quarantine control*, *symptom monitoring*, and *proximity and contact tracing*. The primary interest of this study is in the fourth category, comprising technologies for digital proximity tracing.

Contact-tracing apps (CTAs) have been one of the most frequently implemented technologies in this vein [4-7]. According to most conceptualizations, digital proximity tracing is synonymous with digital contact tracing; that is, it is implied that this technology can only be used for this purpose in a pandemic situation. This perception is exemplified in the typology [3] mentioned earlier as well as in definitions of proximity-tracing solutions provided by the European Centre for Disease Prevention and Control and the World Health Organization, the latter of which says, "Digital proximity tracing refers to a technological approach to public health contact tracing that typically utilizes smartphones or purpose-built devices to capture anonymized interactions between individuals, and subsequently issue alerts, if conditions are met that indicate a period of close proximity to someone who later returns a positive diagnosis of infectious disease" [8].

CTAs are a digital version of the classical manual contact-tracing method commonly used during infectious disease outbreaks. The implementation of CTAs can be voluntary or mandatory. The design of these apps varies in different countries [6] and so does their uptake and acceptance in different populations [7,9-11]. The deployment of contact-tracing technologies has raised many concerns regarding ethics, privacy, and feasibility [12-16].

In contrast to the standard perception of proximity and contact tracing, some studies have suggested that proximity-based technologies can be used for purposes other than contact tracing to support users' situational awareness. For example, they could be used to alert users with a nudge or similar notification, providing them with immediate feedback of their physical distance to another individual [17-19]. As such, this approach exemplifies a design thinking approach to the use of wearable technologies for infectious disease management, where new

areas of application are explored based on existing technologies. This approach emphasizes primary prevention by deploying proximity technologies to achieve a direct impact on behavior and thereby placing the individual user in control of the situation. This differs from the traditional focus of CTAs, which involves the tracking of infectious individuals. However, the practical feasibility of proximity tracing for nudging purposes has not been well studied in the context of infectious diseases. To date, most studies have been theoretical [17,18] or have investigated the effectiveness of this technology in experimental settings [19]. This approach needs to be further studied and include real-world aspects such as behavioral outcomes and user experience [17,19].

Designing Preventive Behavior

A pandemic is a complex system of interacting factors that need to be understood from biological, social, and psychological perspectives, to name a few. A root cause of transmission is the behavior of individuals [20], that is, the human hosts. Therefore, the promotion of preventive behavior is a key aspect in any infectious disease outbreak, and there is a need for evidence-based strategies for the intervention and formation of new patterns of behavior to limit transmission. In the early days of the current pandemic, models and methods from behavioral science were discussed to inform the development of preventive interventions [20-22].

The importance of protective and preventive behavior has been evident in much of the public communication during the COVID-19 pandemic, and simple strategies such as hand washing, wearing a mask, and social and physical distancing have been repeatedly promoted around the world [23,24]. The appropriate use of preventive measures depends on the availability of personal protective equipment, the type of activities, and the physical context. In an experimental study conducted in a laboratory setting, a model system of dolomite dust particles was used to simulate airborne transmission [25]. It was shown that if everyone properly wears a highly efficient and tight-fitting mask, such as FFP2 or KN95, the risk of airborne transmission is very low. The risk increases if only the susceptible person is wearing the FFP2 mask, if surgical masks are used, and if physical distancing is maintained without masking [25]. However, uncertainties remain regarding the significance of various transmission routes [26]. In a retrospective case-control study in community settings in Thailand, it was found that consistent mask wearing, handwashing, and adhering to social distancing were all independently associated with a lower risk of infection, but none of these protective measures alone could provide complete protection from the infection. Complying with all measures was the most effective way to reduce transmission in public gatherings, and it was also found that those who consistently and correctly wore masks were more likely to wash their hands and practice adequate social distancing [27].

However, adhering to preventive recommendations can be challenging. New behaviors will require a change in habit, and cognitive factors such as attention span or memory may prevent full adherence to public health recommendations. Even with information and knowledge, changing subconscious routines or automatic behaviors is challenging [24]. Furthermore, there have been few opportunities for training or support intended to ameliorate these problems or help form new habits [20]. A few novel technologies have emerged that can promote and support new preventive habits such as adherence to handwashing routines [28]. It has also been suggested to add more persuasive design elements to the existing CTAs to increase users' uptake and adherence to public health recommendations [29]. As previously mentioned, technologies that could support physical distancing by increasing users' situation awareness have also been proposed [17-19].

The design of warning systems aimed at improving safety-related behaviors has been studied in a range of contexts. Proximity-based warning systems are often used in dangerous working environments where there is a heightened risk of collisions between people and heavy equipment, such as construction sites, manufacturing plants, and underground mines [30-32]. They have also been used in surgery rooms where extreme attention to details is needed [33,34]. Warning systems in general are also common in various driving assistance systems and semiautonomous vehicles [35-37]. A common feature that has been studied in these various environments is the effectiveness of visual [32], auditory [30], and vibro-tactile [31] signals and multimodal combinations [33-37] of these alerts. Research from dangerous working environments has shown that even with safety standards and planning, workers often fail to recognize or act on safety hazards owing to factors such as lack of attention, cognitive overload, and distractions [30]. These observations have led to an increasing interest in how technologies can help workers recognize and avoid hazards. However, studies on these types of systems also indicate that repeated exposure to warnings may desensitize the user and cause alarm fatigue, particularly if alerts are too frequent, impossible to avoid, or redundant. Recent research efforts have aimed to minimize these types of issues [30,31].

An increasingly popular approach to address cognitive limitations, such as a lack of attention, is nudging. In this study, the term *nudging* [38] refers to an approach that supports behavior change by design rather than behavior change as a result of intention and attitude change. Nudges are small and purposeful design elements that can take the form of, for example, reminders and notifications [38]. A nudge can be defined as “a function of the choice architecture that alters people's behavior in a predictable way that is called for because of cognitive boundaries, biases, routines, and habits in individual and social decision-making and which works by making use of those boundaries, biases, routines, and habits as integral parts of the choice architecture” [39].

A main goal of nudging is to design in a way that is expected to encourage individuals to act in their own best interest [38,40]. The design focus of nudging lies in purposeful changes to the choice architecture, that is, the context in which people act and where the designer can implement nudges. One principle for

designing nudges is to provide feedback that can support individuals by informing them when they make mistakes [38]. This type of nudge can be defined as just-in-time prompts that draw attention to a behavior when it occurs [41]. Therefore, nudges are particularly relevant in situations in which the main goal of the design is to approach the automatic behavior of individuals. In recent years, there has been a heightened interest in using various information technologies to nudge users, which has influenced design approaches in both information systems and human-computer interactions [41].

Nudges can be classified based on multiple perspectives. One model categorizes whether they focus on the reflective or automatic mind, whether they target behavior or choice, and the level of transparency versus nontransparency in the design of nudges [42]. The technology described in this study provides a nudge that would be categorized as focused on the automatic mind and targeting behavior and is transparent for the user. The physical form of the device, called the distance tag, was a credit card-sized wearable that hung around the necks of individuals when they were in the workplace. If the individuals wearing these tags came within 1.5 m of each other, they were alerted by a high-pitched audio signal. This means that the nudge was designed as a just-in-time prompt and provided the user with audio-based feedback on mistakes. It should also be noted that the distance tag could be switched off.

Objective

There has been little research on nudging technologies in the context of infectious disease. The case described in this study is an early example of a large-scale workplace setting in which an emerging nudging technology was used to achieve primary prevention during a pandemic. The technology used on-site, that is, the distance tag, incorporated behavior change by design to improve physical distancing and limit transmission. The main objective of this study was to evaluate the feasibility of using a wearable device to nudge for distance and to gather important insights about how functionality and interaction were experienced by the users. Another goal of this study was to inform future design efforts in this field. The following research questions guided this study:

1. How feasible is it to use proximity-based technology to nudge users for physical distance in a large-scale workplace setting?
2. How did users experience functionality and interaction with this technology?

This study is a retrospective case study [43] that assessed the feasibility and experience of using a distance tag in a workplace for 6 weeks during the production of a television-broadcast music competition (also referred to as *the production* or *production project* in this paper). We were not involved in the production, design of the preventive strategy, or the implementation of the distance tag. The setting of the case offered the unique possibility to learn more about whether wearable proximity technologies can be used in natural settings where individuals meet, work, and socialize during an infectious outbreak. The mixed methods approach in this study contributes with 2 complementary perspectives. The interview participants were those who were responsible for management and safety

during the production; therefore, they could provide a holistic perspective of the case. The survey participants contributed with their individual experience and perception of the technology. The results of this study can contribute to knowledge about the use of preventive and proximity-based technologies and can guide further design efforts in this emerging field.

Methods

Overview

Because the COVID-19 pandemic and technologies used here are both novel, this study was conducted using an explorative approach. This was a retrospective case study [43], in which the research study and all data collection were conducted after the production activities were completed and the music

competition was broadcast. The researchers collected the data based on the experiences of the participants.

The general rationale of a mixed methods study design is that a combination of both qualitative and quantitative approaches provides a better understanding of research problems than only 1 approach [44]. This study had an exploratory sequential design [44], where the first inquiry included qualitative interviews, the results of which informed the second inquiry, that is, the development of the survey and quantitative data collection. This sequential procedure is an example of instrument development [45]. The final analysis combined results from both qualitative and quantitative studies to gain a more comprehensive account of the studied phenomenon, which can be referred to as completeness [45]. The overall study design is presented in [Table 1](#).

Table 1. The mixed methods design used in this study.

Step	Approach	Method	Time	Sampling technique and sample size
1	Qualitative	Preliminary unstructured interview	February 2021	Purposeful sampling (n=1 ^a)
2	Qualitative	Semistructured individual interviews (interview A+B)	March 2021	Snowball sampling (n=2)
3	Qualitative	Semistructured group interview (interview C)	March 2021	Snowball sampling (n=6)
4	Mixed methods quantitative	Web-based questionnaire	April 2021	Total population sampling (n=30)
5	Mixed methods	Final analysis of results from both qualitative and quantitative study	October- December 2021	8 qualitative+30 quantitative (n=8+30; total n=38)

^aThis individual was later included in a more formal group interview (step 3, interview C).

Interviews

The first contact was an informal talk (unstructured interview) in which the researcher contacted the person responsible for the preventive strategy of the production project. The informant described some of the key aspects of the case and their collective experiences, which provided the researcher with initial insight into the case and contributed to the topic guide in the forthcoming interviews. These topics included expectations of the distance tag, how information about the distance tag was distributed, and how the distance tag worked in the setting of the production. The following phase included semistructured interviews with 8 individuals. Interview candidates were selected and contacted using a snowball approach, in which the first respondents suggested others who had insights into the preventive strategy. The participants were strategically selected [46] based on their position in the production project, their experiences from working in the studied setting, and their engagement with the technologies in focus. The selected interviewees had worked at the site of production for 6 weeks and had experience using and helping others to use the distance technology. These interviewees had management roles, technological roles, or safety-related roles in the production project and therefore had a good overview of the perspectives of both participants and management. Thus, they could provide an overview and a holistic perspective on the role of this preventive technology during the entire production process. Two individual interviews and 1 group interview with 6 individuals were conducted in March and April 2021 (interviews A, B, and C) using a videoconferencing software. All interviews

were conducted in Swedish; direct quotes from the interviews were translated by the authors and proofread by a translation service.

Survey

On the basis of the findings of the interviews, a questionnaire was developed to survey the individual experiences of the distance tag. It was based on the key topics and relevant issues from the interviews. The purpose was to reach a larger set of users and thereby gain a broader perspective on experiences. The questionnaire (referred to as the survey) was offered to the participants in Swedish, and the translation of the original text is available in [Multimedia Appendix 1](#). The survey included both closed questions (multiple choice with a set of alternatives or a rating scale) and open-ended questions (text-based responses), the latter of which provided opportunities for participants to expand the answers beyond the set alternatives and to motivate and explain their choices. The exploratory sequential approach and the inclusion of open-ended responses were motivated by the nature of the rapidly evolving pandemic situation and the emerging field of technologies that is being studied in this case. To our knowledge, no preexisting theory or assessment scale has been developed for this specific context. After this 6-week production had ended, a link to the survey was shared via an app used in the production project. A reminder email with a link to the survey was sent so that the survey would be available for everyone who had been involved in the music competition and production (approximately 400 individuals). A total of 30 individuals responded to the questionnaire. A time delay of approximately 2 weeks between the production project

ending and the first distribution of the survey may have affected the response rate because by that time, many of the participants may not have been actively staying up to date with information in the project. Participation in interviews and the survey was voluntary, informed consent was collected, and no personal health data were collected during this study.

Data Analysis

The interview responses were analyzed using thematic analysis. The approach was inductive, as the themes were generated based on the collected data, that is, the users' experiences of a technological approach to support the maintenance of physical distance in their workplace. The coding and theme generation were descriptive and conducted on a semantic level; that is, based on the interviewees' explicit content, no assumptions were made regarding the latent underpinnings of the data [47].

We read the transcribed interviews several times. We highlighted parts of the text that focused on the aim and research questions. The highlighted parts of the text were then condensed into 1 or 2 sentences, and sentences were grouped into categories. As a final step in the qualitative analysis, key topics were identified and used as a basis for quantitative data collection through the questionnaire. The open-ended responses were analyzed based on their semantic content and combined with other results in the final stage. The first subthemes were generated from the qualitative analysis based on data from interviews and open-text responses in the survey, after which the subthemes were merged to create 2 main themes. The quantitative data were subsequently integrated into these themes.

The quantitative data from the survey primarily consisted of categorical choices and ordinal Likert-type scales, survey data of the closed questions are available in [Multimedia Appendix 2](#). The results are presented using descriptive statistics, mainly frequency distributions and the mean or median for the central tendency. A potential change in attitudes about using distance tags before and after the 6-week production project was tested using the Wilcoxon signed-rank test, which is a nonparametric test to account for different measurements of the same individuals for ordinal scale data [48]. The null hypothesis was that there would be no significant difference between attitudes before and after using the distance tag in the workplace.

Ethical Considerations

On the basis of the general guidelines of the Swedish Ethical Review Authority, this study was not classified under the requirements for ethics approval in Sweden [49]. This was a retrospective case study, meaning that we as researchers were not involved in implementing the technology, and the research study began at the end of the music competition to compile the learnings from this case. The respondents were asked to describe their experiences with the technology. From an ethics standpoint, this was an observational study. Furthermore, the study did not set up individual registers or collect personal identifiable information or health data nor did it collect information about religion, sexual orientation, or political preferences. Information about the total number of infected cases was summative and disclosed by the management; no such data were collected from

the study participants, nor could those cases be connected to any specific individual.

All participation was voluntary, requiring opt-in to consent, and respondents could opt out at any time. Before the interviews, the participants were informed about the study both orally and in writing. The survey participants were informed and asked to participate in writing. The informed consent form and oral information were provided in Swedish. The information included the purpose of the study, how the research was to be conducted and reported, and how the collected data were to be stored; the fact that participation was voluntary; the expected duration of the participation; and expected benefits and foreseeable discomforts to the prospective participants. The information provided also stated that the survey respondents were anonymous, and the identities of the interviewees were confidential.

Before the interview began, the participants were asked orally if they had read the information, if they had any questions, and if they consented to participate. The survey participants' consent was ensured by an information text at the end of the survey, stating that if they chose to send the survey, they agreed that their answers could be included in the study. No compensation was offered to participants.

Results

The Case Background

This case involved the production of a publicly funded music competition that took place in Sweden in February and March, lasting for 6 weeks in 2021. In addition to artists, production team members, and technical staff, this yearly competition usually attracts large live audiences, but owing to the pandemic, there was no large live audience in this particular year. Challenges remained because of the large number of people involved in the music competition and production (approximately 400 individuals). To be able to go through with the production and ensure the best possible safety for the participants, a substantial preventive strategy was planned based on national public health recommendations. The preventive strategy included stations for handwashing, forming "social bubbles" with allocated restrooms, measures to ensure physical distancing, and, when the appropriate distance could not be maintained, the use of face masks.

Physical distancing (≥ 1.5 m) was among the most important recommendations in Sweden, and therefore a key issue was finding a way to support distancing behavior and ensure that distance was maintained throughout the workday and production project. The production's COVID-19 strategist decided to test a proximity-based wearable technology that continuously measured the physical distance between individuals. The function is that, if the wearables come within 1.5 m of each other (ie, the individuals wearing them are not keeping their distance), an alarm is tripped and the wearer is notified. Occasionally, the distance tag could be switched off when the sound is inappropriate, for example, during recording sessions.

A pilot test was conducted using a Bluetooth-based device, but the proximity-based detection was not sufficiently specific, so

this technology was deemed to be unfeasible for this production and workplace. A new type of device based on an ultrawide band (UWB), called a distance tag, performed better in proximity detection, so this technology was implemented. The physical form of the device was a card that hung around the necks of individuals entering the facilities. The proximity detection was set to 150 cm, and as the detection error margin was 20 cm, the alarm sounded when the distance between 2 persons was approximately 130 cm to 170 cm. Everyone who entered the facility was strongly advised to wear a distance tag. The device also has a mobile app with more functionalities and an option to enable tracking for contact tracing, but this app was not used during the event because of privacy concerns. In this setup, the UWB device enabled a direct reminder (nudge) for distance but no follow-up or tracking functions.

The total number of participants in the interviews was 8 (ie, 5 men and 3 women); the number of participants in the questionnaire was 30 (ie, 18 men and 12 women). The interviewees were responsible for the production, personnel, or security on-site, whereas most of the questionnaire participants worked in the production. The age distribution of the participants was between 28 and 68 (mean 46.2, SD 1.96) years.

In the analysis, 2 main themes and 7 subthemes were identified: feasibility (*implementation and acceptance—establishing a shared problem; distance tags in context—strategy, environment, and activities; understanding and learning; and accomplishing the purpose*) and design aspects (*a purposefully annoying device; timing, tone, and proximity; and additional functions*).

Feasibility

In this main theme, the focus is on the use of a wearable device to nudge for distance. Four subthemes are described in this section: *implementation and acceptance—establishing a shared problem; distance tags in context—strategy, environment, and activities; understanding and learning; and accomplishing the purpose*.

Implementation and Acceptance—Establishing a Shared Problem

The participants received information about the preventive plan before the distance tags were implemented. Some participants were curious about the technology, while others were skeptical. Several survey respondents noted that their participation was motivated by a desire to prevent infection in themselves and others:

The motivation to prevent transmission during the production as much as possible. [Survey respondent]

The others were not initially assured that this type of device was needed. Several interviewees said that not everyone considered keeping distance to be a problem; that is, they thought that they were already good at maintaining sufficient distance from others. However, using proximity technology to continuously monitor behavior made it clear that this was not the case. Several interviewees described that the device helped them be attentive to distance and taught them how much distance they needed to maintain:

Distance can be challenging. When you start using the tags, you understand more “how much distance.” And you understand that what you had before might not have been a [sufficient] distance... [Interview C]

I’ve heard about “keep the distance.” And of course, we have structured our work to avoid getting too close. But I don’t think I understood how... That you were close until I’ve tried these... I mean, I really got to learn what a good distance is, so even if I believed that I had it before, I have, well, it has opened my eyes for it. [Interview A]

Other interviewees described that when these kinds of devices are implemented, clear and effective communication is important so that the users understand why maintaining a safe distance might be difficult:

Our perception is that we have created the conditions so that people can maintain a distance. But at the end of the day, this is an individual responsibility. [Interview C]

The interviewees also described some practical lessons learned from the implementation. At first, they started by handing out distance tags to everyone and then let users be responsible for their devices. Users were supposed to take the tag home, charge the batteries, and bring it back to the workplace every day. However, this approach did not work well because users forgot their tags at home or forgot to charge the batteries. Instead, a station was set up at the entrance so that distance tags and battery checks were available for everyone entering the site. This step improved the process and made it easier to check that the devices worked and that everyone had a device.

From the managerial perspective, the impression of acceptance was that most people accepted the distance tag sooner or later, even though the feelings were mixed at the beginning. Some felt that the distance tags were slightly disruptive and put them away or switched them off regularly. However, this situation improved after an informational campaign that focused on preventive measures, including information about the need for both face masks and distance tags, asking people questions (self-assessment), and controlling whether the devices were active. From the managerial perspective, the preventive measures, including using technological support to maintain physical distance, were ways to ensure everyone’s safety so that the production project and music competition could be carried out. Security personnel noted that users perceived these measures as a means of protecting their health. After more information was conveyed to users that the preventive measures were meant to protect not only themselves but others as well, the impression was that both adherence and acceptance increased. The interviewees’ perception was that people needed time to get used to the devices but had respect for the situation:

People were a bit unprepared, but acceptance grew over time. [Interview A]

People understood that we need this to be able to go through with an event like this during a pandemic. [Interview B]

The interviewees described another benefit of the devices, which was that the alarm could be a way to avoid conflict or feeling embarrassed about having to ask someone to step back. From the perspective of security personnel, this aspect also improved their working environment, because they did not have to interrupt people in the midst of a conversation to tell them to move back.

In the questionnaire, the participants were asked about their attitudes toward using distance tags, this included to estimate how they felt both before and after the production project. The response option in both questions was a 5-point rating scale (Table 2).

The descriptive statistics show that after the production project, there were no negative responses, the number of neutral responses decreased, and both skeptical and positive responses increased. The median value before and after distribution was *quite positive* (rating 4). The difference between attitudes before and after the intervention was not statistically significant ($W=108$; $P=.80$), so the null hypothesis cannot be rejected.

The overall experience was measured using the questionnaire with 2 questions, where the response was given on a 5-point rating scale (Table 3).

Table 3 shows that the overall experience was mainly positive as the majority of answers (47/60, 78%) were *agree* or *strongly agree*.

Table 2. Response distribution of attitudes toward using distance tags before and after the production project (N=30).

Rating	Response	Before	After
1	Negative—the tags will not work well or did not work well	1	0
2	Skeptical	5	8
3	Neutral—do not know	5	1
4	Quite positive	9	9
5	Positive—the tags will work well or worked well	10	12

Table 3. Overall experience of using the distance tags.

Questions	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
As a whole, I support this effort	0	3	3	14	10
I'd recommend the tag for other workplaces	0	3	4	11	12
Total	0	6	7	25	22

Distance Tags in Context—Strategy, Environment, and Activities

Distance tags were embedded in a larger battery of measures as part of the overall preventive strategy. The connection between the distance tags and the preventive context was a key issue for their feasibility. This holistic approach involved a combination of measures. For instance, face masks were used in situations when distance could not be maintained. In addition to face masks, other preventive measures included handwashing stations, fever scanning at the entrance, and the formation of social bubbles. The latter meant that people working closely together at the event also shared hygiene facilities. The interviewees further described that the combination of distance tags, personal protective equipment, handwashing stations, effective communication, and people overseeing the adherence to these measures was necessary for a successful preventive strategy:

Of course, there are occasions when it won't work, in this kind of production. In these cases, we have other measures. [Interview A]

It was good to continuously get a reminder of the distance, it made you more meticulous with other things such as washing your hands and using face mask. [Survey respondent]

The interviewees emphasized the importance of the event's venue as being sufficiently spacious for individuals to maintain physical distance. One of the key aspects for the successful use of distance tags is that the tags fit in the physical environment. They are particularly useful in places and situations where people meet and tend to stand close to each other or where people's paths cross, for example, by coffee machines. Manual adjustments regarding proximity detection to fit the conditions of the event could be made, but there were situations in which the distance tags were not deemed useful. Some survey respondents commented that the tags were impractical or of no use during work tasks that required close interaction:

If you already have the mindset to keep distance, the tag is not useful. Also, our team could not fully use them since we often worked closely together which made face masks the alternative. [Survey respondent]

Good idea and it probably worked well for teams other than mine, but people became more aware in shared areas. [Survey respondent]

It was also noted by several respondents that these types of devices can cause frustration when the physical setting or type of work does not allow for proper distance and the alarm goes off:

In some environments or during recordings and so forth, this noise can be disturbing. Sometimes it needs

to be switched off when people need to be close or during recording. This is the only thing I see as potentially negative. [Interview C]

It probably works well in situations where you are not dependent on staying close to others during work. [Survey respondent]

Interviewees and survey respondents described many situations in which the distance tags needed to be switched off to avoid the alarm signal, which would disturb key activities such as recording sessions. A few individuals switched off the tag

because they simply preferred to do so, indicating a need to check for adherence to safety recommendations.

In the survey, the participants were asked whether they were in the habit of switching off the distance tag in situations when it should have been used (Table 4).

The descriptive statistics show that a slight majority of the respondents, that is, 57% (17/30) of participants never or only seldom switched off the tag, while 12 did so sometimes or often and 1 person avoided using the tag as much as possible.

Table 4. Response distribution of how often the respondents switched off the tag in situations when it should have been used.

	Never—I attempted to use the tag as much as possible	I have done that very seldom	Sometimes	I often switched it off	I tried to avoid wearing the tag or have it switched on as much as possible
How often did you switch off the tag in situations when it should have been used?	13	4	6	6	1

Understanding and Learning

In its most basic function, the UWB-based distance tag is straightforward to use. It comes in the form of a card that hangs around a user's neck. The interviewees stated that the tag provided an alarm, a high-pitched sound, whenever individuals got too close to other tag users. The sound stopped as soon as the tag users moved away from one another. This behavior did not require much instruction, and the purpose was easy to understand. As the app and contact-tracing functions were not used, the hardware device was the only interface that the users had to learn. The device did not require much from its user, except for keeping the device's batteries charged and ensuring that they were turned on.

Table 5 shows that, in general, the survey respondents found it easy to learn how to use the tag: 90% (27/30) answered *agree* or *strongly agree*.

The distance tag could be switched off when the sound would have been inappropriate, but a downside of the switching-off function is that users sometimes forgot to switch it back on, as the interviewees described. The light indicating whether it was on or off was also confusing to some participants; it was red when it was active, which felt counterintuitive according to 1 survey respondent.

In addition to learning how to use the distance tag, there was also another learning process. The importance of understanding and learning how to maintain distance was described by multiple interviewees. They described a learning curve going through varied phases. In the beginning of the production, the participants had various attitudes, such as curiosity or skepticism. During the first week of use, the alarms went off frequently, which caused some frustration and reduced adherence. However, for some users, this was also a sign that they were not as good at maintaining distance as they thought:

I mean, from the beginning it was very strange, you started with "my God," you stood a bit close, and it beeped. But then you learned to step back and then you started to appreciate it, because then you get it,

you understand that it provides a function. [Interview A]

After the informational campaign organized by the COVID-19 strategist, attitudes toward the tag and adherence to preventive use improved. The interviewees described that both communication and reminders were needed for these types of protective measures to work and to ensure that people used face masks and distance tags and washed their hands. Subsequently, habits started to form, and multiple interviewees described that the alarms sounded less frequently after a while. They explained that people learned to maintain a distance. At the end of the production, people were more meticulous about all the preventive measures. They expressed more appreciation and often asked for new tags. The general impression among the interviewees was that people were generally positive toward distance tags at the end of the production project and had learned how to act while wearing them.

There were also indications that this technology has the potential to facilitate habitual changes. Several interviewees stated that they are better or much better at maintaining their distance now, even when they were not wearing their devices:

My own experience is the best point of reference. You think about it, and even when you walk with a colleague to talk, you automatically keep the distance. [Interview C]

I automatically move backwards now, even when I'm with friends. [Interview A]

I'm more attentive to the behavior of others now. I wasn't before. I thought of it as everyone's own business. Now I react when someone gets too close. I'm so aware now, what 1.5 meters really means and what it can do. [Interview B]

Table 6 shows the survey results concerning whether the tag was helpful when learning to maintain distance.

The overall results showed a slight majority of (57/90, 63%) *agree* and *strongly agree* answers, 24% (22/90) *disagree* or *strongly disagree* answers, and 12% (11/90) *neutral* answers.

Table 5. Ease of learning how to use the distance tags.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
It was easy to learn how to use the tag	0	1	2	12	15

Table 6. The influence of distance tags while learning to maintain distance.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The tag helped me to keep distance	1	7	2	13	7
I am more aware of the distance even when I do not wear the tag	1	8	5	10	6
My colleagues became better at keeping distance due to the tag	0	5	4	14	7
Total	2	20	11	37	20

Accomplishing the Purpose

When the distance tags were first implemented, managerial staff had varying expectations. Some were skeptical and others were not sure what to expect or how the system would work. Some participants also expressed interest and curiosity. After a few weeks of using the distance tags, the general perception of the interviewees and most survey respondents was that the tags had been successful and fulfilled their purpose:

I was skeptical in the beginning but now I think this is a winning concept. [Interview B]

I was a bit hesitant to how it would sound, how disruptive this would be. It showed to be something you adapted to very quickly and it is more helpful than disruptive. [Survey respondent]

I didn't know what to expect...Will this work?...Are they reliable?...But now I think it has worked great. They have absolutely fulfilled the function we thought it would. [Interview A]

In addition to supporting physical distancing, the interviewees' perceptions were that the distance tag helped safeguard the production because using these devices to monitor distance helped ensure safety on-site. The safety personnel said that the technology made sure that they "did not have to be everywhere all the time, to point out the distance" (Interview C). The tag took care of this basic control measure, and the tags were available everywhere. The interviewees expressed that the tags were a good complement to the safety personnel, who only had to ensure that people wore tags and that they had sufficient battery power. Furthermore, if not for the tags, the event would

Table 7. Anticipated positive effects of using the distance tags.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
As a whole, the tag is effective for changing individual behavior	0	3	3	17	7
I believe that the tag has helped us to reduce the spread of virus	0	3	10	12	5
Using the tag contributed to feeling safe and secure	1	7	8	13	1
Total	1	13	21	42	13

Design Aspects

In this main theme, the focus is on the experiences of the distance tag's functionality and interaction. Three subthemes

have required more security personnel to ensure that people maintained sufficient distance. However, 1 interviewee pointed out the following:

Adding more personnel is not always a solution during a pandemic. This would have added more potential sources for infection. [Interview C]

Although interviewees expressed some concerns about the distance tags at the beginning of the production project, several of them experienced that the tags were necessary for maintaining physical distance at the workplace. One interviewee indicated that distance tags were essential for carrying out the event during the pandemic:

If we hadn't used these devices, we wouldn't have been able to go through with the event. [Interview B]

Similar attitudes toward the positive effects of using the tags can be seen in the questionnaire answers (Table 7). The overall results showed of the 90 answers, a slight majority of agree and strongly agree answers 55 (61%), 14 (16%) disagree or strongly disagree answers, and 21 (23%) neutral answers.

The most negative comments related to the functionality of the device concerned battery life, which was deemed unsatisfactory. This point was mentioned by several respondents.

Individuals working at the event were advised to be careful, but they were not isolated from the outside world. Approximately 400 individuals were in place in total, when artists and their entourages were included. During the 6 weeks of the production project, 3 individuals in total were infected with the SARS-CoV-2 according to information from management. No further transmission was found in the workplace.

are described in this section: *a purposefully annoying device; timing, tone, and proximity; and additional functions.*

A Purposefully Annoying Device

Several interviewees described that the device had a high-pitched noise, which caused irritation and frustration, especially in situations in which individuals were prevented from maintaining distance or in situations where the noise caused disruption:

[The device caused] a constant beeping in a large production... [Survey respondent]

The device evoked various emotions among the participants. Some experienced frustration, while others felt that it provided safety by ensuring that everyone on-site maintained a sufficient distance:

It felt safe to use since it is easy to forget about the distance. [Survey respondent]

I am incredibly positive! It has enabled us to be more attentive and, in that way, show respect for others. [Survey respondent]

Table 8 shows how the survey respondents perceived the anticipated negative effects of using the distance tags.

The overall results showed that of the 90 answers, 19 (32%) agree and strongly agree answers, 35 (58%) disagree or strongly

Table 8. Anticipated negative effects of using the distance tags.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The tag contributed to frustration	3	8	2	14	3
The tag contributed to fatigue	18	6	4	2	0
Total	21	14	6	16	3

Table 9. How the audio signal was perceived.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The signal is intolerable, and it should be changed	10	9	7	1	3
The signal serves its purpose	2	1	2	12	13

Table 10. Survey respondents adapting to the audio signal.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Generally, it is easy to avoid the audio signal if you keep sufficient distance	0	5	2	14	9
I got used to the signal and could sometimes ignore it	10	6	0	11	3
Total	10	11	2	25	12

Timing, Tone, and Proximity

The distance tag gave off an alarm lasting a second or so after the distance was breached, for example, after passing another individual in a corridor. Some thought that this was disruptive, while others said that it was an effective way to know that the tag was active.

Table 11 shows opinions from the survey respondents concerning whether the audio signal should be activated when users pass each other or only if they stop too near each other. The result shows preference for the alarm sounding only if 2 individuals stop too close to each other. The option of having

disagree answers, and 6 (10%) neutral answers. The respondents perceived that distance tags contributed to frustration more than they contributed to fatigue. Others commented that the alarm sound must be high pitched and annoying so that it fulfills the required effect, that is, so that people take immediate action:

I'm thinking that, it is good with a "bad" sound, because then you will move away. [Interview C]

It is irritating but it needs to be. [Interview A]

Had we had a more pleasant or a tone of increasing volume, we would not have reacted as quickly. [Interview B]

Most of the survey respondents also found that the alarm signal served its purpose: 83% (25/30) answered agree or strongly agree (Table 9).

Table 10 shows whether survey respondents adapted to the audio signal by keeping sufficient distance or by ignoring the signal. Most of the respondents learned to keep safe distance: 76% (23/30) answered agree or strongly agree. No respondents were neutral about the suggestion that the alarm could be ignored: 46% (14/30) answered agree or strongly agree and 53% (16/30) answered disagree or strongly disagree.

the alarm triggered even when individuals only passed each other got equally supported and opposed by the survey respondents.

Several interviewees and survey respondents asked for a vibration function that could be an alternative when an audio signal was not suitable:

[It works] fine, but it would be good if you could switch to a vibration on occasion such as sound recording. [Survey respondent]

One interviewee said that she got used to the alarm and indicated a risk that one could start ignoring it after a while. It was

proposed that the tone could change every time and avoid getting used to the tone. Others suggested that the tag could provide a prewarning, in the form of a softer sound or vibration, before the proximity is reached where the high-pitched alarm starts. Having a prealarm would let the user know that they were getting close to the threshold, and then they could potentially avoid the disturbing sound. However, most interviewees agreed that it was important for the device to be immediate when a certain proximity was reached. In addition, some respondents stated that the tone should neither be gradually increasing nor pleasant.

Table 12 shows the survey respondents' opinions on 2 suggestions for the enhancement of the audio signal. A slight majority (17/30, 57%) of the respondents *strongly disagreed* or *disagreed* that a weaker signal would have the same effect, 23% (7/30) *strongly agreed* or *agreed* that a weaker signal would have the same effect, and 20% (6/30) were *neutral*. A total of 57% (17/30) of participants *agreed* or *strongly agreed* that there should be a prealarm warning, 30% (9/30) *disagreed* or *strongly disagreed*, and 13% (4/30) were *neutral*.

Table 11. When should the signal be heard?

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
It is appropriate that the signal is heard when we pass each other	4	9	4	9	4
It would be better if the signal is heard only if we stop close to each other, eg. after 2-3 seconds	5	3	3	13	6

Table 12. Suggestions for enhancement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I believe that a weaker signal could have the same effect on my behavior	8	9	6	4	3
There should be a prior warning, eg. beep or vibration, before the audio signal starts	8	1	4	8	9
Total	16	10	10	12	12

Additional Functions

The interviewees and survey respondents agreed that the key function is the immediate warning provided by the distance tag. Some were interested in additional functions, such as data on the number of close interactions, or the use of these devices for contact tracing. Others wanted a battery charge indicator related to the previously mentioned complaints regarding the short battery time. Less than half of the survey respondents were interested in additional functions or in using an app connected to the device (12/30, 40% of respondents). However, attitudes differed and most clearly rejected those ideas (14/30, 46% of respondents) or were hesitant to add more features (3/30, 10% of respondents):

Absolutely NOT an app! But they could have a light indicating when the battery charge is becoming low. [Survey respondent]

No, it should be uncomplicated. It would be good to see the percentage of the battery charge. [Survey respondent]

To fit the settings of the event, the event management set the distance for the tag alarm. The proximity detection was set to 150 cm, and the interviewees agreed that this was a good distance. The detection varied from +20 cm to -20 cm, which means that the alarm would sound from 130 cm to 170 cm, which seemed a reasonable distance in relation to the work tasks and the physical setting. At this alarm distance, most work activities were performed. According to interviewees, 2 full meters would have been too far. Interviewees experienced that the adjusted distance enabled people to talk to each other without raising their voices and they could socialize at the event, although with some distance:

Now, we can have lunch together, but we sit with 1.5 meters apart, and we can be sure that it is exactly that distance. We don't bring an extra chair, then the alarm would set off. But it helps us to hang out, we can sit eye to eye and talk without a screen between us. I think it has assisted us in socializing. [Interview B]

No, I think it (tracking) would be a bit of an invasion of privacy, anyway, even if it had not been legislated. And I believe there is a side effect—in that you deprive the individual of their responsibility. [Interview B]

The primary function is that they keep the distance...But I would have loved to use its full potential with all the functionality and using the app as well. Partly because it would be a fun thing, a gadget, it is making the product more interesting if you can check the app. [Interview C]

It [tracking] could probably be something that would contribute to increased usage. Sounds like a smart idea for contact tracing. [Survey respondent]

Many respondents were satisfied with the functions and did not wish for any additional tracking functions:

I think these are good. They provide us with freedom to move around. And they build on individual responsibility which is particularly important. We cannot organize in a way so that we constantly must check [the distance]. This is a very good support. [Interview B]

Summary of Findings

In total, 7 subthemes were constructed based on the qualitative analysis in this study. Key findings from the survey were added

to these themes, and additional descriptive statistics are provided in [Multimedia Appendix 1](#). The themes are summarized in [Table 13](#), along with the key findings from both interviews (qualitative) and the survey (quantitative and qualitative).

Table 13. Themes, subthemes, and key findings of this study.

Theme, subtheme	Key finding
Feasibility	
Implementation and acceptance—establishing a shared problem	<ul style="list-style-type: none"> • Communication and clarifying the problem is important. • Respondents reported that they had overestimated their capacity to maintain a sufficient distance. • A station at the entrance can simplify the handout process. • Attitudes vary and change over time; most users are positive afterward. • A few are skeptical of the technology, but a majority support the intervention afterward.
Distance tags in context—strategy, environment, and activities	<ul style="list-style-type: none"> • Distance tags are best suited as part of a larger preventive strategy. • They are particularly useful in situations where individuals meet and tend to stand close by habit. • Negative attitudes are often connected to situations where the tag is not useful or too disruptive for the situation. • Usability depends on context and situation. • Alternatives to the high-pitched alarm are needed for certain situations.
Understanding and learning	<ul style="list-style-type: none"> • The distance tag was easy to use and understand. • The indicator light was a little counterintuitive, and individuals sometimes forgot to switch the tag back on. • It took a while to learn a new behavioral pattern; some expressed frustration in the beginning. • There are indications that increased attention to distance remains even when the tag is removed.
Accomplishing the purpose	<ul style="list-style-type: none"> • Most respondents agree that it supported physical distancing and behavior change. • Users report more positive than negative effects. • This technology is feasible when used in the right circumstances. • A short battery life is the most negative aspect. • Few got infected during the 6-week production project, indicating a preventive potential of the strategy as a whole.
Design aspects	
A purposefully annoying device	<ul style="list-style-type: none"> • Multiple respondents agree that the high-pitched alarm is needed for immediate behavioral adaption—this should be the standard setting. • A discrete nudge would not be sufficient in the long term. • There are indications that users might get desensitized even to a high-pitched sound. • The sound can cause frustration, but few people experience fatigue.
Timing, tone, and proximity	<ul style="list-style-type: none"> • Timing and proximity were satisfactory. • The tone fulfills its purpose, but alternatives are requested by some. • Alternatives for certain situations can involve other measures, such as a switching-off function or vibration.
Additional functions	<ul style="list-style-type: none"> • Some are interested in tracking functions and additional data, but it is not deemed necessary. • A battery charge indicator is requested. • Some are negative to tracking for privacy reasons. • The direct warning enables individual responsibility, which is the most valuable function.

Discussion

Principal Findings

A pandemic consists of a complex system of interacting factors, and many factors must be considered to prevent transmission. The behavior of individuals is a key issue in the transmission process [20] and was the focus of attention in the preventive strategy undertaken in this case. This study provides evidence for the rapidly emerging field of mobile preventive technology. Specifically, this study contributes to our understanding of the feasibility, use, and design of a specific wearable technology

meant to maintain or improve physical distancing among its users.

Previous studies have reviewed or suggested a variety of designs that allow proximity detection and encourage physical distancing behavior [17-19,50]. However, few empirical studies have investigated the technical efficacy of proximity detection [19].

The objective of this study was to evaluate the feasibility of using a wearable device that nudges its user to maintain physical distance while also gathering information about how functionality and interactions with this device were experienced

by the users. The following discussion focuses on the feasibility and design of distance tags.

The Feasibility of Proximity-Based Technologies in a Pandemic Context

This case study illustrates that proximity-based nudging technologies are a feasible strategy for infectious disease management in a workplace where many individuals are in physical proximity and move around. Most users in this study supported the technology used in this intervention, would recommend it to others, agreed that it increased the user's awareness of physical distancing, and believed that it was effective in changing individual behavior. A few were skeptical and at least 1 user was dissatisfied. Negative comments mostly concerned a poor fit between the technology and the work tasks or that there were situations in which physical distancing or a high-pitched audio signal were not appropriate. These comments highlight some opportunities for design (discussed in subsequent sections) and show that the usability of this technology is dependent on the context of use and the situation.

A founding idea behind nudging is to provide contextual feedback to facilitate better choices or behavior in distinctive situations [38]. Furthermore, previous studies have shown that context is important for understanding why and how nudges are effective [51]. In this case, the nudge is provided by a technology triggered by its user's close interaction with others. For this nudge to be efficient, the physical context and activities must allow the user to either avoid or rapidly respond to the nudge. Therefore, before implementing this type of technology, it is recommended to undertake a systems thinking approach to understand how a nudge can be designed in relation to its context.

Systems thinking is the fundamental, conceptual pattern that makes it possible to ensure a holistic understanding of the situation [52]. This could involve modeling the problem and designing a preventive strategy as a system, defining the boundaries of the system in relation to the risk of transmission, the stakeholders involved and their perspectives, and their relations and activities in the context.

The choice of UWB instead of a Bluetooth device is in line with other recent studies on the efficacy of Bluetooth devices for contact-tracing purposes, indicating that proximity detection between individuals can be challenging in practice [53,54], and Kindt et al [54] specifically recommended UWB technology as an alternative to Bluetooth. However, as noted by Alo et al [50], UWB is more energy demanding than Bluetooth, and in this study, the battery life of the distance tags was deemed unsatisfactory.

The small number of individuals who became infected during this 6-week project is a possible indication that the overall preventive strategy—including handwashing, social bubbles, wearing face masks, and distance tags—was successful and that transmission in the workplace was minimized. To contextualize these findings, the “background” transmission of SARS-CoV-2 in Stockholm was among the highest in Sweden. At the time of this production project (weeks 5-13 in 2021), the number of infections in the city increased, peaking at 408 new cases per

100,000 individuals in week 13 [55]. This would indicate that approximately 10 new cases would be expected in a group of 400 persons during a 6-week period.

Workplaces have been important sites of transmission of COVID-19 [56], and the fact that this workplace can accommodate several individuals with a low amount of transmission is an encouraging sign. The workplace formed under this production project was not a closed system, meaning that individuals had contact with the outside world, even though they were advised to be careful. To put this in perspective, we also note that the activities undertaken in this case involved teamwork, physical labor, singing, and dancing, which theoretically may have put this workplace at a higher risk of widespread transmission. However, it is important not to draw conclusions about the preventive effects of individual preventive measures, as they were not studied separately. The combined effect of several preventive measures is an effective way to reduce transmission in public gatherings, and it is also likely that individuals that show adherence to one measure, such as consistent mask wearing, are practicing multiple measures [27]. Therefore, more studies are needed to confirm and ensure a sufficient preventive effect of nudging technologies to support preventive behavior.

Design Aspects

A key finding regarding the design of distance tags is the users' experience of the audio signal. The sound can be disruptive, but it did not cause fatigue. Generally, most respondents agreed that the audio signal needs to be annoying to have the intended effect.

Most users also indicated that their attention to distance and their long-term behavior were affected by wearing this device, an interesting finding that might be explained by the design and timing of the audio feature. The just-in-time prompt (the nudging feature) in this technology is in line with the Skinner operant conditioning [57,58], meaning that the high-pitched audio signal provides immediate feedback whenever a user is in close proximity to another user in a systematic and repetitive manner. The annoying signal is a form of positive punishment intended to minimize unwanted user behavior. The signal persists until the user returns, at which point the signal immediately stops. Withdrawal of unwanted signals is a form of negative reinforcement aimed at increasing the prevalence of distancing behavior. This audio feedback model can explain some users' reports that they learned new and automated behavior after using the distance tag. Both approaches to behavioral design (nudging and operant conditioning) can be connected to behavioristic traditions that emphasize the effect of external stimuli on human behavior [57]. Specifically, we propose that nudging is useful in situations in which users fail to act in their best interests because of cognitive limitations, such as memory, attention span, or habits. Models, such as the Skinner operant conditioning, can provide more specific guidance for the design of a nudge's interactive details, such as timing and pitch.

The other design aspects in this study concerned the types of feedback and information provided to users. For some users, having additional tracking functions and data on interactions could potentially be useful or motivational, although these

features were not deemed necessary. At the same time, some users felt quite negatively about tracking and contact-tracing functions due to privacy and ethical concerns. The immediate warning provided sufficient support for this workplace to remain open during times of rising transmission in the region. The value of this technology was the opportunity for individuals to be able to act immediately and by this to be responsible for their own behavior. Compared with previous design proposals [17,18], the distance tag provided the same function, but in a more simplified design. By providing direct feedback and a clear mapping between the behavior and the feedback, it is straightforward and transparent.

There were indications in this study that some users may have become desensitized to the frequent alarm, even with a high-pitched tone. Thaler and Sunstein [38] noted that feedback for warning systems must be designed so that they do not provide warnings too frequently, because users will start to ignore them. The risk of desensitization is supported by previous research on other types of proximity systems used in the construction industry [30,31]. In this study, several participants asked for a vibration function instead of the audio alarm; at the same time, users clearly stated that too discrete a signal would be too easy to ignore. We can assume that a discrete notification used for longer durations would result in rapid desensitization, and the users would start to ignore the feedback entirely. We propose that a vibration should be available as a short-term alternative only when an audio signal is not appropriate or as a prealert to the audio signal.

It is possible that a warning system in 2 stages would be more advantageous, and we propose including a *dual nudge* in the design. In a dual nudge, the device would first vibrate, followed by high-pitched feedback if the situation is not corrected (in this case, if the user does not step back) a few seconds later. This idea is supported by previous research on warning systems in the automobile industry [35,37], which indicated that prealerts could help users gain control of the situation more quickly. Suzuki et al [35] suggested a combination alert, in which a vibration is followed by an audio signal: vibrations were found to be most appropriate for unpredictable conditions, whereas audio signals were best suited in predictable conditions. It has been shown that the response time to vibro-tactile signals is shorter than that of the other modes of alerts [31], which would also support the use of vibrations as the first step of a dual nudge. Furthermore, the user can more easily avoid the audio signal if a prewarning is delivered, which might alleviate some irritation. It may also improve the learning process if the user can quickly adapt to the behavior and avoid negative re-enforcement caused by the annoying sound. Less-frequent alarms would lower the risk of redundant alerts, alarm fatigue, and a desensitized user. The optimal timing of these 2 types of feedback (vibration and audio signal) and the relation to proximity needs to be further researched to optimize the speed of behavioral adaptation as well as the user experience.

Limitations

This study had several limitations. First, 8 interviews and 30 survey responses constituted a small sample size, which limits generalizability and prevents us from undertaking meaningful

statistical comparisons within the sample group. Second, the interviews were conducted in the final week of the production and the survey a few weeks after the production project ended, so there may have been problems such as recall bias and the formulation of accurate recollections from the first weeks of production. Third, although the mixed methods and explorative approach enabled new and important insights into how and why things worked in practice, there were also fewer opportunities to make causal connections (compared with a controlled experiment). A major limitation of this study was the lack of a comparison group. In addition, any potential preventive effect of the distance tag cannot be distinguished from the effects of other preventive measures in place at the time, and it is likely that a combination of measures contributed to the observed low transmission rate. A final potential limitation is that this study was conducted in a Swedish context in a very specific workplace setting; even so, we think that our findings and conclusions may be applicable to other settings, for example, in other workplaces.

Conclusions

This empirical study reports on the feasibility of using wearable technology to nudge individuals to maintain a safe distance in their workplace during a pandemic. The technology is particularly useful in places and situations where people meet and tend to stand close to each other, and it supports the attention to distance. The usability is dependent on the context and situation, which are crucial for the user's ability to adapt. In situations where alarms are unavoidable or unsuitable, distance tags can be experienced as more frustrating than helpful. The study also demonstrated that this type of device is easy to understand and use, and it can be rapidly implemented with a handout station on-site. However, a learning curve needs to be considered in which the user gradually adapts to a new behavior, and users can expect more frequent alarms in the beginning. It is important that managers communicate and clarify the shared *problem* of physical distancing. Information and moral perspectives such as the need to protect others can facilitate acceptance and adherence. Most users agree that the audio signal needs to be irritating, and *purposefully annoying* feedback is suggested to be included in the design, to ensure sufficient behavioral adaptation. Furthermore, we propose a *dual nudge* that involves a vibration followed by a sound to minimize the risk of desensitization. There are indications that the technology facilitates learning how to maintain a greater distance from others and that the behavior change can persist beyond the context of technology use.

This study concludes that nudging technologies based on proximity detection can be used to support this type of preventive behavior, focusing on maintaining physical distance from others. They facilitate physical distancing by providing a just-in-time prompt without the need for tracking contacts. This study provides insights into the emerging field of personal and wearable technologies used for primary preventive purposes during infectious disease outbreaks. Future research is needed to establish their preventive effects.

Furthermore, it would be interesting to explore the feasibility of this technology for outbreaks of other contagious diseases, particularly where transmission is dominated by close contact

or respiratory droplets. Another avenue for future research is to investigate behavior change in more detail and in relation to the different forms of feedback provided by nudging technologies.

Acknowledgments

The authors want to thank the production staff and participants from the music competition in this event, who, with their time and commitment, participated in this study. We are also grateful for the support from Swedish Television, which made this study possible.

Data Availability

All quantitative data generated and analyzed during this study are included in this published article and its supplementary information files. Owing to ethical considerations, the qualitative data generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions.

[[DOCX File, 28 KB - formative_v6i12e39570_app1.docx](#)]

Multimedia Appendix 2

Survey data closed questions.

[[XLSX File \(Microsoft Excel File\), 12 KB - formative_v6i12e39570_app2.xlsx](#)]

References

1. Bedford J, Enria D, Giesecke J, Heymann DL, Ihekweazu C, Kobinger G, WHO Strategic and Technical Advisory Group for Infectious Hazards. COVID-19: towards controlling of a pandemic. *Lancet* 2020 Mar 28;395(10229):1015-1018 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(20\)30673-5](https://doi.org/10.1016/S0140-6736(20)30673-5)] [Medline: [32197103](https://pubmed.ncbi.nlm.nih.gov/32197103/)]
2. Ye J. The role of health technology and informatics in a global public health emergency: practices and implications from the COVID-19 pandemic. *JMIR Med Inform* 2020 Jul 14;8(7):e19866 [[FREE Full text](#)] [doi: [10.2196/19866](https://doi.org/10.2196/19866)] [Medline: [32568725](https://pubmed.ncbi.nlm.nih.gov/32568725/)]
3. Gasser U, Ienca M, Scheibner J, Sleigh J, Vayena E. Digital tools against COVID-19: taxonomy, ethical challenges, and navigation aid. *Lancet Digit Health* 2020 Aug;2(8):e425-e434 [[FREE Full text](#)] [doi: [10.1016/S2589-7500\(20\)30137-0](https://doi.org/10.1016/S2589-7500(20)30137-0)] [Medline: [32835200](https://pubmed.ncbi.nlm.nih.gov/32835200/)]
4. Jacob S, Lawarée J. The adoption of contact tracing applications of COVID-19 by European governments. *Policy Des Pract* 2021;4(1):44-58. [doi: [10.1080/25741292.2020.1850404](https://doi.org/10.1080/25741292.2020.1850404)]
5. Jones K, Thompson R. To use or not to use a COVID-19 contact tracing app: mixed methods survey in Wales. *JMIR Mhealth Uhealth* 2021 Nov 22;9(11):e29181 [[FREE Full text](#)] [doi: [10.2196/29181](https://doi.org/10.2196/29181)] [Medline: [34698645](https://pubmed.ncbi.nlm.nih.gov/34698645/)]
6. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A survey of COVID-19 contact tracing apps. *IEEE Access* 2020 Jul 20;8:134577-134601. [doi: [10.1109/access.2020.3010226](https://doi.org/10.1109/access.2020.3010226)]
7. Zimmermann BM, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. *J Med Internet Res* 2021 Feb 08;23(2):e25525 [[FREE Full text](#)] [doi: [10.2196/25525](https://doi.org/10.2196/25525)] [Medline: [33503000](https://pubmed.ncbi.nlm.nih.gov/33503000/)]
8. Indicator framework to evaluate the public health effectiveness of digital proximity tracing solutions. World Health Organization, European Center for Disease Prevention and Control. Geneva, Switzerland: World Health Organization; 2021. URL: <https://apps.who.int/iris/bitstream/handle/10665/341818/9789240028357-eng.pdf?sequence=1&isAllowed=y> [accessed 2022-03-30]
9. Villius Zetterholm M, Lin Y, Jokela P. Digital contact tracing applications during COVID-19: a scoping review about public acceptance. *Informatics* 2021 Jul 22;8(3):48. [doi: [10.3390/informatics8030048](https://doi.org/10.3390/informatics8030048)]
10. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. *Ir J Med Sci* 2021 Aug;190(3):863-887 [[FREE Full text](#)] [doi: [10.1007/s11845-020-02389-y](https://doi.org/10.1007/s11845-020-02389-y)] [Medline: [33063226](https://pubmed.ncbi.nlm.nih.gov/33063226/)]

11. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mólken M. COVID-19 contact tracing apps: predicted uptake in the Netherlands based on a discrete choice experiment. *JMIR Mhealth Uhealth* 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: [10.2196/20741](https://doi.org/10.2196/20741)] [Medline: [32795998](https://pubmed.ncbi.nlm.nih.gov/32795998/)]
12. Morley J, Cows J, Taddeo M, Floridi L. Ethical guidelines for COVID-19 tracing apps. *Nature* 2020 Jun;582(7810):29-31. [doi: [10.1038/d41586-020-01578-0](https://doi.org/10.1038/d41586-020-01578-0)] [Medline: [32467596](https://pubmed.ncbi.nlm.nih.gov/32467596/)]
13. Sharon T. Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. *Ethics Inf Technol* 2021;23(Suppl 1):45-57 [FREE Full text] [doi: [10.1007/s10676-020-09547-x](https://doi.org/10.1007/s10676-020-09547-x)] [Medline: [32837287](https://pubmed.ncbi.nlm.nih.gov/32837287/)]
14. Vogt F, Haire B, Selvey L, Katelaris AL, Kaldor J. Effectiveness evaluation of digital contact tracing for COVID-19 in New South Wales, Australia. *Lancet Public Health* 2022 Mar;7(3):e250-e258 [FREE Full text] [doi: [10.1016/S2468-2667\(22\)00010-X](https://doi.org/10.1016/S2468-2667(22)00010-X)] [Medline: [35131045](https://pubmed.ncbi.nlm.nih.gov/35131045/)]
15. Grekousis G, Liu Y. Digital contact tracing, community uptake, and proximity awareness technology to fight COVID-19: a systematic review. *Sustain Cities Soc* 2021 Aug;71:102995 [FREE Full text] [doi: [10.1016/j.scs.2021.102995](https://doi.org/10.1016/j.scs.2021.102995)] [Medline: [34002124](https://pubmed.ncbi.nlm.nih.gov/34002124/)]
16. Braithwaite I, Callender T, Bullock M, Aldridge RW. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *Lancet Digit Health* 2020 Nov;2(11):e607-e621 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30184-9](https://doi.org/10.1016/S2589-7500(20)30184-9)] [Medline: [32839755](https://pubmed.ncbi.nlm.nih.gov/32839755/)]
17. Zetterholm M, Elm P, Salavati S. Designing for pandemics: a design concept based on technology mediated nudging for health behavior change. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*. 2021 Jan 5 Presented at: HICSS '21; January 4-7, 2021; Maui, HI, USA p. 3474-3483. [doi: [10.24251/hicss.2021.422](https://doi.org/10.24251/hicss.2021.422)]
18. Chung A, Lessard L, Andreev P, O'Reilly P. Designing information systems to break habits and promote preventive behaviours during large-scale disease outbreaks. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*. 2021 Jan 5 Presented at: HICSS '21; January 4-7, 2021; Maui, HI, USA p. 3484-3493. [doi: [10.24251/hicss.2021.423](https://doi.org/10.24251/hicss.2021.423)]
19. Ho SM, Seraj S, Yee K, Liu X, Erlebacher G. A context-aware social distancing nudge for public health in response to COVID-19 pandemic. In: *Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. 2020 Presented at: SBP-BRIMS '20; October 19-21, 2020; Virtual. [doi: [10.1007/978-3-030-61255-9_6](https://doi.org/10.1007/978-3-030-61255-9_6)]
20. West R, Michie S, Rubin GJ, Amlôt R. Applying principles of behaviour change to reduce SARS-CoV-2 transmission. *Nat Hum Behav* 2020 May;4(5):451-459. [doi: [10.1038/s41562-020-0887-9](https://doi.org/10.1038/s41562-020-0887-9)] [Medline: [32377018](https://pubmed.ncbi.nlm.nih.gov/32377018/)]
21. Lunn PD, Timmons S, Julienne H, Belton CA, Barjaková M, Lavin C, et al. Using decision aids to support self-isolation during the COVID-19 pandemic. *Psychol Health* 2021 Feb;36(2):195-213. [doi: [10.1080/08870446.2020.1849701](https://doi.org/10.1080/08870446.2020.1849701)] [Medline: [33210950](https://pubmed.ncbi.nlm.nih.gov/33210950/)]
22. Lunn PD, Belton CA, Lavin C, McGowan FP, Timmons S, Robertson DA. Using behavioral science to help fight the coronavirus. *J Behav Public Admin* 2020 Mar 29;3(1):1-15. [doi: [10.30636/jbpa.31.147](https://doi.org/10.30636/jbpa.31.147)]
23. Advice for the public: Coronavirus disease (COVID-19). World Health Organization. 2021. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public> [accessed 2022-11-09]
24. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication - a key factor in fighting the COVID-19 pandemic. *Patient Educ Couns* 2020 May;103(5):873-876 [FREE Full text] [doi: [10.1016/j.pec.2020.03.027](https://doi.org/10.1016/j.pec.2020.03.027)] [Medline: [32336348](https://pubmed.ncbi.nlm.nih.gov/32336348/)]
25. Bagheri G, Thiede B, Hejazi B, Schlenczek O, Bodenschatz E. An upper bound on one-to-one exposure to infectious human respiratory particles. *Proc Natl Acad Sci USA* 2021 Dec 07;118(49):e2110117118 [FREE Full text] [doi: [10.1073/pnas.2110117118](https://doi.org/10.1073/pnas.2110117118)] [Medline: [34857639](https://pubmed.ncbi.nlm.nih.gov/34857639/)]
26. Tabatabaeizadeh SA. Airborne transmission of COVID-19 and the role of face mask to prevent it: a systematic review and meta-analysis. *Eur J Med Res* 2021 Jan 02;26(1):1 [FREE Full text] [doi: [10.1186/s40001-020-00475-6](https://doi.org/10.1186/s40001-020-00475-6)] [Medline: [33388089](https://pubmed.ncbi.nlm.nih.gov/33388089/)]
27. Doung-Ngern P, Suphanchaimat R, Panjangampathana A, Janekrongtham C, Ruampoom D, Daochaeng N, et al. Case-control study of use of personal protective measures and risk for SARS-CoV 2 infection, Thailand. *Emerg Infect Dis* 2020 Nov;26(11):2607-2616 [FREE Full text] [doi: [10.3201/eid2611.203003](https://doi.org/10.3201/eid2611.203003)] [Medline: [32931726](https://pubmed.ncbi.nlm.nih.gov/32931726/)]
28. Gasteiger N, Dowding D, Ali SM, Scott AJ, Wilson P, van der Veer SN. Sticky apps, not sticky hands: a systematic review and content synthesis of hand hygiene mobile apps. *J Am Med Inform Assoc* 2021 Aug 13;28(9):2027-2038 [FREE Full text] [doi: [10.1093/jamia/ocab094](https://doi.org/10.1093/jamia/ocab094)] [Medline: [34180527](https://pubmed.ncbi.nlm.nih.gov/34180527/)]
29. Oyibo K, Morita PP. Designing better exposure notification apps: the role of persuasive design. *JMIR Public Health Surveill* 2021 Nov 16;7(11):e28956 [FREE Full text] [doi: [10.2196/28956](https://doi.org/10.2196/28956)] [Medline: [34783673](https://pubmed.ncbi.nlm.nih.gov/34783673/)]
30. Chan K, Louis J, Albert A. Incorporating worker awareness in the generation of hazard proximity warnings. *Sensors (Basel)* 2020 Feb 02;20(3):806 [FREE Full text] [doi: [10.3390/s20030806](https://doi.org/10.3390/s20030806)] [Medline: [32024244](https://pubmed.ncbi.nlm.nih.gov/32024244/)]
31. Huang Y, Hammad A, Zhu Z. Providing proximity alerts to workers on construction sites using Bluetooth Low Energy RTLS. *Autom Constr* 2021 Dec;132:103928. [doi: [10.1016/j.autcon.2021.103928](https://doi.org/10.1016/j.autcon.2021.103928)]
32. Kim Y, Baek J, Choi Y. Smart helmet-based personnel proximity warning system for improving underground mine safety. *Appl Sci* 2021 May 11;11(10):4342. [doi: [10.3390/app11104342](https://doi.org/10.3390/app11104342)]

33. Dixon BJ, Daly MJ, Chan H, Vescan A, Witterick IJ, Irish JC. Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery. *Laryngoscope* 2014 Apr;124(4):853-859. [doi: [10.1002/lary.24385](https://doi.org/10.1002/lary.24385)] [Medline: [24122916](https://pubmed.ncbi.nlm.nih.gov/24122916/)]
34. Schmale IL, Vandelaar LJ, Luong AU, Citardi MJ, Yao WC. Image-guided surgery and intraoperative imaging in rhinology: clinical update and current state of the art. *Ear Nose Throat J* 2021 Dec;100(10):NP475-NP486 [FREE Full text] [doi: [10.1177/0145561320928202](https://doi.org/10.1177/0145561320928202)] [Medline: [32453646](https://pubmed.ncbi.nlm.nih.gov/32453646/)]
35. Suzuki K, Jansson H. An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *JSAE Rev* 2003 Jan;24(1):65-70. [doi: [10.1016/s0389-4304\(02\)00247-3](https://doi.org/10.1016/s0389-4304(02)00247-3)]
36. de Winter JC, Happee R, Martens MH, Stanton NA. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. *Transp Res Part F Traffic Psychol Behav* 2014 Nov;27:196-217. [doi: [10.1016/j.trf.2014.06.016](https://doi.org/10.1016/j.trf.2014.06.016)]
37. van der Heiden RM, Iqbal ST, Janssen CP. Priming drivers before handover in semi-autonomous cars. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017 May Presented at: CHI '17; May 6-11, 2017; Denver, CO, USA p. 392-404. [doi: [10.1145/3025453.3025507](https://doi.org/10.1145/3025453.3025507)]
38. Thaler RH, Sunstein CR. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York, NY, USA: Penguin Publishing Group; 2009.
39. Hansen PG. The definition of nudge and libertarian paternalism: does the hand fit the glove? *Eur J Risk Regul* 2016;7(1):155-174. [doi: [10.1017/s1867299x00005468](https://doi.org/10.1017/s1867299x00005468)]
40. Thaler RH, Sunstein CR. Libertarian paternalism. *Am Econ Rev* 2003 May;93(2):175-179. [doi: [10.1257/000282803321947001](https://doi.org/10.1257/000282803321947001)]
41. Caraban A, Karapanos E, Gonçalves D, Campos P. 23 ways to nudge: a review of technology-mediated nudging in human-computer interaction. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-15. [doi: [10.1145/3290605.3300733](https://doi.org/10.1145/3290605.3300733)]
42. Hansen PG, Jespersen AM. Nudge and the manipulation of choice: a framework for the responsible use of the nudge approach to behaviour change in public policy. *Eur J Risk Regul* 2013;4(1):3-28. [doi: [10.1017/s1867299x00002762](https://doi.org/10.1017/s1867299x00002762)]
43. Street C, Ward K. Retrospective case study. In: Mills AJ, Durepos G, Wiebe E, editors. *Encyclopedia of Case Study Research*. Thousand Oaks, CA, USA: Sage Publications; 2010:824-827.
44. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. 3rd edition. Thousand Oaks, CA, USA: Sage Publications; 2018.
45. Bryman A. Integrating quantitative and qualitative research: how is it done? *Qual Res* 2006 Feb;6(1):97-113. [doi: [10.1177/1468794106058877](https://doi.org/10.1177/1468794106058877)]
46. Parker C, Scott S, Geddes A. Snowball sampling. In: Atkinson P, Delamont S, Cernat A, Sakshaug JW, Williams RA, editors. *Research Design for Qualitative Research*. Thousand Oaks, CA, USA: Sage Publications; 2019.
47. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Qual Quant* 2021 Jun 26;56(3):1391-1412. [doi: [10.1007/s11135-021-01182-y](https://doi.org/10.1007/s11135-021-01182-y)]
48. Coleman JS. Wilcoxon signed ranks test. In: Frey BB, editor. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Thousand Oaks, CA, USA: Sage Publications; 2018:1815-1816.
49. Vanliga frågor: Behöver jag ansöka om etikprövning. Swedish Ethical Review Authority. 2021. URL: <https://etikprovning Smyndigheten.se> [accessed 2021-02-15]
50. Alo UR, Nkwo FO, Nweke HF, Achi II, Okemiri HA. Non-pharmaceutical interventions against COVID-19 pandemic: review of contact tracing and social distancing technologies, protocols, apps, security and open research directions. *Sensors (Basel)* 2021 Dec 30;22(1):280 [FREE Full text] [doi: [10.3390/s22010280](https://doi.org/10.3390/s22010280)] [Medline: [35009822](https://pubmed.ncbi.nlm.nih.gov/35009822/)]
51. Kosters M, Van der Heijden J. From mechanism to virtue: evaluating nudge theory. *Evaluation* 2015 Jul 07;21(3):276-291. [doi: [10.1177/1356389015590218](https://doi.org/10.1177/1356389015590218)]
52. Cabrera D, Colosi L, Lobdell C. Systems thinking. *Eval Program Plann* 2008 Aug;31(3):299-310. [doi: [10.1016/j.evalprogplan.2007.12.001](https://doi.org/10.1016/j.evalprogplan.2007.12.001)] [Medline: [18272224](https://pubmed.ncbi.nlm.nih.gov/18272224/)]
53. Pandl KD, Thiebes S, Schmidt-Kraepelin M, Sunyaev A. How detection ranges and usage stops impact digital contact tracing effectiveness for COVID-19. *Sci Rep* 2021 May 03;11(1):9414 [FREE Full text] [doi: [10.1038/s41598-021-88768-6](https://doi.org/10.1038/s41598-021-88768-6)] [Medline: [33941793](https://pubmed.ncbi.nlm.nih.gov/33941793/)]
54. Kindt PH, Chakraborty T, Chakraborty S. How reliable is smartphone-based electronic contact tracing for COVID-19? *Commun ACM* 2022 Jan;65(1):56-67. [doi: [10.1145/3471933](https://doi.org/10.1145/3471933)]
55. Cases of Covid-19 in Sweden at the Regional level. The Public Health Agency of Sweden. 2022. URL: <https://experience.arcgis.com/experience/19fc7e3f61ec4e86af178fe2275029c5> [accessed 2022-03-30]
56. Preventing and mitigating COVID-19 at work: policy brief, 19 May 2021. World Health Organization. Geneva, Switzerland: World Health Organization; 2021 May 19. URL: <https://www.who.int/publications/i/item/WHO-2019-nCoV-workplace-actions-policy-brief-2021-1> [accessed 2022-03-30]
57. Simon C, Tagliabue M. Feeding the behavioral revolution: contributions of behavior analysis to nudging and vice versa. *J Behav Econ Policy* 2018;2(1):91-97.
58. Skinner BF. Selection by consequences. *Behav Brain Sci* 1984 Feb 04;7(4):477-481. [doi: [10.1017/s0140525x0002673x](https://doi.org/10.1017/s0140525x0002673x)]

Abbreviations

CTA: contact-tracing app

UWB: ultrawide band

Edited by A Mavragani; submitted 14.05.22; peer-reviewed by ER Khalilian, C Kirkland, J Lukan; comments to author 29.08.22; revised version received 19.09.22; accepted 30.09.22; published 12.12.22.

Please cite as:

Villius Zetterholm M, Nilsson L, Jokela P

Using a Proximity-Detection Technology to Nudge for Physical Distancing in a Swedish Workplace During the COVID-19 Pandemic: Retrospective Case Study

JMIR Form Res 2022;6(12):e39570

URL: <https://formative.jmir.org/2022/12/e39570>

doi: [10.2196/39570](https://doi.org/10.2196/39570)

PMID: [36343202](https://pubmed.ncbi.nlm.nih.gov/36343202/)

©My Villius Zetterholm, Lina Nilsson, Päivi Jokela. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Perceptions of, and Obstacles to, SARS-CoV-2 Vaccination Among Adults in Lebanon: Cross-sectional Online Survey

Nadeem Elias Abou-Arraj^{1,2,3*}, MD, MPH; Diana Maddah^{4*}, MPH; Vanessa Buhamdan⁴, BSc; Roua Abbas⁴, BSc; Nadine Kamel Jawad², BA, MPhil; Fatima Karaki^{5*}, MD; Nael H Alami^{4*}, MA, MSc, PhD; Pascal Geldsetzer^{6,7*}, MD, MPH, PhD

¹School of Public Health, University of California, Berkeley, Berkeley, CA, United States

²Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States

³Division of General Medicine, Department of Medicine, University of Michigan Medical School, Ann Arbor, MI, United States

⁴School of Health Sciences, Modern University for Business and Science, Beirut, Lebanon

⁵Refugee and Asylum Seeker Health Initiative (RAHI), Department of Medicine, University of California San Francisco, San Francisco, CA, United States

⁶Division of Primary Care and Population Health, Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States

⁷Chan Zuckerberg Biohub, San Francisco, CA, United States

*these authors contributed equally

Corresponding Author:

Nadeem Elias Abou-Arraj, MD, MPH

School of Public Health

University of California

Berkeley

2121 Berkeley Way

Berkeley, CA, 94704

United States

Phone: 1 510 643 0881

Email: narraj@berkeley.edu

Abstract

Background: The COVID-19 pandemic is an additional burden on Lebanon's fragmented health care system and adds to its ongoing political, economic, and refugee crises. Vaccination is an important means of reducing the impact of the pandemic.

Objective: Our study's aims were to (1) assess the prevalences of intention to vaccinate and vaccine hesitancy in Lebanon; (2) determine how vaccine hesitancy in Lebanon varies by sociodemographic, economic, and geographic characteristics; and (3) understand individuals' motivations for vaccinating as well as concerns and obstacles to vaccination.

Methods: We performed a cross-sectional study from January 29, 2021, to March 11, 2021, using an online questionnaire of open- and closed-ended questions in Arabic via convenience "snowball" sampling to assess the perceptions of adults residing in Lebanon.

Results: Of the 1185 adults who participated in the survey, 46.1% (95% CI: 43.2%-49.0%) intended to receive the SARS-CoV-2 vaccine when available to them, 19.0% (95% CI 16.8%-21.4%) indicated they would not, and 34.0% (95% CI 31.3%-36.8%) were unsure (with an additional 0.9% skipping this question). The most common reasons for hesitancy were concerns about safety, limited testing, side effects, and efficacy. Top motivations for vaccinating were to protect oneself, protect one's family and the public, and end the pandemic. Despite financial hardships in Lebanon, barriers to vaccine access were not frequently described as concerns. Established health care facilities, rather than new temporary vaccination centers, were most frequently selected as preferred vaccination sites.

Conclusions: Vaccine hesitancy appears to be high in Lebanon. Disseminating clear, consistent, evidence-based safety and efficacy information on vaccines may help reduce vaccine hesitancy, especially among the large proportion of adults who appear to be unsure about (rather than opposed to) vaccination.

(*JMIR Form Res* 2022;6(12):e36827) doi:[10.2196/36827](https://doi.org/10.2196/36827)

KEYWORDS

Lebanon; COVID-19; SARS-CoV-2; coronavirus; vaccination; vaccine hesitancy; vaccine acceptance; health care system; misinformation; public health

Introduction

As of November 9, 2021, Lebanon's cumulative COVID-19 case count was 647,778 (95,695 cases per million), and at least 8556 people had died [1]. These numbers were likely underreported due to lack of testing and Lebanon's fragmented health infrastructure [2]. In addition to the pandemic, Lebanon struggles with multiple challenges: a political crisis and economic collapse driven by corruption that began in 2019 and worsened throughout 2020 and 2021, leading to mistrust of the government, inflation, unemployment, poverty, and food insecurity, on top of the strain of being the country with the highest number of refugees per capita in the world due to the protracted Syrian refugee crisis [2-6]. Given Lebanon's compounding crises and limited resources, mass vaccination is a challenging but vital mission. Understanding the population's perceptions of SARS-CoV-2 vaccines is critical for implementing a successful vaccination campaign in the country, providing extra support for vulnerable populations, and bolstering demand for vaccination.

SARS-CoV-2 vaccination began in Lebanon on February 14, 2021 [7]. The Pfizer-BioNTech, Sputnik V, Oxford-AstraZeneca, and Sinopharm vaccines had been approved for use in Lebanon at the time [8]. Despite efforts from the Lebanese government and international aid to secure adequate supply of SARS-CoV-2 vaccines of multiple formulations, as of November 9, 2021, only 1,583,289 people had been fully vaccinated, representing approximately 23% of the total population [1,9-11]. Inadequate demand, possibly due to vaccine hesitancy and logistical challenges, appears to be part of the reason for slow vaccination; as of November 9, 2021, only roughly 43% of the population had registered with the online national vaccine registration tool through the Ministry of Public Health [7]. In December 2021, Lebanon implemented mandatory vaccination orders for civil servants and workers in education, tourism, public transportation, and health [12]. Given the sizeable unvaccinated population, understanding perceptions of the vaccines and perceived obstacles is important to increasing vaccination rates to end the pandemic.

An online survey of Lebanese in April 2020 and May 2020 found that 69% of the convenience sample (predominantly young, university-educated, and male) indicated they would receive a hypothetical vaccine against SARS-CoV-2, though at that time, none had been developed or approved [13]. Another online survey performed in Lebanon from November 2020 to December 2020, prior to administration of SARS-CoV-2 vaccines in Lebanon and across most of the world, found that only 21.4% of respondents wanted to receive the vaccine, but this study did not assess reasons for participants' intentions [14]. These estimates of vaccine acceptance are in the lower end of the range of acceptance rates (13.3%-92%) found in surveys from countries in the Middle East and North Africa, which itself is a region with generally lower acceptance rates than other regions in the world [15].

Surveys assessing SARS-CoV-2 vaccine hesitancy around the world found that reasons for hesitancy include concerns about safety and side effects, doubts about efficacy, unfavorable personal risk and benefits assessments, and general mistrust in science and government [16-18]. Other barriers include distribution and uptake of the vaccines at scale, which are matters of logistics, health care access, and public perception of the vaccine [19,20]. These issues may be exacerbated in the Lebanese context given the deeply seated mistrust of government as well as compounding social crises. To our knowledge, no academic study has been done to elucidate the general public's motivations behind intending to receive or not to receive SARS-CoV-2 vaccination since the vaccination campaign started in Lebanon.

Our study's aims were to (1) assess rates of intention to vaccinate and vaccine hesitancy in Lebanon; (2) determine how vaccine hesitancy in Lebanon varies by sociodemographic, economic, and geographic characteristics; and (3) understand individuals' motivations for vaccinating as well as concerns and obstacles to vaccination.

Methods**Study Design**

We designed a cross-sectional descriptive study using an online Arabic survey. The survey was validated and revised in focus group piloting. We originally intended to distribute it to randomly selected Lebanese phone numbers to obtain an unbiased nationally representative sample, but during piloting, this method needed to be aborted because of low response rates (less than 1%), likely due to mistrust of messages received from an unknown phone number. Given these constraints, we changed our distribution methods to convenience "snowball" sampling, a method described in a later section that had successfully been used elsewhere in the Middle East [14,21]. The recruitment and survey period occurred from January 29, 2021, to March 11, 2021.

Study Population

The target population was all adults living in Lebanon. Although all adults living in Lebanon were eligible, they needed to access the self-administered online survey tool on a mobile phone or computer. The survey was in Arabic, so participants had to be literate in Arabic or assisted by someone who was.

Recruitment and Sampling:

Using a convenience "snowball" sampling method, the research team initiated recruitment by sending a recruitment message in Arabic to their contacts and organizations. Participants interested in the study proceeded to the survey via a link in the recruitment message, which also invited participants to forward the message to their contacts. It could be sent through WhatsApp, SMS, social media, and email. There was no follow-up to determine whether individuals who received the recruitment message

completed or forwarded the survey. No incentives were provided for participation.

Survey Tool and Data Collection

The survey (see [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)) was created using the Qualtrics online survey platform [22]. It was anonymous, self-administered, and required basic literacy in Arabic. The survey content was created by the research team (the majority of whom were fluent in Arabic and English), used several other SARS-CoV-2 vaccine perception studies as a guide [23,24], and was tailored to the Lebanese context. It was piloted in small groups of contacts of the research team of varying educational backgrounds and health literacy and revised prior to survey launch.

The survey included 31 multiple-choice and free-response questions (depending on branch points, participants were not asked each question), divided into an introduction with the informed consent document, followed by questions about screening, demographics, experience with COVID-19, and perceptions of SARS-CoV-2 vaccination. No identifying data were collected. Participants were asked to provide informed consent and affirm that they were 18 years or older, were living in Lebanon, and had heard of “coronavirus” (as SARS-CoV-2 and COVID-19 are referred to in Lebanon). If they declined, the survey automatically ended. If they passed these questions, they proceeded with the survey. After starting the survey, participants had 48 hours before responses were automatically recorded. Participants were prevented from multiple survey attempts on the same device via a Qualtrics feature based on browser cookies.

Ethical Approval and Informed Consent

This study was approved by the Institutional Review Boards of the University of California, Berkeley (Protocol number 2020-11-13811); the Modern University for Business and Science, Lebanon (Project number MU-20201207-19); and Stanford University (by reliance agreement with the University of California, Berkeley, eProtocol number 59832). Informed consent was obtained from all participants prior to the survey; a waiver of written informed consent was granted given that written consent would have been the only identifying information collected.

Data Analysis

Data were downloaded as a .csv file and translated into English using Microsoft Excel [25]. Data cleaning and statistical analysis were performed in R computer software and focused on description rather than identifying causal links [26].

Quality Measures

Several quality measures were implemented. To ensure that participants had baseline familiarity with the survey topics, a

screening question asked whether participants had heard of “coronavirus”; only 5 participants (5/1185, 0.42%) were excluded because of not being aware of “coronavirus.” To identify participants who randomly clicked through the survey, a filter was applied to detect participants who completed the survey in less than 120 seconds; no participants did so.

Quantitative Data Analysis

For some multiple-choice questions, similar categorical responses were consolidated into binary or fewer categories to facilitate interpretation. For binary and categorical variables, the absolute number and relative proportions of participants who selected each response were calculated. Wilson score 95% confidence intervals were calculated for proportions.

Given that the survey period spanned the initiation of SARS-CoV-2 vaccination in Lebanon, a subanalysis was performed in which participants were divided by whether they completed the survey before or after vaccine initiation. For each of these subgroups, sample demographics were recalculated. We did not use sampling weights in our analysis given that this was a nonprobabilistic sample of the Lebanese population that was unlikely to be representative of the general population even after weighting.

Qualitative Data Analysis

If participants indicated that they intended to vaccinate, they were then asked to explain why they intended, did not intend, or were uncertain about vaccinating. These open-ended responses were thematically, iteratively coded using a uniform protocol to facilitate analysis (see [Multimedia Appendix 3](#)).

Results

Sample Characteristics

Of the 1390 participants who initiated the survey, 1185 (85.3%) provided informed consent and passed screening to begin the main survey. Among this group, 1103 (1103/1185, 93.1%) participants completed the entire survey. Selected sample demographic characteristics of those who passed screening are summarized in [Table 1](#). For full sample demographic characteristics, see the expanded table in [Table S1](#) in [Multimedia Appendix 4](#).

Compared with the demographics of residents of Lebanon in general, our sample population had a higher representation of individuals who identified with the Druze religion and were young, female, well-educated, and from the Mount Lebanon region ([Table 1](#)). Underrepresented in our sample were older adults; refugees; non-Lebanese citizens; members of the Sunni and Shi'a religions; and individuals from the less populated governorates of Akkar, Baalbek-Hermel, Nabatieh, North, and South.

Table 1. Participant characteristics (n=1185).

Characteristic ^a	All participants, n (%)	Round 1 participants ^b (n=840), n (%)	Round 2 participants ^c (n=345), n (%)	Estimates for the Lebanese population ^d [6,27-30], %
Gender				
Female	685 (62.1)	479 (60.1)	206 (65.6)	51.6
Male	388 (35.2)	285 (36.1)	103 (32.8)	48.4
Other	7 (0.1)	6 (0.8)	1 (0.3)	N/A ^e
Skip this question	23 (2.1)	19 (2.4)	4 (1.3)	— ^f
Age (years)				
18-34	745 (62.8)	493 (58.7)	252 (73.1)	32.0 ^g
35-54	318 (26.8)	248 (29.5)	70 (20.3)	23.5
≥55	122 (10.3)	99 (11.8)	23 (6.7)	19.8
Governorate				
Beirut	111 (10.1)	89 (11.3)	22 (7.0)	7.1
Mount Lebanon	633 (57.4)	559 (70.8)	74 (23.6)	42.3
Other	342 (31.0)	125 (15.9)	217 (69.1)	50.6
Skip this question	17 (1.5)	16 (2.0)	1 (0.3)	—
Religion				
Christian	242 (21.9)	199 (25.2)	43 (13.7)	32.4
Druze	355 (32.2)	308 (39.0)	47 (15.0)	4.5
Shi'a	136 (12.3)	46 (5.8)	90 (28.7)	31.0
Sunni	124 (11.2)	64 (8.1)	60 (19.1)	31.9
No religion	65 (5.9)	48 (6.1)	17 (5.4)	N/A
Other	6 (0.5)	3 (0.4)	3 (1.0)	0.3
Skip this question	175 (15.9)	121 (15.3)	54 (17.2)	—
Highest education level				
Completed high school, technical school, or less	200 (18.1)	143 (18.1)	57 (18.2)	78.6
Completed some college or more	891 (80.9)	637 (80.8)	254 (80.9)	21.4
Skip this question	11 (1.0)	8 (1.0)	3 (1.0)	—
Employment status				
Employed	586 (51.5)	412 (52.2)	156 (49.7)	N/A
Student	179 (16.2)	117 (14.8)	62 (19.7)	N/A
Unemployed, not seeking work	145 (13.1)	122 (15.5)	23 (7.3)	N/A
Unemployed, seeking work	154 (14.0)	101 (12.8)	53 (16.9)	33.0 ^h
Skip this question	57 (5.2)	37 (4.7)	20 (6.3)	—
Citizenshipⁱ				
Lebanon	1038 (94.1)	757 (96.0)	281 (89.5)	79.8
Not a citizen of Lebanon	53 (4.8)	24 (3.0)	25 (8.0)	20.2
Skip this question	12 (1.1)	8 (1.0)	8 (2.5)	—
Refugee				
Yes	56 (5.1)	29 (3.7)	27 (8.6)	21.9 ^h
No	1015 (92.0)	737 (93.4)	278 (88.5)	78.1 ^h

Characteristic ^a	All participants, n (%)	Round 1 participants ^b (n=840), n (%)	Round 2 participants ^c (n=345), n (%)	Estimates for the Lebanese population ^d [6,27-30], %
Skip this question	32 (2.9)	23 (2.9)	9 (2.9)	—

^aBecause participants were not forced to answer all questions, each question has a different denominator.

^bThose who completed the survey in the first round, prior to initiation of vaccination in Lebanon on February 13, 2021.

^cThose who completed the survey in the second round, after initiation of vaccination in Lebanon.

^dUnless otherwise noted, estimates were obtained from government source that excluded refugees and used 4.84 million (2018) as total population.

^eN/A: not available in the data source.

^fNot applicable.

^g15-34 years old.

^hEstimates were obtained from a source that included refugees and used 6.86 million (2020) as the total population.

ⁱParticipants could select multiple answers, so proportions were calculated using a denominator of all participants who selected an answer for the question.

Quantitative Analysis

Intentions to Vaccinate

We found that 46.1% (95% CI 43.2%-49.0%) of our survey participants intended to receive the SARS-CoV-2 vaccine when

available, 19.0% (95% CI 16.8%-21.4%) indicated they would not get vaccinated, and 34.0% (95% CI 31.3%-36.8%) were unsure about vaccination (with an additional 0.9% skipping this question; [Table 2](#)).

Table 2. Intentions about vaccination by sociodemographic characteristics and experience with COVID-19 for everyone who answered the intention to vaccinate question (n=1185).

Characteristic ^a	Entire sample, n (%)	Intend to vaccinate, % (95% CI)	Do not intend to vaccinate, % (95% CI)	Unsure about vaccine, % (95% CI)
All participants	___ ^b	46.1 (43.2-49.0)	19.0 (16.8-21.4)	34.0 (31.3-6.8)
Gender				
Female	682 (63.6)	42.8 (39.1-46.6)	21.1 (18.1-24.4)	36.1 (32.5-39.8)
Male	385 (35.9)	55.6 (50.5-60.6)	15.6 (12.2-19.7)	28.8 (24.4-33.7)
Other	6 (0.6)	33.3 (6.0-75.9)	50.0 (18.8-81.2)	16.7 (0.9-63.5)
Age (years)				
18-24	364 (31.0)	39.6 (34.5-44.8)	22.5 (18.4-27.2)	37.9 (32.9-43.1)
25-34	372 (31.7)	45.4 (40.3-50.6)	23.1 (19.0-27.8)	31.5 (26.8-36.5)
35-44	191 (16.3)	50.8 (43.5-58.0)	14.1 (9.7-20.0)	35.1 (28.4-42.3)
45-54	125 (10.6)	51.2 (42.1-60.2)	12.0 (7.1-19.3)	36.8 (28.5-45.9)
55-64	88 (7.5)	60.2 (49.2-70.3)	12.5 (6.7-21.7)	27.3 (18.6-38.0)
≥65	34 (2.9)	55.9 (38.1-72.4)	11.8 (3.8-28.4)	32.4 (18.0-50.6)
Governorate				
Baalbek-Hermel	20 (1.9)	50.0 (29.9-70.1)	15.0 (4.0-38.9)	35.0 (16.3-59.1)
Beqaa	98 (9.1)	59.2 (48.8-68.9)	12.2 (6.8-20.8)	28.6 (20.1-38.7)
Beirut	111 (10.3)	53.2 (43.5-62.6)	16.2 (10.1-24.7)	30.6 (22.4-40.2)
Mount Lebanon	632 (58.6)	42.2 (38.4-46.2)	22.8 (19.6-26.3)	35.0 (31.3-38.8)
South	70 (6.5)	57.1 (44.8-68.7)	10.0 (4.5-20.1)	32.9 (22.4-45.2)
Akkar	13 (1.2)	84.6 (53.7-97.3)	7.7 (0.4-37.9)	7.7 (0.4-37.9)
North	58 (5.4)	46.6 (33.5-60.0)	17.2 (9.0-29.9)	36.2 (24.3-49.9)
Nabatieh	77 (7.1)	50.6 (39.1-62.1)	18.2 (10.6-29.0)	31.2 (21.4-42.9)
Religion				
Christian	242 (22.1)	60.3 (53.8-66.4)	9.5 (6.2-14.1)	30.2 (24.5-36.4)
Druze	354 (32.3)	35.6 (30.6-40.9)	25.4 (21.0-30.3)	39.0 (33.9-44.3)
Shi'a	134 (12.2)	48.5 (39.8-57.3)	16.4 (10.8-24.0)	35.1 (27.2-43.8)
Sunni	122 (11.1)	57.4 (48.1-66.2)	21.3 (14.6-29.8)	21.3 (14.6-29.8)
No religion	65 (5.9)	58.5 (45.6-70.3)	16.9 (9.1-28.7)	24.6 (15.1-37.1)
Other	5 (0.5)	20.0 (1.1-70.1)	0.0 (0.0-53.7)	80.0 (29.9-99.0)
Skip this question	173 (15.8)	39.9 (32.6-47.6)	24.9 (18.8-32.1)	35.3 (28.3-42.9)
Education				
Completed high school, technical school, or less	197 (18.2)	42.1 (35.2-49.4)	21.3 (16.0-27.8)	36.5 (30.0-43.7)
Completed some college or more	886 (81.8)	48.8 (45.4-52.1)	18.5 (16.0-21.3)	32.7 (29.7-35.9)
Employment				
Employed	566 (54.3)	51.6 (47.4-55.8)	17.1 (14.2-20.6)	31.3 (27.5-35.3)
Student	179 (17.2)	43.0 (35.7-50.6)	21.2 (15.6-28.1)	35.8 (28.8-43.3)
Unemployed, not seeking work	145 (13.9)	42.1 (34.0-50.6)	18.6 (12.8-26.1)	39.3 (31.4-47.8)
Unemployed, seeking work	152 (14.6)	41.4 (33.6-49.7)	23.0 (16.8-30.7)	35.5 (28.1-30.7)
Annual income (LL^c)				
<1,000,000	121 (11.1)	34.7 (26.4-44.0)	25.6 (18.3-34.5)	39.7 (31.0-49.0)
1,000,000-9,999,999	287 (26.2)	45.6 (39.8-51.6)	18.8 (14.6-23.9)	35.5 (30.1-41.4)

Characteristic ^a	Entire sample, n (%)	Intend to vaccinate, % (95% CI)	Do not intend to vaccinate, % (95% CI)	Unsure about vaccine, % (95% CI)
10,000,000-19,999,999	99 (9.0)	47.5 (37.4-57.7)	24.2 (16.4-34.1)	28.3 (19.9-38.4)
20,000,000-69,999,999	131 (12.0)	63.4 (54.4-71.5)	9.2 (5.0-15.8)	27.5 (20.2-36.1)
≥70,000,000	61 (5.6)	68.8 (55.6-79.8)	6.6 (2.1-16.7)	24.6 (14.8-37.6)
Skip this question	396 (36.2)	42.9 (38.0-48.0)	22.7 (18.8-27.2)	34.3 (29.7-39.3)
Citizenship^d				
Lebanon	1033 (94.7)	47.2 (44.2-50.3)	19.1 (16.7-21.6)	33.7 (30.8-36.7)
Syria	27 (2.5)	40.7 (23.0-61.0)	37.0 (20.1-57.5)	22.2 (9.4-42.7)
Palestine	17 (1.6)	58.8 (33.5-80.6)	11.8 (2.1-37.7)	29.4 (11.4-56.0)
European or North American country	40 (3.7)	62.5 (45.8-76.8)	12.5 (4.7-27.6)	25.0 (13.2-41.5)
Other country	22 (2.0)	31.8 (14.7-54.8)	18.2 (6.0-41.0)	50.0 (30.7-69.3)
Multiple countries	53 (4.9)	54.7 (40.6-68.2)	13.2 (5.9-26.0)	32.1 (20.3-46.4)
Refugee				
Yes	54 (5.1)	50.0 (37.1-62.9)	25.9 (15.4-39.9)	24.1 (13.9-37.9)
No	1009 (94.9)	46.9 (43.8-50.0)	19.0 (16.7-21.6)	34.1 (31.2-37.1)
Completed survey after initiation of vaccination in Lebanon				
Yes	339 (28.9)	56.9 (51.5-62.2)	12.7 (9.4-16.8)	30.4 (25.6-35.6)
No	835 (71.1)	42.3 (38.9-45.7)	21.8 (19.1-24.8)	35.9 (32.7-39.3)
Correct knowledge of transmission^e				
Yes	1076 (99.1)	48.0 (45.1-51.1)	18.1 (15.9-20.6)	33.8 (31.0-36.8)
No	10 (0.9)	50.0 (23.7-76.3)	40.0 (13.7-72.6)	10.0 (0.5-45.9)
Personal history of COVID-19				
Yes	321 (28.9)	42.9 (37.5-48.6)	21.6 (17.3-26.6)	35.4 (30.2-41.0)
No	790 (71.1)	48.8 (45.2-52.3)	18.5 (15.9-21.5)	32.7 (29.4-36.1)
Close acquaintance with history of COVID-19				
Yes	1056 (94.2)	48.0 (44.9-51.0)	18.8 (16.5-21.3)	33.3 (30.4-36.2)
No	65 (5.8)	33.3 (22.2-46.4)	30.2 (19.6-43.2)	36.5 (25.0-49.6)
Mask wearing outside home				
Always or most of the time	1011 (91.2)	50.0 (46.8-53.1)	16.3 (14.1-18.8)	33.7 (30.8-36.8)
Sometimes, rarely, or never	97 (8.8)	18.6 (11.7-28.0)	48.5 (38.3-58.8)	33.0 (24.0-43.4)
Top 3 news sources^d				
Printed newspaper or magazine	86 (7.8)	64.0 (52.8-73.8)	9.3 (4.4-18.0)	26.7 (18.0-37.6)
Radio	33 (3.0)	69.7 (51.1-83.8)	15.2 (5.7-32.7)	15.2 (5.7-32.7)
Television	718 (65.1)	47.2 (43.5-50.9)	17.1 (14.5-20.1)	35.7 (32.2-39.3)
Social media (eg, Facebook, Twitter, YouTube, WhatsApp)	670 (60.1)	43.7 (39.9-47.6)	19.4 (16.5-22.6)	36.9 (33.2-40.7)
Internet but not social media (eg, websites)	609 (55.2)	53.5 (49.5-57.5)	17.4 (14.5-20.7)	29.1 (25.5-32.9)
Talking to friends or family	282 (25.6)	37.6 (32.0-43.6)	24.1 (19.3-29.6)	38.3 (32.6-44.3)

Characteristic ^a	Entire sample, n (%)	Intend to vaccinate, % (95% CI)	Do not intend to vaccinate, % (95% CI)	Unsure about vaccine, % (95% CI)
Religious leaders	6 (0.5)	16.7 (0.9-63.5)	83.3 (36.5-99.1)	0.0 (0.0-48.3)

^aFor the analysis of each characteristic, we omitted participants who skipped the question, unless >10% of participants for that question skipped the question, in which case those who skipped the characteristic question were included in the analysis. We then calculated the proportion of each characteristic subcategory by intention to vaccinate, calculating Wilson score CIs.

^bNot applicable.

^cLL: Lebanese Lira.

^dThe survey allowed participants to choose multiple answers for this characteristic; consequently, the sum of all subcategories does not equal the number of all participants who answered the question.

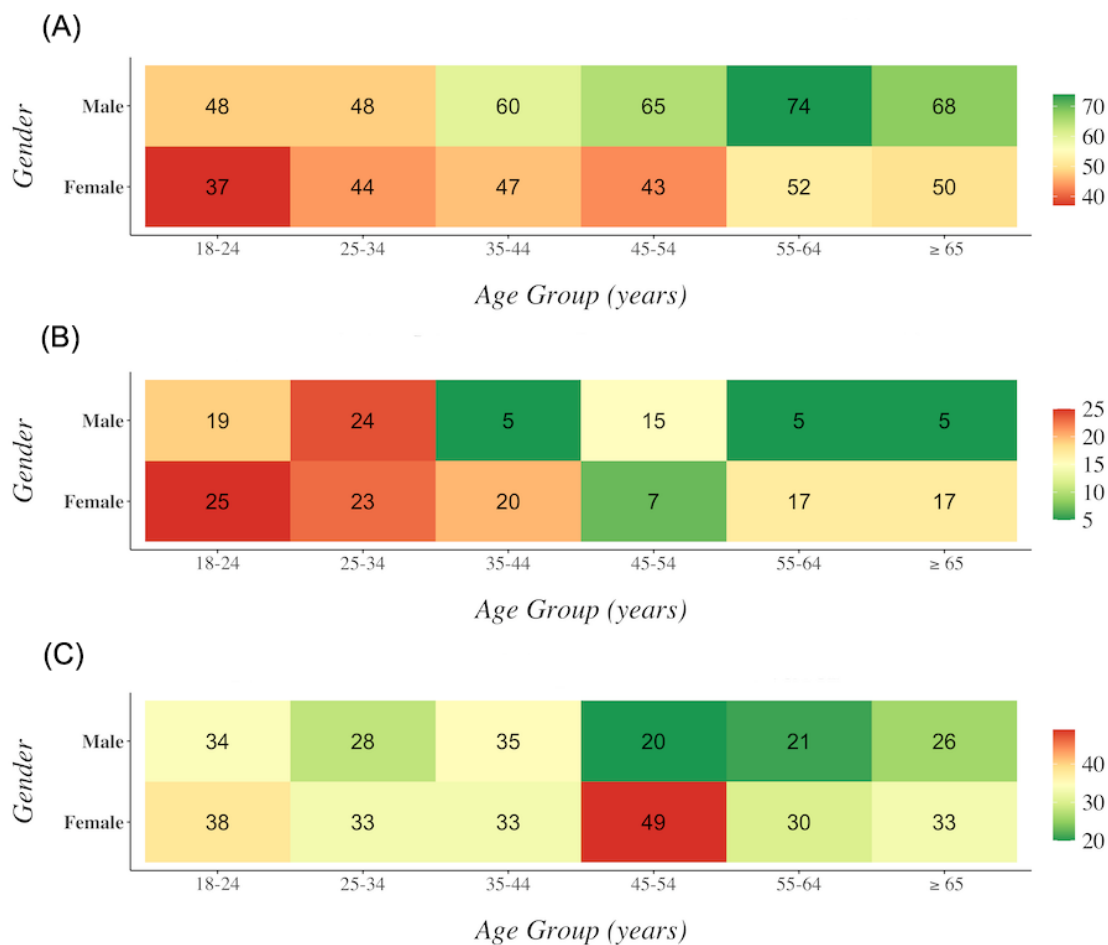
^e“Correct Knowledge” indicated a correct response to a multiple-choice question asking, “In your understanding, how does someone become infected with coronavirus? Choose the best single answer.” The correct response was “Being physically close to an infected person.” Incorrect responses were “Eating raw food or untreated water” and “Being bitten by an insect.”

Intentions to Vaccinate by Demographic Characteristics

Participants were more likely to intend to vaccinate if they identified as male; lived in the Beqaa governorate (Mount Lebanon as reference); were Christian, Sunni, or affiliated with no religion (Druze as reference); and had a higher household income (Table 2 and Figure 1).

There were less strong but still apparent trends toward higher proportions of participants’ intending to vaccinate if they were older, lived in Beirut or Akkar governorates (Mount Lebanon as reference), had attained higher educational status, or were employed. There were no apparent differences nor trends in intention to vaccinate by citizenship or whether participants identified as refugees.

Figure 1. Intention to receive SARS-CoV-2 vaccine by age group and gender for those who provided age, gender, and their intentions about vaccination, reported as the percentage who (A) intended to receive the SARS-CoV-2 vaccine, (B) did not intend to receive the SARS-CoV-2 vaccine, (C) were not sure about receiving the SARS-CoV-2 vaccine. As only 7 participants identified as “other” gender, only participants identifying as “male” or “female” were included.



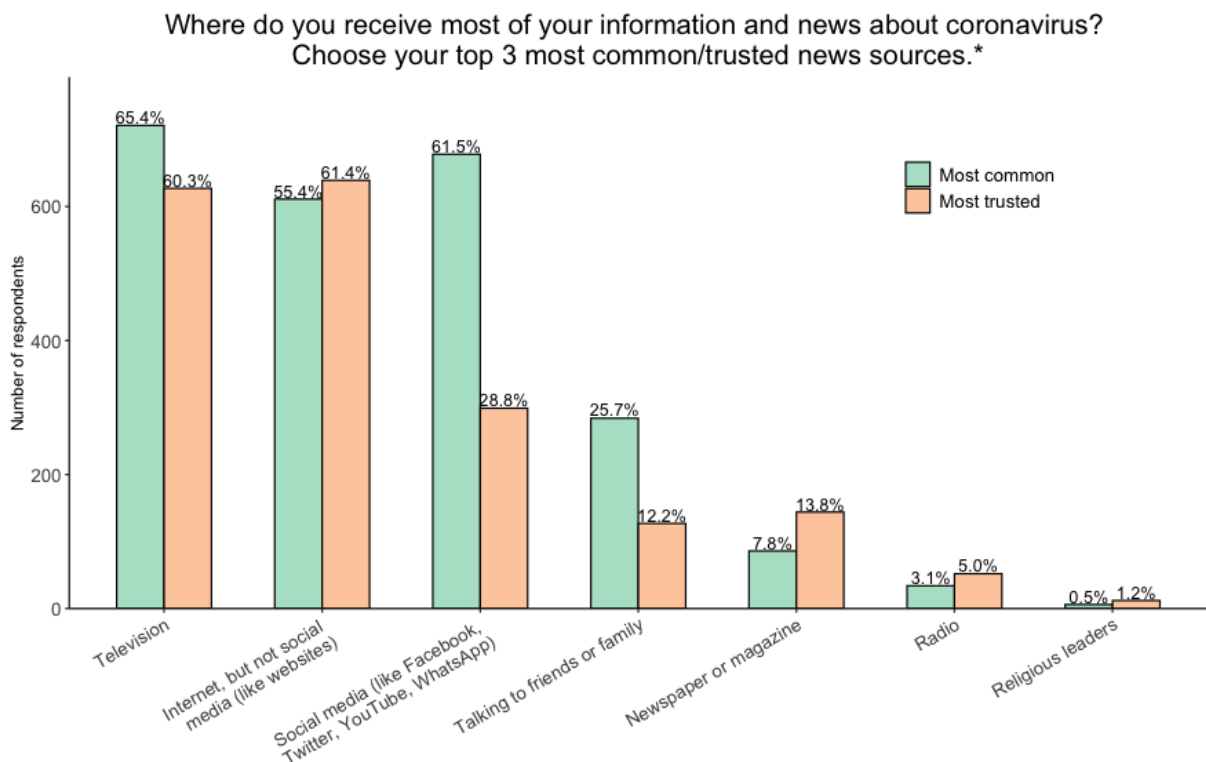
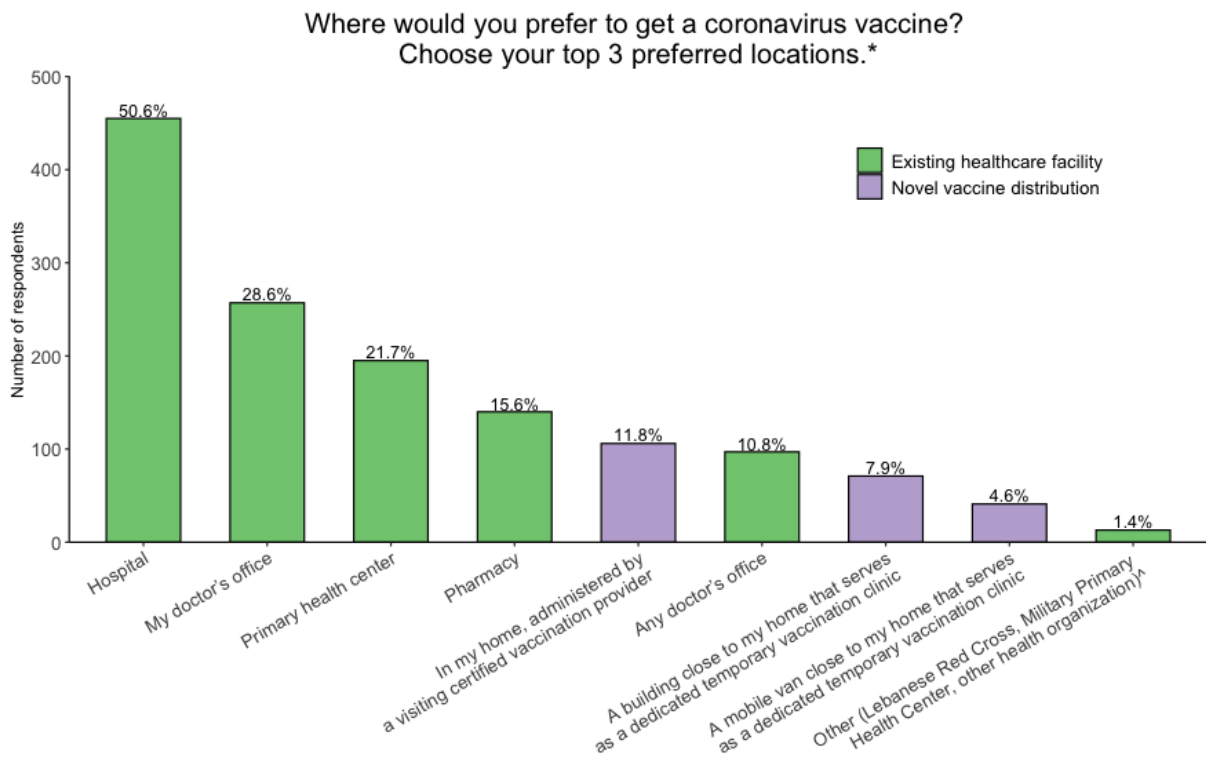
Intentions to Vaccinate Before and After Initiation of Vaccination in Lebanon

Those who completed the survey after initiation of vaccination were more likely to intend to vaccinate (56.9%, 95% CI 51.5%-62.2%) when compared with those who completed the survey before initiation of vaccination (42.3%, 95% CI 38.9%-45.7%; [Table 2](#)). Importantly, these groups were contacted differently and differed in several key demographic characteristics: The group of participants who responded after initiation of vaccination was younger and had a higher proportion of participants who lived outside of the Mount Lebanon region, identified as Shi'a or Sunni rather than Druze or Christian, had European or North American or multiple citizenships, and identified as refugees ([Table 1](#)).

Logistical Considerations About Vaccination

The most commonly selected preferred locations to get vaccinated were hospitals, doctors' offices, primary health centers, and pharmacies ([Figure 2, Multimedia Appendix 5](#)). Temporary vaccination sites were not popular. The most common sources of news about coronavirus were television, social media, and other internet websites ([Figure 2, Multimedia Appendix 5](#)). Very few participants (6/1185, 0.5%) reported commonly relying on religious leaders for coronavirus news ([Multimedia Appendix 5](#)). Participants trusted television and internet websites more than social media ([Multimedia Appendix 5](#)). Among participants who did not intend to vaccinate or were uncertain about vaccination, less than 2% (10/599, 1.7%) stated that a monetary incentive would persuade them to become vaccinated ([Multimedia Appendix 5](#)).

Figure 2. Top 3 ranked responses from each participant for the logistical considerations of where participants would prefer to get a coronavirus vaccine and which news sources were the most commonly used or trusted for coronavirus information. If participants selected “Other;” they could write in their own response. The 3 responses in parentheses (Lebanese Red Cross, Military Primary Health Center, other health organization) encompass the written-in responses.



Qualitative Analysis: Motivations for and Concerns About Vaccination

Most frequently, participants intending to vaccinate cited the following motivations: to protect themselves, their families,

and the public and to end the pandemic and return to normal life (Table 3). Somewhat frequently given reasons for intending to vaccinate included trusting science and research and feeling like there was “no other choice” given the state of the pandemic.

Table 3. Summary of coded open-ended responses describing motivations for intentions about vaccination.

Motivation	Most frequent codes (descending frequency) ^a	Less frequent codes (descending frequency) ^b
Motivations for intending to vaccinate	<ul style="list-style-type: none"> To protect myself To protect and limit spread among the public To end the pandemic and return to normalcy To protect my family 	<ul style="list-style-type: none"> I do not have another better choice or solution. I trust it based on science and research. I want to help achieve “herd immunity.” I have other medical conditions that make my risk of illness higher. My job puts me at risk of contracting COVID-19.
Motivations for not intending to vaccinate	<ul style="list-style-type: none"> I am concerned about safety and limited testing. I am concerned about potential side effects. Nonspecific mistrust of the SARS-CoV-2 vaccines I am worried that the vaccines are not effective. 	<ul style="list-style-type: none"> I do not trust the Lebanese government. I do not trust Lebanese health care or distribution systems. COVID-19 and vaccines are a scam or conspiracy to make money. General vaccine hesitancy; I am against all vaccines. Problems or concerns with ingredients I do not need the vaccine because I already had COVID-19. COVID-19 is not a threat to me or in general. Risks of vaccination are not worth the potential benefits. Society and science in general do not know enough about COVID-19 (causes, symptoms, effects).
Motivations for uncertainty about vaccination	<ul style="list-style-type: none"> I am concerned about safety and limited testing. I am concerned about potential side effects. I am worried that the vaccines are not effective. 	<ul style="list-style-type: none"> I do not think I need the vaccine because I already had COVID-19. I do not trust the Lebanese government. Nonspecific mistrust of the SARS-CoV-2 vaccines I do not have enough information about the vaccines. I want more time or more information before I decide which vaccine to take. I have received conflicting information about the vaccine from one or more sources. I have concerns about mRNA technology. I do not trust Lebanese health care or distribution systems. I think the vaccines in Lebanon will not be the true coronavirus vaccine or that they will be tampered with.

^a“Most frequent codes” were observed in ≥15% of responses to the question.

^b“Less frequent codes” were observed in 3%-15% of responses to the question.

Among participants who did not intend to be vaccinated, the most frequent reasons for vaccine hesitancy were concerns about safety given the fast development and limited testing of the vaccines, fears about side effects, and doubts about efficacy (Table 3). Somewhat frequently cited concerns leading to vaccine hesitancy were mistrust in the Lebanese government.

The most commonly provided reasons for uncertainty about whether participants planned to receive the vaccine were similar: concerns about safety given the fast development and limited testing of the vaccines, fears about side effects, and doubts about efficacy (Table 3). Somewhat frequently mentioned concerns included mistrust of the Lebanese government and health care system, potential fraud in storage or marketing of the vaccines, wanting more information about the vaccines, and needing more time to decide which vaccine to receive.

Discussion

Principal Findings

Our findings of rates of vaccine acceptance and vaccine hesitancy in Lebanon were similar to those of studies in other countries in the Middle East and across the world. The vaccine acceptance of 46.1% in our sample was similar to rates in

Kuwait (53.1%) and Qatar (45%-60%), higher than rates in Jordan (28.4%), and somewhat lower than in Saudi Arabia (64.7%) [21,31-33]. Rates of vaccine acceptance in our study were also similar to a large survey of 15 developed countries across the globe, in which 54% of respondents indicated they intend to vaccinate [34]. However, as systematic reviews of vaccine perception studies showed, vaccine acceptance and hesitancy have varied with time throughout the pandemic, with a trend toward decreasing acceptance throughout 2020 [18,35]. Indeed, our results show lower rates of vaccine acceptance than a survey in April 2020 and May 2020 that included a question about acceptance of hypothetical SARS-CoV-2 vaccines, in which 69.3% of Lebanese residents stated they would be willing to receive a vaccine; it is important to note that, at this time, no vaccines were developed or approved and the survey’s population was predominantly male and younger than ours [13]. In contrast, our study found higher rates of intent to vaccinate than an online convenience survey in November 2020 and December 2020, in which only 21.4% of participants intended to be vaccinated [14]. Notably, the study’s sample was smaller, more predominantly female, and younger than ours, and it took place prior to mass vaccination in most of the world, including in Lebanon. It also did not directly assess motivations for

intention to or not to vaccinate. Our study found lower rates of vaccine acceptance than the vaccine acceptance rate of 87% among 800 university students surveyed in May 2021 and June 2021 at a prominent Lebanese university [36].

Our findings of trends toward increased vaccine hesitancy in women, younger age groups, unemployed individuals, and individuals with lower education attainment are generally consistent with findings in the Middle East and globally, with the exception that, in Kuwait and Qatar, hesitancy was higher in older populations [14,31,32,35,37-39]. Although we attempted to assess the association of vaccine hesitancy with religion and income, 2 important demographic variables in Lebanon, 15.9% and 36.3% of participants skipped questions about religion and income, respectively, demonstrating the topics' sensitive natures. Therefore, we do not recommend inferences about differences in vaccine acceptance or hesitancy by religion or income from our study. Similarly, refugees (56/1185, 5.1%) were underrepresented in our sample. Although no significant difference in vaccine acceptance by refugee status emerged, further focus on vaccination in refugees in Lebanon is merited given their multiple vulnerabilities. One survey among Syrian refugees in Lebanon in January 2021 and February 2021 found that 66% of respondents would accept a "safe and free" vaccine [40].

The timing of our survey period spanned the initiation of vaccination in Lebanon. Although our data suggested increased vaccine acceptance among participants who completed the survey after initiation of vaccination, this must be interpreted cautiously. We believe that the differences in vaccine acceptance before and after initiation of vaccination more likely reflect significant differences in populations surveyed during these periods, given the selection bias inherent in the convenience sampling method.

The logistical considerations about which we asked can provide some guidance for vaccination efforts in Lebanon. The most commonly selected sources of news about COVID-19 were television, social media, and internet websites, with television and internet websites most trusted. Focusing dissemination of vaccine promotion efforts on television, social media, and internet websites could be most efficient to reach those with vaccine hesitancy.

Survey respondents reported preferring to receive the vaccine at familiar, established health care sites. Although temporary dedicated vaccination centers were not popular, a small but significant number of participants (106/899, 11.8%) stated they would prefer vaccination at home by a visiting medical professional, a potentially important means of reaching vulnerable patients unable to travel. We also asked whether a financial incentive would change participants' minds so that they decide to vaccinate; overwhelmingly, they indicated that it would not (589/599, 98.3%).

Despite financial hardships in Lebanon, barriers to vaccine access (cost, transportation, proximity to medical care) were not cited frequently as concerns in our study. This might be explained by the fact that it has become common for governments worldwide to distribute the vaccine free of charge. There were relatively few concerns about vaccine properties

like number of required doses, country of origin, or specific vaccine brands. Also uncommon was opposition to vaccines in general (ie, not specific to SARS-CoV-2 vaccines). This is consistent with previous studies in Lebanon showing moderate uptake of routine vaccinations [41-45]. Although several participants in our study cited conspiracy theories as reasons for not vaccinating, these were relatively uncommon, especially compared with a study in other Arabic-speaking countries, which found rates of belief in conspiracy theories of over 50% [33].

Perhaps the most actionable information generated in this study involves motivations for vaccine acceptance and vaccine hesitancy. The relatively large proportion of participants who were uncertain about vaccination—34.0%—provides a public health opportunity and imperative in the effort to achieve mass vaccination in Lebanon. The majority of concerns about SARS-CoV-2 vaccination involved absence of reliable information and data regarding safety, testing, and efficacy. Increasing public availability in Lebanon of high-quality, accessible information about the SARS-CoV-2 vaccines could assuage such concerns and increase vaccine acceptance. Vaccine promotion campaigns could also use messaging based on the most frequently provided reasons for intending to vaccinate: to protect oneself, protect loved ones, and end the pandemic. Based on our results about common COVID-19 news sources, we recommend disseminating clear, consistent, verifiable safety and efficacy information on television, social media, and news websites. Given the prevalence of mistrust in the government, third parties (like health care organizations) might be most trusted, especially if they are able to provide transparency and reassurance about proper vaccine acquisition and storage to allay concerns about fraud.

Limitations

Our study has several important limitations. First and foremost, our sample is unlikely to be representative of the general population because of the sampling strategy. Although we attempted to mitigate this by collecting important demographic and experiential characteristics, the bias remains, and several important demographic groups were underrepresented, most notably older adults, members of Shi'a and Sunni religions, residents outside of Beirut and Mount Lebanon, non-Lebanese citizens, and those who identify as refugees. Second, participation required literacy in Arabic, internet access, and digital literacy, potentially excluding some populations (though >80% of refugees in Lebanon have access to mobile technology like WhatsApp) [46]. Third, some participants did not answer all questions, possibly causing nonresponse bias. Finally, vaccine hesitancy, perceptions, and concerns may be changing rapidly over time; our results should be interpreted as pertaining to the time period during which the survey was conducted.

Conclusions

This cross-sectional study assessed intentions to vaccinate against SARS-CoV-2 among adults residing in Lebanon, analyzed characteristics associated with vaccine acceptance and hesitancy, and described motivations for and concerns about vaccination. We recommend disseminating clear, consistent, evidence-based safety and efficacy information on vaccines via

most commonly selected news sources: television, social media, and news websites. Repeated assessments of intentions to vaccinate, concerns regarding vaccination, and changes in motivations should be performed, especially with the goal of assessing the perspectives and needs of populations that were underrepresented in this study.

Acknowledgments

We thank our participants for taking the time to share their perspectives; our recruitment and coding assistants (Hanin Abi Ghannam, Jana Al Khatib, Nathalie Al Saady, Ghiwa Nasr); the National Wellness Network and Mariam Fadel for assisting in disseminating the survey; the Stanford Center for Innovation in Global Health and the Refugee and Asylum seeker Health Initiative (RAHI) at the University of California, San Francisco; and Drs. Michele Barry, Cybele Renault, Karen Sokal-Gutierrez, and Anke Hemmerling for their mentorship. PG is a Chan Zuckerberg Biohub investigator.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey tool in English.

[PDF File (Adobe PDF File), 93 KB - [formative_v6i12e36827_app1.pdf](#)]

Multimedia Appendix 2

Survey tool in Arabic.

[PDF File (Adobe PDF File), 125 KB - [formative_v6i12e36827_app2.pdf](#)]

Multimedia Appendix 3

Coding protocol for free-response questions assessing motivations for intent to or not to vaccinate.

[PDF File (Adobe PDF File), 35 KB - [formative_v6i12e36827_app3.pdf](#)]

Multimedia Appendix 4

Table 1S: Unabridged characteristics of all the participants who passed screening.

[DOCX File , 28 KB - [formative_v6i12e36827_app4.docx](#)]

Multimedia Appendix 5

Table 2S: Additional logistical considerations about vaccination.

[DOCX File , 18 KB - [formative_v6i12e36827_app5.docx](#)]

References

1. Ritchie H, Ortiz-Ospina E, Beltekian D, Mathieu E, Hasell J, Macdonald B. Coronavirus Pandemic (COVID-19). Our World in Data. URL: <https://ourworldindata.org/coronavirus> [accessed 2022-11-18]
2. Khoury P, Azar E, Hitti E. COVID-19 response in Lebanon: current experience and challenges in a low-resource setting. JAMA 2020 Aug 11;324(6):548-549 [FREE Full text] [doi: [10.1001/jama.2020.12695](https://doi.org/10.1001/jama.2020.12695)] [Medline: [32722763](https://pubmed.ncbi.nlm.nih.gov/32722763/)]
3. Bizri AR, Khachfe H, Fares M, Musharrafieh U. COVID-19 pandemic: an insult over injury for Lebanon. J Community Health 2021 Jun;46(3):487-493 [FREE Full text] [doi: [10.1007/s10900-020-00884-y](https://doi.org/10.1007/s10900-020-00884-y)] [Medline: [32661861](https://pubmed.ncbi.nlm.nih.gov/32661861/)]
4. Kenner D. The struggle to keep Lebanon healthy. Institute of Current World Affairs. 2021 Apr 19. URL: <https://www.icwa.org/the-struggle-to-keep-lebanon-healthy/> [accessed 2022-11-18]
5. ESCWA warns: more than half of Lebanon's population trapped in poverty. United Nations ESCWA. 2020 Aug 19. URL: <https://www.unescwa.org/news/Lebanon-poverty-2020> [accessed 2020-12-12]
6. Country cooperation strategy: Lebanon: 2019-2023. World Health Organization. Regional Office for the Eastern Mediterranean. 2018. URL: <https://apps.who.int/iris/handle/10665/279422> [accessed 2022-11-18]
7. Corona Virus Vaccination Dashboard. Ministry of Public Health (Lebanon). URL: <https://www.moph.gov.lb/maps/covid19.php> [accessed 2021-11-09]
8. Lebanon. COVID19 Vaccine Tracker. URL: <https://covid19.trackvaccines.org/country/lebanon/> [accessed 2022-04-05]
9. El Deeb S. Lebanon administers 1st vaccines to health workers, elderly. Associated Press News. 2021 Feb 14. URL: <https://apnews.com/article/financial-markets-lebanon-coronavirus-pandemic-coronavirus-vaccine-4fa8a304d97d5203f61f4935304c7737> [accessed 2022-11-18]

10. Mroue B. Lebanon private sector steps in to speed up vaccine campaign. Associated Press News. 2021 Mar 26. URL: <https://apnews.com/article/lebanon-coronavirus-pandemic-russia-financial-markets-coronavirus-vaccine-b81a58fa49d276315c5b6fed0ef8fe53> [accessed 2022-11-18]
11. Lebanon National Deployment and Vaccination Plan for COVID-19 Vaccines. Republic of Lebanon Ministry of Public Health. 2021 Jan 28. URL: <https://www.moph.gov.lb/userfiles/files/Prevention/COVID-19%20Vaccine/Lebanon%20NDVP-%20Feb%2016%202021.pdf> [accessed 2022-11-18]
12. Lebanon reintroduces some COVID-19 prevention measures. Reuters. 2021 Dec 1. URL: <https://www.reuters.com/world/middle-east/lebanon-reintroduces-some-covid-19-prevention-measures-2021-12-01/> [accessed 2022-04-05]
13. Sakr S, Ghaddar A, Sheet I, Eid A, Hamam B. Knowledge, attitude and practices related to COVID-19 among young Lebanese population. BMC Public Health 2021 Apr 06;21(1):653 [FREE Full text] [doi: [10.1186/s12889-021-10575-5](https://doi.org/10.1186/s12889-021-10575-5)] [Medline: [33823826](https://pubmed.ncbi.nlm.nih.gov/33823826/)]
14. Kasrine Al Halabi C, Obeid S, Sacre H, Akel M, Hallit R, Salameh P, et al. Attitudes of Lebanese adults regarding COVID-19 vaccination. BMC Public Health 2021 May 27;21(1):998 [FREE Full text] [doi: [10.1186/s12889-021-10902-w](https://doi.org/10.1186/s12889-021-10902-w)] [Medline: [34044790](https://pubmed.ncbi.nlm.nih.gov/34044790/)]
15. Sallam M, Al-Sanafi M, Sallam M. A global map of COVID-19 vaccine acceptance rates per country: an updated concise narrative review. JMDH 2022 Jan; Volume 15:21-45 [FREE Full text] [doi: [10.2147/jmdh.s347669](https://doi.org/10.2147/jmdh.s347669)]
16. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. The Lancet Public Health 2021 Apr;6(4):e210-e221 [FREE Full text] [doi: [10.1016/s2468-2667\(21\)00012-8](https://doi.org/10.1016/s2468-2667(21)00012-8)]
17. Gadoth A, Halbrook M, Martin-Blais R, Gray A, Tobin N, Ferbas K, et al. Cross-sectional assessment of COVID-19 vaccine acceptance among health care workers in Los Angeles. Ann Intern Med 2021 Jun;174(6):882-885 [FREE Full text] [doi: [10.7326/m20-7580](https://doi.org/10.7326/m20-7580)]
18. Sallam M. COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. Vaccines (Basel) 2021 Feb 16;9(2):A [FREE Full text] [doi: [10.3390/vaccines9020160](https://doi.org/10.3390/vaccines9020160)] [Medline: [33669441](https://pubmed.ncbi.nlm.nih.gov/33669441/)]
19. Guidance on developing a national deployment and vaccination plan for COVID-19 vaccines. World Health Organization. URL: <https://www.who.int/publications/i/item/WHO-2019-nCoV-Vaccine-deployment-2021.1-eng> [accessed 2022-11-18]
20. Brenner H. Focusing COVID-19 vaccinations on elderly and high-risk people. Lancet Reg Health Eur 2021 Mar;2:100044 [FREE Full text] [doi: [10.1016/j.lanepe.2021.100044](https://doi.org/10.1016/j.lanepe.2021.100044)] [Medline: [34173633](https://pubmed.ncbi.nlm.nih.gov/34173633/)]
21. Al-Mohaithef M, Padhi B. Determinants of COVID-19 vaccine acceptance in Saudi Arabia: a web-based national survey. JMDH 2020 Nov; Volume 13:1657-1663 [FREE Full text] [doi: [10.2147/jmdh.s276771](https://doi.org/10.2147/jmdh.s276771)]
22. Qualtrics XM. URL: <https://www.qualtrics.com/> [accessed 2022-11-18]
23. Lazarus JV, Ratzan SC, Palayew A, Gostin LO, Larson HJ, Rabin K, et al. A global survey of potential acceptance of a COVID-19 vaccine. Nat Med 2021 Feb;27(2):225-228 [FREE Full text] [doi: [10.1038/s41591-020-1124-9](https://doi.org/10.1038/s41591-020-1124-9)] [Medline: [33082575](https://pubmed.ncbi.nlm.nih.gov/33082575/)]
24. Fisher KA, Bloomstone S, Walder J, Crawford S, Fouayzi H, Mazor K. Attitudes toward a potential SARS-CoV-2 vaccine. Ann Intern Med 2020 Dec 15;173(12):964-973 [FREE Full text] [doi: [10.7326/m20-3569](https://doi.org/10.7326/m20-3569)]
25. Excel. Microsoft Corporation. URL: <https://www.microsoft.com/en-us/microsoft-365/p/excel/cfq7ttc0hr4r> [accessed 2022-11-18]
26. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. URL: <https://www.r-project.org/> [accessed 2022-11-19]
27. Labour Force and Household Living Conditions Survey (LFHLCS) in Lebanon 2018–2019. International Labour Organization. 2019 Dec 18. URL: https://www.ilo.org/beirut/publications/WCMS_732567/lang--en/index.htm [accessed 2022-11-19]
28. The World Factbook - Lebanon. CIA. URL: <https://www.cia.gov/the-world-factbook/countries/lebanon/> [accessed 2021-05-04]
29. Assessing the Impact of the Economic and COVID-19 Crises in Lebanon. United Nations World Food Programme. 2020 Jun. URL: <https://lebanon.un.org/en/107858-assessing-impact-economic-and-covid-19-crises-lebanon> [accessed 2022-11-19]
30. Lebanon. The World Bank. URL: <https://data.worldbank.org/country/lebanon> [accessed 2021-05-04]
31. Alqudeimat Y, Alenezi D, AlHajri B, Alfouzan H, Almokhaizeem Z, Altamimi S, et al. Acceptance of a COVID-19 vaccine and its related determinants among the general adult population in Kuwait. Med Princ Pract 2021;30(3):262-271 [FREE Full text] [doi: [10.1159/000514636](https://doi.org/10.1159/000514636)] [Medline: [33486492](https://pubmed.ncbi.nlm.nih.gov/33486492/)]
32. Alabdulla M, Reagu S, Al-Khal A, Elzain M, Jones R. COVID-19 vaccine hesitancy and attitudes in Qatar: A national cross-sectional survey of a migrant-majority population. Influenza Other Respir Viruses 2021 May;15(3):361-370 [FREE Full text] [doi: [10.1111/irv.12847](https://doi.org/10.1111/irv.12847)] [Medline: [33605010](https://pubmed.ncbi.nlm.nih.gov/33605010/)]
33. Sallam M, Dababseh D, Eid H, Al-Mahzoum K, Al-Haidar A, Taim D, et al. High rates of COVID-19 vaccine hesitancy and its association with conspiracy beliefs: a study in Jordan and Kuwait among other Arab countries. Vaccines (Basel) 2021 Jan 12;9(1):A [FREE Full text] [doi: [10.3390/vaccines9010042](https://doi.org/10.3390/vaccines9010042)] [Medline: [33445581](https://pubmed.ncbi.nlm.nih.gov/33445581/)]
34. Institute of Global Health Innovation. Global attitudes towards a COVID-19 vaccine. Imperial College London. 2021 May. URL: https://www.imperial.ac.uk/media/imperial-college/institute-of-global-health-innovation/GlobalVaccineInsights_ICL-YouGov-Covid-19-Behaviour-Tracker_20210520_v2.pdf [accessed 2022-11-19]

35. Lin C, Tu P, Beitsch L. Confidence and receptivity for COVID-19 vaccines: a rapid systematic review. *Vaccines (Basel)* 2020 Dec 30;9(1):A [FREE Full text] [doi: [10.3390/vaccines9010016](https://doi.org/10.3390/vaccines9010016)] [Medline: [33396832](https://pubmed.ncbi.nlm.nih.gov/33396832/)]
36. Bou Hamdan M, Singh S, Polavarapu M, Jordan T, Melhem N. COVID-19 vaccine hesitancy among university students in Lebanon. *Epidemiol. Infect* 2021 Nov 02;149:1 [FREE Full text] [doi: [10.1017/s0950268821002314](https://doi.org/10.1017/s0950268821002314)]
37. Murphy J, Vallières F, Bentall R, Shevlin M, McBride O, Hartman T, et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. *Nat Commun* 2021 Jan 04;12(1):29 [FREE Full text] [doi: [10.1038/s41467-020-20226-9](https://doi.org/10.1038/s41467-020-20226-9)] [Medline: [33397962](https://pubmed.ncbi.nlm.nih.gov/33397962/)]
38. Machida M, Nakamura I, Kojima T, Saito R, Nakaya T, Hanibuchi T, et al. Acceptance of a COVID-19 vaccine in Japan during the COVID-19 pandemic. *Vaccines (Basel)* 2021 Mar 03;9(3):210 [FREE Full text] [doi: [10.3390/vaccines9030210](https://doi.org/10.3390/vaccines9030210)] [Medline: [33802285](https://pubmed.ncbi.nlm.nih.gov/33802285/)]
39. Lindholt MF, Jørgensen F, Bor A, Petersen M. Public acceptance of COVID-19 vaccines: cross-national evidence on levels and individual-level predictors using observational data. *BMJ Open* 2021 Jun 15;11(6):e048172 [FREE Full text] [doi: [10.1136/bmjopen-2020-048172](https://doi.org/10.1136/bmjopen-2020-048172)] [Medline: [34130963](https://pubmed.ncbi.nlm.nih.gov/34130963/)]
40. Salibi N, Abdulrahim S, El Haddad M, Bassil S, El Khoury Z, Ghattas H, et al. COVID-19 vaccine acceptance in older Syrian refugees: Preliminary findings from an ongoing study. *Prev Med Rep* 2021 Dec;24:101606 [FREE Full text] [doi: [10.1016/j.pmedr.2021.101606](https://doi.org/10.1016/j.pmedr.2021.101606)] [Medline: [34692378](https://pubmed.ncbi.nlm.nih.gov/34692378/)]
41. El Khoury G, Salameh P. Influenza vaccination: a cross-sectional survey of knowledge, attitude and practices among the Lebanese adult population. *Int J Environ Res Public Health* 2015 Dec 05;12(12):15486-15497 [FREE Full text] [doi: [10.3390/ijerph121215000](https://doi.org/10.3390/ijerph121215000)] [Medline: [26690189](https://pubmed.ncbi.nlm.nih.gov/26690189/)]
42. Romani MH, Musharrafieh U, Lakkis N, Hamadeh G. Family physicians beliefs and attitudes regarding adult pneumococcal and influenza immunization in Lebanon. *Fam Pract* 2011 Dec;28(6):632-637 [FREE Full text] [doi: [10.1093/fampra/cmr038](https://doi.org/10.1093/fampra/cmr038)] [Medline: [21746697](https://pubmed.ncbi.nlm.nih.gov/21746697/)]
43. Kmeid M, Azouri H, Aaraj R, Bechara E, Antonios D. Vaccine coverage for Lebanese citizens and Syrian refugees in Lebanon. *Int Health* 2019 Nov 13;11(6):568-579 [FREE Full text] [doi: [10.1093/inthealth/ihz023](https://doi.org/10.1093/inthealth/ihz023)] [Medline: [31089721](https://pubmed.ncbi.nlm.nih.gov/31089721/)]
44. Mansour Z, Brandt L, Said R, Fahmy K, Riedner G, Danovaro-Holliday M. Home-based records' quality and validity of caregivers' recall of children's vaccination in Lebanon. *Vaccine* 2019 Jul 09;37(30):4177-4183 [FREE Full text] [doi: [10.1016/j.vaccine.2019.05.032](https://doi.org/10.1016/j.vaccine.2019.05.032)] [Medline: [31221562](https://pubmed.ncbi.nlm.nih.gov/31221562/)]
45. Robertson T, Weiss W, Jordan Health Access Study Team, Lebanon Health Access Study Team, Doocy S. Challenges in estimating vaccine coverage in refugee and displaced populations: results from household surveys in Jordan and Lebanon. *Vaccines (Basel)* 2017 Aug 12;5(3):1 [FREE Full text] [doi: [10.3390/vaccines5030022](https://doi.org/10.3390/vaccines5030022)] [Medline: [28805672](https://pubmed.ncbi.nlm.nih.gov/28805672/)]
46. Below the Surface: Results of a WhatsApp Survey of Syrian Refugees and Host Communities in Lebanon. United Nations Development Programme. 2018. URL: <https://data2.unhcr.org/en/documents/details/67579> [accessed 2021-05-07]

Edited by G Eysenbach; submitted 26.01.22; peer-reviewed by M Sallam, I Mircheva; comments to author 18.03.22; revised version received 06.04.22; accepted 14.09.22; published 14.12.22.

Please cite as:

Abou-Arraj NE, Maddah D, Buhamdan V, Abbas R, Jawad NK, Karaki F, Alami NH, Geldsetzer P

Perceptions of, and Obstacles to, SARS-CoV-2 Vaccination Among Adults in Lebanon: Cross-sectional Online Survey

JMIR Form Res 2022;6(12):e36827

URL: <https://formative.jmir.org/2022/12/e36827>

doi: [10.2196/36827](https://doi.org/10.2196/36827)

PMID: [36383635](https://pubmed.ncbi.nlm.nih.gov/36383635/)

©Nadeem Elias Abou-Arraj, Diana Maddah, Vanessa Buhamdan, Roua Abbas, Nadine Kamel Jawad, Fatima Karaki, Nael H Alami, Pascal Geldsetzer. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 14.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Twitter Data to Estimate the Prevalence of Symptoms of Mental Disorders in the United States During the COVID-19 Pandemic: Ecological Cohort Study

Ruilie Cai¹, MSc; Jiajia Zhang^{1,2,3}, PhD; Zhenlong Li^{2,3,4}, PhD; Chengbo Zeng^{2,3,5}, PhD; Shan Qiao^{2,3,5}, PhD; Xiaoming Li^{2,3,5}, PhD

¹Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

²South Carolina SmartState Center for Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

³University of South Carolina Big Data Health Science Center, Columbia, SC, United States

⁴Geoinformation and Big Data Research Lab, Department of Geography, University of South Carolina, Columbia, SC, United States

⁵Department of Health Promotion, Education and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

Corresponding Author:

Ruilie Cai, MSc

Department of Epidemiology and Biostatistics

Arnold School of Public Health

University of South Carolina

921 Assembly Street

Columbia, SC, 29208

United States

Phone: 1 8039556789

Email: rcai@email.sc.edu

Abstract

Background: Existing research and national surveillance data suggest an increase of the prevalence of mental disorders during the COVID-19 pandemic. Social media platforms, such as Twitter, could be a source of data for estimation owing to its real-time nature, high availability, and large geographical coverage. However, there is a dearth of studies validating the accuracy of the prevalence of mental disorders on Twitter compared to that reported by the Centers for Disease Control and Prevention (CDC).

Objective: This study aims to verify the feasibility of Twitter-based prevalence of mental disorders symptoms being an instrument for prevalence estimation, where feasibility is gauged via correlations between Twitter-based prevalence of mental disorder symptoms (ie, anxiety and depressive symptoms) and that based on national surveillance data. In addition, this study aims to identify how the correlations changed over time (ie, the temporal trend).

Methods: State-level prevalence of anxiety and depressive symptoms was retrieved from the national Household Pulse Survey (HPS) of the CDC from April 2020 to July 2021. Tweets were retrieved from the Twitter streaming application programming interface during the same period and were used to estimate the prevalence of symptoms of mental disorders for each state using keyword analysis. Stratified linear mixed models were used to evaluate the correlations between the Twitter-based prevalence of symptoms of mental disorders and those reported by the CDC. The magnitude and significance of model parameters were considered to evaluate the correlations. Temporal trends of correlations were tested after adding the time variable to the model. Geospatial differences were compared on the basis of random effects.

Results: Pearson correlation coefficients between the overall prevalence reported by the CDC and that on Twitter for anxiety and depressive symptoms were 0.587 ($P<.001$) and 0.368 ($P<.001$), respectively. Stratified by 4 phases (ie, April 2020, August 2020, October 2020, and April 2021) defined by the HPS, linear mixed models showed that Twitter-based prevalence for anxiety symptoms had a positive and significant correlation with CDC-reported prevalence in phases 2 and 3, while a significant correlation for depressive symptoms was identified in phases 1 and 3.

Conclusions: Positive correlations were identified between Twitter-based and CDC-reported prevalence, and temporal trends of these correlations were found. Geospatial differences in the prevalence of symptoms of mental disorders were found between the northern and southern United States. Findings from this study could inform future investigation on leveraging social media

platforms to estimate symptoms of mental disorders and the provision of immediate prevention measures to improve health outcomes.

(*JMIR Form Res* 2022;6(12):e37582) doi:[10.2196/37582](https://doi.org/10.2196/37582)

KEYWORDS

mental health; anxiety disorder; depressive disorder; COVID-19; national survey; social media; Twitter; mixed model; anxiety; National Household Pulse survey; geospatial

Introduction

Mental disorders are a significant public health challenge in the United States. It was estimated that approximately 30% of US adults experience a disorder once in a 12-month span, and the prevalence of lifetime disorder could be as high as 50% [1,2]. Anxiety disorder is one of the most common mental disorders. The lifetime prevalence of generalized anxiety disorder ranged from 3.6% to 5.1% [2,3]. One general comorbidity of anxiety is major depression, which is another common mental disorder. The estimated prevalence of major depressive episodes was 8.12% in 2019-2020 [4]. Some previous national surveys reported that nearly 7% of adults had a symptom of depressive disorder [5-7]. Mental disorders increase the burden of socioeconomics and health care usage [8-10]. Depressive disorders and anxiety disorders were reported to be the 13th and 24th leading contributors to the global burden in 2019, respectively [11].

The COVID-19 pandemic posed significant challenges to people's psychosocial well-being and mental health. On the one hand, the rapid transmission and variation of SARS-CoV-2 resulted in a high disease incidence and considerable mortality [12,13], which may increase fear toward COVID-19. On the other hand, myths and misinformation on social media, mandatory government responses such as citywide lockdowns, and discrimination and stigma against people infected with COVID-19 would exacerbate the mental health problems in the public [14,15]. Therefore, during the COVID-19 pandemic, anxiety and depressive disorders were common health problems among not only patients with COVID-19 [16,17] and those with pre-existing psychiatry disorders [18,19] but also the general population [20,21]. Furthermore, as a result of the increased workload in the health care system caused by the COVID-19 pandemic, an increasing number of health care providers are experiencing mental disorders [22-24]. Thus, during the outbreak period, monitoring and estimating the prevalence of mental disorders is critical to inform immediate prevention measures, reduce health care costs, and improve population health.

Social media platforms, such as Twitter, can be a potentially useful tool for estimating the prevalence of mental disorders. Messages posted on social media platforms (eg, tweets) expressing negative emotions may provide researchers with hints about the prevalence of mental disorders. Given the time-intensiveness of and high costs in recruiting participants and conducting surveys or screening using existing scales, a social media-based approach can provide timely estimation and prediction at a lower cost. This timeliness is especially a desperate need during the COVID-19 pandemic for immediate public health practices.

Social media data can provide population-wide information for estimating the prevalence of mental disorders. Previous research has taken advantage of the large volume of social media data for mental health research on public sentiment [25], estimating suicide risk in a particular population [26], or identifying risk factors and linguistic features [27,28]. Although social media data have been used for the estimation of various health issues [6,28,29], there is a dearth of research investigating whether social media can be a good data source for estimating the mental health status of the public during the COVID-19 pandemic. If it is feasible to use social media as a good instrument for estimating symptoms of mental disorders, time and resources could be saved, and more action could be taken for intervention.

To address this knowledge gap, we conducted an ecological cohort study that aims to (1) verify the feasibility of Twitter-based prevalence of the symptoms of mental disorders being an instrument for prevalence estimation, where the feasibility is gauged via the correlations between Twitter-based prevalence of symptoms (ie, anxiety and depressive symptoms) and those reported by the national surveillance data; and (2) examine the temporal trends of these correlations during different phases of the COVID-19 pandemic. Twitter is used because of its "open data" policy so that its database can be easily accessed by the official Twitter streaming application programming interface [30].

Methods

Data Resources and Preprocessing

The national prevalence of the symptoms of mental disorders was collected through the Household Pulse Survey (HPS), which was conducted by the National Center for Health Statistics and the Census Bureau [31]. Briefly, the HPS is a 20-minute web-based survey using a probability-based sample design to measure the social and economic impact of the COVID-19 pandemic, including the anxiety and depression status of individuals in US households; samples were chosen from the Census Bureau Master Address File Data, housing units were randomly selected to participate, and one respondent from each housing unit was selected to respond for him- or herself [32]. Measures of depressive and anxiety symptoms were based on the validated 2-item Patient Health Questionnaire and 2-item Generalized Anxiety Disorder scale, respectively. Adults were confirmed as having disorder symptoms if items included in the questionnaire generally occurred on more than half of the days or nearly every day during the past 7 days [33]. In this study, the data were aggregated by state, and collection periods consisted of 4 phases that occurred weekly from April 23 to July 21, 2020 (phase 1), biweekly from August 19 to October 26, 2020 (phase 2), biweekly from October 28 to December 21,

2020 (phase 3), biweekly from January 6 to March 29, 2021 (phase 3), and biweekly from April 14 to July 5, 2021 (phase 3.1), and the data were recorded at the end of a week.

Social media data were collected from the Twitter platform using the official Twitter streaming application programming interface from April 23, 2020, to July 5, 2021. The data were collected in JSON format. The following information was extracted: tweet ID, user ID, posted date, message, and location (to determine which state a tweet originated from). In total, approximately 5 million (N=5,099,770) Twitter users were sampled. Nonhuman user accounts (bots) were excluded

following the procedure detailed in a previous study [34]. According to previous research [29,35], keywords related to anxiety and depressive symptoms were used to label the tweets and estimate the weekly or biweekly prevalence of symptoms of anxiety, depression, and overall mental disorders (anxiety or depressive symptoms) for each state at each phase provided in the HPS. The total number of users after keyword filtering was 538,801. Table 1 shows the keywords used for labeling. The prevalence of symptoms during the 4 phases was calculated as the proportion of Twitter users who had keyword-filtered anxiety and depressive symptoms tweets.

Table 1. List of keywords^a.

Keywords related to anxiety symptoms	Keywords related to depressive symptoms
Anxiety	depression
Anxious	depressed
Irritable	depressive
Restless	feeling blue
Feared	insomnia
Scared	clonazepam
Nervous	imipramine
Outrage	prozac
Dread	sertraline
panic	zoloft
	hopeless

^aThe keyword list is based on Jashinsky et al [29] and Homan et al [35].

Statistical Analysis

After data preprocessing, each state had 2 sets of data (ie, both CDC-reported prevalence and Twitter-based prevalence) for symptoms of anxiety, depression, and overall mental disorders (anxiety or depressive symptoms) at each time point. The associations between CDC-reported and Twitter-based prevalence were examined using Pearson correlation analysis. The temporal trends between them were further examined through linear mixed models with random intercepts. In linear mixed models, we used CDC-reported prevalence in each state as the outcome and Twitter-based prevalence in each state and time (in weeks) as covariates. The different measures in each state were accounted for by the normal random intercept at the state level. The interaction between Twitter-based prevalence and time was further investigated and determined with the best-subset selection based on the Akaike information criterion. Restricted maximum likelihood was used for parameter estimation. Geospatial differences were compared on the basis

of random effects. Analysis was conducted in R (version 4.0.3; R Foundation for Statistical Computing) and GeoPandas. The significant level was set as $\alpha=.05$ (2-sided).

Results

Descriptive Statistics

Overall, 49 contiguous states, except for Alaska and Hawaii, were included for the analysis. Table 2 summarizes the CDC-reported and Twitter-based prevalence of symptoms of anxiety, depression, and overall mental disorders across the 4 phases. According to the CDC report, nearly 30% of the participants in different states experienced symptoms of anxiety, while 24% of participants experienced symptoms of depression, and 35% of participants experienced overall symptoms of mental disorders. On Twitter, 1% of Twitter users discussed the topic of anxiety, while 0.27% of them discussed depression, and 1.3% of them discussed overall mental disorders.

Table 2. Descriptive statistics for the symptoms of mental disorders.

	Centers for Disease Control and Prevention			Twitter		
	Anxiety (%), median (IQR)	Depression (%), median (IQR)	Overall ^a (%), median (IQR)	Anxiety (%), median (IQR)	Depression (%), median (IQR)	Overall ^a (%), median (IQR)
Phase 1 ^b	30.60 (27.70-33.70)	24.75 (22.10-27.60)	35.05 (32.60-38.60)	1.21 (1.10-1.34)	0.29 (0.25-0.33)	1.47 (1.34-1.61)
Phase 2 ^b	31.50 (29.20-33.70)	23.90 (21.80-26.10)	36.10 (33.30-38.40)	1.20 (1.11-1.30)	0.30 (0.26-0.34)	1.47 (1.35-1.59)
Phase 3 ^b	34.40 (31.60-37.10)	27.30 (24.70-29.80)	39.60 (36.60-42.30)	1.03 (0.93-1.16)	0.27 (0.23-0.31)	1.26 (1.15-1.41)
Phase 3.1 ^b	25.20 (23.00-27.50)	20.90 (18.70-22.90)	29.75 (29.00-32.20)	0.92 (0.81-1.02)	0.22 (0.19-0.24)	1.11 (1.01-1.22)

^aAnxiety or depressive symptoms.

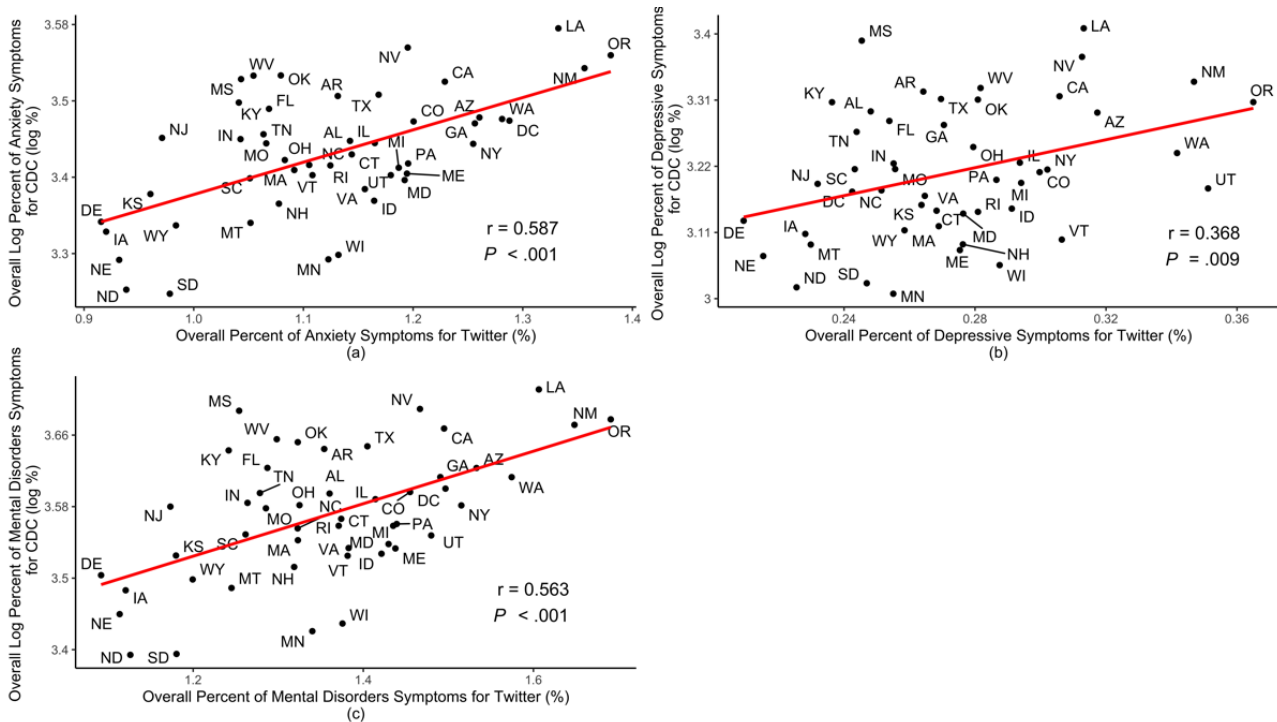
^bPhase 1: April 23 to July 21, 2020; phase 2: August 19 to October 26, 2020; phase 3: October 28 to December 21, 2020, and January 6 to March 29, 2021; and phase 3.1: April 14 to July 5, 2021.

Correlation Analyses

Among the 49 states, the overall (anxiety or depressive symptoms) Twitter-based prevalence of symptoms of mental disorders showed a significant positive correlation with

CDC-reported prevalence. Specifically, as shown in Figure 1, Pearson correlation coefficients for the prevalence of symptoms of anxiety, depression, and mental disorders reported by the CDC and estimated using Twitter were 0.587 ($P < .001$), 0.368 ($P < .001$), and 0.563 ($P < .001$), respectively.

Figure 1. Correlations between Twitter-based prevalence and the CDC-reported prevalence of symptoms of (A) anxiety, (B) depression, and (C) overall mental disorders. CDC: Centers for Disease Control and Prevention.



Regression Analysis

Models with predictors including time, Twitter-based prevalence, and the interaction between them had the lowest Akaike information criterion for symptoms of anxiety, depression, and overall mental disorders, which was used as the final model. Detailed coefficients and numbers of repeat-measured observations of each model are reported in Multimedia Appendix 1.

The 3-D plots (Figures 2-4) were used to illustrate the interplay among the fitted CDC-reported prevalence, Twitter-based prevalence, and time. In Figures 2-4, the x-axis, “Time,” indicates that the study period that was from May 5, 2020, to July 5, 2021, and the y-axis, “Twitter,” showed the range of Twitter-based prevalence for anxiety symptoms (as %); the z-axis, “Effect,” represents the mean value of model-estimated prevalence of anxiety symptoms (in %). The solid surface in the middle was the mean value, and translucent surfaces above and below show the upper and lower bounds of the 95% CI.

Figure 2. Model-based mean function of CDC-reported prevalence of anxiety symptoms. The “Time” axis shows the study period; the “Twitter” axis shows the range of Twitter-based prevalence for anxiety symptoms; and the “Effect” axis represents the mean value of model-estimated prevalence of anxiety symptoms. The solid surface in the middle was the mean value and translucent surfaces above and below showed the upper and lower bounds of the 95% CI.

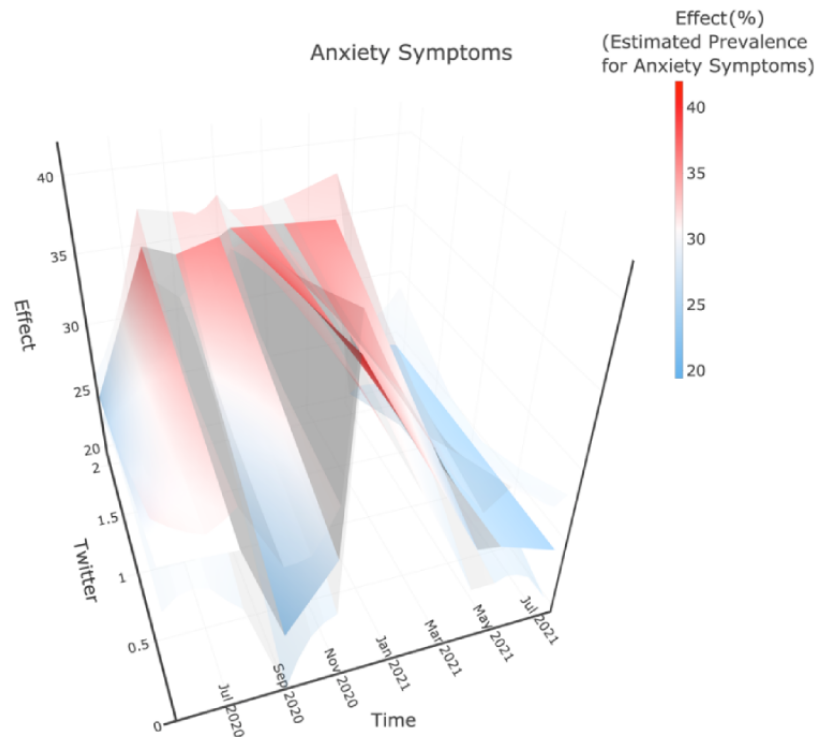


Figure 3. Model-based mean function of CDC-reported prevalence of depressive symptoms. The “Time” axis shows the study period; the “Twitter” axis shows the range of Twitter-based prevalence for depressive symptoms; and the “Effect” axis represents the mean value of model-estimated prevalence of depressive symptoms. The solid surface in the middle was the mean value and translucent surfaces above and below showed the upper and lower bounds of the 95% CI.

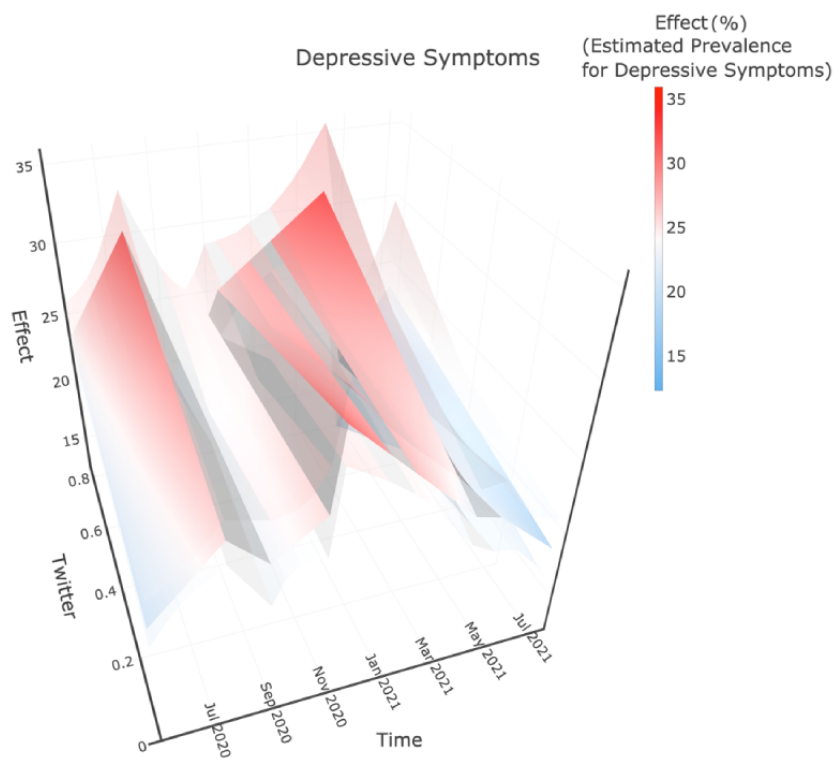
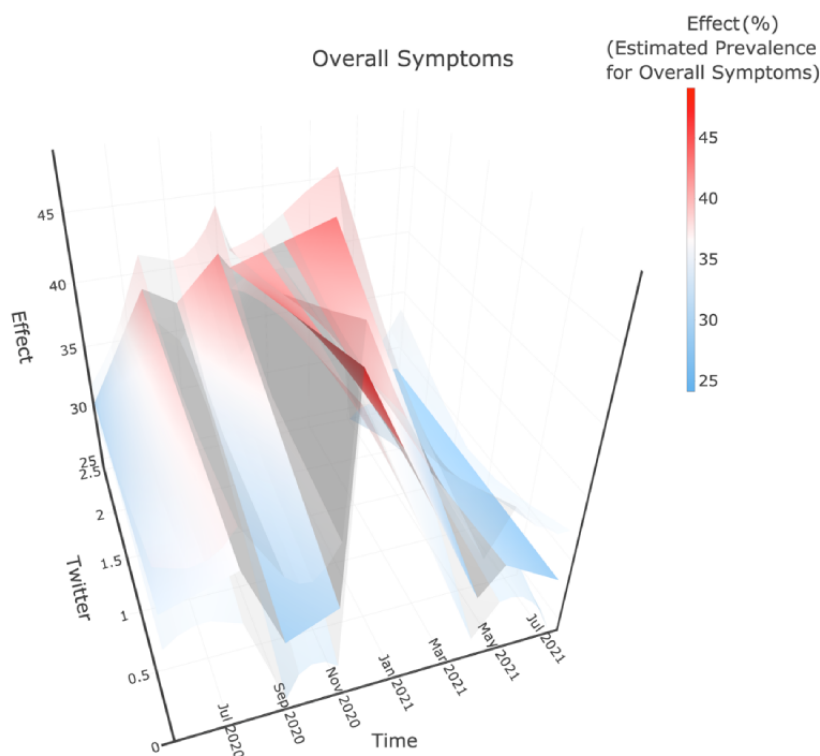


Figure 4. Model-based mean function of CDC-reported prevalence of symptoms of overall mental disorders. The “Time” axis shows the study period; the “Twitter” axis shows the range of Twitter-based prevalence of symptoms of mental disorders; and the “Effect” axis represents the mean value of model-estimated prevalence of symptoms. The solid surface in the middle was the mean value and translucent surfaces above and below showed the upper and lower bounds of the 95% CI.



Anxiety Symptoms

In general, anxiety symptoms increased until November 2020 and then decreased rapidly. The largest fitted prevalence was found at the beginning of phase 3, which is shown as the reddest surface. To clarify the relationship, the 2D association of Twitter-based prevalence and estimated CDC-reported prevalence is drawn in [Multimedia Appendix 2](#). We obtained negative but insignificant coefficients at the beginning of phases 1 and 3, while significant positive coefficients were obtained at phase 2 and the rest of phases 1 and 3.

Depressive Symptoms

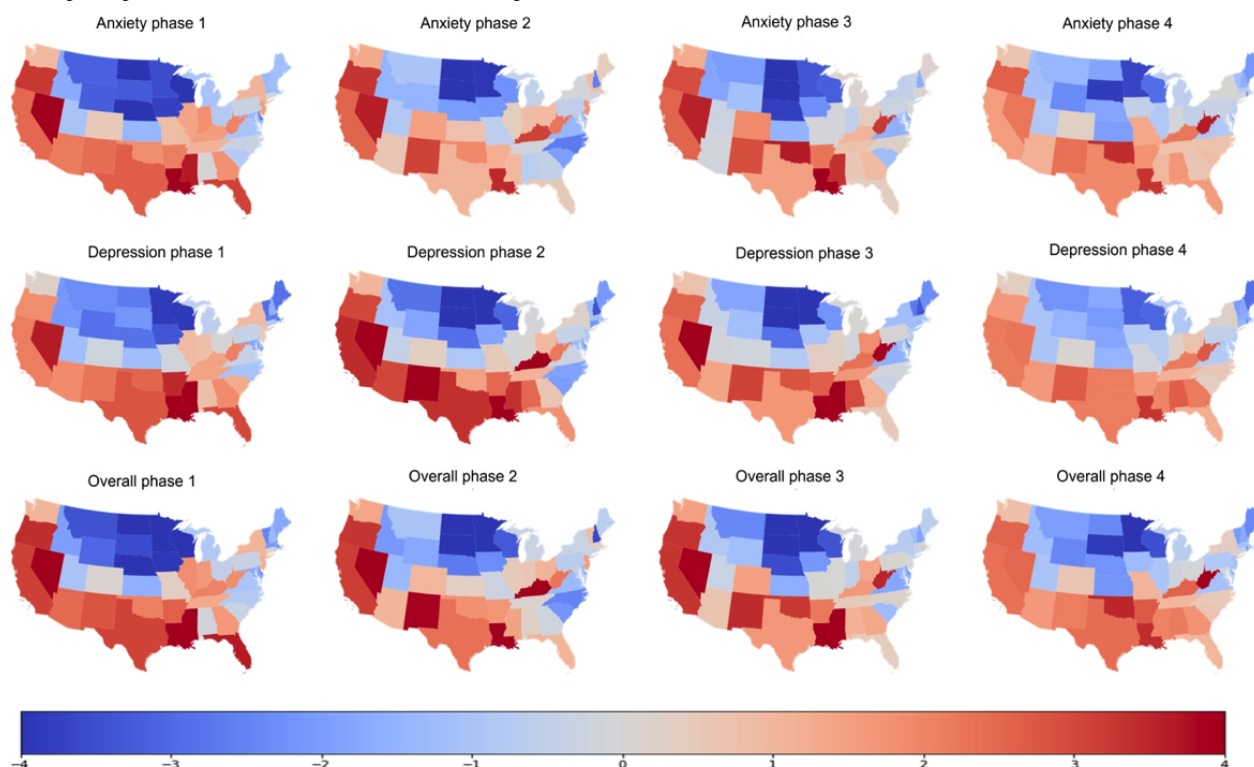
The model-estimated prevalence was initially low, increased rapidly in phase 1, but decreased slightly in phase 2. In phase 3, the prevalence was maintained at a high level but plummeted at the beginning of phase 3.1. The pattern is further shown in [Multimedia Appendix 3](#), which indicates that Twitter-based prevalence had a significant and positive correlation with CDC-reported prevalence only in phase 1 and the second part of phase 3. Correspondingly, in [Figure 3](#), temporal trends in correlations are shown through color variation from white to red in phase 1 and in the second part of phase 3.

Symptoms of Overall Mental Disorders

The surfaces of fitted CDC-reported prevalence of the symptoms of overall mental disorders were similar to those for anxiety symptoms. The marginal coefficient of the Twitter-based prevalence had a decreasing trend for anxiety symptoms but had an increasing trend for symptoms of overall mental disorders in phase 2, while both of them showed positive correlations ([Multimedia Appendix 4](#)).

Spatial Patterns of the Random Effects

Geospatial disparities in the prevalence of symptoms of mental disorders across the states were characterized by the random effects of linear mixed models. A larger random effect indicated a larger deviance of a state from the nationwide average. To show the geospatial disparities, maps with random effects were drawn ([Figure 5](#)). A clear geospatial difference in the prevalence of symptoms of mental disorders between the northern and southern United States was found. Specifically, states located in the north-central and northeast United States had relatively lower mental health prevalence than those in the southern and southwestern United States across the 4 phases and 2 types of symptoms of mental disorders. As shown in the fourth column in [Figure 5](#), the colors are lighter and the random effects are closer to 0, which indicates that the disparity decreased in phase 3.1.

Figure 5. Spatial patterns of the random effects for different phases.

Discussion

Principal Findings

This study examined the correlations in the prevalence of symptoms of mental disorders between those reported by the CDC and those estimated using Twitter data, as well as the temporal trends of these correlations. We note that the results of our study only capture the relationship among symptoms of mental disorders and do not imply mental disorders, since the CDC-reported prevalence was based on the 2-item Patient Health Questionnaire and the 2-item Generalized Anxiety Disorder scale. These findings indicate medium positive correlations between CDC-reported and Twitter-based prevalence for anxiety and overall mental disorder symptoms, and a weaker positive correlation for depressive symptoms. These correlations varied by different phases of the COVID-19 pandemic. For example, the correlation was significantly positive in phase 2 and late-phase 3 but was insignificant at other periods for both anxiety and overall mental disorders. Clear geospatial disparities in the prevalence of symptoms of mental disorders between the northern and southern United States were also found.

Our findings are consistent with those of previous research; Hussain [36] also identified the correlation between social media-based estimates and survey-reported prevalence by estimating public sentiment toward COVID-19 vaccines, using data from Facebook and Twitter platforms in United States and the United Kingdom and comparing the results with nationwide surveys [36]. Therefore, our findings could provide another piece of empirical evidence on the feasibility of using Twitter to estimate the prevalence of symptoms of mental disorders.

The correlations between Twitter-based and CDC-reported prevalence varied by different phases during the COVID-19 pandemic. The possible explanation is that Twitter-based prevalence is more time sensitive than the CDC-reported prevalence. The emotions potentially reflected through posting tweets are sensitive to daily events (eg, lifestyle change and staying at home) and social events (eg, citywide lockdown) [36,37]. However, the CDC-reported prevalence may lag behind the Twitter-based prevalence, which results in temporal trends in the correlation between them. Another explanation is that there may be a time lag between social events and some emotional responses. For example, Zhang et al [37] found that Twitter users with depression would respond to the pandemic and post tweets later than those without depression. This may indicate the heterogeneity in the emotional responses regarding posting tweets, which could affect the accuracy of Twitter-estimated symptoms of mental disorders at different time points and induce its temporal correlation with CDC-reported prevalence. Future studies are needed to consider the time lag and the heterogeneity in emotional responses when using tweets to estimate the prevalence of symptoms of mental disorders.

Maps with random effects showed clear spatiotemporal disparities in the prevalence of symptoms of mental disorders in states located in the north-central and northeast United States, which had relatively lower prevalence of mental disorder symptoms than those from other parts of the United States. The disparities may be related to variations in the COVID-19 pandemic in different parts of the United States. On the one hand, different locations have a varying burden of COVID-19. For example, Miller et al [38] mapped the disease and hospitalization burden and found that areas located in the western United States had a high concentration of the cumulative

number of severe COVID-19 cases. The overwhelming public health or health care system may fail to control disease transmission, which may increase panic in the public. On the other hand, the geospatial disparities in socioeconomic status (SES) may also contribute to the pattern found in our study. People living in low-SES regions may be at high risk of financial difficulty during the COVID-19 pandemic owing to stay-at-home orders, citywide lockdowns, and social distancing measures. Low SES has been proved to be associated with a greater risk of psychopathology after the COVID-19 pandemic [39]. However, previous studies did not show a consistent spatiotemporal pattern. For example, socioeconomic attributes (eg, education level, income, and occupation) have been confirmed to have a significant association with sentiment among Twitter users, but the number of COVID-19 cases in a region did not show a significant sentiment association [40]; however, Whitney and Peterson [41] found a similar pattern among US youth to that reported in this study in which areas in the southern United States had a worse mental health status than those in the northern United States.

Implementation

Our model could be used as a quick instrument for estimating symptoms of mental health disorders. In the future, once an estimate of the prevalence of the symptoms of a certain mental health disorder is obtained from Twitter (via keywords or any other advanced estimation models), the Twitter-based estimate and corresponding time can be factored into our model, and an estimate of surveillance-based prevalence can be obtained.

Limitations

There are some limitations in our study, which need to be acknowledged. First, although we retrieved representative keywords from previous research, there were still other keywords that were not considered in our study, which may result in the underestimated prevalence of symptoms of mental disorders. Second, since tweets with our predefined keywords may also refer to unrelated topics [42], the keywords used in this study may misclassify Twitter users into different groups

regarding the presence of symptoms of mental disorders, which may affect the accuracy of estimation and impair the validity of our findings. Future studies are needed to improve our strategy of refining tweets to obtain accurate estimations. Other advanced analytic approaches (eg, machine learning approaches) could be used in addition to more comprehensive keywords. Third, as an ecological study, the individual sociodemographic characteristics of Twitter users were not considered when estimating the prevalence of symptoms of mental disorders. Sociodemographic characteristics have a close relationship with the outcomes of mental disorders. Thus, incorporating confounders (eg, characteristics) into model training could improve the accuracy of model estimation [43-45]. Besides, other ecological confounders such as political proceedings and socioeconomic downfall are also not included in our study, and more efforts should be made to incorporate those confounders in future studies. Finally, the sociodemographics in Twitter data have a bias [30]. For example, compared to HPS participants aged 18 years and older, as most of the Twitter users tend to be young (18-29 years old) and in considerable SES [46], the mental health problems among them may be different from those of individuals who did not use Twitter. For instance, young high-SES adults may have better access to mental health resources or seek help proactively. Therefore, these populations may have a lower risk of experiencing symptoms of mental disorders than older lower-SES adults [47].

Conclusions

Positive correlations were found between CDC-reported and Twitter-based symptom prevalence, and temporal trends of these correlations were identified. Spatiotemporal disparities were also observed between the northern and southern United States. Findings from this study could inform future research to improve the accuracy of estimating the prevalence of symptoms of mental disorders using social media platforms. Public health practitioners and policy makers could use Twitter-based prevalence to inform immediate prevention measures and mental health services to cope with mental disorders during the COVID-19 pandemic and future public health emergencies.

Acknowledgments

This study was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (award R01AI127203-4S1). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Both the National Institute of Allergy and Infectious Diseases and the National Institutes of Health had no role in the design of the study or in the collection, analysis, and interpretation of the data.

Data Availability

Twitter data were collected using Twitter's public streaming application programming interface from the public domain in accordance with Twitter's Developer Agreement. Following Twitter's policy on "Redistribution of Twitter content" [48], individual tweets cannot be redistributed.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mixed Models Coefficients.

[[DOCX File , 16 KB - formative_v6i12e37582_app1.docx](#)]

Multimedia Appendix 2

Marginal Coefficient of Twitter-based Prevalence for Anxiety Symptoms.

[[PNG File , 55 KB - formative_v6i12e37582_app2.png](#)]

Multimedia Appendix 3

Marginal Coefficient of Twitter-based Prevalence for Depressive Symptoms.

[[PNG File , 54 KB - formative_v6i12e37582_app3.png](#)]

Multimedia Appendix 4

Marginal Coefficient of Twitter-based Prevalence for Overall Mental Disorder Symptoms.

[[PNG File , 57 KB - formative_v6i12e37582_app4.png](#)]

References

1. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994 Jan 01;51(1):8-19. [doi: [10.1001/archpsyc.1994.03950010008002](https://doi.org/10.1001/archpsyc.1994.03950010008002)] [Medline: [8279933](https://pubmed.ncbi.nlm.nih.gov/8279933/)]
2. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun 01;62(6):617-627 [FREE Full text] [doi: [10.1001/archpsyc.62.6.617](https://doi.org/10.1001/archpsyc.62.6.617)] [Medline: [15939839](https://pubmed.ncbi.nlm.nih.gov/15939839/)]
3. Wittchen H, Zhao S, Kessler RC, Eaton WW. DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1994 May 01;51(5):355-364. [doi: [10.1001/archpsyc.1994.03950050015002](https://doi.org/10.1001/archpsyc.1994.03950050015002)] [Medline: [8179459](https://pubmed.ncbi.nlm.nih.gov/8179459/)]
4. 2019-2020 National Survey On Drug Use And Health: Model-Based Prevalence Estimates (50 States And The District Of Columbia). Substance Abuse and Mental Health Services Administration. 2021. URL: <https://www.samhsa.gov/data/report/2019-2020-nsduh-state-prevalence-estimates> [accessed 2022-12-06]
5. Terlizzi E, Schiller J. Estimates of Mental Health Symptomatology, by Month of Interview: United States, 2019. National Center for Health Statistics. 2021. URL: <https://www.cdc.gov/nchs/data/nhis/mental-health-monthly-508.pdf> [accessed 2022-12-06]
6. McClellan C, Ali M, Mutter R, Kroutil L, Landwehr J. Using social media to monitor mental health discussions - evidence from Twitter. *J Am Med Inform Assoc* 2017 May 01;24(3):496-502 [FREE Full text] [doi: [10.1093/jamia/ocw133](https://doi.org/10.1093/jamia/ocw133)] [Medline: [27707822](https://pubmed.ncbi.nlm.nih.gov/27707822/)]
7. Results from the 2013 National Survey on Drug Use and Health: Mental Health Findings. US Department of Health and Human Services. 2014. URL: <https://www.samhsa.gov/data/sites/default/files/NSDUHmhfr2013/NSDUHmhfr2013.pdf> [accessed 2022-12-06]
8. Gold SM, Köhler-Forsberg O, Moss-Morris R, Mehnert A, Miranda JJ, Bullinger M, et al. Comorbid depression in medical diseases. *Nat Rev Dis Primers* 2020 Aug 20;6(1):69. [doi: [10.1038/s41572-020-0200-2](https://doi.org/10.1038/s41572-020-0200-2)] [Medline: [32820163](https://pubmed.ncbi.nlm.nih.gov/32820163/)]
9. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. *Nat Rev Dis Primers* 2016 Sep 15;2(1):16065. [doi: [10.1038/nrdp.2016.65](https://doi.org/10.1038/nrdp.2016.65)] [Medline: [27629598](https://pubmed.ncbi.nlm.nih.gov/27629598/)]
10. Wittchen H. Generalized anxiety disorder: prevalence, burden, and cost to society. *Depress Anxiety* 2002 Dec 18;16(4):162-171. [doi: [10.1002/da.10065](https://doi.org/10.1002/da.10065)] [Medline: [12497648](https://pubmed.ncbi.nlm.nih.gov/12497648/)]
11. GBD 2019 DiseasesInjuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9)] [Medline: [33069326](https://pubmed.ncbi.nlm.nih.gov/33069326/)]
12. Liu Y, Gayle A, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med* 2020 Mar 13;27(2) [FREE Full text] [doi: [10.1093/jtm/taaa021](https://doi.org/10.1093/jtm/taaa021)] [Medline: [32052846](https://pubmed.ncbi.nlm.nih.gov/32052846/)]
13. Rocklöv J, Sjödin H, Wilder-Smith A. COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *J Travel Med* 2020 May 18;27(3) [FREE Full text] [doi: [10.1093/jtm/taaa030](https://doi.org/10.1093/jtm/taaa030)] [Medline: [32109273](https://pubmed.ncbi.nlm.nih.gov/32109273/)]
14. Bao Y, Sun Y, Meng S, Shi J, Lu L. 2019-nCoV epidemic: address mental health care to empower society. *Lancet* 2020 Feb;395(10224):e37-e38. [doi: [10.1016/S0140-6736\(20\)30309-3](https://doi.org/10.1016/S0140-6736(20)30309-3)]
15. Ye Z, Yang X, Zeng C, Wang Y, Shen Z, Li X, et al. Resilience, social support, and coping as mediators between COVID-19-related stressful experiences and acute stress disorder among college students in China. *Appl Psychol Health Well Being* 2020 Dec 15;12(4):1074-1094 [FREE Full text] [doi: [10.1111/aphw.12211](https://doi.org/10.1111/aphw.12211)] [Medline: [32666713](https://pubmed.ncbi.nlm.nih.gov/32666713/)]
16. Bo H, Li W, Yang Y, Wang Y, Zhang Q, Cheung T, et al. Posttraumatic stress symptoms and attitude toward crisis mental health services among clinically stable patients with COVID-19 in China. *Psychol Med* 2020 Mar 27;51(6):1052-1053. [doi: [10.1017/s0033291720000999](https://doi.org/10.1017/s0033291720000999)]

17. Zhang J, Lu H, Zeng H, Zhang S, Du Q, Jiang T, et al. The differential psychological distress of populations affected by the COVID-19 pandemic. *Brain Behav Immun* 2020 Jul;87:49-50 [FREE Full text] [doi: [10.1016/j.bbi.2020.04.031](https://doi.org/10.1016/j.bbi.2020.04.031)] [Medline: [32304883](https://pubmed.ncbi.nlm.nih.gov/32304883/)]
18. Fernández-Aranda F, Casas M, Claes L, Bryan DC, Favaro A, Granero R, et al. COVID-19 and implications for eating disorders. *Eur Eat Disord Rev* 2020 May 28;28(3):239-245 [FREE Full text] [doi: [10.1002/erv.2738](https://doi.org/10.1002/erv.2738)] [Medline: [32346977](https://pubmed.ncbi.nlm.nih.gov/32346977/)]
19. Zhou J, Liu L, Xue P, Yang X, Tang X. Mental Health Response to the COVID-19 Outbreak in China. *Am J Psychiatry* 2020 Jul 01;177(7):574-575. [doi: [10.1176/appi.ajp.2020.20030304](https://doi.org/10.1176/appi.ajp.2020.20030304)] [Medline: [32375540](https://pubmed.ncbi.nlm.nih.gov/32375540/)]
20. Guo Y, Cheng C, Zeng Y, Li Y, Zhu M, Yang W, et al. Mental health disorders and associated risk factors in quarantined adults during the COVID-19 outbreak in China: cross-sectional study. *J Med Internet Res* 2020 Aug 06;22(8):e20328 [FREE Full text] [doi: [10.2196/20328](https://doi.org/10.2196/20328)] [Medline: [32716899](https://pubmed.ncbi.nlm.nih.gov/32716899/)]
21. Sønderskov KM, Dinesen PT, Santini ZI, Østergaard SD. The depressive state of Denmark during the COVID-19 pandemic. *Acta Neuropsychiatr* 2020 Aug 22;32(4):226-228 [FREE Full text] [doi: [10.1017/neu.2020.15](https://doi.org/10.1017/neu.2020.15)] [Medline: [32319879](https://pubmed.ncbi.nlm.nih.gov/32319879/)]
22. Chen Y, Zhou H, Zhou Y, Zhou F. Prevalence of self-reported depression and anxiety among pediatric medical staff members during the COVID-19 outbreak in Guiyang, China. *Psychiatry Res* 2020 Jun;288:113005 [FREE Full text] [doi: [10.1016/j.psychres.2020.113005](https://doi.org/10.1016/j.psychres.2020.113005)] [Medline: [32315886](https://pubmed.ncbi.nlm.nih.gov/32315886/)]
23. Mo Y, Deng L, Zhang L, Lang Q, Liao C, Wang N, et al. Work stress among Chinese nurses to support Wuhan in fighting against COVID-19 epidemic. *J Nurs Manag* 2020 Jul 20;28(5):1002-1009 [FREE Full text] [doi: [10.1111/jonm.13014](https://doi.org/10.1111/jonm.13014)] [Medline: [32255222](https://pubmed.ncbi.nlm.nih.gov/32255222/)]
24. Xiao H, Zhang Y, Kong D, Li S, Yang N. The effects of social support on sleep quality of medical staff treating patients with coronavirus disease 2019 (COVID-19) in January and February 2020 in China. *Med Sci Monit* 2020 Mar 05;26. [doi: [10.12659/msm.923549](https://doi.org/10.12659/msm.923549)]
25. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comput Biol* 2011 Oct 13;7(10):e1002199 [FREE Full text] [doi: [10.1371/journal.pcbi.1002199](https://doi.org/10.1371/journal.pcbi.1002199)] [Medline: [22022249](https://pubmed.ncbi.nlm.nih.gov/22022249/)]
26. Pestian JP, Matykiewicz P, Linn-Gust M, South B, Uzun O, Wiebe J, et al. Sentiment analysis of suicide notes: a shared task. *Biomed Inform Insights* 2012 Jan 30;5(Suppl 1):3-16 [FREE Full text] [doi: [10.4137/bii.s9042](https://doi.org/10.4137/bii.s9042)] [Medline: [22419877](https://pubmed.ncbi.nlm.nih.gov/22419877/)]
27. Le X, Lancashire I, Hirst G, Jokel R. Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Lit Linguistics Comput* 2011 May 24;26(4):435-461. [doi: [10.1093/litc/fqr013](https://doi.org/10.1093/litc/fqr013)]
28. Choudhury MD, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27-May 2, 2013; Paris. [doi: [10.1145/2470654.2466447](https://doi.org/10.1145/2470654.2466447)]
29. Jashinsky J, Burton SH, Hanson CL, West J, Giraud-Carrier C, Barnes MD, et al. Tracking suicide risk factors through Twitter in the US. *Crisis* 2014 Jan 01;35(1):51-59. [doi: [10.1027/0227-5910/a000234](https://doi.org/10.1027/0227-5910/a000234)] [Medline: [24121153](https://pubmed.ncbi.nlm.nih.gov/24121153/)]
30. Jiang Y, Li Z, Ye X. Understanding demographic and socioeconomic biases of geotagged Twitter users at the county level. *Cartogr Geogr Inf Sci* 2018 Feb 09;46(3):228-242. [doi: [10.1080/15230406.2018.1434834](https://doi.org/10.1080/15230406.2018.1434834)]
31. Anxiety and Depression. National Center for Health Statistics. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm> [accessed 2022-12-06]
32. Berkowitz SA, Basu S. Unemployment insurance, health-related social needs, health care access, and mental health during the COVID-19 pandemic. *JAMA Intern Med* 2021 May 01;181(5):699-702 [FREE Full text] [doi: [10.1001/jamainternmed.2020.7048](https://doi.org/10.1001/jamainternmed.2020.7048)] [Medline: [33252615](https://pubmed.ncbi.nlm.nih.gov/33252615/)]
33. Vahratian A, Blumberg SJ, Terlizzi EP, Schiller JS. Symptoms of anxiety or depressive disorder and use of mental health care among adults during the COVID-19 pandemic - United States, August 2020-February 2021. *MMWR Morb Mortal Wkly Rep* 2021 Apr 02;70(13):490-494 [FREE Full text] [doi: [10.15585/mmwr.mm7013e2](https://doi.org/10.15585/mmwr.mm7013e2)] [Medline: [33793459](https://pubmed.ncbi.nlm.nih.gov/33793459/)]
34. Li Z, Huang X, Ye X, Jiang Y, Martin Y, Ning H, et al. Measuring global multi-scale place connectivity using geotagged social media data. *Sci Rep* 2021 Jul 19;11(1):14694 [FREE Full text] [doi: [10.1038/s41598-021-94300-7](https://doi.org/10.1038/s41598-021-94300-7)] [Medline: [34282241](https://pubmed.ncbi.nlm.nih.gov/34282241/)]
35. Homan C, Johar R, Liu T, Lytle M, Silenzio V, Alm CO. Toward macro-insights for suicide prevention: analyzing fine-grained distress at scale. 2014 Presented at: Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; 2014; Baltimore, MD. [doi: [10.3115/v1/W14-3213](https://doi.org/10.3115/v1/W14-3213)]
36. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. *J Med Internet Res* 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: [10.2196/26627](https://doi.org/10.2196/26627)] [Medline: [33724919](https://pubmed.ncbi.nlm.nih.gov/33724919/)]
37. Zhang Y, Lyu H, Liu Y, Zhang X, Wang Y, Luo J. Monitoring depression trends on Twitter during the COVID-19 pandemic: observational study. *JMIR Infodemiology* 2021;1(1):e26769 [FREE Full text] [doi: [10.2196/26769](https://doi.org/10.2196/26769)] [Medline: [34458682](https://pubmed.ncbi.nlm.nih.gov/34458682/)]
38. Miller I, Becker A, Grenfell B, Metcalf CJE. Mapping the burden of COVID-19 in the United States. medRxiv. Preprint posted online April 11, 2020. [doi: [10.1101/2020.04.05.20054700](https://doi.org/10.1101/2020.04.05.20054700)]
39. Goldmann E, Galea S. Mental health consequences of disasters. *Annu Rev Public Health* 2014 Mar 18;35(1):169-183. [doi: [10.1146/annurev-publhealth-032013-182435](https://doi.org/10.1146/annurev-publhealth-032013-182435)] [Medline: [24159920](https://pubmed.ncbi.nlm.nih.gov/24159920/)]

40. Rahman MM, Ali GN, Li XJ, Samuel J, Paul KC, Chong PH, et al. Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data. *Heliyon* 2021 Feb;7(2):e06200 [FREE Full text] [doi: [10.1016/j.heliyon.2021.e06200](https://doi.org/10.1016/j.heliyon.2021.e06200)] [Medline: [33585707](https://pubmed.ncbi.nlm.nih.gov/33585707/)]
41. Whitney DG, Peterson MD. US National and state-level prevalence of mental health disorders and disparities of mental health care use in children. *JAMA Pediatr* 2019 Apr 01;173(4):389-391 [FREE Full text] [doi: [10.1001/jamapediatrics.2018.5399](https://doi.org/10.1001/jamapediatrics.2018.5399)] [Medline: [30742204](https://pubmed.ncbi.nlm.nih.gov/30742204/)]
42. Kelly R. Twitter study reveals interesting results about usage. 2009. URL: <http://www.pearanalytics.com/2009/twitter-study-reveals-interesting-results-40-percent-pointless-babble/> [accessed 2022-12-06]
43. Skapinakis P, Bellos S, Koupidis S, Grammatikopoulos I, Theodorakis PN, Mavreas V. Prevalence and sociodemographic associations of common mental disorders in a nationally representative sample of the general population of Greece. *BMC Psychiatry* 2013 Jun 4;13(1). [doi: [10.1186/1471-244x-13-163](https://doi.org/10.1186/1471-244x-13-163)]
44. Jin Y, Zhang Y, Zhang Q, Rao W, Zhang L, Cui L, et al. Prevalence and socio-demographic correlates of poor mental health among older adults in agricultural areas of China. *Front Psychiatry* 2020 Nov 5;11:549148 [FREE Full text] [doi: [10.3389/fpsy.2020.549148](https://doi.org/10.3389/fpsy.2020.549148)] [Medline: [33250790](https://pubmed.ncbi.nlm.nih.gov/33250790/)]
45. Patel RM, Singh US. Prevalence study of cognitive impairment and its associated sociodemographic variables using Mini-Mental Status Examination among elderly population residing in field practice areas of a medical college. *Indian J Community Med* 2018;43(2):113-116 [FREE Full text] [doi: [10.4103/ijcm.IJCM_102_17](https://doi.org/10.4103/ijcm.IJCM_102_17)] [Medline: [29899611](https://pubmed.ncbi.nlm.nih.gov/29899611/)]
46. Smith A, Brenner J. Twitter use 2012. Pew Research Center. 2012. URL: <https://www.pewresearch.org/internet/2012/05/31/twitter-use-2012/> [accessed 2022-12-06]
47. Goldman N, Gleit DA, Weinstein M. Declining mental health among disadvantaged Americans. *Proc Natl Acad Sci U S A* 2018 Jul 10;115(28):7290-7295 [FREE Full text] [doi: [10.1073/pnas.1722023115](https://doi.org/10.1073/pnas.1722023115)] [Medline: [29915079](https://pubmed.ncbi.nlm.nih.gov/29915079/)]
48. Twitter Developer Platform. URL: <https://developer.twitter.com/en/developer-terms/moreon-restricted-use-cases> [accessed 2022-12-08]

Abbreviations

CDC: Centers for Disease Control and Prevention

HPS: Household Pulse Survey

SES: socioeconomic status

Edited by A Mavragani; submitted 25.02.22; peer-reviewed by N Esfandiari, N Nekliudov; comments to author 05.10.22; revised version received 29.11.22; accepted 30.11.22; published 20.12.22.

Please cite as:

Cai R, Zhang J, Li Z, Zeng C, Qiao S, Li X

Using Twitter Data to Estimate the Prevalence of Symptoms of Mental Disorders in the United States During the COVID-19 Pandemic: Ecological Cohort Study

JMIR Form Res 2022;6(12):e37582

URL: <https://formative.jmir.org/2022/12/e37582>

doi: [10.2196/37582](https://doi.org/10.2196/37582)

PMID: [36459569](https://pubmed.ncbi.nlm.nih.gov/36459569/)

©Ruilie Cai, Jiajia Zhang, Zhenlong Li, Chengbo Zeng, Shan Qiao, Xiaoming Li. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

IT and the Quality and Efficiency of Mental Health Care in a Time of COVID-19: Case Study of Mental Health Providers in England

Frederick Hassan Konteh¹, MA, PhD; Russell Mannion¹, BA, PhD; Rowena Jacobs², PhD

¹Health Services Management Centre, School of Social Policy, University of Birmingham, Birmingham, United Kingdom

²Centre for Health Economics, University of York, York, United Kingdom

Corresponding Author:

Frederick Hassan Konteh, MA, PhD
Health Services Management Centre
School of Social Policy
University of Birmingham
40 Edgbaston Park Road
Park House
Birmingham, B15 2RT
United Kingdom
Phone: 44 7401415960
Fax: 44 121 414 7050
Email: f.konteh@sky.com

Abstract

Background: In England, COVID-19 has significantly affected mental health care and tested the resilience of health care providers. In many areas, the increased use of IT has enabled traditional modes of service delivery to be supported or even replaced by remote forms of provision.

Objective: This study aimed to assess the use and impact of IT, in remote service provision, on the quality and efficiency of mental health care during the pandemic. We drew on sociotechnical systems theory as a conceptual framework to help structure the gathering, analysis, and interpretation of data.

Methods: We conducted a national scoping survey that involved documentary analysis and semistructured interviews with 6 national stakeholders and case studies of 4 purposefully selected mental health providers in England involving interviews with 53 staff members.

Results: Following the outbreak of COVID-19, mental health providers rapidly adjusted their traditional forms of service delivery, switching to digital and telephone consultations for most services. The informants provided nuanced perspectives on the impact on the quality and efficiency of remote service delivery during the pandemic. Notably, it has allowed providers to attend to as many patients as possible in the face of COVID-19 restrictions, to the convenience of both patients and staff. Among its negative effects are concerns about the unsuitability of remote consultation for some people with mental health conditions and the potential to widen the digital divide and exacerbate existing inequalities. Sociotechnical systems theory was found to be a suitable framework for understanding the range of systemic and sociotechnical factors that influence the use of technology in mental health care delivery in times of crisis and normalcy.

Conclusions: Although the use of IT has boosted mental health care delivery during the pandemic, it has had mixed effects on quality and efficiency. In general, patients have benefited from the convenience of remote consultation when face-to-face contact was impossible. In contrast, patient choice was often compromised, and patient experience and outcomes might have been affected for some people with mental health conditions for which remote consultation is less suitable. However, the full impact of IT on the quality and efficiency of mental health care provision along with the systemic and sociotechnical determinants requires more sustained and longitudinal research.

(*JMIR Form Res* 2022;6(12):e37533) doi:[10.2196/37533](https://doi.org/10.2196/37533)

KEYWORDS

COVID-19; mental health care; information technology; digital; inequalities; sociotechnical systems

Introduction

COVID-19 has negatively impacted mental health worldwide, and the full ramifications of the pandemic on population mental health may last much longer than the initial phases of the pandemic [1,2]. Deterioration in existing mental health and new presentations are likely to contribute to the rising demand for mental health services, including depressive and anxiety disorders, which may stem from the psychological and economic effects of the pandemic linked to factors such as bereavement, restrictions in movement, lack of social interactions, domestic violence, and unemployment [2,3]. In England, as in most parts of the world, the pandemic, especially in the initial phase, severely disrupted health and social care provision, including mental health services [1,4,5]. By April 2021, the demand for mental health services in England was at a record high without a commensurate increase in funding for the sector [6]. With the initial focus being primarily on physical health (fighting the pandemic), COVID-19 has challenged the resilience of mental health providers, requiring them to adapt and transform their service delivery.

COVID-19 hit shortly after the National Health System (NHS) Long Term Plan for England was published, at the heart of which was a strong commitment by the UK government to use digital technology to transform health and social care [7,8]. In the United Kingdom, especially England, the use of technology in health care was relatively low before the pandemic—remote consultations were predominantly via telephone [9-12] with community mental health services experiencing the least investment in new technology [13]. At the national policy level, there was an acknowledgment that poor technological infrastructure in health and social care needs to be addressed [14].

The NHS Long Term Plan articulates the government's strategic intention to digitalize health and care in England, offering patients the choice to access web-based services as an alternative to face-to-face consultations [7]. System interoperability is a key aspiration of the Long Term Plan, with the potential for significant cost savings for the NHS, saving time and money for service users, and the convenience of telephone and video consultation. The NHS Mental Health Implementation Plan, an offshoot of the NHS Long Term Plan, outlines the provision of digital options to service users and the rolling out of "digitally-enabled models of therapy" for "specific mental health pathways" by 2021-2022 and for local systems to make use of "digital clinical decision-making tools" by 2023-2024 [13]. COVID-19 has helped to fast-track the NHS Long Term Plan's digital technology actualization. In a policy statement at the start of the pandemic (July 30, 2020), the Department of Health and Social Care in England urged providers to effect a wholesale switch to remote service provision through the use of telephone and digital technology [10]. Mental health providers appeared to have responded to this call. Thus, the pandemic has accelerated the rapid rollout of IT in mental health service delivery [1,4,11,15,16].

Research has highlighted the unprecedented use of remote (digital and telephone) consultations in mental health care during

the pandemic [1,4,5,9-12,15-20]. The greatest benefit was that it enabled providers to respond flexibly to the needs of service users while adjusting to pandemic-enforced challenges and restrictions. In the absence of face-to-face consultations, the use of IT facilitated a higher frequency of contact between professionals and patients than would have hitherto been possible [1,4,5,11,16]. Evidence that telephone has remained the most common means of remote consultation for mental health care providers in England, throughout the pandemic [10,11], suggests that the country has some catching up to do with respect to digital technology in mental health.

One of the benefits of remote consultation is its potential to overcome geographic barriers (the friction of distance), allowing access to populations in remote locations with serious accessibility problems, thereby reducing inequalities. However, a serious disadvantage is the potential for digital exclusion among vulnerable and disadvantaged population groups [1,10,11,18,21]. For example, technology and environment-related factors, such as problems with broadband or internet connectivity, continue to render many people living in remote locations relatively disadvantaged. Linked to digital exclusion is digital poverty, which encompasses issues of affordability (of technological devices and facilities including internet connection), accessibility, skills, and motivation [18].

The COVID-19 pandemic has witnessed a growing research interest in the use of IT in health care delivery. However, as Ellis and others [5] observed, empirical research has focused largely on physical health, whereas the mental health sector remains underresearched. There has been a call for research to examine the factors that influence the optimal use of technology in health care delivery, generally [9]. The purpose of this study was to understand how mental health providers in England used IT, including internet-enabled digital applications and traditional forms of remote service delivery, during the pandemic and the implications of this shift for quality and efficiency. It also seeks to enhance the understanding of the complexity of factors (social or human, system-related, and technical) affecting the use of IT in remote mental health care delivery and to apply the sociotechnical systems (STS) framework [22-25] to a case study design.

Methods

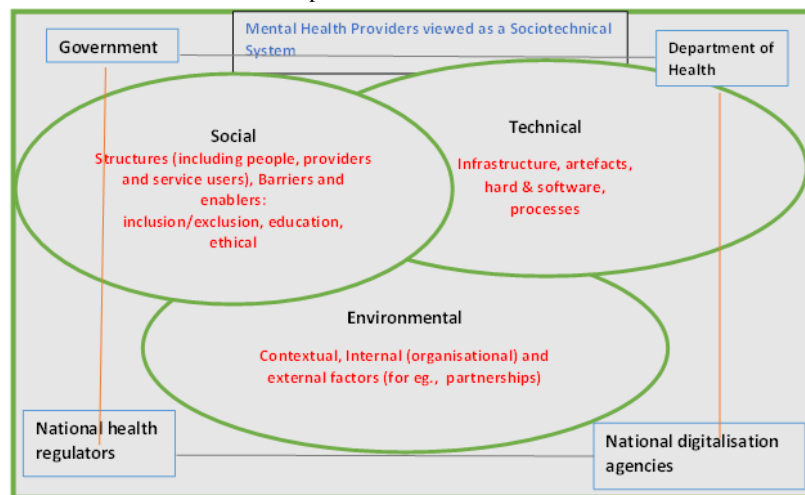
Theoretical Framework

An important notion related to the use of IT in health care is that, given their multiple objectives and complexity, health care organizations are best conceived as STS [22-25]. We applied the STS theory during data analysis, helping to interpret the research findings and to enhance the understanding of the major factors at work in the use of technology in mental health service delivery [22,25]. On the basis of the STS model, a health care system, like any complex and dynamic system, comprises a mix of interacting social and technical subsystems that affect each other [22-26]. Figure 1 illustrates the adaptation of the STS framework used in this study. Scholars have proposed a variety of STS models [27]. One model suggests that STS is composed of 2 sets of variables or subsystems, namely, social and technical structures [26]. The technical subsystem comprises the process

and technology components, whereas the social structures represent people (including program specialists, service providers, and service users) and organizations (including groups, teams, and departments). Another conceptualization of STS is that it comprises social, technical, and environmental components [28]. STS aligns with the key tenets of general systems theory and the notion of a nonlinear complex system [28]. STS theories emphasize the interdependence of the subsystems and the need for joint optimization, which requires

an alignment of social and technical constituents with a wider environment or system to achieve the desired transformational goal of an organization [22-28]. The proposed framework (Figure 1), an adaptation of the STS, includes social, technical, environmental, and contextual factors as well as the wider systemic influences, including government, the Department of Health, and regulatory or facilitating agencies (notably NHS Digital).

Figure 1. Sociotechnical systems and IT in mental health care provision.



Research Questions

This study is derived from a larger study that aimed to understand how mental health trusts in England adapted and adjusted services to cope with the COVID-19 pandemic [29]. Our interrelated research questions were as follows:

- How have mental health providers adapted service provision and used technology in remote service delivery during the pandemic?
- What were each provider's technological capabilities, such as allowing adjustments to service provision?
- How has the use of technology affected the quality and efficiency of mental health care delivery during the pandemic?

Study Design

We used a qualitative research approach in a multiple case study design, comprising 4 mental health providers, renamed here, to protect anonymity, as providers A, B, C, and D. The case study has long been considered suitable for the in-depth investigation of a contemporaneous subject [30]. The multiple case studies and qualitative research approach allowed for comparisons of similarities and differences to be explored between cases, with the scope to uncover new variables in a complex social context, where the researchers could not "manipulate" the subjects of inquiry [30-32]. At the outset, we conducted a scoping survey, comprising a review of relevant documentation and key informant interviews with key staff from national stakeholder organizations.

Data Collection

Data were collected between March and December 2021. In the scoping phase, we emailed 10 national stakeholder organizations, explaining the study and requesting contact with a suitable informant. These were purposively selected based on their role in the national health system or mental health care oversight. With 60% success, this resulted in interviews with people who held very senior or executive-level positions in the following organizations: NHS England (2), the Care Quality Commission (CQC; 1), the Mental Health Commissioners Network (MHCN; 1), the Health Care Financial Management Association (1), and the *Get it Right First-Time* program (1). Not only do their perspectives help in grounding our research within extant national policy developments and interests around the use of digital technology in health care delivery in England, but they also provide complementary evidence, especially for the key generic findings of the case study.

The 4 case study sites were purposefully selected to represent the opposite ends of the performance spectrum (based on our earlier work in which the research team used relevant secondary NHS data to rank all mental health providers in England into low- and high-performance categories). The selected providers offer a range of services and reflect different population catchments, staff sizes, and geographic spread (2 [A and B] mainly rural, 1 [C] metropolitan, and 1 [D] rural or urban based). The research team conducted 53 semistructured interviews with 72 potential informants (18 at each site) who were invited to participate (with a response rate of 74%). Participants, who were purposively sampled, comprised executive team members, senior managers and service directors or clinicians, patient representatives, and clinical commissioning groups responsible for commissioning services for each provider. For a breakdown

of study participants by organization refer to [Multimedia Appendix 1](#). Emails were sent out by the designated research support officer at each site, inviting eligible staff members to the study and introducing the researcher (interviewer). A web-based meeting was arranged for those who expressed a willingness to participate in the study. A consent form, signed by the researcher, was emailed within 6 to 24 hours before the interview, and the informant was encouraged to read, sign, and return via email before the interview. Most of the interviews were conducted using Microsoft Teams, and in a few cases (mainly with patient representatives), interviews were conducted using Zoom (Zoom Video Communications). All informants were asked for permission to record the interview (even though they would have indicated their approval on the consent form). Interviews lasted between 30 and 45 minutes; the questions for national stakeholders and case study informants cover similar themes but with slight modification for the latter; the themes or questions span beyond the focus of this paper—being part of a larger study (refer to [Multimedia Appendix 2](#) for the topic guides).

Data Analysis

As reported in our previous publications [33], this was part of a larger study in which we used administrative and patient survey data for each mental health provider in England to construct a composite performance indicator (ranking providers into low- and high-performing categories). Documents from the providers' websites and relevant national reports were also reviewed, providing useful background and policy-related material to complement the data from the case study interviews.

We used NVivo (QSR International) to code the transcripts and analyze the data. We coded and analyzed all transcripts from the interviews as one file, both within and across cases, enabling us to explore differences and synergies between and across providers and national stakeholders [31,32]. We followed the 5 stages of the framework method (familiarization, theme identification, indexing, charting, and interpretation) to analyze and structure the data [34]. On the basis of abductive theorizing and pattern matching, we explored the views and experiences of participants regarding our study objectives, allowing us to draw out and integrate the common themes from the interview data. The codebook was developed by FK and RM and reviewed by RJ, and coding was carried out by FK, with coding outputs shared with other authors for discussion. For the NVivo codebook, please refer to [Multimedia Appendix 3](#).

Ethical Considerations

We obtained research ethics approval from the NHS Health Research Authority (reference number: 21/PR/0047) and the participating organizations. There was an undertaking (stated in the research passport) that the research does not involve “regulated activity with children and/or adults as defined in the Safeguarding Vulnerable Groups Act 2006.” In addition, in the Integrated Research Application System form, it was stated that the study has nothing to do with “human tissue samples and data.”

Informed and written consent was obtained before every interview signed by the researcher and countersigned by study

participant. The participant's information sheet assured every participant about confidentiality and anonymity of the information they would provide and their identity. It stated further that participation was voluntary, and as the research was not considered to be harmful, there were no compensation arrangements. Participants were also advised about where to report if they had any concerns or complaints regarding the study.

Results

Providers Adapting Service Provision by Switching to Remote Service Delivery Following the Outbreak of COVID-19

There was a convergence of views from informants that following the outbreak of COVID-19, mental health providers switched to remote service provision via telephone and video consultations at an unprecedented speed and scale. Providers adopted web-based platforms and digital apps to ensure contact with patients and web-based interaction among staff, particularly during periods when COVID-19-imposed restrictions made physical interaction impossible. Safety considerations for both health professionals and patients were part of the motivation for the rapid switch to remote service provision. Study respondents were of the view that service users have benefited from services being “more responsive and immediate”:

It seemed to us that people adopted very quickly, much more quickly than usual. Normally if you're going to roll out a new IT system they're talking about it for four years...to get an electronic note system up and running and everybody using it. But this actually appeared to work really quickly... [CQC Official]

...We took action pretty quickly to, as I'm sure others did, a) protect our staff by ensuring that they were working from home if they could but b) open up digital online and telephone channels... [Provider A, Director of Strategy]

Digital consultation was not considered a viable option for some services until COVID-19 hit. For example, the Manager for Older People's Community Mental Health Team in provider C noted that they have had “to offer e-consultations as an option which we weren't really doing prior to COVID.” However, COVID-19 saw an upsurge in the use of digital apps such as Attend Anywhere, eConsult, and Kooth and web-based meeting platforms, particularly Zoom and MS Teams. It was suggested that providers adopt remote service provision as a first or default option when face-to-face contact was not possible. Informants made repeated references to “digital first” and “digital by default,” which resonated with the emphasis on the development of technological solutions set out in the NHS Long Term Plan. Despite what appeared to be a remarkable effort of mental health providers to escalate remote service provision, including digital options, during the pandemic, there was a perception among informants that telephone, which was generally viewed as inferior to video consultation, remained the most common medium of interaction between mental health providers and service users:

The adaption we've done, we do more video consultations now, Attend Anywhere...in the early stages we did group things on Zoom, on Teams...we've got Attend Anywhere now for one-to-one interventions [Provider B, Head of Mental Health & Learning Disability]

The message we have given really has been digital first, not digital only. So, if it can move online, move it online... [Provider A, Director of Nursing]

There is a bit of a myth about digital health, digital mental health. What they mean is they have been using the phone a lot. I mean 95% of this is by phone. [NHS England Official (2)]

Some services, notably Increasing Access to Psychological Therapy, were able to deploy technology during the pandemic more rapidly than others because before the pandemic, they had already been ahead in the use of telephone and digital technology.

Providers' Technological Capabilities in Coping With COVID-19—Systemic Factors

The fact that telephone consultations have remained the most common form of remote interaction with service users means that there are still opportunities for digitalizing health care. However, technological capabilities varied across providers both before and during the pandemic. How technologically prepared a provider was depended greatly on its digital maturity, which in turn was a function of system-wide factors. Two of the providers were a global digital exemplar (GDE) and a GDE “fast follower” D and C, respectively (both high-performing providers). These were national awards that came with extra funding, which the National Health System set up in 2017 to support the development and use of digital technology among a selected number of NHS providers, following a bidding process. A GDE is similar to a pacesetter in the use of digital technology, and a GDE fast follower becomes affiliated with a GDE to collaborate and share best practices in the use of technology to deliver health services. This meant that providers D and C were investing optimally in technology and were relatively digitally mature and technologically prepared when COVID-19 arrived. In contrast, the rollout of digital technology was slower for the low-performing providers, A and B, who were less prepared technologically. Some informants in these 2 organizations reported a few challenges with the adoption of digital technology during the pandemic:

We're really fortunate to be quite a digitally enabled organisation. So we already had numbers of virtual platforms...We were able to move probably more swiftly than others to a virtual platform. [Provider D, Director of Nursing]

I think we weren't terribly well prepared. I don't think there was much...suddenly there was need for all clinicians to have a laptop, to have an “Attend Anywhere” account and be trained up in Attend Anywhere. We weren't prepared or that and we were on the back foot. [Provider B, Medical Director]

It emerged that all mental health providers, including those who were technologically less prepared, were boosted in their quest to repurpose services through additional funding from the government during the pandemic. This system-wide support has enabled providers to invest in different areas and may help minimize existing disparities with regard to resource allocation and the development of innovation and technology:

We have managed to get more funding into capital spend...in a lot of ways we have actually got more money going into mental health and we are more attuned to what is going on in the country, and that will carry on with the official investment over the next year. [NHS England Official (2)]

A more flexible approach to contracting, in which the provider was allowed freedom to take quick decisions around spending, has also helped. Informants agreed that the collaboration and cooperation with partners, especially between providers and their Clinical Commissioning Group, was part of the reasons they managed to adapt service provision so quickly, including in digitalization:

I think the reason they got better was because we had a shared common purpose and we didn't have the luxury of time to debate and discuss and argue and write papers and develop plans, and all that stuff. It was a matter of, we need to act and we need to act now and we need to act together...It was a matter of, we need to do something and do something different and do something quickly. [Provider B, CCG_P1 Deputy Director]

Drawing further on system-wide collaboration, at least three of the case study providers benefited from private sector support. For example, provider B applied for and received a grant from Barclays Bank, which they used to supplement investments in digital technology. Providers A, B, and C worked with the voluntary sector in setting up internet cafes at strategic locations (to allow patients access to services remotely and free of cost), supporting others with digital gadgets, and with the basic skill of using digital technology.

Technology Impact on the Quality and Efficiency of Mental Health Care Delivery During COVID-19—Sociotechnical Determinants

National stakeholders and case study informants provided mixed views about the impact of COVID-19 and the influence of technology on quality and efficiency. They pointed to the speed with which a health professional would make contact with several patients within a short period without having to travel. For example, the Director of Finance in provider B suggested that patients did not have to wait for long before being attended to, and this, along with the convenience of accessing service remotely, has had a positive effect on patient experience. This was said to have allowed mental health professionals time for other activities, for example, administration and work. In addition, it enabled health providers to reduce their carbon footprint and contribute meaningfully toward the government's commitment to achieving a net zero NHS by 2045. Organization-wide meetings and internal team meetings quickly

switched to video conferences, and these were reported to be very well attended.

With respect to quality, the perspectives of case study informants centered on the following: quality of mental health service has not been (adversely) affected; quality might have been affected, if not compromised, in certain instances; and it would take time for services to determine the actual impact on quality using the conventional indicators or measures. Informants were more assured in their assessment of how the use of IT during the pandemic has positively reflected on mental health care efficiency:

I think the quality and the ability to be able to see people, and digitally see people, has been an improvement, because we've probably been able to contact people maybe more frequently than we would have done if we were doing face-to-face. [Provider C, Director of Nursing]

I think there's probably a mixed picture. So, for those people who engaged well with digital their service was probably enhanced because it was a lot more responsive and immediate, and for those that didn't or couldn't then their provision would have been reduced especially in the early days. We kept the wards open the whole time for high acute and emergencies, but for community cases if you weren't able to engage digitally then you probably did experience some deterioration or you certainly weren't getting the same level of support that you were used to. [Provider A, Associate Director of People]

What I would caveat on that is we don't know the true quality benefits of digital working yet because we've got no objective measures at the moment to really understand that. [Provider A, Director of Operations]

There was a notion that web-based interfacing, including conducting ward rounds via video technology, prevented people from congregating in confined spaces and therefore helped to minimize the spread of COVID-19. In addition, providers organized web-based visitations (meetings) between patients in inpatient wards and their families, which served as a good substitute for physical visits and helped to enhance patient experience during the pandemic. Some informants suggested that the inherent social and technical barriers, as presented below, would have compromised the quality or minimized the effectiveness of digital consultation. However, while drawing on funds provided by NHS England or grants from donor organizations or working with the voluntary sector, mental health providers did their best to address the problem of digital inequalities.

A range of social and technical factors can be seen to have a bearing on IT use during the pandemic, which is relevant to the STS framework. First, informants made repeated references to “digital poverty,” “digital divide,” “digital exclusion,” and “digital inequality,” suggesting that the scaling up of IT in service provision during the pandemic might have exacerbated inequalities. It has been suggested that the low socioeconomic status of some service users (or sheer poverty) affects their

ability to afford or own a digital device or connect to the internet or Wi-Fi. Informants, especially of providers A and B, also noted that access to digital consultation might be a huge challenge for service users who reside in very remote locations. Digital illiteracy emerged as another factor driving digital exclusion; there was a suggestion that this was mainly a function of age and that older service users were less “tech-savvy” than their younger counterparts:

...Access has been quite problematic during the pandemic and certainly during lockdown because largely this has been virtual access...and not everybody finds that easy and there is a group of people who are digitally disadvantaged who find that problematic. [MHCN Official]

Now, for the digitally excluded they [digital options] are useless...I am a consultant psychiatrist for a homeless team in [City X]. I have got two patients over the last year and I have been using the phone to speak to them. Most of the rest haven't got a phone. [NHS England Official (2)]

When people have said “I don't like it, I can't use it” and actually we still do have pockets of the county where the wifi is rubbish and actually there are still large groups of people that don't have any access to any digital technology, and we have to accept that. [Provider A, Director of Nursing]

Second, there was a view that remote services, which were the only option in certain periods of the pandemic, were less appropriate, if not ineffective, for some mental health conditions, such as autism and emotionally unstable personality disorders. This brings to the fore the role of technology experts (designers and programmers) and the need to ensure that devices are developed to offer optimum utility to every service user. Third, and closely related to the second point, is the ethical aspect of patient choice, which was not always guaranteed during the peak of COVID-19. The informants agreed that remote consultation was not a true substitute for in-person interaction. It was suggested, like for patients, that mental health staff, who could only attend web-based meetings during the peak of the pandemic, dearly missed the face-to-face (social) interaction with colleagues in the office:

There's been real pros to the new approach and real limitations...I think for example, some of the personality disorder services, people with EUPD; those would have always been face to face previously, they're now online but they're going back to face to face and some of that works for that cohort of patients, some of it doesn't. If you're paranoid and schizophrenic and worried about computers, you know, your experience is going to be very different isn't it? [Provider C, CEO]

Pretty much, autism assessments as well which was something that we struggled with to start with...It's difficult to do those assessments; it's not ideal to do it virtually because you're not picking up on everybody's body language. [Provider A, CEO]

Fourth, informants also highlighted the challenges with privacy, particularly for patients requiring a private “safe space” to discuss personal and confidential issues without the potential for other members of the household to overhear. For example, victims of domestic violence, children experiencing neglect, and people who are not confident in talking about their mental health issues.

Digital illiteracy was not only an issue for older service users. It was reported that some services, particularly in providers A and B, were adopting digital consultations during the pandemic for the first time, and many of their staff members needed “skilling-up” to be able to use the technology. As noted above, case study providers made concerted attempts to mitigate some of the sociotechnical challenges, including drawing on government and private sector funding support to upgrade their technology infrastructure and collaborating with voluntary organizations to address the problem of digital inequality. Finally, providers were quick in providing their staff with the requisite support—equipping them with the tools and skills needed to make the switch to digital technology and enabling staff to be accredited to conduct remote assessments. An equally important factor was the keen sense of commitment to care for patients on the part of mental health professionals during the pandemic. There was a convergence of views that mental health services have shown remarkable resilience and that staff were doing their utmost, often to the point of burnout or exhaustion, to provide optimum quality service during the pandemic.

Discussion

Principal Findings

We examined the perspectives of national stakeholders and staff from 4 mental health providers and found evidence that the pandemic served as a catalyst for the rapid uptake and deployment of IT in mental health care delivery during the pandemic. This finding is consistent with earlier studies on the ability of mental health providers and services to adapt and respond to the needs of service users during the pandemic, mainly through remote service provision and the use of technology [1,2,4,5,9-12,15-19]. We have also provided additional evidence on the mixed effects on the quality and efficiency of the radical switch to remote (mental health) service provision during the pandemic. Very few studies in this area had focused on the mental health sector. The study has further, perhaps for the first time, explored, using qualitative data and an STS lens, the mix of sociotechnical and system-wide factors affecting remote mental health provision in England.

Provider’s Adaptation of Service Provision and Technology Capabilities

It has been suggested that a health care organization, being a sociotechnical system, needs to be resilient [34] not least during a public health emergency. We found that the case study providers demonstrated a remarkable resilience in responding to the needs of service users during the pandemic by rapidly adapting service provision. The mental health sector in the United Kingdom and elsewhere can potentially build on the momentum in the rollout of IT created by the pandemic

[4,11,13]. When COVID-19–imposed restrictions proscribed face-to-face consultations, mental health providers were compelled to adapt their services to continue to respond to the needs of patients. Telephone and video consultations quickly emerged as the default strategy. There is evidence of improvement in the use of digital technology in mental health care delivery, following the outbreak of COVID-19, with some services adopting the use of digital platforms for the first time. Despite the progress made, the case studies have provided additional evidence that telephone communication has remained the main form of contact used by mental health providers in England, including the provision of 24/7 crisis support lines by mental health services during the pandemic [4,10,13]. Nevertheless, its effectiveness is disputed, with reports that the telephone lines have not been operational 24/7 [19]. This means that despite the reported progress, the digitalization of mental health care has remained less than optimal and that mental health services in England are still not yet at the level of digital capability that the providers would like them to be.

The case study findings revealed variability within and across provider organizations with respect to the pace at which IT was rolled out to deliver remote services. Some services that had earlier adopted remote service provision before the pandemic, notably talking therapies (Increasing Access to Psychological Therapy services), were quicker in adapting or expanding the use of digital technology. Although all providers appeared to roll out IT during the pandemic relatively quickly, we found evidence of disparities between providers in technological preparedness, with the high-performing providers, C and D, being better prepared (being a GDE and a GDE fast follower, respectively) than their low-performing counterparts, A and B. In contrast, the low-performing providers were neither a GDE nor a fast follower, and both required greater investments and more time and effort, including equipping clinicians with digital devices and the requisite training or skills to use the tools and adapt to the innovative ways of delivering services. Therefore, how soon and how much both staff and service users realized the benefits of IT during the outbreak depended largely on how advanced the provider’s technology infrastructure was (or their digital maturity).

Impact of IT Use on the Quality and Efficiency of Mental Health Care During COVID-19

The case study findings were mixed regarding the impact of telephone and digital consultations on the quality and efficiency of service delivery during the pandemic. Conclusive evidence regarding the true impact on the quality of care is lacking [7]. Given that providers found it difficult to process routine quality measures during the peak of the pandemic, informants’ perspectives were either based on anecdotal evidence or inferred from what they saw as the benefits and drawbacks of remote service provision. However, from the case studies, the rapid rollout of IT to deliver mental health care during the pandemic appeared to have been accompanied by a number of benefits and disadvantages for patients and staff, which is consistent with the findings of other studies [1,6,11-13,16-20]. The greatest benefit was that mental health services managed to respond remotely to the needs of patients when face-to-face consultation was not possible, serving as a convenient and cost-saving

strategy for interaction between patients and mental health professionals. Service efficiency, in particular, has been positively affected by the ability of mental health professionals to attend to as many patients as possible within a short period. In addition to financial savings (from reduced travel and use of office space and facilities) with some providers rethinking how to use estate facilities more efficiently, the mental health sector's contribution to the wider environment in terms of reduced carbon emissions from less driving during COVID-19 has been unprecedented. However, linked to the major downsides of video and telephone consultations were serious ethical and clinical concerns. The concerns centered on the notion that telephone or digital technology (more so telephone) cannot be a true substitute for face-to-face consultation [10], not only because it rules out the sociopsychological utility of physical interaction but also because some study participants suggested that remote service provision could be inappropriate for some patients with certain mental health conditions such as personality disorders and autism. Concerns have also been raised about the potential for worsening inequalities stemming from digital exclusion. This brings into focus the critical underlying sociotechnical factors at play, affecting the uptake and effectiveness of IT in the delivery of mental health services and the implications for quality and patient experience.

Determinants of IT Use in Mental Health Care Delivery—STS Framework

Mental health providers, as all health care organizations, are composed of complex macro and micro systems or subsystems, including departments, services, and units manned by humans and communities of service users—all with their dynamic and unique characteristics—and technical and other social elements [22-25,35]. Drawing on STS theory, we found that a range of factors are important for understanding the use (and optimization) of IT to deliver mental health services during and beyond the pandemic.

From the case study findings, it is evident that system-wide factors are key to IT uptake during the pandemic. The July 30, 2020, statement by the Secretary of Health and Social Care, in which he stressed the urgent need for technology in remote service provision [11], provided the national impetus and system-wide platform for a switch to digitalization of health services, including mental health. Nevertheless, the degree of prepandemic digital maturity influenced the speed with which a provider managed to switch to innovative digital options in delivering mental health services [7-13], with providers C and D being ahead of A and B in that regard. Digital maturity was, in turn, a function of historical investments in technology, which saw provider D becoming a GDE and provider C a GDE fast follower 3 years before the COVID-19 outbreak. A number of macrolevel and mesolevel factors were fundamental in enabling mental health providers, including those who were hitherto less digitally enabled, to meaningfully invest in their technology infrastructure and adopt more innovative ways of delivering services. These include new funding from the government, a more flexible approach to contracting and disbursements, donations from the private sector, and collaboration with voluntary and charitable organizations to empower service users to access digital service offers. There were concerns, however,

about how providers could ensure that nationally mandated “digital governance” policy and standards are in place and being followed to protect the rights of all individuals using digital technology to access service. Digital governance at the system and provider levels, particularly regarding privacy and data protection for patients, is a fundamental aspect of the digitalization of health care service delivery [11-13].

It has been suggested that collaboration between the relevant agencies and actors in the health care system—government, Department of Health, NHS England and Improvement, NHS Digital, producers of technologies for mental health, mental health providers, and patient groups—is critical to enhancing information and digital technology use in improving mental health care delivery [22-25,35]. This could help, for example, to enhance the effectiveness of digital options through a careful design of digital tools to meet the special needs of patients with very challenging mental health conditions, such as autism and personality disorders.

Alongside the system-wide factors, a range of social and technical factors have been at play, influencing the use of information and digital technology. These factors are important in any attempt to optimize the benefits and minimize the disadvantages associated with the use of IT in mental health care. The social status of service users, notably digital poverty or exclusion, as highlighted above, has emerged as an important underlying factor of digital inequality. Providers have been mindful of this and have been making efforts to address the problem and promote digital inclusiveness within available resources and support from the system (notably provision of digital devices to service users who could not afford them and free internet cafes at strategic locations). However, this raises issues around how providers could expand and sustain such effort as a “digital library” scheme as well as have safeguarding and data protection under control [9,12,36-38].

We found that patients with mental health issues, like other members of the population, responded to the blanket national policy pronouncements around COVID-19 safety restrictions including the “stay at home order,” with many presenting when their conditions were already deteriorating. Perhaps a realization that remote service provision, which became the default option, is never a true substitute for face-to-face consultations and that some mental health cases are less suited to remote consultation might have called for a more nuanced messaging and approach to mental health patients during the peak of the pandemic. An important finding is that remote service provision or interaction means that service users and staff missed the positive element of “socializing” via face-to-face meetings, with negative implications for their well-being. It is clear that staff have become exhausted and bored with the monotony of web-based meetings. In scaling up telephone and digital consultations there is a need to pay close attention to the needs and preferences of staff and patients and services before deciding which option is best for them.

As Greenhalgh et al [9] have reported, there is an implicit dilemma or contradiction in the notion that telephone and digital solutions were available to all service users as an option. This is because not all patients were in a position to take up the offer,

for example, because of technology and Wi-Fi accessibility issues or being uncomfortable with technology as a medium for receiving care. Thus, patient choice remains paramount to clinical decision-making, and it is vitally important that providers quickly adjust their mode of service provision by switching to face-to-face consultation whenever possible.

Strengths and Limitations

The study participants included key national stakeholders; senior managers; and executive team members of case study providers, including clinicians, allowing a rich mix of perspectives to reflect both the national landscape of mental health care provision and localized service provider contexts. Another strength of this study is its exploration of the key underlying factors influencing the uptake of IT through the application of STS theory. However, this study has some limitations. First, typical of a case study of this nature, it is difficult to generalize the findings to the rest of the mental health sector in England or the United Kingdom, let alone to other health care systems. However, the involvement of national stakeholders has helped to provide some useful generic insights into the role of technology in mental health delivery during the pandemic. Second, the study was purely qualitative, drawing on the subjective views and perspectives of the study participants, without complementary quantitative data to explore causal linkages, such as between the use of technology and clinical outcomes. Third, the study was limited to senior officials at the case study sites, representatives of commissioning groups, and a very small number of patient representatives, without room for lower-level health professionals or frontline staff.

Conclusions

The study highlights that digitally enabled remote mental health service provision has accelerated during the COVID-19 pandemic in England, with many services still heavily reliant on telephone for remote service delivery. Mixed messages emerged regarding the effects of IT on service quality and efficiency during the pandemic. Among the positive aspects is the perception that patients have generally benefited from the convenience of remote consultation when face-to-face meeting was impossible. Mental health professionals have been able to attend to many patients more rapidly than would have been possible with face-to-face consultation, saving travel time costs for both providers and service users. The major downsides of remote consultation include reduced levels of quality for people with specific types of mental health conditions requiring face-to-face contact, patient choice being compromised, and the likelihood that inequalities have been exacerbated because of a growing “digital divide.” A full assessment of the impact of IT use on the quality and efficiency of mental health care provision after the pandemic will require further and a more sustained longitudinal research. Given that telephone remains the most common means of remote consultation, there is a need for more targeted effort to support service users and health professionals (in need of such support) in catching up with the pace of digital transformation. In any future public health emergency, it is recommended that decision makers and service providers carefully consider the most appropriate service delivery option for specific services, ensuring that service provision options remain responsive to patient needs and support patient choice and retaining face-to-face provision when preferred.

Acknowledgments

The Health Foundation’s Efficiency Research Program provided financial support for this study. The Health Foundation is an independent charity committed to bringing about better health and health care for people in the United Kingdom. The funding agreement ensured the authors’ independence in designing the study, interpreting the data, and writing and publishing the report. The authors would like to thank the participating mental health providers, the research governance staff who supported the recruitments, and all those participating as key informants.

Data Availability

To ensure anonymity and confidentiality the team had undertaken (in the Integrated Research Application System form) that data, in the form of transcribed manuscripts, will be stored securely on the university system and accessible only to members of the research team via a secure network for a period of 12 months to 3 years, following the end of data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Breakdown of study participants by organization (N=53).

[\[DOCX File, 17 KB - formative_v6i12e37533_app1.docx\]](#)

Multimedia Appendix 2

Quality and efficiency of mental health care provision in a time of COVID-19: topic guides.

[\[DOCX File, 43 KB - formative_v6i12e37533_app2.docx\]](#)

Multimedia Appendix 3

EMHeP COVID-19 case study codebook.

[\[DOCX File , 29 KB - formative_v6i12e37533_app3.docx \]](#)**References**

1. Johnson S, Dalton-Locke C, Vera San Juan N, Foye U, Oram S, Papamichail A, COVID-19 Mental Health Policy Research Unit Group. Impact on mental health care and on mental health service users of the COVID-19 pandemic: a mixed methods survey of UK mental health care staff. *Soc Psychiatry Psychiatr Epidemiol* 2021 Jan;56(1):25-37 [FREE Full text] [doi: [10.1007/s00127-020-01927-4](https://doi.org/10.1007/s00127-020-01927-4)] [Medline: [32857218](https://pubmed.ncbi.nlm.nih.gov/32857218/)]
2. Mental health services and COVID-19: preparing for the rising tide. NHS Confederation. 2020 Aug 12. URL: <https://www.nhsconfed.org/publications/mental-health-services-and-covid-19> [accessed 2021-12-16]
3. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* 2021 Nov 06;398(10312):1700-1712 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7)] [Medline: [34634250](https://pubmed.ncbi.nlm.nih.gov/34634250/)]
4. Gkeredakis M, Lifshitz-Assaf H, Barrett M. Crisis as opportunity, disruption and exposure: exploring emergent responses to crisis through digital technology. *Inf Organ* 2021 Mar;31(1):100344. [doi: [10.1016/j.infoandorg.2021.100344](https://doi.org/10.1016/j.infoandorg.2021.100344)]
5. Ellis LA, Meulenbroeks I, Churrua K, Pomare C, Hatem S, Harrison R, et al. The application of e-mental health in response to COVID-19: scoping review and bibliometric analysis. *JMIR Ment Health* 2021 Dec 06;8(12):e32948 [FREE Full text] [doi: [10.2196/32948](https://doi.org/10.2196/32948)] [Medline: [34666306](https://pubmed.ncbi.nlm.nih.gov/34666306/)]
6. Bunn S, Dias MC. Mental health impacts of the COVID-19 pandemic on adults. United Kingdom Parliament. 2021 Jul 21. URL: <https://post.parliament.uk/research-briefings/post-pn-0648/> [accessed 2021-12-16]
7. The NHS long term plan. National Health Service. 2019 Jan. URL: <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf> [accessed 2021-11-10]
8. Digital Transformation in Community Health Services. NHS Confederation & NHS Providers. 2021 Nov. URL: <https://tinyurl.com/2chmkht4> [accessed 2021-11-10]
9. Fisk M, Livingstone A, Pit SW. Telehealth in the context of COVID-19: changing perspectives in Australia, the United Kingdom, and the United States. *J Med Internet Res* 2020 Jun 09;22(6):e19264 [FREE Full text] [doi: [10.2196/19264](https://doi.org/10.2196/19264)] [Medline: [32463377](https://pubmed.ncbi.nlm.nih.gov/32463377/)]
10. Greenhalgh T, Rosen R, Shaw SE, Byng R, Faulkner S, Finlay T, et al. Planning and evaluating remote consultation services: a new conceptual framework incorporating complexity and practical ethics. *Front Digit Health* 2021 Aug 13;3:726095 [FREE Full text] [doi: [10.3389/fdgth.2021.726095](https://doi.org/10.3389/fdgth.2021.726095)] [Medline: [34713199](https://pubmed.ncbi.nlm.nih.gov/34713199/)]
11. Liberati E, Richards N, Parker J, Willars J, Scott D, Boydell N, et al. Remote care for mental health: qualitative study with service users, carers and staff during the COVID-19 pandemic. *BMJ Open* 2021 Apr 22;11(4):e049210 [FREE Full text] [doi: [10.1136/bmjopen-2021-049210](https://doi.org/10.1136/bmjopen-2021-049210)] [Medline: [33888531](https://pubmed.ncbi.nlm.nih.gov/33888531/)]
12. Wherton J, Greenhalgh T, Shaw SE. Expanding video consultation services at pace and scale in Scotland during the COVID-19 pandemic: national mixed methods case study. *J Med Internet Res* 2021 Oct 07;23(10):e31374 [FREE Full text] [doi: [10.2196/31374](https://doi.org/10.2196/31374)] [Medline: [34516389](https://pubmed.ncbi.nlm.nih.gov/34516389/)]
13. The NHS Mental Health Implementation Plan 2019/20 – 2023/24. National Health Service. 2019 Jul. URL: <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/07/nhs-mental-health-implementation-plan-2019-20-2023-24.pdf> [accessed 2021-11-10]
14. The future of healthcare: our vision for digital, data and technology in health and care. Department of Health and Social Care. 2018 Oct 17. URL: <https://www.gov.uk/government/publications/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care> [accessed 2021-11-10]
15. Budhwani S, Fujioka JK, Chu C, Baranek H, Pus L, Wasserman L, et al. Delivering mental health care virtually during the COVID-19 pandemic: qualitative evaluation of provider experiences in a scaled context. *JMIR Form Res* 2021 Sep 21;5(9):e30280 [FREE Full text] [doi: [10.2196/30280](https://doi.org/10.2196/30280)] [Medline: [34406967](https://pubmed.ncbi.nlm.nih.gov/34406967/)]
16. Zeilig H, Gee B, Teague B, Wilson J, Hackmann C. Foregrounding the perspectives of mental health services users during the COVID-19 pandemic. *Ment Health Soc Inclusion* 2020 Jul 03;24(3):131-133. [doi: [10.1108/mhsl-05-2020-0028](https://doi.org/10.1108/mhsl-05-2020-0028)]
17. Mazziotti R, Rutigliano G. Tele-mental health for reaching out to patients in a time of pandemic: provider survey and meta-analysis of patient satisfaction. *JMIR Ment Health* 2021 Jul 29;8(7):e26187 [FREE Full text] [doi: [10.2196/26187](https://doi.org/10.2196/26187)] [Medline: [34114956](https://pubmed.ncbi.nlm.nih.gov/34114956/)]
18. Spanakis P, Peckham E, Mathers A, Shiers D, Gilbody S. The digital divide: amplifying health inequalities for people with severe mental illness in the time of COVID-19. *Br J Psychiatry* 2021 Oct;219(4):529-531 [FREE Full text] [doi: [10.1192/bjp.2021.56](https://doi.org/10.1192/bjp.2021.56)] [Medline: [35048887](https://pubmed.ncbi.nlm.nih.gov/35048887/)]
19. Gullstlett MK, Kristiansen E, Nilsen ER. Therapists' experience of video consultation in specialized mental health services during the COVID-19 pandemic: qualitative interview study. *JMIR Hum Factors* 2021 Jul 15;8(3):e23150 [FREE Full text] [doi: [10.2196/23150](https://doi.org/10.2196/23150)] [Medline: [34096505](https://pubmed.ncbi.nlm.nih.gov/34096505/)]

20. Venville A, O'Connor S, Roeschlein H, Ennals P, McLoughlan G, Thomas N. Mental health service user and worker experiences of psychosocial support via telehealth through the COVID-19 pandemic: qualitative study. *JMIR Ment Health* 2021 Aug 12;8(8):e29671 [FREE Full text] [doi: [10.2196/29671](https://doi.org/10.2196/29671)] [Medline: [34182461](https://pubmed.ncbi.nlm.nih.gov/34182461/)]
21. The Topol Review - Preparing the healthcare workforce to deliver the digital future: An independent report on behalf of the Secretary of State for Health and Social Care. National Health Service. 2019 Feb. URL: <https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf> [accessed 2021-12-12]
22. Marjanovic S, Altenhofer M, Hocking L, Chataway J, Ling T. Innovating for improved healthcare: sociotechnical and innovation systems perspectives and lessons from the NHS. *Sci Public Policy* 2020 Apr;47(2):283-297. [doi: [10.1093/scipol/scaa005](https://doi.org/10.1093/scipol/scaa005)]
23. Lawler EK, Hedge A, Pavlovic-Veselinovic S. Cognitive ergonomics, socio-technical systems, and the impact of healthcare information technologies. *Int J Ind Ergon* 2011 Jul;41(4):336-344. [doi: [10.1016/j.ergon.2011.02.006](https://doi.org/10.1016/j.ergon.2011.02.006)]
24. Ackerman M, Goggins S, Hermann T, Prilla M, Stary C. *Designing Healthcare That Works: A Sociotechnical Approach*. Cambridge, MA, USA: Academic Press; 2017.
25. Waterson P. Health information technology and sociotechnical systems: a progress report on recent developments within the UK National Health Service (NHS). *Appl Ergon* 2014 Mar;45(2):150-161. [doi: [10.1016/j.apergo.2013.07.004](https://doi.org/10.1016/j.apergo.2013.07.004)] [Medline: [23895916](https://pubmed.ncbi.nlm.nih.gov/23895916/)]
26. Carayon P. Human factors of complex sociotechnical systems. *Appl Ergon* 2006 Jul;37(4):525-535. [doi: [10.1016/j.apergo.2006.04.011](https://doi.org/10.1016/j.apergo.2006.04.011)] [Medline: [16756937](https://pubmed.ncbi.nlm.nih.gov/16756937/)]
27. Pasmore W, Winby S, Mohrman SA, Vanasse R. Reflections: sociotechnical systems design and organization change. *J Change Manag* 2019;19(2):67-85. [doi: [10.1080/14697017.2018.1553761](https://doi.org/10.1080/14697017.2018.1553761)]
28. Walker GH, Stanton NA, Salmon PM, Jenkins DP. A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theor Issues Ergon Sci* 2008 Nov;9(6):479-499. [doi: [10.1080/14639220701635470](https://doi.org/10.1080/14639220701635470)]
29. Mannion R, Konteh FH, Jacobs R. Impact of COVID-19 in mental health trusts. *J Health Serv Res Policy* (forthcoming) 2022 Aug 08:13558196221116298 [FREE Full text] [doi: [10.1177/13558196221116298](https://doi.org/10.1177/13558196221116298)] [Medline: [35938487](https://pubmed.ncbi.nlm.nih.gov/35938487/)]
30. Yin RK. *Case Study Research: Design and Methods - Vol. 5*. 4th edition. Thousand Oaks, CA, USA: Sage Publications; 2009.
31. Biancone P, Secinaro S, Marseglia R, Calandra D. E-health for the future. Managerial perspectives using a multiple case study approach. *Technovation* (forthcoming) 2021 Oct:102406. [doi: [10.1016/j.technovation.2021.102406](https://doi.org/10.1016/j.technovation.2021.102406)]
32. Baxter P, Jack S. Qualitative case study methodology: study design and implementation for novice researchers. *Qual Rep* 2008;13(4):544-559. [doi: [10.46743/2160-3715/2008.1573](https://doi.org/10.46743/2160-3715/2008.1573)]
33. Jacobs R, Konteh FH, Mannion R. Changing leadership, management and culture in mental health trusts. *Ment Health Rev J* (forthcoming) 2022 Sep 01. [doi: [10.1108/mhrj-03-2022-0018](https://doi.org/10.1108/mhrj-03-2022-0018)]
34. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013 Sep 18;13:117 [FREE Full text] [doi: [10.1186/1471-2288-13-117](https://doi.org/10.1186/1471-2288-13-117)] [Medline: [24047204](https://pubmed.ncbi.nlm.nih.gov/24047204/)]
35. Bostrom RP, Gupta S, Thomas D. A Meta-theory for understanding information systems within sociotechnical systems. *J Manag Inf Syst* 2009;26(1):17-47. [doi: [10.2753/mis0742-1222260102](https://doi.org/10.2753/mis0742-1222260102)]
36. Taysom E, Crilly N. Resilience in sociotechnical systems: the perspectives of multiple stakeholders. *J Des Econ Innov* 2017;3(3):165-182. [doi: [10.1016/j.sheji.2017.10.011](https://doi.org/10.1016/j.sheji.2017.10.011)]
37. The Health and Social Care Committee's Expert Panel: Evaluation of the Government's progress against its policy commitments in the area of mental health services in England. House of Commons. 2021 Nov 30. URL: <https://committees.parliament.uk/publications/8156/documents/83466/default/> [accessed 2021-12-12]
38. Awad A, Trenfield SJ, Pollard TD, Ong JJ, Elbadawi M, McCoubrey LE, et al. Connected healthcare: improving patient care using digital health technologies. *Adv Drug Deliv Rev* 2021 Nov;178:113958. [doi: [10.1016/j.addr.2021.113958](https://doi.org/10.1016/j.addr.2021.113958)] [Medline: [34478781](https://pubmed.ncbi.nlm.nih.gov/34478781/)]

Abbreviations

CQC: Care Quality Commission
GDE: global digital exemplar
MHCN: Mental Health Commissioners Network
NHS: National Health System
STS: sociotechnical systems

Edited by A Mavragani, J Torous; submitted 24.02.22; peer-reviewed by A Selick, O Nguyen; comments to author 07.08.22; revised version received 26.10.22; accepted 24.11.22; published 29.12.22.

Please cite as:

Konteh FH, Mannion R, Jacobs R

IT and the Quality and Efficiency of Mental Health Care in a Time of COVID-19: Case Study of Mental Health Providers in England
JMIR Form Res 2022;6(12):e37533

URL: <https://formative.jmir.org/2022/12/e37533>

doi: [10.2196/37533](https://doi.org/10.2196/37533)

PMID: [36423321](https://pubmed.ncbi.nlm.nih.gov/36423321/)

©Frederick Hassan Konteh, Russell Mannion, Rowena Jacobs. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 29.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Pilot Evaluation of Possible Airborne Transmission in a Geriatric Care Facility Using Carbon Dioxide Tracer Gas: Case Study

Yo Ishigaki¹, PhD; Shinji Yokogawa^{1,2}, PhD; Yuki Minamoto³, PhD; Akira Saito⁴, MT; Hiroko Kitamura⁵, MD, PhD; Yuto Kawauchi¹, BE

¹Graduate School of Informatics and Engineering, University of Electro-communications, Tokyo, Japan

²Info-powered Energy System Research Center, University of Electro-communications, Tokyo, Japan

³School of Engineering, Tokyo Institute of Technology, Tokyo, Japan

⁴Department of Clinical Laboratory Medicine, Miyagi Anti-Tuberculosis Association, Miyagi, Japan

⁵Occupational Health Training Center, University of Occupational and Environmental Health, Fukuoka, Japan

Corresponding Author:

Yo Ishigaki, PhD

Graduate School of Informatics and Engineering

University of Electro-communications

1-5-1, Chofu-gaoka, Chofu

Tokyo, 182-8585

Japan

Phone: 81 42 443 5662

Email: ishigaki@uec.ac.jp

Abstract

Background: Although several COVID-19 outbreaks have occurred in older adult care facilities throughout Japan, no field studies focusing on airborne infections within these settings have been reported. Countermeasures against airborne infection not only consider the air change rate (ACR) in a room but also the airflow in and between rooms. However, a specific method has not yet been established by Japanese public health centers or infectious disease-related organizations.

Objective: In April 2021, 59 COVID-19 cases were reported in an older adult care facility in Miyagi, Japan, and airborne transmission was suspected. The objective of this study was to simultaneously reproduce the ACR and aerosol advection in this facility using the carbon dioxide (CO₂) tracer gas method to elucidate the specific location and cause of the outbreak. These findings will guide our recommendations to the facility to prevent recurrence.

Methods: In August 2021, CO₂ sensors were placed in 5 rooms where airborne infection was suspected, and the CO₂ concentration was intentionally increased using dry ice, which was subsequently removed. The ACR was then estimated by applying the Seidel equation to the time-series changes in the CO₂ concentration due to ventilation. By installing multiple sensors outside the room, advection outside the room was monitored simultaneously. Aerosol advection was verified using computer simulations. Although the windows were closed at the time of the outbreak, we conducted experiments under open-window conditions to quantify the effects of window opening.

Results: The ACR values at the time of the outbreak were estimated to be 2.0 to 6.8 h⁻¹ in the rooms of the facility. A low-cost intervention of opening windows improved the ventilation frequency by a factor of 2.2 to 5.7. Ventilation depended significantly on the window-opening conditions (*P* values ranging from .001 to .03 for all rooms). Aerosol advection was detected from the private room to the day room in agreement with the simulation results. Considering that the individual who initiated the infection was in the private room on the day of infection, and several residents, who later became secondarily infected, were gathered in the day room, it was postulated that the infectious aerosol was transmitted by this air current.

Conclusions: The present results suggest that secondary infections can occur owing to aerosol advection driven by large-scale flow, even when the building design adheres to the ventilation guidelines established in Japan. Moreover, the CO₂ tracer gas method facilitates the visualization of areas at a high risk of airborne infection and demonstrates the effectiveness of window opening, which contributes to improved facility operations and recurrence prevention.

(*JMIR Form Res* 2022;6(12):e37587) doi:[10.2196/37587](https://doi.org/10.2196/37587)

KEYWORDS

nursing home; care home; airborne transmission; ventilation frequency; air change rate; ACR; computational fluid dynamics; CFD; mobile phone

Introduction

Background

COVID-19 should be prevented in health care and long-term care facilities because of the high risk of mortality [1]. In Japan, a series of COVID-19 outbreaks has been reported in nursing care facilities [2]. However, no field epidemiology case studies have assessed airborne infection by measuring the air change rate (ACR) and airflow in Japanese older adult care facilities.

Ventilation plays an important role in controlling the airborne transmission of COVID-19, with mass transmission of its etiologic pathogen, SARS-CoV-2, reported in poorly ventilated rooms. This was demonstrated in Ishigaki et al [3], which noted that ventilation was impeded by excessive shielding of an office space with plastic sheeting where a 7-person outbreak occurred. Li et al [4] reported a 9-person outbreak that occurred locally in a poorly ventilated restaurant, and Jang et al [5] reported that 112 people were infected in a small fitness studio with very poor ventilation in South Korea. Furthermore, Menzies et al [6] reported that in various hospital settings, an average of less than 2 ventilation cycles per hour (ACR per hour) in examination rooms represented a determinant of secondary tuberculosis transmission, that is, tuberculin conversion. On the basis of this study, the Centers for Disease Control and Prevention (CDC) in the United States established a standard for negative pressure room ventilation to isolate patients with infectious diseases, which included an ACR recommendation of 6 per hour (for existing buildings) to 12 per hour (for new buildings) with a safety factor [7]. Subsequently, based on the CDC standard of 12 (h^{-1}) ACR, the World Health Organization (WHO) established a standard of $576 \text{ m}^3/\text{hour}$ per person, with a doubled safety factor, for natural ventilation in health facilities treating patients with infectious diseases, assuming that each patient occupies a space of $4 \times 2 \times 3 \text{ m}^3$ [8]. In Japan, the Ministry of Health, Labor and Welfare (MHLW) has suggested a similar value of $576 \text{ m}^3/\text{hour}$ per person for health facilities. Furthermore, the MHLW has recommended a ventilation rate of $30 \text{ m}^3/\text{hour}$ per person in general commercial facilities to prevent indoor aerosol transmission of COVID-19 [9].

Fine or dry droplets can remain airborne for several minutes to several hours [10-13]. Therefore, secondary infections can occur when these infectious aerosols are transported by air currents, regardless of the distance from the infected person. Airborne transmission of COVID-19 was suspected in a shopping mall in China, as shoppers not in direct contact with each other became infected at the same time [14]. In Australia, a secondary infection occurred in a church with minimal ventilation, from a choir to 12 attendees, with a reported airborne distance of up to 15 m [15]. Moreover, 14 people were infected in a wide range of seats evenly distributed from rows 7 to 29 on both sides of a plane during a domestic short aisle flight in Japan [16]. Hence, when designing countermeasures against airborne infection, it

is necessary to consider not only the number of times a room is ventilated but also the airflow within and between rooms.

Pathogens such as *Mycobacterium tuberculosis* (etiologic pathogen of tuberculosis), measles, and varicella-zoster virus (chicken pox) are airborne [17]. The long-range transmission of infectious aerosols containing these pathogens can be controlled by buoyancy because of temperature differences and airflow control by natural winds and fans [18]. However, several reports have indicated that the effective reach of airborne COVID-19 transmission is at least 2-10 m. For instance, Anchordoqui et al [19] used computer simulations to analyze the aerodynamic properties of particulate matter containing SARS-CoV-2 and noted that it could propagate farther than the recommended social distance of 1.8 m. Morawska et al [20] reported multiple cases of propagation beyond a distance of 1-2 m. Hunziker [21] focused on the behavior of aerosols in an air-conditioned hospital room and reported aerosols with micron-order particle size propagating up to a distance of 5-6 m as jet passengers. However, according to Guven et al [22], micron-order particles propagate over long distances of 2-8 m. Moreover, Anderson et al [23] found that airborne viruses can remain active for up to 27 hours depending on the conditions of temperature and humidity, whereas infectious aerosols can travel up to 7.0-8.2 m. More specifically, infectious aerosols, or gas clouds, formed by sneezing can reach a distance of 7-8 m and remain in the air for hours, depending on the ventilation system [24]. In addition, Lima et al [25] surveyed 10 studies and concluded that aerosol particles containing viruses can reach distances of up to 10 m and survive for several hours in air. Furthermore, Azimi et al [26] investigated a large outbreak on a cruise ship in which 712 of 3711 crew members and passengers were infected, concluding that reinhalation of infectious aerosols was the primary route of transmission.

Goal of This Study

In this study, an outbreak site of SARS-CoV-2 was investigated in a nursing home in Miyagi Prefecture, Japan, to measure the ACR and visualize the behavior of infectious aerosol distribution using a carbon dioxide (CO_2) tracer gas method—that can estimate the actual ACR of a room and is used to prevent airborne transmission of COVID-19 [27]—at various window-opening conditions. Aerosol advection originating from the room was quantified by simultaneously monitoring the tracer gas leaking from the room. In addition, the aerosol distribution was analyzed via computational fluid dynamics (CFD) based on a 3D model of the facility to ensure that the 3D behavior corresponded with the measured CO_2 tracer gas results. CFD simulations are effective for detailing the airborne behavior of infectious aerosols [28]. However, its implementation requires specialized knowledge, detailed surveying, and computational resources. In contrast, the CO_2 tracer gas method can be implemented with dry ice and a sensor, and if its effectiveness is confirmed, it can be used in numerous facilities to prevent recurrent outbreaks.

The primary purpose of this study was not to identify the direct cause of the outbreak but to identify the factors responsible for, and the origin locations of, outbreaks in older adult care facilities in terms of ACR and aerosol dispersion. Furthermore, this study seeks to establish a CO₂ tracer gas method that can simultaneously monitor ventilation and aerosol behavior in hospitals and older adult care facilities to contribute to the prevention of recurrent airborne infections.

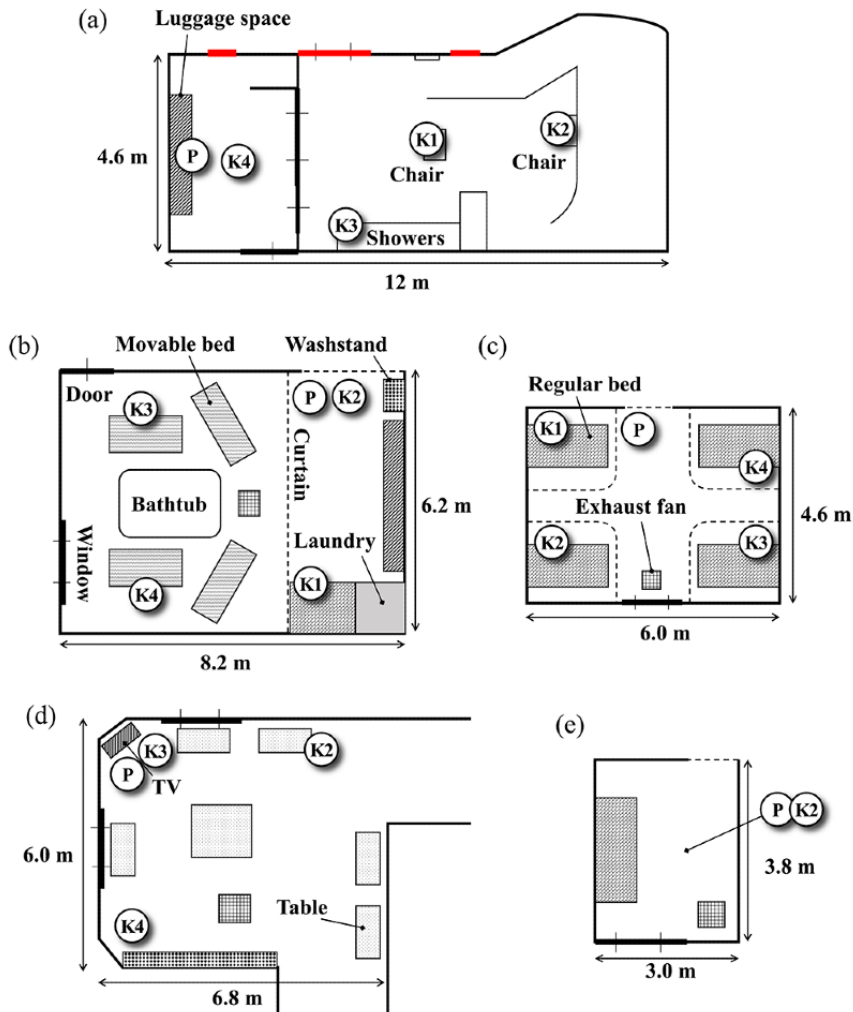
Methods

Facility Overview

The older adult care facility investigated in this study was located in Miyagi Prefecture, Japan, where 59 cases were reported in the same building from April to May 2021. Of these, 36 were users of the facility (29 residents and 7 daily visitors) and the other 23 were facility staff members. As 2 positive cases of infection were obtained on the first day using polymerase chain reaction (PCR) tests, and the subsequent PCR tests performed 2 days later further confirmed the positivity of 14 people on the same floor, airborne transmission was strongly suspected to be responsible for the propagation of the infection. Other causes include contact and droplet infections, making it almost impossible to identify a direct cause. However, as masks and face shields were worn and hand sanitizers were used by staff members who spread the secondary infection, and as the infection spread in a short period among those who were not in close contact with each other, it is logical to suspect airborne transmission in these circumstances. [Multimedia Appendix 1](#) summarizes the time course of the emergence of infected patients in April. After the outbreak, the Disaster Medical Assistance Team was dispatched to the zone and critically ill patients were hospitalized, after which the infection was considered under control. The field survey reported in this study was conducted in August 2021.

On the basis of the interviews with facility staff, the following five separate areas in the building, where large-scale secondary infection occurred, were extensively studied: (1) regular bathrooms, (2) nursing bathrooms, (3) shared rooms, (4) private rooms, and (5) day rooms, as shown in [Figure 1](#). Residents used both regular and nursing bathrooms, and staff members accompanied them for assistance. The risk of aerosol infection is considered to increase in bathrooms, as neither the residents nor staff wear a mask because of the high-humidity environment. Furthermore, a care recipient using a nursing bathroom requires high-level care, and the staff must talk to them while making contact, which is expected to increase the risk of transmission. The shared room had a maximum of 4 beds, 3 of which were for residents, whereas the other was unused at the time of the outbreak. Although residents were instructed to wear masks at all times, it was difficult to enforce them under certain conditions, such as dementia. A private room is the one for a single resident; therefore, the risk of aerosol infection is relatively low. It was segregated from the corridor using a curtain. The first resident who tested positive for COVID-19 was within a private room next to the day room; subsequently, a mass infection outbreak occurred. The day room is a place for relaxation; has free access for all residents and other visitors temporarily visiting the facility between 6:00 AM and 8:00 PM daily; and is furnished with chairs, tables, and televisions. As the day room is frequently used by multiple people, including care recipients and staff, the risk of infection is expected to be relatively high. These rooms were first investigated individually during the primary measurements, as described later. Therefore, the layout of these rooms in the entire floor plan was irrelevant. Furthermore, several transmission routes were involved in the studied mass infection outbreak event. This study focused on airborne transmission via aerosols; however, other transmission routes, such as direct contact, were not excluded.

Figure 1. Floor plan of each room investigated during the primary measurement. The locations of sensors P and K1-K4 are also shown. (a) Regular bathroom, (b) nursing bathroom, (c) shared room, (d) day room, and (e) private room. The ceiling heights of these rooms were 2.6, 2.4, 2.7, 2.4, and 2.4 m, respectively. K: TR-76Ui sensor used; P: SCD-30 sensor used; TV: television.



Ethics Approval

This study was approved by the Ethics Committee on Experiments on Human Subjects at the University of Electro-Communications, Chofugaoka 1-5-1, Chofu, Tokyo, Japan (approval number: 21,005).

Measurements

On-site measurements were conducted under the guidance of an industrial physician to ensure safety while confirming that the results of the PCR test of all residents, staff, and researchers were negative, and safety measures, such as the use of personal protective equipment and disinfection, were taken. In this study, the following 2 types of experiments were performed: primary and secondary measurements.

Primary Measurement

In the primary experiment conducted on August 13, 2021, the tracer gas method was used to measure the ACR in 5 areas. CO₂, which was obtained by vaporizing dry ice in the study room, was used as the tracer gas. Two types of nondispersive infrared-type CO₂ sensors were used. The first was a mobile CO₂ sensor (Yaguchi Electric Corporation) equipped with a nondispersive infrared sensor SCD-30 (Sensirion AG). Any

measurements performed with this sensor are hereafter designated “P.” The second sensor was a TR-76Ui (T&D Corporation), designated “K,” with an index of measurement locations. A total of 5 sensors (1 P sensor and 4 K sensors, K1-K4) were used during the measurements. Figure 1 shows the arrangement of the sensors in each room. All the sensors were placed at a height of approximately 1 m from the ground.

Measurements were conducted in each room as follows:

1. The CO₂ sensors were installed at the locations shown in Figure 1, and the measurements were initiated.
2. The mechanical ventilation system in the room of interest was turned off, and the windows and doors, if any, were closed to create a closed room.
3. Dry ice was placed in the room of interest to achieve a CO₂ concentration sufficiently high compared with the background (approximately 400 ppm). The concentration must be at least 2000 ppm but should not exceed the permissible concentration of 5000 ppm (at 8 h of exposure), as specified by the Japanese Industrial Safety and Health Law. In addition, a blower was used to sufficiently mix the room air with the generated CO₂ gas, as the gas evaporated

- from the dry ice yielded a low temperature and tended to remain close to the floor.
4. When the CO₂ concentration was sufficiently increased, dry ice was removed from the room, and the mechanical ventilation system and conditions of the windows and doors were set according to the target measurement conditions. The time point was denoted as t_0 .
 5. The room was immediately vacated to avoid CO₂ addition from breathing. This time point was defined as the start of ventilation (t_{sta})
 6. CO₂ concentration was monitored remotely from outside the room.
 7. Once the CO₂ concentration decreased sufficiently, the measurement was completed; this time point was denoted as t_{end} . The time-series measurement data from t_{sta} to t_{end} were saved for analysis and used to estimate ACR.
 8. If the CO₂ concentration remained sufficiently high, step 3 could be omitted, and we proceeded to step 4.

The measured CO₂ concentration data were processed to calculate the ACR in each room. The Seidel equation is as follows:

$$\frac{dC}{dt} = \frac{Q}{V}(C_0 - C) + M$$

where C_{end} is the CO₂ concentration (ppm) at $t = t_{end}$, C_0 is the steady-state value of the CO₂ concentration in the absence of pollution (ppm), C_{sta} is the concentration (ppm) at $t = t_{sta}$, V is the room volume (m³), Q is the ventilation rate (m³/h), and M is the rate of pollutant generation (ppm m³/h). The measurement

start and end times, t_{sta} and t_{end} , are in hours. Furthermore, C_0 was assumed to be 400 ppm. Because the room was vacant during the measurement, $M = 0$ in equation 1, which means the following:

$$ACR = \frac{C_0 - C_{end}}{C_{sta} - C_{end}}$$

The ACR value, calculated as Q / V (1/h), was the decrease in the CO₂ concentration from t_{sta} until t_{end} .

Secondary Measurement

In addition to measuring the 5 rooms individually, secondary measurements (Figure 2) were performed on August 13, 2021, to investigate the effect of the interplay between airflow in the private and day rooms, as these rooms are spatially connected via a 2-3 m long corridor. This dynamic fluid aspect suggests that aerosols leaked from the private room to the day room. A resident who was COVID-19-positive in the early stage of the mass outbreak was isolated in a private room. For secondary measurement, the sensors were placed as shown in Figure 2. The private room was filled with CO₂ gas, measurement steps 1 to 6 were followed, and gas leakage from the private room to the day room was assessed.

Figure 3 provides a photograph of the CO₂ smoke leaking from the private room shown in Figure 2 to the corridor leading to the day room, captured from the camera angle shown in Figure 2. A substantial amount of CO₂ smoke leaked from the gap between the curtain and the floor and advected toward the ceiling. By the time it reached the day room, a high concentration of CO₂ smoke was observed near the head height of a person sitting in a wheelchair.

Figure 2. Floor plan including the private and day rooms connected by the short corridor and sensor locations for the secondary measurement. The orientation axes for the numerical simulation are shown. The first COVID-19-positive resident was found in a private room next to the day room. The locations of different sensors P, K1, K3, and K4 are also shown. K: TR-76Ui sensor used; P: SCD-30 sensor used.

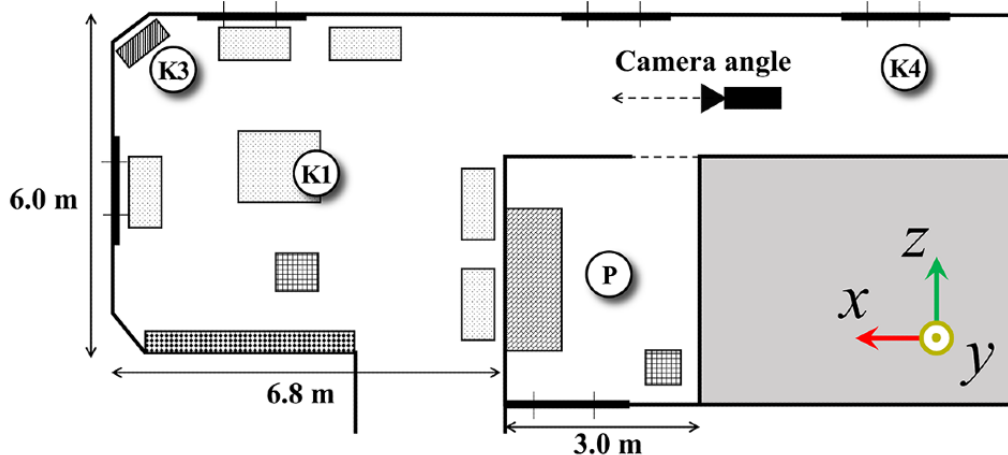


Figure 3. Direct photograph of the carbon dioxide smoke leaking from the private room of the first COVID-19–positive patient into the corridor leading to the day room containing multiple lounging tenants (taken from the camera angle shown in Figure 2).



Numerical Simulation

Given that the spatiotemporal distribution of CO₂ concentration is not readily determined from the secondary measurement experiment, a numerical simulation of gas leakage from the private rooms to the day rooms was performed. This simulation corresponded to a secondary measurement. The simulation was performed using FlowSquare+ (version 2021R1.0, Nora Scientific Co. Ltd), which solves the transport equations for mass (density, ρ):



for momentum ρu_i :



for energy (temperature, T):



and for the mass fraction of the CO₂ gas generated by the dry ice Y_{CO_2} :



in a large eddy simulation context. The subscript indexes $i=, 2,$ and 3 correspond to the directions $x, y,$ and $z,$ respectively; $\nu, \alpha,$ and D are the kinematic viscosity, thermal diffusivity, and molecular diffusivity, respectively; and $\rho v = 2.0 \times 10^{-6}$ (kg/m/s), $\alpha = \nu / Pr,$ and $D = \nu / Sc,$ where the Prandtl number (Pr) = 0.7 and the unity Schmidt number are considered. The subgrid-scale stress tensor, $\tau_{ij},$ and scalar fluxes, $\phi_{T,b}$ and $\phi_{Y,b}$ were modeled based on static Smagorinsky and gradient diffusion models, respectively. The transport equations were discretized onto $N_x \times N_y \times N_z = 175 \times 90 \times 30$ uniform mesh points using a second-order finite-difference scheme in space and advanced in time using the explicit Euler method. Only the advection term was calculated using the first-order upwind scheme. The domain dimensions were $L_x \times L_y \times L_z = 17.5 \times 3.0 \times 9.0$ (m³). Physical boundaries that did not coincide with the Cartesian mesh were expressed using a second-order immersed boundary method.

The fluid in the entire domain was initialized at 18 °C and $Y_{CO_2} = 0.0.$ Warm air supplied by the air conditioner in the room produced a temperature of 30 °C. The fluid velocity of the air conditioner installed in the private room was $(u_x, u_y, u_z) = (0, -1, 2)$ m/s, whereas that of the air conditioner installed in the day room was $(0, -1, 1)$ m/s. The ventilation fan in the private room was turned off, whereas that in the air from the day room was discharged at velocities of $(0, 10,$ and $0)$ m/s. These settings were based on measurement conditions, considering the situation during mass infection outbreaks. To mimic the aerosol dispersion from a COVID-19–positive person in a private room, a small inflow boundary was considered on the bed with a constant (CO₂) gas flow at 36 °C at a velocity of $(0, 0.707, 0.707)$ m/s; therefore, its magnitude was 1.0 m/s. Furthermore, the fluid flow issued at this inflow boundary yielded $Y_{CO_2} = 1.0.$

During the numerical simulation, measurement probes were located for P, K1, and K3, as in the measurements (Figure 2), which facilitated the validation of the simulation results by comparing the measured data. The measurement at K4 was not performed in the simulation because of its proximity to the computational boundary.

Results

ACR Estimation

Table 1 summarizes the ACR values calculated using equation 2 from the measured CO₂ concentration for each sensor in each room, as shown in Figure 1. We conducted a factorial effect analysis using a general linear model with the estimated ACR as the objective variable and window-opening conditions, sensor location, and sensor model (P or K) as explanatory variables.



where y_i is the ACR estimated from the experiments, x_{i1} is the variable indicating the window conditions, x_{i2} is the variable indicating the sensor locations (Figure 1), x_{i3} is the variable related to the sensor model, a_0, \dots, a_3 are the regression coefficients, and ϵ_i is the error that is independent and identically distributed in normal distribution. As the explanatory variables

are nominal measures, they were transformed into dummy variables for the analysis. This analysis, using a general linear model, can be regarded as factorial effect analysis with a 3-factor analysis of variance. JMP Pro 16.2.0 (SAS Institute Inc) was

used for the analysis. The estimated ACR used in the analysis was 15 h⁻¹ for the regular bathroom, 11 h⁻¹ for the nursing bathroom, and 10 h⁻¹ for the shared room.

Table 1. Air change rate (ACR) values and per capita ventilation volumes calculated based on the primary measurement results in each room shown in Figure 1.

Room and window-opening condition ^a	Measured ACR							Ventilation volume ^b m ³ /person
	P ^c (h ⁻¹)	K1 ^d (h ⁻¹)	K2 (h ⁻¹)	K3 (h ⁻¹)	K4 (h ⁻¹)	Ave. ^e (h ⁻¹)	Ratio (r _{ACR}) ^f	
Regular bathroom								
Closed ^g	5.13	3.56	2.56	6.14	1.25	3.73	1.00 ^h	56.7 ⁱ
Open1 ^j	1.00	1.00	7.99	3.81	1.00	5.90	1.58	70.6 ^{i,k}
Open2 ^{j,l}	1.00	5.38	7.5	1.00	11.9	8.25	2.21	84.0 ^{i,k}
Open3 ^{j,l,m}	9.41	9.37	9.62	9.77	11.2	9.86	2.65	125 ^{i,k}
Nursing bathroom								
Closed ^g	2.38	1.56	3.17	2.55	2.27	2.39	1.00	36.8 ⁱ
Open	16.6	11.8	11.0	16.1	18.9	14.9	5.74	211 ^{i,k}
Shared room								
Closed ^g	3.67	1.05	2.34	1.09	1.81	1.99	1.00	38.0 ⁱ
Open	4.74	10.4	7.40	6.68	7.20	7.29	3.59	136 ^{i,k}
Private room								
Closed ^g	3.74	1.00	6.17	1.00	1.00	4.96	1.00	67.8 ⁱ
Open	16.2	8.66	6.81	11.5	16.8	12.0	2.18	148 ^{i,k}
Day room								
Closed ^g	7.79	1.00	4.26	6.73	8.47	6.81	1.00	30.6 ^{i,k}

^aState of window opening and closing set in step 4 in the *Methods* section.

^bVentilation volume per person per hour calculated by multiplying the average air change rate (ACR) by the volume of the room and dividing it by room capacity.

^cP: SCD-30 sensor.

^dK: TR-76Ui sensor.

^eArithmetic mean of the calculated ACR values recorded by different sensors in each room.

^fRatio of ACR values between “Open” and “Closed” conditions.

^gConditions during mass infection outbreaks.

^hSince “Closed” is the denominator in determining the ratio, it is expressed as 1.00 as a ratio to itself.

ⁱMeasured hourly ventilation volume compliant with the Ministry of Health, Labor and Welfare guidelines (2020).

^jTwo windows near the chairs (highlighted in red in Figure 1) were also opened.

^kMeasured hourly ventilation volume compliant with the Centers for Disease Control and Prevention guidelines (2003).

^lA window near the luggage space (highlighted in red in Figure 1) was also opened.

^mThe windows facing the hallway outside the bathroom were additionally opened.

Effects of Window-Opening

The results showed that ventilation had a significant effect on window-opening conditions in all rooms ($P=.027$, $.001$, and $.03$ for the regular bathroom, nursing bathroom, and shared room, respectively). This is consistent with the box plots of the ACR merged with the sensor locations and models for the opening and closing of the windows, as shown in Figure 4. In contrast,

no significant effect was observed at any of the sensor locations. Therefore, ventilation in the room was likely uniformly improved by opening the window in the regular and nursing bathrooms and the shared room. There was no significant dependence on the sensor model. It is well known that the analysis of variance is somewhat robust regarding normality (the population distribution of the observed values in each group is normally distributed) and equality of variance (the population

variance of each group is equal). Consistent results were also obtained using Welch 2-tailed t test, which was used to test the hypothesis that 2 populations have equal means regarding 2 samples that may have unequal variances. Therefore, the analysis of the effect of window conditions is considered valid.

A more detailed analysis was conducted for the private room, where the first resident who tested positive for COVID-19 was staying. In the private room, the P and K2 sensors were installed close together because of the limited size of the room. Thus, we used a generalized linear mixed model (GLMM) to estimate and test the effects of ACR and window-opening conditions (with ventilation time and window-opening set as fixed effects) and the difference between sensor models (with the sensor model as a random effect). A GLMM was assumed for the model in equation 2 with the sensor model difference as a variable effect, and an analysis of the change in CO₂ concentration with ventilation time was conducted. On the assumption of normal distributed random effects, y can be expressed in the form of a GLMM as follows:

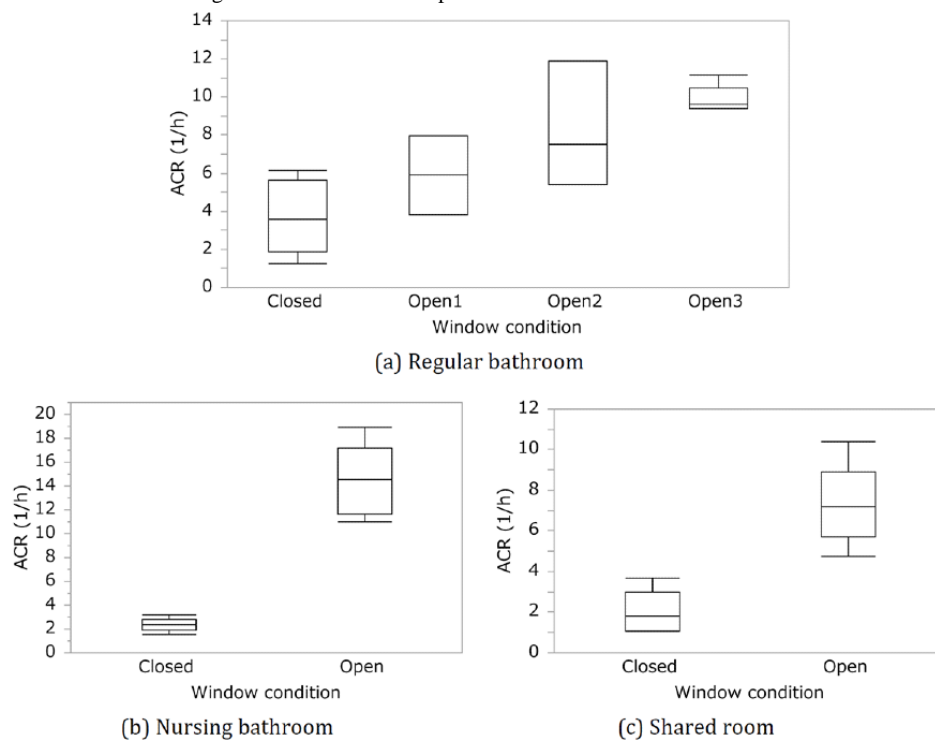
$$y = x\beta + Zu + e(7)$$

where β is the fixed effect, u and e are the normally distributed random effects, and X and Z are the design matrices. The fixed effect of the GLMM, which indicates a time-series improvement, can be summarized as $E(y) = X\beta$. Zu is the random effect of the individual variations. Here, e represents the intraindividual random effect. JMP (version 16.2) was used for the analysis

using the maximum likelihood method. During the covariance parameter estimate of the random effect, the Wald P value of the sensor's measurement was not significant ($P=.96$). Therefore, no difference in the sensor measurements was observed. However, the estimation of the fixed effects showed that the effects of ventilation time and window status were both significant, and their interaction was also highly significant ($P<.001$). The CO₂ reduction (Q/V in equation 2) was 3.90/h without window opening and 11.52/h with window opening. In other words, it was confirmed that window opening improved the ACR by a factor of 3.

Although the windows were closed during the mass infection outbreak, even under these conditions, all rooms met the MHLW standard for ventilation ($>30 \text{ m}^3/\text{h}$ per person) [6], suggesting that additional ventilation measurements were unnecessary. Furthermore, the r_{ACR} values in Table 1 indicate that in all rooms, except the day room where the window-open experiment was not conducted, the ACR substantially improved by 2.2 to 5.7 times by opening the window to meet the CDC standard [4]. There was no significant dependence on the location or model of the sensor in all rooms where the window-open experiments were conducted, suggesting that the proposed method for estimating the ACR is efficient. Even under open-window conditions, no room met the WHO criteria [5]. However, because this facility is not classified as an infectious disease ward, this level of ventilation capacity is not strictly required.

Figure 4. Box plots of estimated air change rates (ACRs) showing the effects of opening the window in a (a) regular bathroom, (b) nursing bathroom, and (c) shared room. The window condition agrees with the definition presented in Table 1.



Fluid Advection

The temporal variation of the measured CO₂ concentrations in a larger area, including the private room and the day room (Figure 2), is shown in Figure 5. After the generation of CO₂

gas was stopped at step 4 (as per the primary measurement subsection) at $t = t_0$, the concentration at sensor P steadily decreased with time. Although the concentration of CO₂ measured at K1 to K3 fluctuated, it exhibited a gradual increase toward $t - t_0 = 9$ minutes. The concentration at K3 was

substantially higher than that of the other 2 K sensors in the early measurement stage ($t - t_0 = 2$ min), suggesting that the generated CO₂ was predominantly advected from the private room by a large flow pattern rather than a simple diffusion process. This is also supported by the results shown in Figure 4, suggesting that ventilation performance solely depended on the window-opening condition, which dictates airflow. Later in the measurement ($t - t_0 > 9$ min), the CO₂ concentrations at K1 and K4 surpassed that of K3, with a relatively steep slope in variation. This rapid change in the rate of concentration increase was due to advection, rather than diffusion.

The evolution of temporal CO₂ concentrations for the positions of the numerical simulations at the positions of the P, K1, and K3 probes is shown in Figure 5. Despite attempts to mimic the environment in a nursing home facility, certain uncertainties in the experimental setup, such as the initial field and precise boundary conditions for ventilation or air conditioning devices, were unavoidable. Although this causes a discrepancy between the simulation results and measured data, the CO₂ concentration obtained from the numerical simulation showed levels quantitatively similar to the measured values. Furthermore, the evolution of the CO₂ concentration shows similar trends after multiplying the time axis of the simulation results by an a posteriori factor of 2.5, that is, $t^* = 2.5(t_{sim} - t_{sim,0})$, where the subscript “sim” denotes the physical time in the simulation.

Regarding the evolution of the CO₂ concentration in computationally simulated P, a peak was observed at $t^* = 0$, and its value was approximately 10^4 ppm. Subsequently, a monotonic decrease mode occurred until the end of the simulation. This mode transition is similar to the measurement

result for the P sensor, although the peak is four times larger than the measurement value, and its decrease is locally ($t^* \sim 1$) nonmonotonic. The simulation and measurement results differed owing to the uncertainties of the initial and boundary conditions in the simulation, mimicking the experimental setup. For example, the airflow through the small gaps between the wall, door, and curtains, the flow direction of the air conditioning devices, and the working staff in the test room cannot be accurately considered in the 3D model. However, the overall trend observed in the measurements is well reproduced in the numerical simulation.

Figure 6 provides instantaneous snapshots of the CO₂ isosurface at 1000 ppm at $t^* = 0.0, 0.9, 2.8,$ and 8.4 minutes. A high concentration of CO₂ was observed throughout the private room when the CO₂ generation was cut off at $t^* = 0$. The maximum CO₂ concentration was 8.8×10^5 ppm. Owing to turbulent and molecular diffusion effects, the maximum CO₂ concentration was monotonically reduced to 2.9×10^5 ppm ($t^* = 0.9$ min; Figure 6), 5.1×10^4 ppm ($t^* = 2.8$ min; Figure 6), and 1.11×10^4 ppm ($t^* = 8.4$ min; Figure 6). Furthermore, the region with a CO₂ concentration > 1000 ppm (high-concentration zone), which is substantially greater than the background value, C_0 , is spread across the day room and corridor. In particular, the high-concentration zone occupies the corridor in front of the private room at $t^* = 2.8$ minutes. The local fluid velocity, overlaid on the CO₂ isosurface, shows that the leading edge of the CO₂ isosurface moves toward the day room (red arrows in Figure 6) and produces a relatively high velocity. Therefore, large-scale flow dictated the spread of the gas mixture containing infectious aerosols rather than the molecular diffusion process.

Figure 5. Temporal variation of the measured carbon dioxide (CO₂) concentration during the secondary measurement (thin line) and the numerical simulation (thick line) at the sensor positions shown in Figure 2. Note that the modified time axis t^* is used for comparison with the simulation results. The vertical lines indicate $t - t_0 (=t^*) = 2$ and $t - t_0 (=t^*) = 9$. K: TR-76Ui sensor used; P: SCD-30 sensor used.

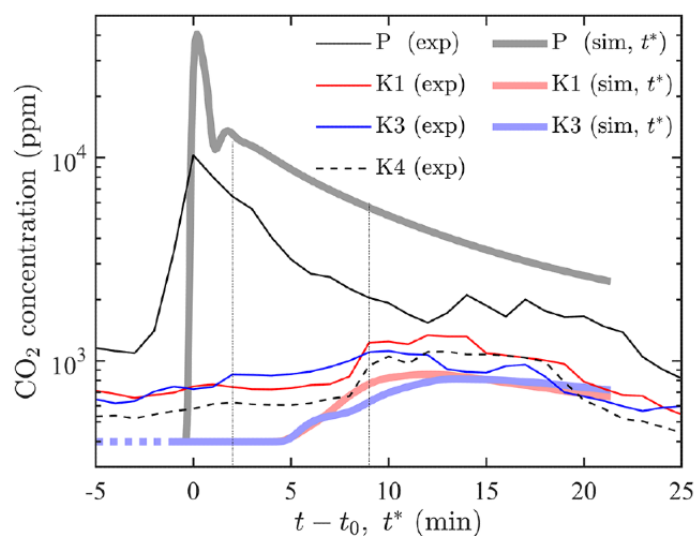
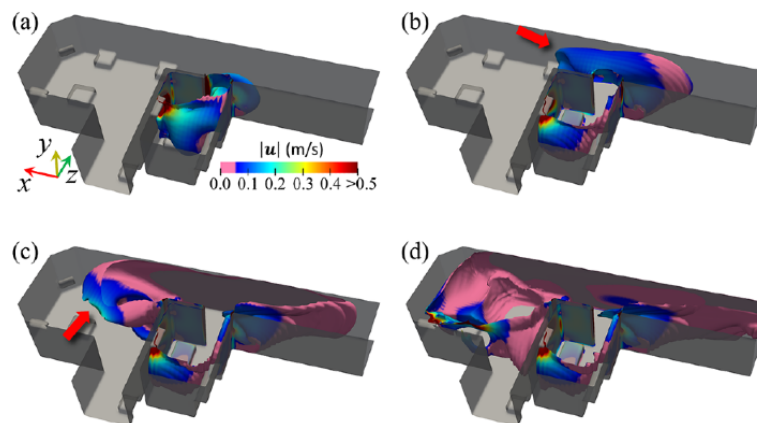


Figure 6. Temporal evolution of the carbon dioxide concentration isosurfaces at 1000 ppm at $t^* =$ (a) 0.0, (b) 0.9, (c) 2.8, and (d) 8.4 minutes. The local color of the isosurface represents the local velocity magnitude in m/s (rainbow: high-velocity region, pink: low-velocity region).



Discussion

Principal Findings

The CO₂ tracer gas method was used in 6 areas where airborne infection was suspected, and the ACR was successfully calculated using the primary measurement. On the basis of these results, the ventilation rate in each room at the time of the outbreak met the MHLW guidelines and did not correspond to poorly ventilated space. However, the secondary measurement revealed advection from the private room to the day room by visualizing the CO₂ tracer gas. The overall trend in gas concentrations detected via secondary measurements agreed well with the results of the numerical simulation. Hence, given that the person who initiated the infection occupied the private room on the day of the infection and that several occupants were gathered in the day room, it was postulated that the infectious aerosol was transmitted by this airflow. These findings confirm the usefulness of the CO₂ tracer gas method for estimating the ACR and visualizing airflow.

The experimental data confirmed that the opening of windows in this facility promoted natural ventilation and significantly improved the ACR from 2.2 to 6.2 times, thus satisfying the CDC criteria. These findings prompted the facility staff to improve their operations by opening windows when appropriate. Furthermore, in this facility, regular nursing bathrooms were equipped with jalousie windows to ensure the privacy of residents. Therefore, opening windows for additional ventilation does not negatively affect privacy. However, in many older adult care facilities, opening windows is avoided because of concerns regarding dangerous behavior. Residents diagnosed with behavioral and psychological symptoms of dementia can attempt to use the window to leave the facility. Therefore, opening windows as a measure to improve air ventilation should be conducted when considering these safety aspects. For example, a simple locking mechanism can be installed to prevent windows from opening beyond a certain point, thereby effectively ensuring safety and ventilation. In addition, future studies should clarify the effects of open windows on heating and cooling.

The results of the secondary measurements discussed in Figures 5 and 6 suggest that aerosols can be advected between different

rooms by the relatively strong fluid flows created by ventilation fans and air conditioning devices. The required ventilation volume (per hour) was determined based on the capacity of the room. For smaller rooms, such as the private or shared rooms in this study, the required ventilation volume is smaller, and vice versa for larger rooms, such as the day room. Therefore, in open or semiopen buildings, where multiple rooms are spatially connected, such as the present facility, advective flows from private to common areas may exist, and adhering to the ventilation volume standard for each room individually may not be sufficient to prevent mass infection outbreaks, as in the present case. This is a challenge faced by many nursing homes where private and common spaces are directly or indirectly connected. However, the facility learned about the risk of aerosol advection from the private room to the day room and subsequently took steps to prohibit its use of this private room. Future facility modifications are planned to reverse the pressure difference by improving the ventilation system to limit or prevent aerosol leakage.

Limitations

We found that local flows could advect infectious aerosols that could not be predicted solely by the ACR of the individual rooms. This finding was based on a combination of experimental (CO₂ tracer) and numerical (CFD) examinations, which were performed at this facility. Therefore, a limitation of this study is that it infers the likelihood and route of airborne transmission only from such circumstantial evidence. In the future, it will be possible to trace the order of viral transmission and the route of infection more directly for each individual by analyzing the entire genome of infected people at the time of an outbreak. Furthermore, because various hospitals and offices have video surveillance cameras and entry or exit records, the combination of these security logs, epidemiological data, and 3D models will enable a clear reconstruction of the infection scenario. In addition, a digital contact tracing system [29] and a web-based smartphone app tenant management tool [30] would be useful in evidence building.

Another limitation is the seasonal reproducibility. April 2020, when the outbreak was first reported, was in spring; August 2021, when the experiment was conducted, was in summer. Although mechanical conditions, such as window opening and closing, ventilator, and air conditioner operating conditions,

could be reproduced, the amount of buoyant ventilation owing to the difference in indoor and outdoor temperatures, especially when windows were opened, would have varied with the season. Although the effect of buoyant ventilation by season on natural ventilation is another important topic, and a detailed analysis by CFD is desirable, it is outside the scope of this study.

Furthermore, although this study analyzed cases in which the infection spread locally, the results did not indicate a general causal relationship between ventilation volume and infection rate. Future field studies and meta-analyses of a large number of outbreak cases will clarify whether there is a correlation between the amount of ventilation and the risk of COVID-19 infection.

Comparison With Prior Work

Anderson et al [31] determined that typical older adult care facilities may be vulnerable to COVID-19 because they are designed to promote social interaction and collaboration among residents via common spaces (eg, day rooms and areas for social activities) and hallways without partitions [32,33]. This aspect is important for residents' social interactions and daily monitoring by staff. Furthermore, in Japan, the deregulation of the Law for Partial Revision of the Building Standards (enacted on September 25, 2018), which exempted the floor area of common corridors from the calculation of the floor area ratio for nursing homes, may have provided an impetus for the active use of corridors as common relaxation areas. However, from the perspective of infection control, there is room for improvement in these open-plan architectural guidelines. With these precedents, practical guidelines should be formulated specifically to address the operational patterns of older adult care facilities. As an example of a temporary guideline during a pandemic, downwind transmission can easily be prevented by discontinuing the use of private rooms close to common rooms in older adult care facilities. The excessive installation of vinyl partitions may also contribute to mass infection [3] due to the stagnation of fluid flow, which is essential for active ventilation. Therefore, care should be taken when designing partitions such that they do not interfere with ventilation. A more quantitative measure would be recommended to monitor the pressure difference between the room and hallway [34], as recommended in health care settings. If the possibility of downwind transmission created by the pressure difference

becomes apparent, other measures (eg, transparent partitions or air curtains) can be taken to prevent large-scale airflow, such as that observed in the present physical or fluid-dynamic numerical simulation. However, these measures must consider accessibility and visibility to ensure residents' quality of life.

Conclusions

In this study, a real-world mass infection outbreak, which occurred in an older adult care facility, was simulated experimentally and numerically to investigate the controlling factors and quantify the effectiveness of various natural ventilation settings using the ACR, assuming that airborne transmission has occurred. Using the CO₂ tracer gas method, we determined that the low-cost intervention of opening windows could improve the ventilation frequency by a factor of 2.2 to 5.7. This implies that advective fluid flows are key to controlling the spread of zones with high CO₂ concentrations. Therefore, this CO₂ tracer gas method, implemented with dry ice and a sensor, can quantify the ACR and airflow and contribute to the prevention of recurrent airborne infections. Moreover, this method can be performed within a relatively short time, even in the presence of patients or residents.

A numerical simulation was performed to obtain the spatiotemporal evolution in such high CO₂ concentration zones under conditions similar to those of the present experiment. The development of zones with high CO₂ concentrations occurred in the first few minutes. Furthermore, the leading edge of such zones toward the day room, where multiple residents gather for activities, yields a relatively high fluid velocity, suggesting that large-scale advective flow dictates the spread of such high CO₂ concentration zones. These results suggest that secondary infections could occur because of aerosol advection driven by a large-scale flow topology, even if the ventilation is sufficient. Furthermore, this phenomenon may be influenced by architectural design specific to typical older adult care facilities.

To prevent or deter outbreaks of mass infections in older adult care facilities, policies for guidelines on architectural design and reviews of related laws will be necessary, considering both the quality of life of the residents and suppression of large-scale flow toward communal areas. In addition, quantitative studies and interventions are required to avoid downwind contamination of existing buildings.

Acknowledgments

This study was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (grant 21K19820) and the KDDI Foundation.

Data Availability

The data supporting the findings of this study are available from the corresponding author, YI, upon reasonable request.

Authors' Contributions

YI conceptualized and designed the study and contributed to writing, reviewing, and approving of the manuscript. SY performed the formal analysis and contributed to writing, reviewing, and approving the manuscript. YM performed data visualization and

contributed to writing, reviewing, and approving the manuscript. AS, HK, and YK contributed to the data extraction and investigation. All authors have approved the final manuscript as submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Time course of the emergence of infected patients in April.

[[XLSX File \(Microsoft Excel File\), 29 KB - formative_v6i12e37587_app1.xlsx](#)]

References

1. Lee SW, Yuh WT, Yang JM, Cho YS, Yoo IK, Koh HY, et al. Nationwide results of COVID-19 contact tracing in South Korea: individual participant data from an epidemiological survey. *JMIR Med Inform* 2020 Aug 25;8(8):e20992 [[FREE Full text](#)] [doi: [10.2196/20992](https://doi.org/10.2196/20992)] [Medline: [32784189](https://pubmed.ncbi.nlm.nih.gov/32784189/)]
2. The Yomiuri Shimbun. Elderly care facilities under pressure amid COVID surge. *The Japan News*. 2022 Aug 17. URL: <https://japannews.yomiuri.co.jp/society/general-news/20220817-52136/> [accessed 2022-10-23]
3. Ishigaki Y, Kawauchi Y, Yokogawa S, Saito A, Kitamura H, Moritake T. Experimental investigation to verify if excessive plastic sheeting shielding produce micro clusters of SARS-CoV-2. *medRxiv Preprint* posted online on May 27, 2021. [doi: [10.1101/2021.05.22.21257321](https://doi.org/10.1101/2021.05.22.21257321)]
4. Li Y, Qian H, Hang J, Chen X, Cheng P, Ling H, et al. Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. *Build Environ* 2021 Jun;196:107788 [[FREE Full text](#)] [doi: [10.1016/j.buildenv.2021.107788](https://doi.org/10.1016/j.buildenv.2021.107788)] [Medline: [33746341](https://pubmed.ncbi.nlm.nih.gov/33746341/)]
5. Jang S, Han SH, Rhee JY. Cluster of coronavirus disease associated with fitness dance classes, South Korea. *Emerg Infect Dis* 2020 Aug;26(8):1917-1920 [[FREE Full text](#)] [doi: [10.3201/eid2608.200633](https://doi.org/10.3201/eid2608.200633)] [Medline: [32412896](https://pubmed.ncbi.nlm.nih.gov/32412896/)]
6. Menzies D, Fanning A, Yuan L, FitzGerald JM. Hospital ventilation and risk for tuberculous infection in canadian health care workers. *Canadian Collaborative Group in Nosocomial Transmission of TB. Ann Intern Med* 2000 Nov 21;133(10):779-789. [doi: [10.7326/0003-4819-133-10-200011210-00010](https://doi.org/10.7326/0003-4819-133-10-200011210-00010)] [Medline: [11085840](https://pubmed.ncbi.nlm.nih.gov/11085840/)]
7. Guidelines for Environmental Infection Control in Health-Care Facilities: Recommendations of CDC and the Healthcare Infection Control Practices Advisory Committee (HICPAC). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. 2019. URL: <https://www.cdc.gov/infectioncontrol/guidelines/environmental/index.html> [accessed 2022-01-09]
8. Natural ventilation for infection control in health care settings. World Health Organization (WHO). 2009. URL: <https://apps.who.int/iris/handle/10665/44167> [accessed 2022-01-09]
9. Ventilation to improve "poorly ventilated closed spaces" in commercial facilities. Ministry of Health, Labour and Welfare, Japan (MHLW). 2020. URL: <https://www.mhlw.go.jp/content/10900000/000616069.pdf> [accessed 2022-01-09]
10. Scientific Brief: SARS-CoV-2 Transmission. Centers for Disease Control and Prevention. 2021. URL: <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmission.html> [accessed 2022-12-16]
11. Position paper of the Gesellschaft für Aerosolforschung on understanding the role of aerosol particles in SARS-CoV-2 infection. *Gesellschaft für Aerosolforschung*. 2020 Dec 17. URL: https://ae00780f-bbdd-47b2-aa10-e1dc2cdeb6dd.files.usr.com/ugd/fab12b_0b691414cfb344fe96d4b44e6f44a5ab.pdf [accessed 2022-12-16]
12. Jones RM, Brosseau LM. Aerosol transmission of infectious disease. *J Occup Environ Med* 2015 May;57(5):501-508. [doi: [10.1097/JOM.0000000000000448](https://doi.org/10.1097/JOM.0000000000000448)] [Medline: [25816216](https://pubmed.ncbi.nlm.nih.gov/25816216/)]
13. Wang CC, Prather KA, Sznitman J, Jimenez JL, Lakdawala SS, Tufekci Z, et al. Airborne transmission of respiratory viruses. *Science* 2021 Aug 27;373(6558):eabd9149 [[FREE Full text](#)] [doi: [10.1126/science.abd9149](https://doi.org/10.1126/science.abd9149)] [Medline: [34446582](https://pubmed.ncbi.nlm.nih.gov/34446582/)]
14. Cai J, Sun W, Huang J, Gamber M, Wu J, He G. Indirect virus transmission in cluster of COVID-19 cases, Wenzhou, China, 2020. *Emerg Infect Dis* 2020 Jun;26(6):1343-1345 [[FREE Full text](#)] [doi: [10.3201/eid2606.200412](https://doi.org/10.3201/eid2606.200412)] [Medline: [32163030](https://pubmed.ncbi.nlm.nih.gov/32163030/)]
15. Katelaris AL, Wells J, Clark P, Norton S, Rockett R, Arnott A, et al. Epidemiologic evidence for airborne transmission of SARS-CoV-2 during church singing, Australia, 2020. *Emerg Infect Dis* 2021 Jun;27(6):1677-1680 [[FREE Full text](#)] [doi: [10.3201/eid2706.210465](https://doi.org/10.3201/eid2706.210465)] [Medline: [33818372](https://pubmed.ncbi.nlm.nih.gov/33818372/)]
16. Toyokawa T, Shimada T, Hayamizu T, Sekizuka T, Zukeyama Y, Yasuda M, et al. Transmission of SARS-CoV-2 during a 2-h domestic flight to Okinawa, Japan, March 2020. *Influenza Other Respir Viruses* 2022 Jan;16(1):63-71 [[FREE Full text](#)] [doi: [10.1111/irv.12913](https://doi.org/10.1111/irv.12913)] [Medline: [34605181](https://pubmed.ncbi.nlm.nih.gov/34605181/)]
17. Tellier R, Li Y, Cowling BJ, Tang JW. Recognition of aerosol transmission of infectious agents: a commentary. *BMC Infect Dis* 2019 Jan 31;19(1):101 [[FREE Full text](#)] [doi: [10.1186/s12879-019-3707-y](https://doi.org/10.1186/s12879-019-3707-y)] [Medline: [30704406](https://pubmed.ncbi.nlm.nih.gov/30704406/)]
18. Tang JW, Li Y, Eames I, Chan PK, Ridgway GL. Factors involved in the aerosol transmission of infection and control of ventilation in healthcare premises. *J Hosp Infect* 2006 Oct;64(2):100-114 [[FREE Full text](#)] [doi: [10.1016/j.jhin.2006.05.022](https://doi.org/10.1016/j.jhin.2006.05.022)] [Medline: [16916564](https://pubmed.ncbi.nlm.nih.gov/16916564/)]
19. Anchordoqui LA, Chudnovsky EM. A physicist view of COVID-19 airborne infection through convective airflow in indoor spaces. *Sci Med J* 2020 Aug 28;2:68-72. [doi: [10.28991/scimedj-2020-02-si-5](https://doi.org/10.28991/scimedj-2020-02-si-5)]

20. Morawska L, Milton DK. It is time to address airborne transmission of coronavirus disease 2019 (COVID-19). *Clin Infect Dis* 2020 Dec 03;71(9):2311-2313 [FREE Full text] [doi: [10.1093/cid/ciaa939](https://doi.org/10.1093/cid/ciaa939)] [Medline: [32628269](https://pubmed.ncbi.nlm.nih.gov/32628269/)]
21. Hunziker P. Minimising exposure to respiratory droplets, 'jet riders' and aerosols in air-conditioned hospital rooms by a 'Shield-and-Sink' strategy. *BMJ Open* 2021 Oct 12;11(10):e047772 [FREE Full text] [doi: [10.1136/bmjopen-2020-047772](https://doi.org/10.1136/bmjopen-2020-047772)] [Medline: [34642190](https://pubmed.ncbi.nlm.nih.gov/34642190/)]
22. Guven Y, Sezgin BI, Çelik EG. The importance and improvement of indoor air quality in dental clinics within the context of Covid-19. *EC Dent Sci* 2021;20(1):32-50 [FREE Full text]
23. Anderson EL, Turnham P, Griffin JR, Clarke CC. Consideration of the aerosol transmission for COVID-19 and public health. *Risk Anal* 2020 May;40(5):902-907 [FREE Full text] [doi: [10.1111/risa.13500](https://doi.org/10.1111/risa.13500)] [Medline: [32356927](https://pubmed.ncbi.nlm.nih.gov/32356927/)]
24. Bourouiba L. Turbulent gas clouds and respiratory pathogen emissions: potential implications for reducing transmission of COVID-19. *JAMA* 2020 May 12;323(18):1837-1838. [doi: [10.1001/jama.2020.4756](https://doi.org/10.1001/jama.2020.4756)] [Medline: [32215590](https://pubmed.ncbi.nlm.nih.gov/32215590/)]
25. Lima AA, Nunes IC, da Silva Duarte JL, Meili L, de Carvalho Nagliate P, dos Santos Almeida AG. Characteristics of SARS-CoV-2 aerosol dispersion in indoor air: scoping review. *Res Soc Dev* 2021 Apr 17;10(4):e44310414300. [doi: [10.33448/rsd-v10i4.14300](https://doi.org/10.33448/rsd-v10i4.14300)]
26. Azimi P, Keshavarz Z, Cedeno Laurent JG, Stephens B, Allen JG. Mechanistic transmission modeling of COVID-19 on the Diamond Princess cruise ship demonstrates the importance of aerosol transmission. *Proc Natl Acad Sci U S A* 2021 Feb 23;118(8):e2015482118 [FREE Full text] [doi: [10.1073/pnas.2015482118](https://doi.org/10.1073/pnas.2015482118)] [Medline: [33536312](https://pubmed.ncbi.nlm.nih.gov/33536312/)]
27. Kitamura H, Ishigaki Y, Ohashi H, Yokogawa S. Ventilation improvement and evaluation of its effectiveness in a Japanese manufacturing factory. *Sci Rep* 2022 Oct 21;12(1):17642 [FREE Full text] [doi: [10.1038/s41598-022-22764-2](https://doi.org/10.1038/s41598-022-22764-2)] [Medline: [36271253](https://pubmed.ncbi.nlm.nih.gov/36271253/)]
28. Mehade Hussain S, Goel S, Kadapa C, Aristodemou E. A short review of vapour droplet dispersion models used in CFD to study the airborne spread of COVID19. *Mater Today Proc* 2022;64:1349-1356 [FREE Full text] [doi: [10.1016/j.matpr.2022.03.724](https://doi.org/10.1016/j.matpr.2022.03.724)] [Medline: [35495177](https://pubmed.ncbi.nlm.nih.gov/35495177/)]
29. Wilmink G, Summer I, Marsyla D, Sukhu S, Grote J, Zobel G, et al. Real-time digital contact tracing: development of a system to control COVID-19 outbreaks in nursing homes and long-term care facilities. *JMIR Public Health Surveill* 2020 Aug 25;6(3):e20828 [FREE Full text] [doi: [10.2196/20828](https://doi.org/10.2196/20828)] [Medline: [32745013](https://pubmed.ncbi.nlm.nih.gov/32745013/)]
30. Echeverría P, Mas Bergas MA, Puig J, Isnard M, Massot M, Vedia C, et al. COVIDApp as an innovative strategy for the management and follow-up of COVID-19 cases in long-term care facilities in Catalonia: implementation study. *JMIR Public Health Surveill* 2020 Jul 17;6(3):e21163 [FREE Full text] [doi: [10.2196/21163](https://doi.org/10.2196/21163)] [Medline: [32629425](https://pubmed.ncbi.nlm.nih.gov/32629425/)]
31. Anderson DC, Grey T, Kennelly S, O'Neill D. Nursing home design and COVID-19: balancing infection control, quality of life, and resilience. *J Am Med Dir Assoc* 2020 Nov;21(11):1519-1524 [FREE Full text] [doi: [10.1016/j.jamda.2020.09.005](https://doi.org/10.1016/j.jamda.2020.09.005)] [Medline: [33138934](https://pubmed.ncbi.nlm.nih.gov/33138934/)]
32. Chmielewski E. Excellence in Design: Optimal Living Space for People with Alzheimer's Disease and Related Dementias. Perkins Eastman/Alzheimer's foundation of America. 2014 Jun. URL: <https://alzfdn.org/wp-content/uploads/2017/11/Excellence-in-Design-white-paper-June-2014.pdf> [accessed 2022-01-09]
33. Wrublowsky R. Design Guide for Long Term Care Homes. Document edition. 2018.01. Robert Wrublowsky. 2018. URL: https://www.fgiguilines.org/wp-content/uploads/2018/03/MMP_DesignGuideLongTermCareHomes_2018.01.pdf [accessed 2022-01-09]
34. Jensen PA, Lambert LA, Iademarco MF, Ridzon R, CDC. Guidelines for preventing the transmission of Mycobacterium tuberculosis in health-care settings, 2005. *MMWR Recomm Rep* 2005 Dec 30;54(RR-17):1-141 [FREE Full text] [Medline: [16382216](https://pubmed.ncbi.nlm.nih.gov/16382216/)]

Abbreviations

- ACR:** air change rate
- CDC:** Centers for Disease Control and Prevention
- CFD:** computational fluid dynamics
- GLMM:** generalized linear mixed model
- CO₂:** carbon dioxide
- K:** TR-76Ui sensor
- MHLW:** Ministry of Health, Labor and Welfare
- P:** SCD-30 sensor
- PCR:** polymerase chain reaction
- WHO:** World Health Organization

Edited by A Mavragani; submitted 01.03.22; peer-reviewed by R Fang, P Dunn; comments to author 07.10.22; revised version received 27.10.22; accepted 07.12.22; published 30.12.22.

Please cite as:

Ishigaki Y, Yokogawa S, Minamoto Y, Saito A, Kitamura H, Kawauchi Y

Pilot Evaluation of Possible Airborne Transmission in a Geriatric Care Facility Using Carbon Dioxide Tracer Gas: Case Study

JMIR Form Res 2022;6(12):e37587

URL: <https://formative.jmir.org/2022/12/e37587>

doi: [10.2196/37587](https://doi.org/10.2196/37587)

PMID: [36583933](https://pubmed.ncbi.nlm.nih.gov/36583933/)

©Yo Ishigaki, Shinji Yokogawa, Yuki Minamoto, Akira Saito, Hiroko Kitamura, Yuto Kawauchi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 30.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Corrigenda and Addenda

Addendum: The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees

Joris Elmar van Gend¹, MSc; Jan Willem Jaap Roderick van 't Klooster¹, IR, PhD; Catherine Adriana Wilhelmina Bolman², PhD; Julia Elisabeth Wilhelmina Cornelia van Gemert-Pijnen³, PhD

¹The BMS Lab, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

²Department of Psychology, Open University of the Netherlands, Heerlen, Netherlands

³Section of Psychology Health & Technology, Department of Technology, Human and Institutional Behavior, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

Corresponding Author:

Joris Elmar van Gend, MSc

The BMS Lab

Faculty of Behavioural, Management and Social Sciences

University of Twente

De Zul 10

Enschede, 7522 NB

Netherlands

Phone: 31 0534892525

Email: bmslab@utwente.nl

Related Article:

Correction of: <https://formative.jmir.org/2022/11/e38904>

(*JMIR Form Res* 2022;6(12):e44497) doi:[10.2196/44497](https://doi.org/10.2196/44497)

In “The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees” (*JMIR Form Res* 2022;6(11):e38904), the authors made the following changes:

1. An *Acknowledgments* section was not included in the initial manuscript and the authors desire this to be added to the published article. The following statement has been added under the new *Acknowledgments* section:

The authors would like to thank the following groups and individuals for their contribution to the research project. First, Prof Dr Wolfgang Ebbers in his role as head of the research team CoronaMelder. Second, the researchers would like to acknowledge all the anonymous participants from the end users and MHS (“GGD”) employee groups. Lastly, the team would like to thank those employees of the Dutch MHS communication and CM app development teams that

helped reflect on and refine the findings and the local newspaper that helped with the recruitment of participants.

2. The incorrect city was added to affiliation 2. The original affiliation was:

²Department of Psychology, Open University of the Netherlands, Utrecht, Netherlands

The corrected version is:

²Department of Psychology, Open University of the Netherlands, Heerlen, Netherlands

The correction will appear in the online version of the paper on the JMIR Publications website on December 2, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 21.11.22; this is a non-peer-reviewed article; accepted 22.11.22; published 02.12.22.

Please cite as:

van Gend JE, van 't Klooster JWJR, Bolman CAW, van Gemert-Pijnen JEW

Addendum: The Dutch COVID-19 Notification App: Lessons Learned From a Mixed Methods Evaluation Among End Users and Contact-Tracing Employees

JMIR Form Res 2022;6(12):e44497

URL: <https://formative.jmir.org/2022/12/e44497>

doi: [10.2196/44497](https://doi.org/10.2196/44497)

PMID: [36459694](https://pubmed.ncbi.nlm.nih.gov/36459694/)

©Joris Elmar van Gend, Jan Willem Jaap Roderick van 't Klooster, Catherine Adriana Wilhelmina Bolman, Julia Elisabeth Wilhelmina Cornelia van Gemert-Pijnen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 02.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Monitoring and Identifying Emerging e-Cigarette Brands and Flavors on Twitter: Observational Study

Qihang Tang^{1*}; Runtao Zhou^{1*}; Zidian Xie², MSc, PhD; Dongmei Li², MSc, PhD

¹Goergen Institute for Data Science, University of Rochester, Rochester, NY, United States

²Department of Clinical & Translational Research, University of Rochester Medical Center, Rochester, NY, United States

*these authors contributed equally

Corresponding Author:

Dongmei Li, MSc, PhD

Department of Clinical & Translational Research

University of Rochester Medical Center

265 Crittenden Boulevard CU 420708

Rochester, NY, 14642

United States

Phone: 1 808 554 2956

Email: Dongmei_Li@urmc.rochester.edu

Abstract

Background: Flavored electronic cigarettes (e-cigarettes) have become very popular in recent years. e-Cigarette users like to share their e-cigarette products and e-cigarette use (vaping) experiences on social media. e-Cigarette marketing and promotions are also prevalent online.

Objective: This study aims to develop a method to identify new e-cigarette brands and flavors mentioned on Twitter and to monitor e-cigarette brands and flavors mentioned on Twitter from May 2021 to December 2021.

Methods: We collected 1.9 million tweets related to e-cigarettes between May 3, 2021, and December 31, 2021, by using the Twitter streaming application programming interface. Commercial and noncommercial tweets were characterized based on promotion-related keywords. We developed a depletion method to identify new e-cigarette brands by removing the keywords that already existed in the reference data set (Twitter data related to e-cigarettes from May 3, 2021, to August 31, 2021) or our previously identified brand list from the keywords in the target data set (e-cigarette-related Twitter data from September 1, 2021, to December 31, 2021), followed by a manual Google search to identify new e-cigarette brands. To identify new e-cigarette flavors, we constructed a flavor keyword list based on our previously collected e-cigarette flavor names, which were used to identify potential tweet segments that contain at least one of the e-cigarette flavor keywords. Tweets or tweet segments with flavor keywords but not any known flavor names were marked as potential new flavor candidates, which were further verified by a web-based search. The longitudinal trends in the number of tweets mentioning e-cigarette brands and flavors were examined in both commercial and noncommercial tweets.

Results: Through our developed methods, we identified 34 new e-cigarette brands and 97 new e-cigarette flavors from commercial tweets as well as 56 new e-cigarette brands and 164 new e-cigarette flavors from noncommercial tweets. The longitudinal trend of the e-cigarette brands showed that JUUL was the most popular e-cigarette brand mentioned on Twitter; however, there was a decreasing trend in the mention of JUUL over time on Twitter. Menthol flavor was the most popular e-cigarette flavor mentioned in the commercial tweets, whereas mango flavor was the most popular e-cigarette flavor mentioned in the noncommercial tweets during our study period.

Conclusions: Our proposed methods can successfully identify new e-cigarette brands and flavors mentioned on Twitter. Twitter data can be used for monitoring the dynamic changes in the popularity of e-cigarette brands and flavors.

(*JMIR Form Res* 2022;6(12):e42241) doi:[10.2196/42241](https://doi.org/10.2196/42241)

KEYWORDS

e-cigarettes; brand; flavor; Twitter

Introduction

Electronic cigarettes (e-cigarettes) or vapes are electronic devices that heat and aerosolize liquids containing propylene glycol, vegetable glycerin, flavorings, and in most cases, nicotine [1]. Although e-cigarettes have been marketed as safer and healthier alternatives to help adult smokers quit smoking, e-cigarettes have become very popular among teenagers who have never smoked [2]. Based on data from the Centers for Disease Control and Prevention, 11.3% of high school students and 2.8% of middle school students in the United States reported using e-cigarettes in 2021. More than a quarter of the current e-cigarette users in high school reported daily e-cigarette use [3]. Although previous studies have shown that e-cigarettes might be less harmful than smoking due to the low level of toxic chemicals released during their usage, they can still harm adolescents' brains [4,5]. The nicotine usually contained in e-cigarettes is highly addictive and can affect the development of the parts of the brain that control attention, learning, mood, and impulse control, which can continue affecting up to the early age of 25 years [6,7]. In addition, nicotine use may increase the risk of adolescents' addictions to other drugs in the future [7]. A previous study has found that for people with chronic obstructive pulmonary disease or asthma [8], using e-cigarettes regularly can potentially accelerate their disease progression.

To protect public health in the general population, especially in youth, different tobacco policies have been announced to intervene and regulate e-cigarette products to prevent excessive e-cigarette use among teenagers. On December 20, 2019, a federal law was enacted to raise the federal minimum legal sales age from 18 years to 21 years for all tobacco products, including e-cigarettes, across the United States [9]. This policy makes it harder for teenagers to legally access e-cigarettes. One study has reported that adolescents and young adults described appealing flavors as their leading reason to use e-cigarette products [10]. Thus, several policies have been implemented to regulate e-cigarette flavors. On January 2, 2020, the Food and Drug Administration (FDA) announced the flavor enforcement policy, which was later implemented on February 6, 2020, to ban the sale of all unauthorized flavored cartridge-based e-cigarette products, except tobacco and menthol-flavored products [11]. Following the FDA flavor enforcement policy on May 18, 2020, the New York state banned all e-cigarette flavors, except the tobacco flavor [12].

Although many tobacco regulation policies are in place to reduce the vaping epidemic, especially among youth in recent years, the e-cigarette industry is trying to find a loophole to increase the sales of e-cigarette products through aggressive marketing on social media. Therefore, it is of the utmost importance to monitor the real-time dynamic changes in e-cigarette brands and flavors available in the market. With timely information on the prevalence of popular e-cigarette brands and flavors, regulatory authorities and policy makers could take immediate actions to help reduce the vaping epidemic, such as by issuing warning letters to unauthorized emerging e-cigarette products that are becoming popular. Twitter is one of the largest and most influential social media platforms, which had around 48.35

million active users in the United States in 2019 [13]. More importantly, Twitter allows e-cigarette users to share their experiences with e-cigarette products and discuss the government's e-cigarette-related policies and regulations [14]. In addition, vaping products have been actively promoted online, especially on social media by the vaping industry. With aggressive social media marketing, JUUL has successfully grown from a little-known brand in 2015 to the largest e-cigarette retailer in 2017, and similar social media marketing strategies have been used by many other e-cigarette brands [15]. It was estimated that in 2021, about 7 in 10 middle school and high school students were exposed to e-cigarette advertising [16]. Previous studies have leveraged social media data to study e-cigarette-related topics. For example, Sun et al [17] demonstrated that Twitter data could be used to study the promotions of various disposable e-cigarette flavors and related discussion topics. Lu et al [18] showed the feasibility of utilizing Reddit data to monitor the popularity of e-cigarette flavors based on a list of known e-cigarette flavors. A named-entity recognition model was proposed to identify e-cigarette brands and flavors from Instagram posts [19]. However, there is no effective approach to identify new e-cigarette brands or flavors in the market. With the dynamic changes in e-cigarette products, it is very important to identify and monitor new e-cigarette brands or flavors that are emerging and becoming popular in the market. The newly identified e-cigarette brands and flavors along with the popularity of different brands and flavors can be a useful asset for the government to make informed policy to reduce the vaping epidemic.

In this study, using Twitter data, we monitored the longitudinal trend of mention of e-cigarette brands and flavors in commercial and noncommercial tweets. More importantly, we identified new e-cigarette brands and flavors mentioned on Twitter. Therefore, our findings provide important insights into the dynamic e-cigarette landscape with new e-cigarette brands and flavors, which will help policymakers or tobacco regulators to enforce better regulations for public health protection.

Methods

Data Collection and Preprocessing

Through the Twitter streaming application programming interface, we collected e-cigarette-related Twitter posts (tweets) between May 3, 2021, and December 31, 2021, by using a set of e-cigarette-related keywords such as e-cigarette, e-cig, vaping, and e-liquid [18,20]. Duplicated tweets were removed from the data set. There were 1,932,707 e-cigarette-related tweets collected in total. By filtering with different keywords (such as e-cig, vaping, supply, and deal) from the names of the users who posted the tweets and promotion-related keywords (such as discount and free shipping) from the contents of the tweets, the tweets were characterized as commercial tweets (tweets that are from retailers or contain promotion information). The remaining tweets were characterized as noncommercial tweets (general discussions of e-cigarette-related topics). During this process, the language of the content of the tweets was evaluated, and only English tweets were kept. During the period between May 3, 2021, and December 31, 2021, there were

285,170 commercial tweets and 1,472,928 noncommercial tweets. From September 1, 2021, to December 31, 2021, there were 139,199 commercial tweets and 728,284 noncommercial tweets. Among all the tweets between May 3, 2021, and December 31, 2021, 890,209 tweets were retweets (tweets that were originally composed by other users and were reposted). These retweets were used for new brand and flavor identification but not for monitoring the popularity of the recorded brands and flavors.

Ethics Approval

This study has been reviewed and approved by the Office for Human Subject Protection Research Subjects Review Board at the University of Rochester (study ID: STUDY00006570).

New e-Cigarette Brand Identification

We collected and constructed an identified list of the available e-cigarette brands and flavors by searching web-based stores, which contained 129 e-liquid brands and 1198 e-liquid flavors [18]. To identify possible new e-cigarette brands mentioned on Twitter, we developed a depletion approach. For this, e-cigarette Twitter data from May 3, 2021, to August 31, 2021, were used as the reference data set and Twitter data from September 1, 2021, to December 31, 2021, were used as the target data set to identify new e-cigarette brands. As shown in Figure S1 in [Multimedia Appendix 1](#), a reference single-word list was constructed using the reference data set (e-cigarette Twitter data from May 3, 2021, to August 31, 2021) by tokenizing the words and recording the frequency of appearance of each of the unique word tokens. Two target single-word lists (commercial and noncommercial single-word lists) were constructed using the same approach from the target Twitter data set (Twitter data from September 1, 2021, to December 31, 2021). By removing the word tokens of the target single-word lists that appeared in the reference single-word list or the identified brand list, 2 depleted lists (commercial and noncommercial lists) were generated. The 2 depleted lists were then manually examined through Google search to check if each word token referred to an e-cigarette brand. As there were more than 50,000 keywords that were mentioned less than 10 times, only keywords mentioned at least 10 times were manually checked considering the time constraints. We considered the word token as the mentioning of a brand if the word was either a brand name or a product name within the brand. If the word token was only a part of the brand name and the word token was referring to the brand, we also considered the word as the mention of the brand. To verify newly identified brands, we conducted a Google search to ensure that these brands have been listed in e-cigarette-related websites or web-based stores.

New e-Cigarette Flavor Identification

As shown in Figure S2 in [Multimedia Appendix 1](#), new flavor identification was done differently from brand identification,

and the whole collected data set from May 3, 2021, to December 31, 2021, was used to identify new e-cigarette flavors. The type of tweets, either commercial or noncommercial, was tracked during the identification. All tweets containing the key phrase “new flavor” were directly saved as new flavor candidates. A flavor keyword list was then constructed by using unique single words that appeared in the identified list of our previously collected e-cigarette flavors [18]. Tweets were split into different segments by punctuation and the keyword “and” to deal with the possibility of multiple flavors mentioned in the same tweet. As flavor names were likely to contain certain common flavor keywords, segments that contained one or more flavor keywords were more likely to contain e-cigarette flavors. By removing segments that contained flavors that were already existing in our previous identified flavor list, the remaining tweet segments became the candidates for potential new flavors. We conducted a similar Google search to make sure that these newly identified flavors are existing e-cigarette flavors. The newly identified e-cigarette brands and flavors determined by the methods mentioned above were added to our e-cigarette brand or flavor list.

e-Cigarette Brand and Flavor Monitoring

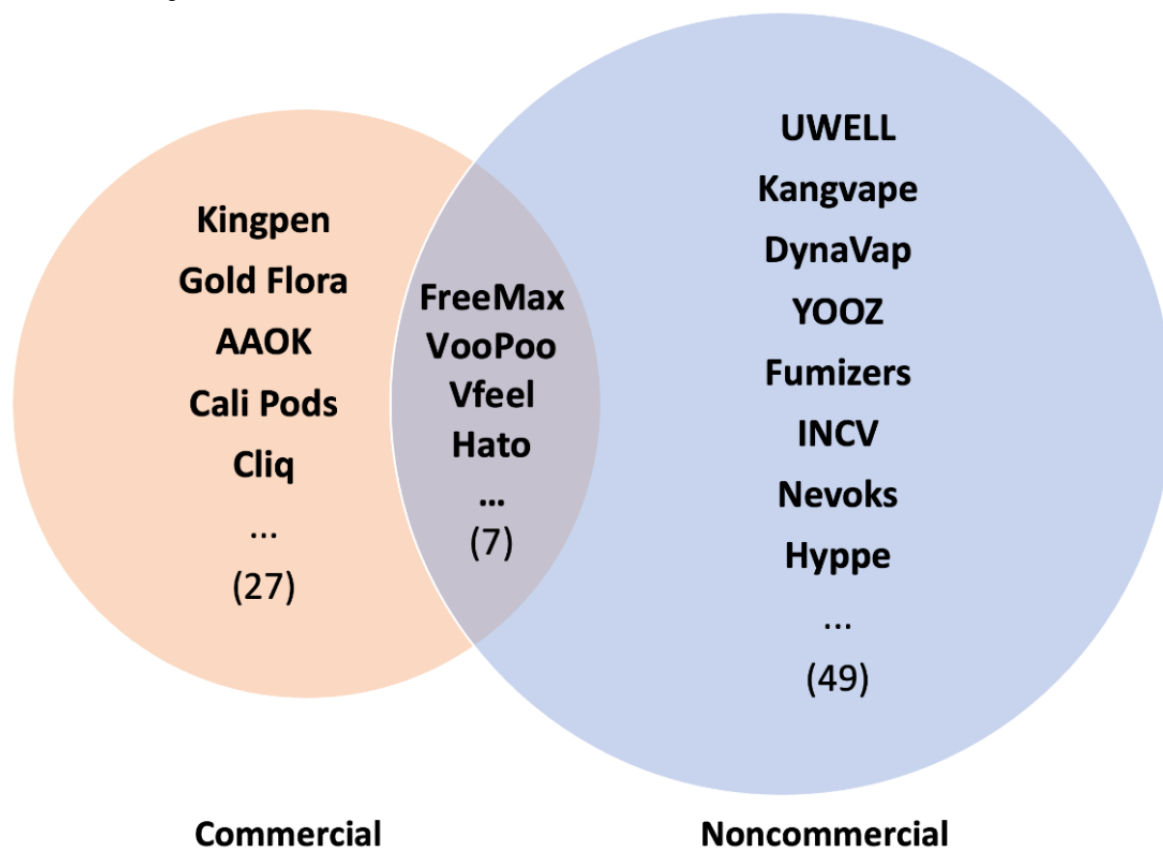
To determine if each e-cigarette-related tweet contains an e-cigarette brand or flavor, each tweet was split into segments based on punctuations. Then, we searched each e-cigarette brand name or flavor name in our constructed brand or flavor database that contained brands and flavors from both the original database and the newly identified ones from the target Twitter data set in each segment based on exact match (case insensitivity). We then calculated the number of tweets mentioning each e-cigarette brand and flavor in the commercial and noncommercial tweets each month.

Results

Identification of New e-Cigarette Brands on Twitter

Using the depletion methods, we identified 27,012 candidate words from commercial tweets and 62,117 candidate words from noncommercial tweets between September 1, 2021, and December 31, 2021, which were not present in tweets between May 3, 2021, and August 31, 2021. Among these candidate words, 831 commercial candidate words and 1966 noncommercial candidate words appeared at least 10 times. After manually searching those words with frequencies over 10 times online, we identified 34 new e-cigarette brands (such as Kingpen and Gold Flora) from the commercial tweets and 56 new e-cigarette brands (such as Uwell, Kangvape, and DynaVap) from the noncommercial tweets. Among the newly identified brands, 7 e-cigarette brands (ie, FreeMax, VooPoo, Vfeel, Hato, Innokin, Vapresso, and Vaptex) were identified from both the commercial and the noncommercial tweets ([Figure 1](#) and [Table S1](#) in [Multimedia Appendix 1](#)).

Figure 1. New electronic cigarette brands mentioned on Twitter.

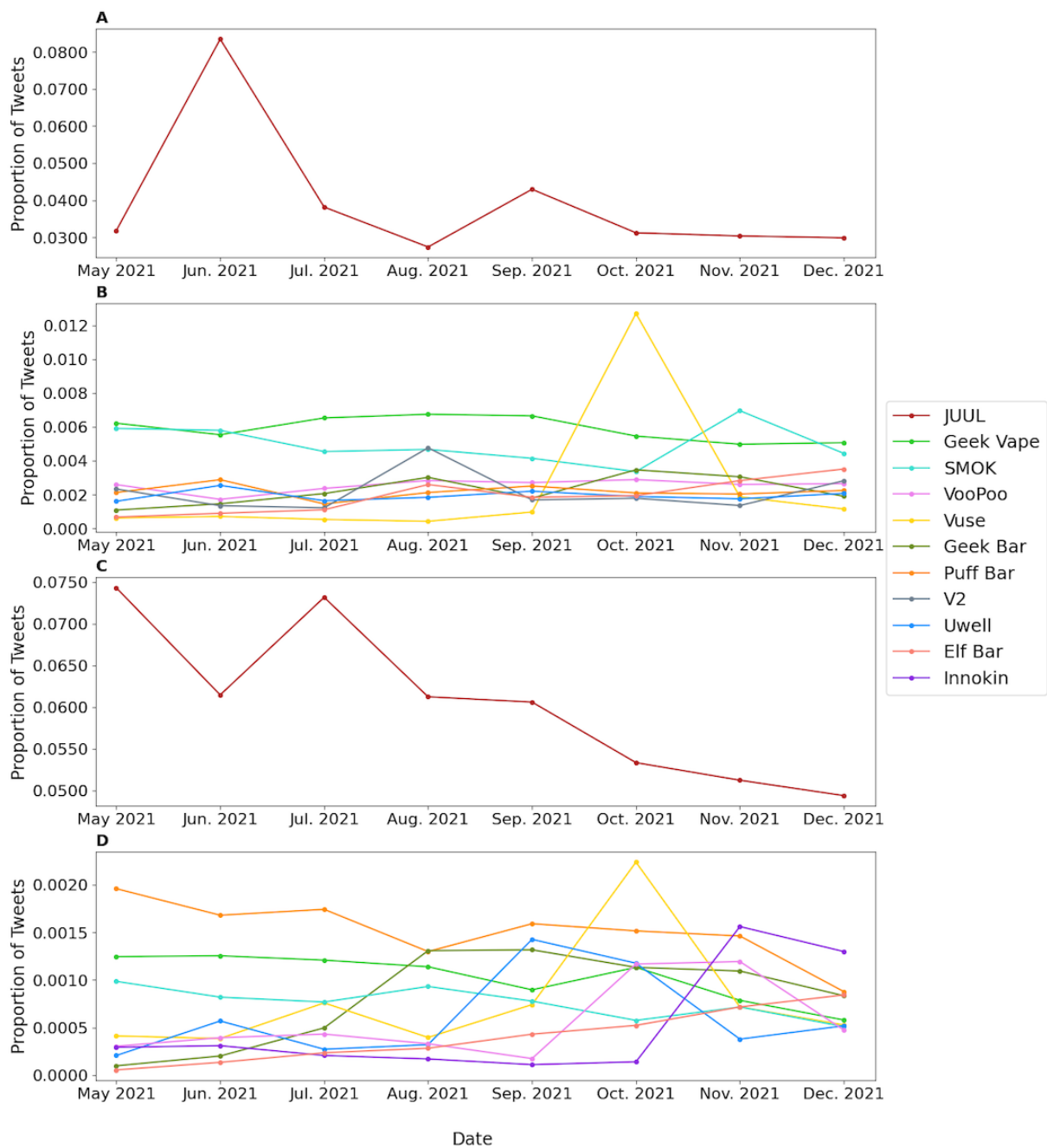


Longitudinal Monitoring of e-Cigarette Brands Mentioned on Twitter

To monitor the dynamic changes of each e-cigarette brand mentioned on Twitter, we calculated the proportion of tweets mentioning each e-cigarette brand over time in both commercial and noncommercial tweets. Due to the significant difference between the popularity of JUUL and other e-cigarette brands, we put JUUL in separate graphs to better show the trends of JUUL and other brands (Figure 2). As shown in Figure 2A and Figure 2C, the proportion of e-cigarette-related tweets mentioning JUUL was the highest among all e-cigarette brands over time in both commercial and noncommercial tweets. In

the commercial tweets, the monthly mentions of JUUL were steady between 3% and 5%, with an exceptionally high proportion of over 8% in June 2021 (Figure 2A). In the noncommercial tweets, the monthly mentions of JUUL showed a decreasing trend (Figure 2C). In both commercial and noncommercial tweets (Figure 2B and 2D), the monthly mentions of Vuse were roughly steady, with an exceptional peak around 1.27% (commercial) and 0.2% (noncommercial) in October 2021. Although it was not a top mentioned brand in commercial tweet, Innokin seemed to be a popular brand mentioned by e-cigarette users in noncommercial tweets, which was obviously popular in November and December 2021.

Figure 2. The longitudinal trends of electronic cigarette brands mentioned on Twitter. A: JUUL in commercial tweets. B: The other top 9 electronic cigarette brands in commercial tweets. C: JUUL in noncommercial tweets. D: The other top 9 electronic cigarette brands in noncommercial tweets.

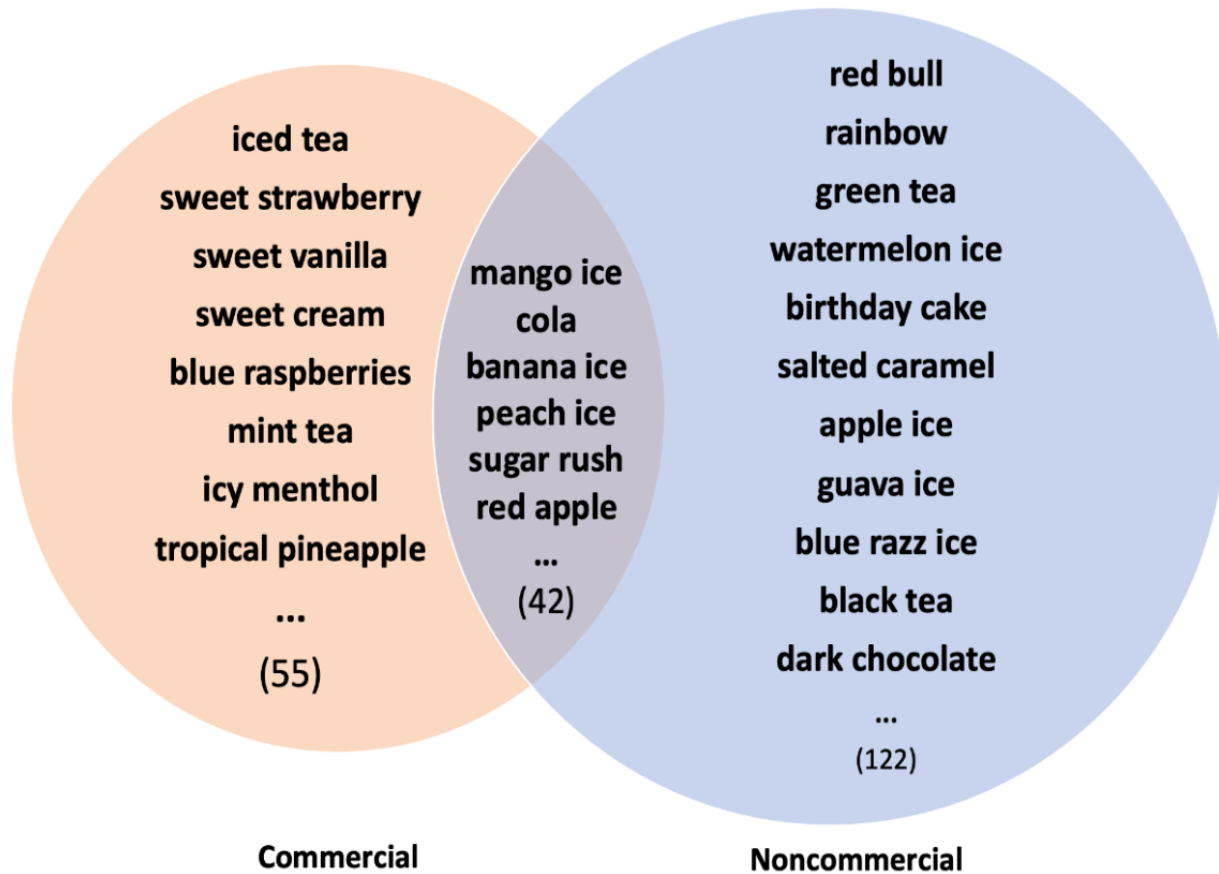


Identification of New e-Cigarette Flavors on Twitter

With the quick evolvement in the e-cigarette market due to different flavor regulations and marketing needs, it is important to identify new e-cigarette flavors appearing in the market. By comparing to our previously collected e-cigarette flavor lists, from our e-cigarette Twitter data set (May 3 to December 31, 2021), we identified 1145 flavor candidates from commercial tweets and 3736 candidates from noncommercial tweets. By

manually verifying all the flavor candidates, we identified 97 new e-cigarette flavors (such as iced tea, sweet strawberry, and sweet vanilla) from commercial tweets and 164 new flavors (such as red bull, rainbow, and green tea) from noncommercial tweets (Figure 3 and Table S2 in Multimedia Appendix 1). Among them, 42 new e-cigarette flavors (such as mango ice, cola, and banana ice) were identified from both commercial and noncommercial tweets (Figure 3 and Table S3 in Multimedia Appendix 1).

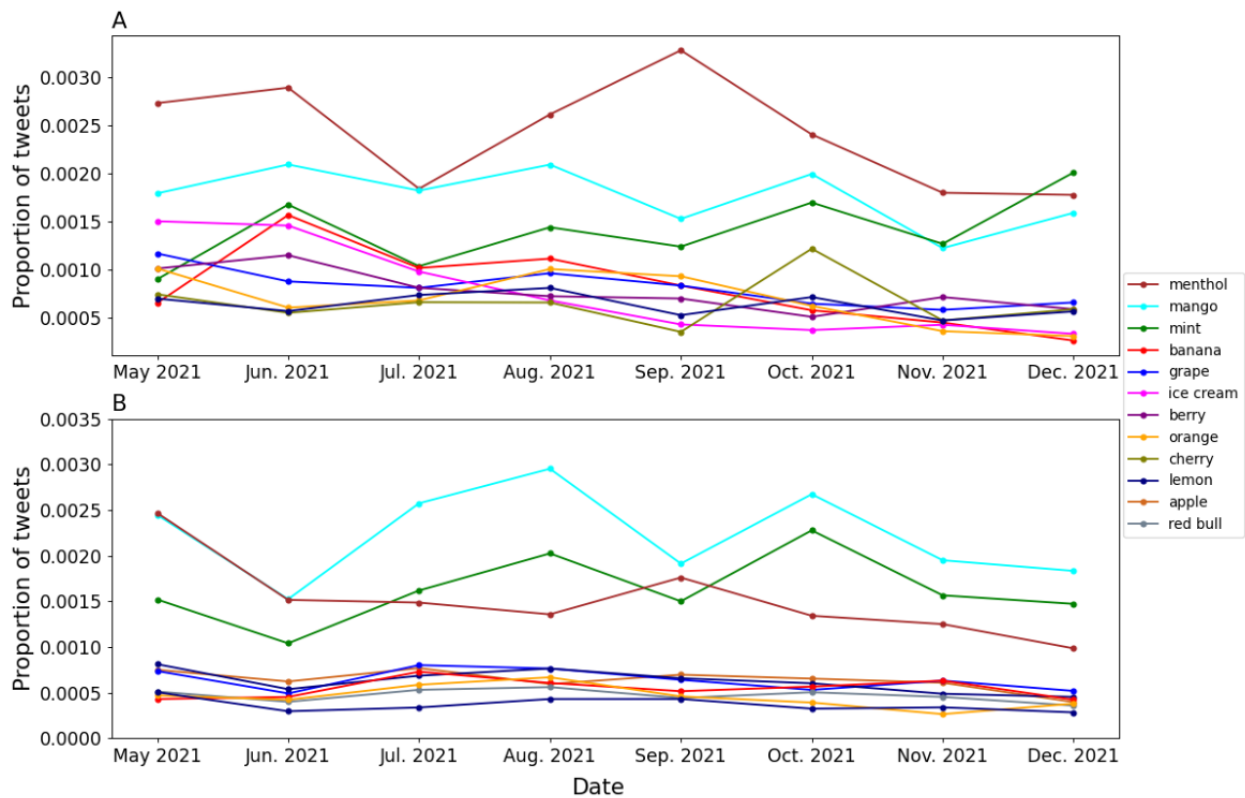
Figure 3. New electronic cigarette flavors mentioned on Twitter.



Longitudinal Monitoring of e-Cigarette Flavors Mentioned on Twitter

With different flavor regulations on e-cigarettes and the dynamic changes in the e-cigarette market, the mentions (or even the use) of different e-cigarette flavors on Twitter might be evolving. As shown in Figure 4, in both commercial and noncommercial tweets, mango, menthol, and mint were the most popular e-cigarette flavors. The trends in e-cigarette flavors were represented as the proportions of tweets mentioning each

flavor in the total number of e-cigarette-related tweets within each month. The top 10 most popular e-cigarette flavors mentioned in both commercial and noncommercial tweets generally had the same popularity with some variation from May 3, 2021, to December 31, 2021. In the commercial tweets, the menthol flavor was the most popular e-cigarette flavor, followed by mango and mint flavors (Figure 4A). In the noncommercial tweets, the mango flavor was the most popular e-cigarette flavor mentioned, followed by menthol and mint flavors (Figure 4B).

Figure 4. The longitudinal trend of electronic cigarette flavors mentioned on Twitter. A: Commercial tweets. B: Noncommercial tweets.

Discussion

Principal Findings

With the rapid change in the e-cigarette landscape, it is important to monitor not only the existing but also the newly emerging e-cigarette brands and flavors. In this study, using Twitter data, we developed a depletion method to identify newly emerging e-cigarette brands and a keyword identification method to identify newly emerging e-cigarette flavors based on Twitter data. Using the developed methods, we identified 76 new e-cigarette brands that were not included in our previously collected e-cigarette brands [18]. In addition, we identified 177 new e-cigarette flavors that were not present in our previously collected e-cigarette flavors. Furthermore, we monitored the trends of e-cigarette brands and flavors mentioned in both commercial and noncommercial tweets from May 3, 2021, to December 31, 2021, separately. In both commercial and noncommercial tweets, the top mentioned e-cigarette brands and flavors were similar. JUUL was the most prevalent brand, and mango, menthol, and mint were the most popular flavors mentioned in both commercial tweets and noncommercial tweets. With the dynamic changes in the e-cigarette brands in the market, it is necessary to identify not only the existing but also the newly emerging e-cigarette brands and flavors to monitor the changes in the e-cigarette market. However, as the brand names can be any word and phrase without any obvious common pattern, it is very difficult to identify potential brand names directly from tweets. As a result, instead of directly identifying potential brand names, our method tries to exclude all the word tokens, which were neither related to any e-cigarette brand nor related to e-cigarette brands that are already recorded, during the depletion process. All the word tokens left may

therefore be related with e-cigarette brands that are not in our recorded list. As the list of the general English words checked and the list of the known e-cigarette brands are being constantly updated and enlarged, our method will become more accurate and thus will require less human labor. Different from e-cigarette brand names, e-cigarette flavor names are more likely to contain certain keywords, especially when new flavors are more likely to be the mixture of single flavors that are mostly likely included in our previously collected flavor list.

Comparison With Prior Work

Previous studies have taken different approaches to monitor e-cigarette-related activities and entities on social media platforms. Several studies have developed named-entity recognition models to identify emerging brands and flavors on Instagram. For example, Chew et al [19] were able to identify brands with an F1-score of 0.815 and flavors with an F1-score of 0.828 with fine-tuned distilled bidirectional encoder representations from transformers network. Kavuluru and Sabbir [21], on the other hand, developed a proponent classification model to identify e-cigarette proponents (users who advocate e-cigarette publicly with their account) on Twitter. Another study developed a cluster model to identify the co-occurrence of different e-cigarette-related hashtags and the geographical information where e-cigarette-related discussions occurred [22]. However, all the known prior studies took the approach to either monitor only the known e-cigarette brands and flavors or to identify the related information (proponents, hashtags, locations, etc). To the best of our knowledge, no study has tried to identify newly emerging e-cigarette brands and flavors. Our research has introduced a novel approach to identify emerging e-cigarette brands and flavors in the market. More importantly, we created

a method that systematically monitors the popularity of different e-cigarette brands and flavors longitudinally. As Chew et al [19] suggested, new brands are entering the marketplace by first promoting on social media, and therefore, relying only on the traditional data source (such as Nielsen sales data) may not be enough. In this study, we developed some useful approaches to identify new e-cigarette brands and flavors mentioned on Twitter, which will help us better understand the dynamic changes in the e-cigarette market. Different from the classification methods or clustering models developed in a prior study [19], our methods do not include any machine learning approach and therefore do not have evaluation metrics (such as F1-score) that can be used to directly compare with prior work. Theoretically, our methods could also be easily adapted to identify new brands or flavors for other tobacco products (such as waterpipe and oral nicotine pouches) or other nontobacco products by using social media data. Although our methods require human involvement at the final step, human involvement is likely to significantly decrease as more Twitter data are analyzed and added into the reference Twitter data set.

The motivation for monitoring e-cigarette brands and flavors is to demonstrate the effect of government policy changes on people's perception of e-cigarettes. One direct way to detect the people's perception changes on e-cigarettes is by looking at the monthly popularity trends of e-cigarette brands and flavors. Many previous studies have focused on the public perceptions of government regulation on e-cigarettes. On January 2, 2020, the FDA released an e-cigarette enforcement policy, which prohibited the sale of all flavored cartridge-based e-cigarettes, except for menthol and tobacco flavors [11]. One study showed that the government's flavor ban might have changed the public perceptions of e-cigarettes by using Twitter data [23]. Another study concentrated public reaction on New York state e-cigarette flavor ban on September 17, 2019. Both studies indicated that after the policy was announced, people tended to have a more negative attitude toward e-cigarettes on Twitter [12]. Those studies demonstrated that public reactions to different governments' e-cigarette flavor policies can be detected using tweets. In this study, we showed that although JUUL is the most popular e-cigarette brand in terms of the number of tweets mentioned per month by a large margin, its number of mentions showed a general declining trend. By contrast, e-cigarette brands such as Elf Bar, Geek Bar, and Vuse appeared to have an increasing number of tweets mentioned per month. One reason to explain the downward trend is JUUL's limited flavor options for their products. JUUL has been criticized by the public for directly target advertising toward youth since 2015 [24]. Studies have shown that at its peak in 2018, 31% of the JUUL users were younger than 18 years [25]. Part of JUUL's popularity among teenagers comes from the fact that JUUL used to produce a wide variety of nontraditional-flavored e-cigarette products [24]. Nontraditional-flavored e-cigarettes are more appealing to teenagers compared to the traditional-flavored e-cigarette counterparts, as they provide a sensory perception of sweetness and coolness [26]. However, due to increasing public pressure against JUUL's appealing flavor products, which attract adolescents, in November 2018, JUUL voluntarily restricted their sales of mango, fruit medley, and cucumber pods on its

website [27]. In October 2019, JUUL lab announced that it would suspend sales of all nontobacco- and nonmenthol-based flavors in the United States [28]. In January 2020, FDA's flavor enforcement policy on all unauthorized cartridge-based flavored e-cigarettes made it impossible for JUUL to bring back their most loved fruit-flavored products like mango JUUL pods [11,29]. Studies have shown that ever since JUUL announced an end to the sale of their flavored e-cigarette products, their relative search volume on Google dropped steadily at a pace of 8.8 relative search volume per week. However, the relative search volume of disposable e-cigarette brands like Elf Bar, Puff Bar, and Geek Bar showed an increasing trend over time [30], which roughly corresponds to the trend shown in Figure 2B. These trends demonstrated that people are gradually replacing cartridge-based e-cigarette brands like JUUL with disposable e-cigarette brands like Puff Bar, Geek Bar, and Elf Bar to circumvent the FDA's flavor ban. Another way for e-cigarette companies to circumvent the FDA flavor regulations is to introduce the "concept flavor" (vague noncharacterizing descriptions on the packaging that do not expressly refer to real flavors) to rename their products [31]. For example, the popular disposable e-cigarette brand BIDI Stick unveiled a dozen new flavors, including Zest (formerly Jungle Juice), Arctic (formerly Mint freeze), and Solar (Berry Blast), among others. On June 23, 2022, the FDA issued market denial orders to JUUL Lab Inc for all their products currently marketed in the United States due to JUUL's lack of sufficient evidence regarding the toxicological profile of their products to demonstrate that they meet the public health standards [32].

Although JUUL was dominant on Twitter during our study period, Vuse is a top e-cigarette brand, and its monthly trends showed an exceptionally high peak in October 2021 in contrast to the stable or slight downward trend of other brands. One potential reason for this peak is the approval announcement of the FDA in October 2021. The FDA announced on October 12, 2021, that it would allow the selling of Vuse Solo e-cigarette in the United States, making Vuse the first authorized e-cigarette brand for sale in the United States [33]. In response to this announcement, it was very likely that people were discussing Vuse and the decision of FDA in the same month. In addition, e-cigarette retailers and other commercial Twitter accounts were posting promotion tweets about Vuse, leading to the exceptionally high frequencies of mention in both commercial and noncommercial tweets. Although Twitter data may not fully reflect the e-cigarette users and retailer groups, our results demonstrated that the effect of FDA's announcements and policies on e-cigarette brands and flavors on general users and retailers can be monitored through monitoring the existing and newly emerging e-cigarette brands and flavors on Twitter.

Our data showed that almost all the top 10 e-cigarette flavors, including menthol, grape, orange, and banana, experienced a general downward trend in terms of number of mentions in commercial tweets. This indicates that flavored e-cigarettes have been promoted lesser on Twitter in response to the FDA e-cigarette flavor regulation policies [34]. FDA has restricted the Premarket Tobacco Product Applications approval since February 2020, which greatly reduced the number of new flavored e-cigarette products entering the e-cigarette market. In

total, the FDA agencies have issued 263 marketing denial orders to more than 1 million flavored e-cigarette products [35]. However, despite the FDA e-cigarette flavor regulation enforced in February 2020, there are still flavored e-cigarette product promotions on Twitter. In an effort to circumvent the FDA e-cigarette flavor regulation, 6 big manufacturers of 98 different brands, including Buff Bar and Geek Bar, claimed that their products contain synthetic nicotine [36]. FDA did not enforce any flavor regulation policies on e-cigarette products containing synthetic nicotine. Thus, it is reasonable for vaping companies or stores to promote their flavored e-cigarette products containing synthetic nicotine on Twitter. Another way to avoid FDA e-cigarette regulations is to sell disposable e-cigarettes. Since the FDA flavor regulations did not include disposable e-cigarette devices [11], vaping companies or vaping stores can still promote their flavored disposable e-cigarette on Twitter. The Twitter data used in this study includes disposable e-cigarettes. Thus, e-cigarette flavors mentioned in commercial tweets may come from disposable e-cigarette product promotions.

Limitations

Although the Twitter data analyzed in this study have a large sample size, the trends in e-cigarette brands and flavors on Twitter may not reflect the trends in the real world, as Twitter users cannot represent all the population around the globe. The e-cigarette brands and flavors that have never been mentioned on Twitter will not be captured. For both methods identifying new e-cigarette brands and flavors proposed in this study, there are several limitations. First, both methods require sufficient prior knowledge. New e-cigarette brand identification requires enough tweets being labeled in order to be used in the depletion method to reduce the noise in the resulting candidate list. In this

study, the first half of the e-cigarette-related Twitter data set is assumed to be the reference data and has not been manually checked. Unidentified e-cigarette brands in the reference data set may not be captured if these brands are never mentioned in the target data set. As many general English words are removed through depletion, names of e-cigarette brands that are formed by general English words may not be captured. Flavor identification requires enough known flavors to capture enough flavor keywords. For any new flavor wherein the name does not have any overlap with the names in our previously recorded list, if they are mentioned in tweets without the keyword “new flavor,” they will not be captured. Second, both methods generate long candidate lists and require a large amount of human labor, which need to be further improved with the help of deep learning techniques. Although there is an increasing popularity of “concept flavor” for e-cigarettes to avoid the flavor regulatory policies, we do not capture these flavors as they do not fall into the standard flavor names, which need to be further examined in future studies.

Conclusion

In this paper, we proposed a new method to identify new e-cigarette brands and flavors. This study is one of the first attempts to use Twitter data to discover new e-cigarette brands and flavors. We have successfully identified 62 new e-cigarette brands and 305 new e-cigarette flavors mentioned on Twitter. More importantly, we were able to monitor the dynamic changes in e-cigarette brands and flavors mentioned on Twitter over time (both commercial and noncommercial). Our results demonstrate that Twitter data can be used to discover new e-cigarette brands and flavors as well as provide real-time surveillance on e-cigarette brands and flavors in the market.

Acknowledgments

This study was supported by the Western New York-based Center for Research on Flavored Tobacco Products under cooperative agreement U54CA228110 from the National Cancer Institute of the National Institutes of Health and the US Food and Drug Administration. This study was also supported by the National Cancer Institute of the National Institutes of Health (2P01CA200512).

Data Availability

The data sets generated during or analyzed in this study are available from the corresponding author on reasonable request.

Authors' Contributions

ZX and DL conceived and designed the study. QT, RZ, and ZX analyzed the data. QT, RZ, ZX, and DL wrote the manuscript. All authors edited and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables and figures.

[[DOCX File, 59 KB - formative_v6i12e42241_app1.docx](#)]

References

1. McRobbie H, Phillips A, Goniewicz ML, Smith KM, Knight-West O, Przulj D, et al. Effects of Switching to Electronic Cigarettes with and without Concurrent Smoking on Exposure to Nicotine, Carbon Monoxide, and Acrolein. *Cancer Prev Res (Phila)* 2015 Sep;8(9):873-878. [doi: [10.1158/1940-6207.CAPR-15-0058](https://doi.org/10.1158/1940-6207.CAPR-15-0058)] [Medline: [26333731](https://pubmed.ncbi.nlm.nih.gov/26333731/)]
2. Oriakhi M. Vaping: An Emerging Health Hazard. *Cureus* 2020 Mar 26;12(3):e7421 [FREE Full text] [doi: [10.7759/cureus.7421](https://doi.org/10.7759/cureus.7421)] [Medline: [32351806](https://pubmed.ncbi.nlm.nih.gov/32351806/)]
3. Park-Lee E, Ren C, Sawdey MD, Gentzke AS, Cornelius M, Jamal A, et al. Notes from the Field: E-Cigarette Use Among Middle and High School Students - National Youth Tobacco Survey, United States, 2021. *MMWR Morb Mortal Wkly Rep* 2021 Oct 01;70(39):1387-1389 [FREE Full text] [doi: [10.15585/mmwr.mm7039a4](https://doi.org/10.15585/mmwr.mm7039a4)] [Medline: [34591834](https://pubmed.ncbi.nlm.nih.gov/34591834/)]
4. Farsalinos K, Romagna G, Tsiapras D, Kyrzopoulos S, Voudris V. Characteristics, perceived side effects and benefits of electronic cigarette use: a worldwide survey of more than 19,000 consumers. *Int J Environ Res Public Health* 2014 Apr 22;11(4):4356-4373 [FREE Full text] [doi: [10.3390/ijerph110404356](https://doi.org/10.3390/ijerph110404356)] [Medline: [24758891](https://pubmed.ncbi.nlm.nih.gov/24758891/)]
5. Dwyer JB, Broide RS, Leslie FM. Nicotine and brain development. *Birth Defects Res C Embryo Today* 2008 Mar;84(1):30-44. [doi: [10.1002/bdrc.20118](https://doi.org/10.1002/bdrc.20118)] [Medline: [18383130](https://pubmed.ncbi.nlm.nih.gov/18383130/)]
6. Marynak KL, Gammon DG, Rogers T, Coats EM, Singh T, King BA. Sales of Nicotine-Containing Electronic Cigarette Products: United States, 2015. *Am J Public Health* 2017 May;107(5):702-705. [doi: [10.2105/ajph.2017.303660](https://doi.org/10.2105/ajph.2017.303660)]
7. Taylor G, McNeill A, Girling A, Farley A, Lindson-Hawley N, Aveyard P. Change in mental health after smoking cessation: systematic review and meta-analysis. *BMJ* 2014 Feb 13;348:g1151 [FREE Full text] [doi: [10.1136/bmj.g1151](https://doi.org/10.1136/bmj.g1151)] [Medline: [24524926](https://pubmed.ncbi.nlm.nih.gov/24524926/)]
8. Bozier J, Rutting S, Xenaki D, Peters M, Adcock I, Oliver BG. Heightened response to e-cigarettes in COPD. *ERJ Open Res* 2019 Feb;5:1 [FREE Full text] [doi: [10.1183/23120541.00192-2018](https://doi.org/10.1183/23120541.00192-2018)] [Medline: [30820437](https://pubmed.ncbi.nlm.nih.gov/30820437/)]
9. Newly signed legislation raises federal minimum age of sale of tobacco products to 21. US Food and Drug Administration. URL: <https://www.fda.gov/tobacco-products/ctp-newsroom/newly-signed-legislation-raises-federal-minimum-age-sale-tobacco-products-21> [accessed 2022-11-12]
10. Soneji SS, Knutzen KE, Villanti AC. Use of Flavored E-Cigarettes Among Adolescents, Young Adults, and Older Adults: Findings From the Population Assessment for Tobacco and Health Study. *Public Health Rep* 2019 Mar 12;134(3):282-292. [doi: [10.1177/0033354919830967](https://doi.org/10.1177/0033354919830967)]
11. Yingst JM, Bordner CR, Hobkirk AL, Hoglen B, Allen SI, Krebs NM, et al. Response to Flavored Cartridge/Pod-Based Product Ban among Adult JUUL Users: "You Get Nicotine However You Can Get It". *Int J Environ Res Public Health* 2020 Dec 30;18(1):207 [FREE Full text] [doi: [10.3390/ijerph18010207](https://doi.org/10.3390/ijerph18010207)] [Medline: [33396598](https://pubmed.ncbi.nlm.nih.gov/33396598/)]
12. Sun L, Lu X, Xie Z, Li D. Public Reactions to the New York State Policy on Flavored Electronic Cigarettes on Twitter: Observational Study. *JMIR Public Health Surveill* 2022 Feb 03;8(2):e25216 [FREE Full text] [doi: [10.2196/25216](https://doi.org/10.2196/25216)] [Medline: [35113035](https://pubmed.ncbi.nlm.nih.gov/35113035/)]
13. Salman A. Twitter by the Numbers: Stats, Demographics & Fun Facts. Omnicore. URL: <https://www.omnicoreagency.com/twitter-statistics/> [accessed 2022-11-12]
14. McCausland K, Maycock B, Leaver T, Wolf K, Freeman B, Jancey J. E-Cigarette Advocates on Twitter: Content Analysis of Vaping-Related Tweets. *JMIR Public Health Surveill* 2020 Oct 14;6(4):e17543 [FREE Full text] [doi: [10.2196/17543](https://doi.org/10.2196/17543)] [Medline: [33052130](https://pubmed.ncbi.nlm.nih.gov/33052130/)]
15. Huang J, Duan Z, Kwok J, Binns S, Vera LE, Kim Y, et al. Vaping versus JUULing: how the extraordinary growth and marketing of JUUL transformed the US retail e-cigarette market. *Tob Control* 2019 Mar;28(2):146-151 [FREE Full text] [doi: [10.1136/tobaccocontrol-2018-054382](https://doi.org/10.1136/tobaccocontrol-2018-054382)] [Medline: [29853561](https://pubmed.ncbi.nlm.nih.gov/29853561/)]
16. Lechner WV, Janssen T, Kahler CW, Audrain-McGovern J, Leventhal AM. Bi-directional associations of electronic and combustible cigarette use onset patterns with depressive symptoms in adolescents. *Prev Med* 2017 Mar;96:73-78 [FREE Full text] [doi: [10.1016/j.ypmed.2016.12.034](https://doi.org/10.1016/j.ypmed.2016.12.034)] [Medline: [28024859](https://pubmed.ncbi.nlm.nih.gov/28024859/)]
17. Sun L, Tao C, Xie Z, Li D. Promotion of Disposable Electronic Cigarette Flavors and Topics on Twitter. *Int J Environ Res Public Health* 2020 Dec 10;17(24):9221 [FREE Full text] [doi: [10.3390/ijerph17249221](https://doi.org/10.3390/ijerph17249221)] [Medline: [33321714](https://pubmed.ncbi.nlm.nih.gov/33321714/)]
18. Lu X, Chen L, Yuan J, Luo J, Luo J, Xie Z, et al. User Perceptions of Different Electronic Cigarette Flavors on Social Media: Observational Study. *J Med Internet Res* 2020 Jun 24;22(6):e17280 [FREE Full text] [doi: [10.2196/17280](https://doi.org/10.2196/17280)] [Medline: [32579123](https://pubmed.ncbi.nlm.nih.gov/32579123/)]
19. Chew R, Wenger M, Guillory J, Nonnemaker J, Kim A. Identifying Electronic Nicotine Delivery System Brands and Flavors on Instagram: Natural Language Processing Analysis. *J Med Internet Res* 2022 Jan 18;24(1):e30257 [FREE Full text] [doi: [10.2196/30257](https://doi.org/10.2196/30257)] [Medline: [35040793](https://pubmed.ncbi.nlm.nih.gov/35040793/)]
20. Chen L, Lu X, Yuan J, Luo J, Luo J, Xie Z, et al. A Social Media Study on the Associations of Flavored Electronic Cigarettes With Health Symptoms: Observational Study. *J Med Internet Res* 2020 Jun 22;22(6):e17496 [FREE Full text] [doi: [10.2196/17496](https://doi.org/10.2196/17496)] [Medline: [32568093](https://pubmed.ncbi.nlm.nih.gov/32568093/)]
21. Kavuluru R, Sabbir AKM. Toward automated e-cigarette surveillance: Spotting e-cigarette proponents on Twitter. *J Biomed Inform* 2016 Jun;61:19-26 [FREE Full text] [doi: [10.1016/j.jbi.2016.03.006](https://doi.org/10.1016/j.jbi.2016.03.006)] [Medline: [26975599](https://pubmed.ncbi.nlm.nih.gov/26975599/)]
22. Allem J, Ferrara E, Uppu SP, Cruz TB, Unger JB. E-Cigarette Surveillance With Social Media Data: Social Bots, Emerging Topics, and Trends. *JMIR Public Health Surveill* 2017 Dec 20;3(4):e98. [doi: [10.2196/publichealth.8641](https://doi.org/10.2196/publichealth.8641)] [Medline: [29263018](https://pubmed.ncbi.nlm.nih.gov/29263018/)]

23. Lu X, Sun L, Xie Z, Li D. Perception of the Food and Drug Administration Electronic Cigarette Flavor Enforcement Policy on Twitter: Observational Study. *JMIR Public Health Surveill* 2022 Mar 29;8(3):e25697 [FREE Full text] [doi: [10.2196/25697](https://doi.org/10.2196/25697)] [Medline: [35348461](https://pubmed.ncbi.nlm.nih.gov/35348461/)]
24. Fadus MC, Smith TT, Squeglia LM. The rise of e-cigarettes, pod mod devices, and JUUL among youth: Factors influencing use, health implications, and downstream effects. *Drug Alcohol Depend* 2019 Aug 01;201:85-93 [FREE Full text] [doi: [10.1016/j.drugalcdep.2019.04.011](https://doi.org/10.1016/j.drugalcdep.2019.04.011)] [Medline: [31200279](https://pubmed.ncbi.nlm.nih.gov/31200279/)]
25. Kaplan B, Cherukupalli R, Welding K, Kennedy R, Cohen J. The youth e-cigarette epidemic: New estimates of JUUL Labs' revenue from youth users in the US. *Tob Induc Dis* 2021;19:33 [FREE Full text] [doi: [10.18332/tid/133874](https://doi.org/10.18332/tid/133874)] [Medline: [33935613](https://pubmed.ncbi.nlm.nih.gov/33935613/)]
26. Goldenson NI, Leventhal AM, Simpson KA, Barrington-Trimis JL. A Review of the Use and Appeal of Flavored Electronic Cigarettes. *Curr Addict Rep* 2019 Jun;6(2):98-113 [FREE Full text] [doi: [10.1007/s40429-019-00244-4](https://doi.org/10.1007/s40429-019-00244-4)] [Medline: [31453046](https://pubmed.ncbi.nlm.nih.gov/31453046/)]
27. Morean ME, Bold KW, Kong G, Camenga DR, Jackson A, Simon P, et al. High school students' use of JUUL pod flavors before and after JUUL implemented voluntary sales restrictions on certain flavors in 2018. *PLoS One* 2020;15(12):e0243368 [FREE Full text] [doi: [10.1371/journal.pone.0243368](https://doi.org/10.1371/journal.pone.0243368)] [Medline: [33320885](https://pubmed.ncbi.nlm.nih.gov/33320885/)]
28. Juul labs suspends sale of non-tobacco, non-menthol-based flavors in the US. Juul Labs. URL: <https://www.juulabs.com/juul-labs-suspends-sale-of-non-tobacco-non-menthol-based-flavors-in-the-u-s/> [accessed 2022-11-12]
29. Groom AL, Vu TT, Kesh A, Hart JL, Walker KL, Giachello AL, et al. Correlates of youth vaping flavor preferences. *Prev Med Rep* 2020 Jun;18:101094 [FREE Full text] [doi: [10.1016/j.pmedr.2020.101094](https://doi.org/10.1016/j.pmedr.2020.101094)] [Medline: [32373447](https://pubmed.ncbi.nlm.nih.gov/32373447/)]
30. Dai H, Hao J. Online popularity of JUUL and Puff Bars in the USA: 2019-2020. *Tob Control* 2022 Jan;31(1):7-10. [doi: [10.1136/tobaccocontrol-2020-055727](https://doi.org/10.1136/tobaccocontrol-2020-055727)] [Medline: [33051277](https://pubmed.ncbi.nlm.nih.gov/33051277/)]
31. Laestadius L, Vassey J, Kim M, Ozga J, Li D, Stanton C, et al. Themes in e-liquid concept names as a marketing tactic: evidence from Premarket Tobacco Product Applications in the USA. *Tob Control* 2022 Sep 28:2022. [doi: [10.1136/tc-2022-057657](https://doi.org/10.1136/tc-2022-057657)] [Medline: [36171148](https://pubmed.ncbi.nlm.nih.gov/36171148/)]
32. FDA denies authorization to market Juul products. US Food and Drug Administration. URL: <https://www.fda.gov/news-events/press-announcements/fda-denies-authorization-market-juul-products> [accessed 2022-11-12]
33. Berk CC, Lucas A. FDA approves Vuse vaping products for sale, marking a first for the e-cigarette industry. *CNBC News*. URL: <https://www.cnn.com/2021/10/12/fda-approves-vuse-vaping-products-for-sale-marking-a-first-for-the-e-cigarette-industry-htm> [accessed 2022-11-12]
34. FDA finalizes enforcement policy on unauthorized flavored cartridge-based e-cigarettes that appeal to children, including fruit and mint. US Food and Drug Administration. URL: <https://www.fda.gov/news-events/press-announcements/fda-finalizes-enforcement-policy-unauthorized-flavored-cartridge-based-e-cigarettes-appeal-children> [accessed 2022-11-12]
35. Tobacco products marketing orders. US Food and Drug Administration. URL: <https://www.fda.gov/tobacco-products/market-and-distribute-tobacco-product/tobacco-products-marketing-orders> [accessed 2022-11-12]
36. Ramamurthi D, Chau C, Lu Z, Rughoobur I, Sanaie, K, Krishna P, et al. Marketing of "tobacco-free" and "synthetic nicotine" products. *Stanford Research*. URL: <https://tobacco-img.stanford.edu/wp-content/uploads/2022/03/13161808/Synthetic-Nicotine-White-Paper-3-8-2022F.pdf> [accessed 2022-11-12]

Abbreviations

e-cigarette: electronic cigarette

FDA: Food and Drug Administration

Edited by A Mavragani; submitted 28.08.22; peer-reviewed by L Laestadius, X Lu; comments to author 06.10.22; revised version received 24.10.22; accepted 14.11.22; published 05.12.22.

Please cite as:

Tang Q, Zhou R, Xie Z, Li D

Monitoring and Identifying Emerging e-Cigarette Brands and Flavors on Twitter: Observational Study

JMIR Form Res 2022;6(12):e42241

URL: <https://formative.jmir.org/2022/12/e42241>

doi: [10.2196/42241](https://doi.org/10.2196/42241)

PMID: [36469415](https://pubmed.ncbi.nlm.nih.gov/36469415/)

©Qihang Tang, Runtao Zhou, Zidian Xie, Dongmei Li. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 05.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete

bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Sample Bias in Web-Based Patient-Generated Health Data of Dutch Patients With Gastrointestinal Stromal Tumor: Survey Study

Anne Dirkson¹, MSc; Dide den Hollander^{2,3,4}, MD, PhD; Suzan Verberne¹, PhD; Ingrid Desar⁴, MD, PhD; Olga Husson^{2,3,5}, MSc, PhD; Winette T A van der Graaf^{2,6}, MD, PhD; Astrid Oosten⁶, MD, PhD; Anna K L Reyners⁷, MD, PhD; Neeltje Steeghs², MD, PhD; Wouter van Loon⁸, MSc; Gerard van Oortmerssen^{1,9}, PhD; Hans Gelderblom¹⁰, Prof Dr; Wessel Kraaij^{1,11}, Prof Dr Ir

¹Leiden Institute of Advanced Computer Science, Leiden University, Leiden, Netherlands

²Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, Netherlands

³Department of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, Netherlands

⁴Department of Medical Oncology, Radboud University Medical Center, Nijmegen, Netherlands

⁵Department of Surgical Oncology, Erasmus Medical Center, Rotterdam, Netherlands

⁶Department of Medical Oncology, Erasmus Medical Center, Rotterdam, Netherlands

⁷Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

⁸Department of Methodology and Statistics, Leiden University, Leiden, Netherlands

⁹Sarcoma Patient Advocacy Global Network, Wölfersheim, Germany

¹⁰Department of Medical Oncology, Leiden University Medical Center, Leiden, Netherlands

¹¹The Netherlands Organisation for Applied Scientific Research, Den Haag, Netherlands

Corresponding Author:

Suzan Verberne, PhD

Leiden Institute of Advanced Computer Science

Leiden University

Niels Bohrweg 1

Leiden, 2333 CA

Netherlands

Phone: 31 71 527 7096

Email: s.verberne@liacs.leidenuniv.nl

Abstract

Background: Increasingly, social media is being recognized as a potential resource for patient-generated health data, for example, for pharmacovigilance. Although the representativeness of the web-based patient population is often noted as a concern, studies in this field are limited.

Objective: This study aimed to investigate the sample bias of patient-centered social media in Dutch patients with gastrointestinal stromal tumor (GIST).

Methods: A population-based survey was conducted in the Netherlands among 328 patients with GIST diagnosed 2-13 years ago to investigate their digital communication use with fellow patients. A logistic regression analysis was used to analyze clinical and demographic differences between forum users and nonusers.

Results: Overall, 17.9% (59/328) of survey respondents reported having contact with fellow patients via social media. Moreover, 78% (46/59) of forum users made use of GIST patient forums. We found no statistically significant differences for age, sex, socioeconomic status, and time since diagnosis between forum users (n=46) and nonusers (n=273). Patient forum users did differ significantly in (self-reported) treatment phase from nonusers ($P=.001$). Of the 46 forum users, only 2 (4%) were cured and not being monitored; 3 (7%) were on adjuvant, curative treatment; 19 (41%) were being monitored after adjuvant treatment; and 22 (48%) were on palliative treatment. In contrast, of the 273 patients who did not use disease-specific forums to communicate with fellow patients, 56 (20.5%) were cured and not being monitored, 31 (11.3%) were on curative treatment, 139 (50.9%) were being monitored after treatment, and 42 (15.3%) were on palliative treatment. The odds of being on a patient forum were 2.8 times as high for a patient who is being monitored compared with a patient that is considered cured. The odds of being on a patient forum were 1.9 times as high for patients who were on curative (adjuvant) treatment and 10 times as high for patients who were in the

palliative phase compared with patients who were considered cured. Forum users also reported a lower level of social functioning (84.8 out of 100) than nonusers (93.8 out of 100; $P=.008$).

Conclusions: Forum users showed no particular bias on the most important demographic variables of age, sex, socioeconomic status, and time since diagnosis. This may reflect the narrowing digital divide. Overrepresentation and underrepresentation of patients with GIST in different treatment phases on social media should be taken into account when sourcing patient forums for patient-generated health data. A further investigation of the sample bias in other web-based patient populations is warranted.

(*JMIR Form Res* 2022;6(12):e36755) doi:[10.2196/36755](https://doi.org/10.2196/36755)

KEYWORDS

social media; patient forum; sample bias; representativeness; pharmacovigilance; rare cancer

Introduction

Background

Web-based patient forums provide patients with both emotional and informational support [1]. In recent years, social media (defined as a web-based communication channel where information and messages are exchanged) has also been investigated as a potential complementary information source for patient-generated health data, for example, for pharmacovigilance [2-6]. The main advantage of social media is that it offers uncensored information [7] in large quantities [8]. Moreover, patients are more likely to share information with fellow patients than with their physicians [9]. Thus, social media may contain information that is not collected in clinical trials or reported in spontaneous reporting systems.

Postmarket surveillance is necessary as clinical trials are of limited duration and suffer from sample bias; they often exclude older patients, patients with comorbidities, and pregnant women [10,11]. Current postmarket medication surveillance systems rely mostly on spontaneous reports of adverse events, medical literature, and observational databases. Most of these spontaneous reports are made by health professionals. In fact, in the Dutch surveillance system Lareb, only 26.3% of all reports between 2010 and 2015 were made by patients [12].

Reliance on spontaneous reports alone results in a severe underreporting of adverse drug responses (ADRs) [13]. According to the work by Lopez-Gonzalez et al [14], underreporting is associated with reporting of severe ADRs only, fear of ridicule for reporting suspected ADRs, lethargy, and indifference and complacency by professionals (ie, the idea that only safe drugs are allowed onto the market). Although previous work has shown that the ADRs reported on social media are often less serious than those reported via official channels, they do affect the quality of life of the patient [6]. In fact, social media would be able to provide a more patient-centric view of which ADRs are most salient to patients on a day-to-day basis [15].

However, researchers as well as patients have expressed concern about sample bias on social media [6,16-22]. Previous research on social media use in general shows that young people, women, and people of a higher socioeconomic class are generally highly represented [23-26]. Although there has been some work that shows that these differences persist over time [26,27], other work indicates that some factors such as age are becoming less influential as the overall adoption of social media is growing.

According to a recent report of the Pew Research Centre, 72% of all Americans were using social media in 2021 including 45% of adults aged >65 years [28].

On the basis of studies of the general population of social media users [23-27], it appears that those demographic groups that consume more medication (ie, older patients, people of low socioeconomic status, and patients with chronic conditions) are generally not highly represented on social media platforms [14]. However, it remains unclear whether these findings generalize to the specific case of web-based patient-to-patient communication.

Although there is a large literature base on patient communication forums and the extraction of adverse drug effects, to date, the work on sample bias in web-based patient-to-patient communication is limited to 2 studies. Prior work on American patients with breast cancer [29,30] using action logs of forum activity in an artificial setting has shown that users are relatively more likely to be Caucasian than African American. No other significant demographic differences were found between users and nonusers. A more comprehensive overview of the literature on patient communication forums for patients with gastrointestinal stromal tumor (GIST) on broader topics than bias can be found in the recent work of den Hollander et al [31] and our own prior work [32].

Other studies addressed another bias that is relevant when mining social media for patient-generated health data: so-called activity bias [33] or the fact that only some users actively post messages. In this paper, we will use the term *passive users* for forum users that do not post messages and *active users* for forum users that do post messages. Passive users are also commonly referred to as *lurkers* in previous research. Among patients with breast cancer, Han et al [29] found that active users were more likely to be younger, Caucasian, living alone, and have a greater information need than passive users. Another study [34] specifically compared passive to active community members for breast cancer, arthritis, and fibromyalgia and corroborated that posters are younger on average. They also found that active users had a longer disease history and a higher self-reported mental well-being than passive users. In this paper, we do not compare active and passive users because of the small sample size.

As Baeza-Yates [33] noted, “any remedy of bias starts with awareness of its existence.” Thus, to provide a starting point for mitigating bias for the use of patient-generated health data from social media in the future, we conducted a survey to

investigate sample bias in social media use among patients with GIST in the Netherlands relative to the survey sample. GIST is a rare form of cancer, which often has a long palliative care trajectory in which patients are treated with chronic, oral medication (tyrosine kinase inhibitors [TKIs]) for many years. If caught early, GIST can be cured. Treatment with TKIs can improve survival for patients with GIST both in adjuvant and palliative setting but often also lead to adverse drug events [31]. Patient reports from social media may be especially valuable for rare disorders where patients are sparse and spread out geographically.

Objectives

In this study, we investigated (1) what proportion of patients have contact with fellow patients on social media, (2) why patients abstain from engaging with web-based patient communities, and (3) to what extent there are significant demographic and clinical differences between those that use social media to converse with patients and those that do not. This study did not assess general social media use but focused specifically on the web-based communication with other patients. We defined social media as a web-based communication channel where information and messages are exchanged. When referring to *web-based patient communities*, we mean web-based groups on social media where the main purpose of the group is for (certain) patients (eg, patients with breast cancer) to communicate with one another. We use the term web-based patient communities and patient forums interchangeably.

On the basis of general social media, we hypothesized that forum users will differ in demographic factors including age, sex, and socioeconomic status from nonusers. We also hypothesized that forum users will differ in marital status and have a lower level of social functioning than nonusers, in line with the social compensation model [35] (ie, those who have less real-life [offline] social support make more use of

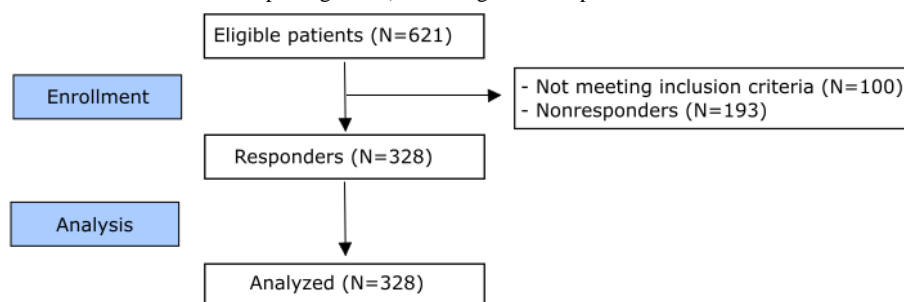
web-based digital communities). We also expect that forum users will differ from nonusers in their treatment status and that their symptom burden may be higher, whereas their global health scale may be lower. Overall, we expect patients with worse outcomes to be web-based more often to ask for and receive advice than their peers with better health outcomes.

Methods

Study Design and Participants

A cross-sectional study was conducted among Dutch patients with GIST aged ≥ 18 years at diagnosis, diagnosed between January 1, 2008, and December 31, 2018, in 5 GIST reference centers. Patients were selected from the Netherlands Cancer Registry (NCR), a population-based registry, which is maintained by the Netherlands Comprehensive Cancer Organization (in Dutch: Integraal Kankercentrum Nederland or IKNL) and collects patient and tumor characteristics on all newly diagnosed patients with cancer in the Netherlands. Exclusion criteria were cognitive impairment or being too ill at the time of the study according to advice from a (former) treating specialist. Eligible patients were invited by their (former) treating physician by a letter explaining the study. Upon consent of the patient, including permission to link the survey data with NCR data, patients could complete the survey on the web or on paper upon request. Refer to [Figure 1](#) for a diagram of the response rate. Survey administration was done within the Patient-Reported Outcomes Following Initial Treatment and Long-term Evaluation of Survivorship registry [36], a data management system set up for the study of the physical and psychosocial impact of cancer and its treatment. Patient-Reported Outcomes Following Initial Treatment and Long-term Evaluation of Survivorship registry contains a large web-based component and is linked directly to clinical data from the NCR. Data were collected from September 2020 to June 2021.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of response rate.



Ethics Approval

Ethics approval for the cross-sectional study was provided by the medical ethical committee of the Radboud University Medical Centre (2019-5888). According to the Dutch law, approval of one ethical committee for questionnaire research is valid for all participating centers. Patients gave informed consent, including permission to link the survey data with NCR data, before completing the survey.

Survey

Participants completed questions regarding their participation in social media and web-based patient communities. These questions were developed by the authors. Respondents were asked whether and how patients use digital platforms to have contact with other patients. Possible answers (translated to English) were “Generic social media (like Facebook or Twitter),” “General forum or discussion group,” “Specific online patient forum,” “Other, namely...,” or “I do not use digital communication.” Patients were provided with the following

definition for a digital medium (translated to English): a web-based communication channel where information and messages are exchanged between participants. Patients were allowed to give multiple answers.

Respondents having contact with other patients on the web were subsequently asked about their motivations for going on the web and about their frequency of posting messages. Both questions were adapted from a Dutch survey designed by van Uden-Kraan et al [34] in collaboration with medical experts and patient representatives. Survey respondents were allowed to provide multiple reasons for engaging with web-based forums as well as additional reasons in an open text field. Respondents who did not have contact with other patients on specific web-based patient forums were asked for their reasons for not doing so. Survey respondents were allowed to provide multiple reasons for abstaining from forum use as well as additional reasons in an open text field.

Demographic variables (ie, age, sex, and socioeconomic status) as well as clinical variables (ie, tumor type, tumor stage, time since diagnosis, whether surgery was performed, and whether targeted therapy was part of treatment) of survey respondents were collected from the NCR. Survey respondents were additionally asked about their marital status, their current treatment phase, whether they presently use medication, their most recent medication (if any), and the presence of the 14 comorbid conditions measured in the Charlson comorbidity index [37] (heart condition, stroke, high blood pressure, asthma, chronic bronchitis, chronic obstructive pulmonary disease, diabetes, stomach ulcer, liver disorder, blood disorder, thyroid disease, depression, arthritis, and back pain). Patients were allowed to fill in "Other" for the most recent targeted medication received for treating GIST. This option was intended for new or experimental TKIs, but because patients frequently used this option for other types of medication such as antacids, it was removed for post hoc analysis.

The options patients can choose for self-reported treatment phase are defined as follows: "Cured and not monitored" ("I am cured and no longer need to be monitored") refers to patients who are considered cured after surgery with or without adjuvant imatinib, "On curative treatment" ("I am being treated and can still be cured") refers to patients who are undergoing adjuvant imatinib treatment, "Follow-up after treatment" ("I am not being treated but am only being monitored") refers to patients who are being monitored after surgery with or without adjuvant imatinib and are not undergoing treatment at this time, "On palliative treatment" ("I am being treated but cannot be cured") refers to patients undergoing palliative treatment with thyroid kinase inhibitors, and "Best supportive care" ("I cannot be cured but am not being treated") refers to patients who are palliative but are not receiving TKIs.

To measure overall health-related quality of life (QoL), social functioning, and symptom burden, participants completed the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire C30 version 3.0 [38,39]. Health-related QoL was measured with 2 items on a scale from 1 to 7 (from "very poor" to "excellent"). Social functioning was measured with 2 items on a scale from 1 to 4 (1, "not at all"; 2,

"a little"; 3, "quite a bit"; and 4, "very much"). Eight symptom-specific items were evaluated on the same scale (ie, dyspnea, pain, insomnia, appetite loss, nausea, constipation, diarrhea, and fatigue). Each symptom was measured with 1-3 items. The scores for a single symptom from multiple items were averaged. Symptom burden was measured by averaging the 8 symptom scales. For 17 respondents, symptom burden was not assessed, as there were missing data for at least one symptom. All scales were linearly transformed to a "0-100" scale in line with the standard scoring manual [40]. A higher score on the functional scales and global QoL means better functioning and QoL, whereas a higher score on the symptom scales means more complaints.

Any questions that were not previously validated were pretested with patients and changed according to their feedback (cognitive debriefing). The questionnaires cannot be shared because of copyright restrictions.

Data Analysis

The reasons for abstaining and engaging with web-based patient-to-patient communication were analyzed manually by the first author. In total, 15.8% (52/328) of the cases contain missing data. As none of these cases are forum users, the data are not missing completely at random. As we do not observe any other patterns in the missing data that cannot be explained by the variables on which we have full information, the data are missing at random. As the missing data occur in multiple variables, we used Multivariate Imputation by Chained Equations [41,42] to impute these values, which is valid under the assumption of missing at random. We generated 20 imputed data sets that include all survey respondents (N=328).

We aimed to analyze whether there were statistically significant differences in demographic and clinical characteristics as well as the QoL measures between forum users and nonusers. For each imputed data set, a multiple logistic regression analysis was performed with forum use as the dependent variable and demographic and clinical factors as the independent variables (refer to Surveys). The effects of one variable on forum use are thus conditional on the other variables in the model. We report the average and SD of the 20 imputed data sets, as this provides a more reliable result than a single run. We use the mean as the average for all variables except the *P* value where we use the median [43].

For this analysis, the number of variables was restricted by the small size of the user population. We checked for multicollinearity using Variance Inflation Factor tests. If the Variance Inflation Factor value was >3, we removed one of the collinear explanatory variables. In total, we removed 2 variables accordingly: the most recent medication and whether the patient is on systemic treatment currently ("On systemic treatment currently"). Note that whether the patient received targeted therapy at some point in time ("Targeted therapy") is included. Moreover, 2 categories of self-reported treatment phase, namely, on palliative treatment and on best supportive care, needed to be merged into one palliative category, as only one patient was receiving best supportive care.

Benjamini-Hochberg correction [44] was used to adjust for multiple testing (controlling the false discovery rate or type I errors at 0.05). Analyses were conducted using *statsmodels* (version 0.12.2) and *scipy* (version 1.4.1) in Python 3.7. Graphs were created with *plotly* (version 5.3.1) in Python 3.7.

Results

Participants

In total, 328 patients with GIST responded to the survey (response rate 64%). The median age of the participants was 67 (range 28-91) years, and 53.8% (174/328) of the participants were male (Table 1). On average, they had been diagnosed with

GIST for 5 years, ranging from 1 to 12 years since diagnosis. In total, 49.3% (162/328) of the participants are in follow-up after treatment with curative intent, 18.5% (61/328) were considered cured and are not in follow-up, and 30.4% (100/328) receive systematic treatment either with curative (n=34) or palliative intent (n=66). Moreover, 1 patient received the best supportive care only.

Overall, 9 patients did not answer the question about forum use, and their forum use is thus unknown. Consequently, the sum of the reported numbers under forum use (Yes and No) does not equal the number reported for all respondents. The percentages were calculated based on the counts per category, that is, 54.9% (150/273) of nonusers are male.

Table 1. Demographic characteristics of survey respondents.

Demographic characteristic	All (N=328)	Forum user ^a	
		No (n=273)	Yes (n=46)
Age (years), median (range)	67 (28-91)	68 (28-91)	65 (47-83)
Sex, n (%)			
Male	174 (53)	150 (54.9)	21 (45.7)
Female	154 (47)	123 (45.0)	25 (54.3)
Socioeconomic status, n (%)			
Low (1-3)	90 (27.4)	74 (27.1)	13 (28.3)
Intermediate (4-7)	132 (40.2)	113 (41.4)	16 (34.8)
High (8-10)	106 (32.3)	86 (31.5)	17 (37)
Marital status, n (%)			
Married or living together	246 (7)	202 (74)	38 (82.6)
Single	79 (24.1)	68 (24.9)	8 (17.4)
Missing	4 (1.2)	3 (1.1)	0 (0)
Time since diagnosis (years), median (range)	5 (1-12)	5 (1-12)	5 (2-11)
Tumor stage, n (%)			
I	121 (36.9)	109 (39.9)	8 (17.4)
II	61 (18.6)	51 (18.7)	10 (21.7)
III	66 (20.1)	53 (19.4)	10 (21.7)
IV	55 (16.8)	38 (13.9)	16 (34.8)
Missing	25 (7.6)	22 (8.1)	2 (4.3)
Surgery, n (%)			
Yes	287 (87.5)	244 (89.4)	36 (78.3)
No	41 (12.5)	29 (10.6)	10 (21.7)
Targeted therapy, n (%)			
Yes	214 (65.2)	170 (62.3)	39 (84.8)
No	114 (34.8)	103 (37.7)	7 (15.2)
Self-reported current treatment status, n (%)			
Cured and not monitored	61 (18.6)	56 (20.5)	2 (4.3)
On curative treatment	34 (10.4)	31 (11.4)	3 (6.5)
Follow-up after treatment	162 (49.4)	139 (50.9)	19 (41.3)
On palliative treatment	66 (20.1)	42 (15.4)	22 (47.8)
Best supportive care	1 (0.3)	1 (0.4)	0 (0)
Missing	4 (1.2)	4 (1.5)	0 (0)
On systemic treatment currently, n (%)			
Yes	208 (63.4) ^b	181 (66.3)	25 (54.3)
No	108 (32.9)	83 (30.4)	21 (45.7)
Missing	12 (3.7)	9 (3.3)	0 (0)
Most recent medication, n (%)			
Imatinib	178 (54.3)	140 (51.3)	31 (67.4)
Sunitinib	9 (2.7)	7 (2.6)	2 (4.3)
Regorafenib	6 (1.8)	4 (1.5)	2 (4.3)

Demographic characteristic	All (N=328)	Forum user ^a	
		No (n=273)	Yes (n=46)
Other	15 (4.6)	8 (2.9)	4 (8.7)
No therapy	114 (34.8)	103 (37.7)	7 (15.2)
Missing	14 (4.3)	11 (4)	0 (0)
Number of comorbid conditions, n (%)			
0	109 (33.2)	92 (33.7)	14 (30.4)
1	71 (21.6)	59 (21.6)	10 (21.7)
2+	146 (44.5)	120 (44.0)	22 (47.8)
Missing	2 (0.6)	2 (0.7)	0 (0)
Global health scale (0-100), mean (SD)	78.6 (18.1)	79.0 (17.7)	76.1 (20.1)
Symptom burden (0-100), mean (SD)	12.1 (12.8)	11.4 (12.6)	15.6 (13)
Social functioning (0-100), mean (SD)	92.4 (18.9)	93.8 (17.1)	84.8 (26)

^aNine participants did not answer this question.

^bIt appears that patients who are currently being monitored may have misunderstood this question, inflating the number of patients who are currently on targeted medication for gastrointestinal stromal tumor.

Social Media Use

Among the participating Dutch patients with GIST, 81% do not have contact with other patients via any social media platform (Table 2). We distinguished between specific social media, such as patient forums, and general social media, such as Twitter or Facebook. Although it is possible for patient communities to exist as groups on general social media platforms (in fact: the biggest GIST forum is a Facebook group), general social media

refers to communication with peers outside of GIST-specific communities on these general social media platforms. Of the patients who communicate with peers via social media, the majority (46/59, 78%) make use of specific web-based patient forums focused on GIST. Only 6 respondents make use of general social media platforms to communicate with other patients with GIST, and only 7 respondents use more general cancer-related forums or discussion groups for this purpose.

Table 2. Descriptive statistics for the use of social media to have contact with other patients (N=328).

Which of the following digital media do you use to have contact with other patients? ^a (Indicate all that apply)	Values, n (%)
General social media (such as Facebook or Twitter)	6 (1.8)
General cancer-related forum or discussion group	7 (2.1)
GIST ^b -specific web-based patient forum	46 (14)
Any social medium	59 (18)
None or via another medium than social media	265 (80.8)
Missing	4 (1.2)

^aRespondents can give multiple answers to this question.

^bGIST: gastrointestinal stromal tumor.

Reasons for Abstaining From Web-Based Patient-to-Patient Communication

Table 3 presents the reasons the 265 nonusers report for not using any digital medium to communicate with fellow patients. Patients were allowed to report multiple reasons. A total of 20 patients did not fill in the question. The most common reason

reported for abstaining from using a digital medium to communicate with peers was that they felt no need to do so (78/265, 29.4%), followed by finding it too confronting (33/265, 12.5%) and not knowing where to find web-based communities (30/265, 11.3%). Only 8 participants reported not using social media to communicate with other patients because they lacked the skills or access to do so.

Table 3. The reasons nonusers report for not using social media to communicate with other patients (N=265).

Self-reported reason ^a	Values, n (%)
Feel no need to communicate (digitally) with other patients	78 (29.4)
I find it too confronting or burdensome	33 (12.5)
I do not know where to find online communities	30 (11.3)
There are too many negative comments	26 (9.8)
I do not have the time	23 (8.7)
The information shared is useless or less valuable	20 (7.5)
I communicate with enough patients personally or via another nondigital medium	18 (6.8)
I do not use social media, lack a computer or digital skills, or do not like obtaining information digitally	8 (3)
I obtain sufficient information via my medical specialist or by searching online	7 (2.6)
I no longer have symptoms or do not like to consider myself a patient	5 (1.9)
I have privacy concerns	3 (1.1)
They do not exist in my language	2 (0.8)
No particular reason	1 (0.4)
Missing	20 (7.5)

^aMultiple answers were possible.

Reasons for Engaging With Patient Forums

Survey respondents most frequently used patient forums to communicate with other patients. The number of survey responders that made use of other web-based platforms was too small to analyze how they compare with nonusers. Thus, we will focus on analyzing the sample bias of GIST-specific patient forums. Hereafter, when we refer to “forum users,” we mean users of GIST-specific patient forums.

Table 4. The reasons users report for visiting the patient forum (N=45).

Self-reported reason ^a	Value, n (%)
When I have a question about my illness	18 (40)
When I have heard new information about my illness	18 (40)
When I am curious about how other members are doing	16 (36)
When I observe new symptoms	14 (31)
When I have a lot of symptoms	6 (13)
When I feel insecure	5 (11)
Before making a medical choice	4 (9)
Because I enjoy the company	4 (9)
Because other members expect me to be there	2 (4)
When I feel lonely	1 (2)
It is part of my daily routine	1 (2)
I never use the forum anymore	1 (2)

^aMultiple answers were possible.

Characteristics of the Patient Forum Users

In total, 85.8% (273/328) of the participants were not making use of specialized GIST patient forums (Table 1). The difference in model fit between the multiple logistic regression model and

Table 4 presents the reasons users reported for engaging with a disease-specific patient forum. The most prevalent reasons were having a question about their illness (18/45, 40%), having heard new information about their illness (18/45, 40%), and being curious about how the other members are doing (16/45, 36%). Another prevalent trigger was experiencing new symptoms (14/45, 31%).

the null model was found to be statistically significant in all 20 imputed data sets (likelihood ratio [LR]=47.0, SD 1.48, $df=20$; $P<.001$). LR tests between the full model and the full model without the variable were used to test significance of individual variables.

Table 5 reports the average results of 20 runs of multiple logistic regression models of which factors influence forum use. Our analysis shows that self-reported treatment status differs significantly between forum users and nonusers for each run (LR=10.6; $P=.001$). The odds of being on a patient forum were 2.8 times as high for a patient who is being monitored compared with a patient who is considered cured. The odds of being on a patient forum were 1.9 times as high for patients who were on curative (adjuvant) treatment and 10 times as high for patients who were in the palliative phase compared with patients who were considered cured.

We did not find significant differences between forum users and nonusers for other disease-related characteristics when they were adjusted for covariates. We also did not find significant differences in key demographic variables such as age, sex, socioeconomic status, and marital status. However, we did find a significant difference in level of social functioning in 7 of 20 runs (LR=6.8; $P=.008$). Forum users on average reported a lower level of social functioning than nonusers (84.8 vs 93.8 out of 100). These scores were normalized according to the scoring manual [40]. Converting the normalized values back to the mean raw score gives a 1.19 for forum users and a 1.46 for nonusers, where 1 translates to the highest possible value for self-reported social functioning on the survey items.

Table 5. Average results (with SD) of a logistic regression of demographic and clinical characteristics of patient forum users and nonusers using Multivariate Imputation by Chained Equations with 20 runs.

	Coefficient, mean (SD)	SE, mean (SD)	df ^a	LR, mean (SD)	P value, median (SD)	Odds ratio		
						5% (SD)	Mean (SD)	95% (SD)
Intercept	-2.80 (0.54)	2.08 (0.03)	N/A ^b	N/A	N/A	N/A	N/A	N/A
Age	-0.02 (0.004)	0.02 (0.0002)	1	1.32 (0.54)	.26 (.10)	0.95 (0.004)	0.98 (0.004)	1.02 (0.004)
Sex	0.62 (0.04)	0.37 (0.004)	1	2.86 (0.35)	.09 (.02)	0.90 (0.03)	1.86 (0.07)	3.86 (0.16)
Socioeconomic status	N/A	N/A	2	1.37 (0.49)	.25 (.08)	N/A	N/A	N/A
Low (1-3) ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Intermediate (4-7)	-0.39 (0.10)	0.44 (0.006)	N/A	N/A	N/A	0.20 (0.03)	0.68 (0.07)	1.62 (0.16)
High (8-10)	0.05 (0.10)	0.44 (0.005)	N/A	N/A	N/A	0.45 (0.04)	1.06 (0.10)	2.50 (0.26)
Marital status	-0.32 (0.09)	0.47 (0.006)	1	0.52 (0.25)	.47 (.11)	0.29 (0.04)	0.73 (0.06)	1.82 (0.06)
Time since diagnosis	0.02 (0.02)	0.07 (0.001)	1	0.12 (0.16)	.85 (.15)	0.88 (0.02)	1.02 (0.02)	1.17 (0.02)
Tumor type	0.57 (0.06)	0.38 (0.003)	1	2.29 (0.52)	.13 (.04)	0.84 (0.05)	1.77 (0.11)	3.70 (0.24)
Tumor stage	N/A	N/A	3	2.60 (0.92)	.12 (.07)	N/A	N/A	N/A
I ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
II	0.51 (0.13)	0.55 (0.009)	N/A	N/A	N/A	0.57 (0.07)	1.67 (0.21)	4.89 (0.63)
III	0.21 (0.21)	0.63 (0.01)	N/A	N/A	N/A	0.37 (0.09)	1.27 (0.29)	4.31 (0.94)
IV	0.86 (0.17)	0.66 (0.01)	N/A	N/A	N/A	0.66 (0.12)	2.41 (0.43)	8.83 (1.61)
Surgery	0.04 (0.12)	0.57 (0.01)	1	0.05 (0.10)	.89 (.10)	0.34 (0.04)	1.05 (0.12)	3.24 (0.42)
Targeted therapy	0.12 (0.10)	0.57 (0.01)	1	0.07 (0.08)	.83 (.10)	0.37 (0.03)	1.13 (0.11)	3.49 (0.38)
Self-reported current treatment status	N/A	N/A	3	10.67 (1.10)	<i>.001*</i> (<.001)	N/A	N/A	N/A
Cured and not monitored ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
On curative treatment	0.59 (0.26)	1.07 (0.05)	N/A	N/A	N/A	0.23 (0.04)	1.86 (0.45)	15.56 (4.65)
Follow-up after treatment	1.03 (0.26)	0.87 (0.06)	N/A	N/A	N/A	0.52 (0.08)	2.88 (0.69)	16.18 (5.10)
Palliative	2.29 (0.23)	0.97 (0.06)	N/A	N/A	N/A	1.50 (0.23)	10.11 (2.21)	68.68 (19.84)
Number of comorbid conditions	N/A	N/A	2	0.42 (0.26)	.53 (.14)	N/A	N/A	N/A
0 ^c	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1	0.28 (0.11)	0.50 (0.007)	N/A	N/A	N/A	0.50 (0.06)	1.33 (0.14)	3.51 (0.36)
2+	0.21 (0.08)	0.45 (0.005)	N/A	N/A	N/A	0.51 (0.04)	1.23 (0.09)	2.99 (0.24)
Global health scale or QoL ^d	0.03 (0.002)	0.01 (0.0001)	1	4.38 (0.69)	.04 (.02)	1.00 (0.002)	1.04 (0.002)	1.06 (0.002)
Symptom burden	-0.0003 (0.005)	0.02 (0.0004)	1	0.09 (0.10)	.83 (.11)	0.96 (0.006)	1.00 (0.005)	1.04 (0.005)
Social functioning	-0.03 (0.002)	0.01 (0.0002)	1	6.87 (0.90)	<i>.008 (.005)</i>	0.96 (0.002)	0.98 (0.001)	0.99 (0.002)

^aThere are separate *df* values for each variable depending on the number of categories in that variable.

^bN/A: not applicable.

^cThese categories were taken as the reference categories for calculating the influence of different categories of the variable on forum use

^dQoL: quality of life.

*Italicized values indicate statistically significant values.

Discussion

Principal Findings

A survey was conducted among 328 patients with GIST in the Netherlands. Our results show that most survey respondents do not have contact with other patients via social media. They indicate a large heterogeneity of reasons of why they abstain from doing so, with the most prevalent being they feel no need, find it too confronting, or do not know where to find such web-based communities. Of the minority who do use social media for this purpose, most use disease-specific patient forums. The most prevalent reasons for accessing a patient forum are (1) having a question about their illness, (2) having heard new information, (3) experiencing new symptoms, or (4) wondering how other patients are doing. Patient forum users differ significantly in (self-reported) treatment phase from nonusers. Patients in the palliative phase are 10 times more likely to be forum users than patients who are cured. Patients who are monitored approximately 3 times and patients undergoing curative treatment approximately 2 times are more likely to be users than cured patients. For 7 of 20 data imputations, forum users also have a significantly lower level of social functioning.

Comparison With Existing Literature

In contrast to the general population of social media users, patient forum users do not appear to differ in age, sex, and socioeconomic status from nonusers. On the one hand, this may be an effect of the increasingly more widespread adoption of social media. This idea is supported by the small number of patients that indicate they lack the skills or access to be on social media (8/265, 3.3%). On the other hand, it is also possible that there is less demographic bias on patient forums than in general social media. This may be related to the widely different goals that users have with their participation. Although a feeling of community and social support may overlap, patients report motivations such as questions around their illness and the experience of new symptoms that normal social media users are unlikely to share.

Prior work [29] on forum use among patients with breast cancer did not find significant differences between forum users and nonusers in terms of clinical characteristics, that is, stage of cancer and QoL. We similarly did not find any significant differences for these characteristics, although we did find significant differences for clinical characteristics that prior work did not investigate, that is, treatment phase. Prior work also found that among patients with breast cancer, nonusers and passive users had greater offline social support than posters. Their results supported the social compensation model [35], that is, those who have less real-life (offline) social support use and engage on the web with digital communities. The lower offline support of forum users compared with nonusers in our data also supports this theory. However, passive users appear to have a lower offline support than active users among patients with GIST. This would support the competing theory: the social engagement model [45], that is, those that have more social resources will use and benefit from web-based social communities more. Consequently, our data offer support for the social compensation model for those who use a forum (ie,

those with less real-life support are more likely to be on a forum) and social engagement theory for those who actually actively engage with the forum community (ie, users with sufficient social resources will be active and benefit more). Demographic differences in terms of age, marital status (ie, living alone or not), and disease duration between passive and active users that were found in previous work were not evident from our data.

Limitations

First and foremost, we only studied a specific patient population in a single country, and thus, further research is needed to elucidate to what extent our results are generalizable. Patients in other countries may have lower digital access or skills or may not wish to use social media for patient-to-patient communication for other reasons (eg, other privacy laws or country-specific customs).

Our choice of patients with GIST as a target population may also impact to which disorders our results generalize to. Patients with GIST have a median age of mid-60 years [46], meaning that it is on average an older population than the general population that is often studied for social media use. Our results may consequently also generalize better to conditions that are prevalent in an older population. GIST is also characterized by a long palliative phase in which patients receive treatment. Thus, our results may also generalize better to conditions that similarly have a long treatment duration (eg, metastasized breast cancer). As GIST is a rare type of cancer, our results may also generalize better to rare conditions than common conditions. Further research into other patient populations should be able to provide more insight into the differences in forum use between rare conditions and common conditions. The fact that GIST is a rare condition makes it an interesting first case. Patient-generated health data from social media are particularly promising for rare conditions because of their dispersed patient communities and the scarcity of research [47].

A second limitation of this study is the small sample size. Among the 328 respondents, only 46 (14.1%) indicate that they make use of patient forums. Nonetheless, given the low incidence of GIST at 12.7 per million [48], this is a substantial number of participants. A third limitation is the sample bias of the survey itself. There may be 2 underlying factors, namely, selection bias and responder bias. Patients who were too ill or had cognitive impairment were excluded, leading to selection bias. A nonresponder analysis was conducted using the database of the NCR to assess the extent of the responder bias. After correcting for multiple testing, no significant differences were found in terms of age, sex, socioeconomic status, time since diagnosis, tumor stage, and primary treatment between responders and nonresponders. Moreover, it was possible to fill in the survey on paper, which prevents the exclusion of less digitally adept patients on these grounds.

Future Work and Recommendations

On the basis of this work, a number of recommendations can be made. First, of the possible digital resources that can be used to source complementary real-world evidence, for instance, for pharmacovigilance, patient forums should be preferred over other social media. Our results reveal that patients with GIST

strongly prefer disease-specific patient forums over general social media for communicating with fellow patients. However, most research in this field currently focuses on general social media such as Twitter [4,5]. Our results are in line with previous work that estimates ADR reports to be more prevalent in patient forums than on Twitter [49].

Although we find that there is sample bias in patient forum users and, thus, the sample is not wholly representative for the patient population, sample bias is also a concern for other sources of patient reports. Understanding which patients are overrepresented and underrepresented on web-based forums is the first step to using web-based patient reports as a complementary resource, for instance, for pharmacovigilance, which is seen as a realistic first use case. For pharmacovigilance specifically, it is not of great concern that patients who are considered cured and not undergoing treatment currently are underrepresented. Future work into comparing the sample bias of clinical trials with that of web-based patient forums would be beneficial to further explore its complementary value in detail. It would also be valuable to gain more insight into the different types of forum users.

Second, it may be beneficial to create awareness among medical professionals that patients are more likely to search for information in web-based patient communities when they have questions, have been given new information, or have new symptoms. Medical professionals could try to aid patients in their information need by pointing them toward such resources in these cases. This may also take away the barrier mentioned

by patients that they do not know where to find such web-based communities.

Third, future work into the sample bias of patient forums for other patient populations is necessary, as this study was limited to a single population in a single country. Nonetheless, our work is a stepping stone toward dissuading the concerns that researchers have expressed regarding the sample bias of social media [6,16-22] by unraveling on which characteristics users differ significantly from the overall patient population. Future work could also investigate how compensatory measures can be implemented to statistically correct for sample bias. As these factors may not be known for the participants of a forum, it would also be worthwhile to consider to what extent correcting for sample bias is possible without this information.

Conclusions

In this study, we investigated how representative participants in patient forums are for the general patient population by conducting a survey among patients with GIST in the Netherlands. We found statistically significant differences in terms of treatment phase and offline social support between forum users and nonusers. The consequent overrepresentation and underrepresentation of certain types of patients should be considered when sourcing patient forums for patient-generated health data. As our study was limited to a single patient population, a further investigation of the sample and activity bias in other web-based patient populations is warranted. Sample bias is inherent to any information source, and only through awareness of these biases can these resources be used as a source for complementary real-world evidence in the future.

Acknowledgments

This survey was partly funded by a research grant from Novartis (grant 006.18). AD is supported by a research grant from the Stichting Internet Domeinregistratie Nederland fund (grant 174054). DdH is supported by a research grant from the European Organisation for Research and Treatment of Cancer Quality of Life Group (grant 003-2018). The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The authors would like to thank Carla Fokkema-Vlooswijk and Esther Derksen-Peters for their support with the recruitment and data collection through Patient-Reported Outcomes Following Initial Treatment and Long-term Evaluation of Survivorship in the survey study.

Data Availability

The data are not publicly available.

Authors' Contributions

AD, DdH, SV, and OH contributed to conceptualization. AD, DdH, SV, OH, WTAvdG, and GvO contributed to methodology. AD, DdH, ID, OH, WTAvdG, AO, AKLR, NS, WvL, and HG contributed to formal analysis and investigation. AD wrote the original paper. DdH, SV, ID, OH, WTAvdG, AO, AKLR, NS, WvL, GvO, HG, and WK reviewed and edited the article.

Conflicts of Interest

None declared.

References

1. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. *Qual Health Res* 2008 Mar 01;18(3):405-417. [doi: [10.1177/1049732307313429](https://doi.org/10.1177/1049732307313429)] [Medline: [18235163](https://pubmed.ncbi.nlm.nih.gov/18235163/)]

2. Caster O, Dietrich J, Kürzinger ML, Lerch M, Maskell S, Norén GN, et al. Assessment of the utility of social media for broad-ranging statistical signal detection in pharmacovigilance: results from the WEB-RADR project. *Drug Saf* 2018 Dec;41(12):1355-1369 [FREE Full text] [doi: [10.1007/s40264-018-0699-2](https://doi.org/10.1007/s40264-018-0699-2)] [Medline: [30043385](https://pubmed.ncbi.nlm.nih.gov/30043385/)]
3. Golder S, Smith K, O'Connor K, Gross R, Hennessy S, Gonzalez-Hernandez G. A comparative view of reported adverse effects of statins in social media, regulatory data, drug information databases and systematic reviews. *Drug Saf* 2021 Feb 01;44(2):167-179 [FREE Full text] [doi: [10.1007/s40264-020-00998-1](https://doi.org/10.1007/s40264-020-00998-1)] [Medline: [33001380](https://pubmed.ncbi.nlm.nih.gov/33001380/)]
4. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: a review. *J Biomed Inform* 2015 Apr;54:202-212 [FREE Full text] [doi: [10.1016/j.jbi.2015.02.004](https://doi.org/10.1016/j.jbi.2015.02.004)] [Medline: [25720841](https://pubmed.ncbi.nlm.nih.gov/25720841/)]
5. Lardon J, Abdellaoui R, Bellet F, Asfari H, Souvignet J, Texier N, et al. Adverse drug reaction identification and extraction in social media: a scoping review. *J Med Internet Res* 2015 Jul 10;17(7):e171 [FREE Full text] [doi: [10.2196/jmir.4304](https://doi.org/10.2196/jmir.4304)] [Medline: [26163365](https://pubmed.ncbi.nlm.nih.gov/26163365/)]
6. Audeh B, Bellet F, Beyens M, Lillo-Le Louët A, Bousquet C. Use of social media for pharmacovigilance activities: key findings and recommendations from the Vigi4Med project. *Drug Saf* 2020 Sep 16;43(9):835-851. [doi: [10.1007/s40264-020-00951-2](https://doi.org/10.1007/s40264-020-00951-2)] [Medline: [32557179](https://pubmed.ncbi.nlm.nih.gov/32557179/)]
7. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. *Clin Pharmacol Ther* 2012 Jun 02;91(6):1010-1021 [FREE Full text] [doi: [10.1038/clpt.2012.50](https://doi.org/10.1038/clpt.2012.50)] [Medline: [22549283](https://pubmed.ncbi.nlm.nih.gov/22549283/)]
8. Brosch S, de Ferran A, Newbould V, Farkas D, Lengsavath M, Tregunno P. Establishing a framework for the use of social media in pharmacovigilance in Europe. *Drug Saf* 2019 Aug 30;42(8):921-930 [FREE Full text] [doi: [10.1007/s40264-019-00811-8](https://doi.org/10.1007/s40264-019-00811-8)] [Medline: [30929141](https://pubmed.ncbi.nlm.nih.gov/30929141/)]
9. Davison KP, Pennebaker JW, Dickerson SS. Who talks? The social psychology of illness support groups. *Am Psychol* 2000;55(2):205-217. [doi: [10.1037/0003-066x.55.2.205](https://doi.org/10.1037/0003-066x.55.2.205)]
10. Shenoy P, Harugeri A. Elderly patients' participation in clinical trials. *Perspect Clin Res* 2015;6(4):184-189 [FREE Full text] [doi: [10.4103/2229-3485.167099](https://doi.org/10.4103/2229-3485.167099)] [Medline: [26623388](https://pubmed.ncbi.nlm.nih.gov/26623388/)]
11. Stricker BH, Psaty BM. Detection, verification, and quantification of adverse drug reactions. *BMJ* 2004 Jul 03;329(7456):44-47 [FREE Full text] [doi: [10.1136/bmj.329.7456.44](https://doi.org/10.1136/bmj.329.7456.44)] [Medline: [15231627](https://pubmed.ncbi.nlm.nih.gov/15231627/)]
12. van Hunsel F, de Waal S, Härmark L. The contribution of direct patient reported ADRs to drug safety signals in the Netherlands from 2010 to 2015. *Pharmacoepidemiol Drug Saf* 2017 Aug 19;26(8):977-983. [doi: [10.1002/pds.4236](https://doi.org/10.1002/pds.4236)] [Medline: [28524293](https://pubmed.ncbi.nlm.nih.gov/28524293/)]
13. Hazell L, Shakir SA. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf* 2006;29(5):385-396. [doi: [10.2165/00002018-200629050-00003](https://doi.org/10.2165/00002018-200629050-00003)] [Medline: [16689555](https://pubmed.ncbi.nlm.nih.gov/16689555/)]
14. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. *Drug Saf* 2009;32(1):19-31. [doi: [10.2165/00002018-200932010-00002](https://doi.org/10.2165/00002018-200932010-00002)] [Medline: [19132802](https://pubmed.ncbi.nlm.nih.gov/19132802/)]
15. Mao JJ, Chung A, Benton A, Hill S, Ungar L, Leonard CE, et al. Online discussion of drug side effects and discontinuation among breast cancer survivors. *Pharmacoepidemiol Drug Saf* 2013 Mar;22(3):256-262 [FREE Full text] [doi: [10.1002/pds.3365](https://doi.org/10.1002/pds.3365)] [Medline: [23322591](https://pubmed.ncbi.nlm.nih.gov/23322591/)]
16. Tufekci Z. Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media* 2014:505-514. [doi: doi.org/10.1609/icwsm.v8i1.14517]
17. Stieglitz S, Mirbabaie M, Ross B, Neuberger C. Social media analytics – Challenges in topic discovery, data collection, and data preparation. *Int J Inform Manag* 2018 Apr;39:156-168. [doi: [10.1016/j.ijinfomgt.2017.12.002](https://doi.org/10.1016/j.ijinfomgt.2017.12.002)]
18. Sloane R, Osanlou O, Lewis D, Bollegala D, Maskell S, Pirmohamed M. Social media and pharmacovigilance: a review of the opportunities and challenges. *Br J Clin Pharmacol* 2015 Oct 02;80(4):910-920 [FREE Full text] [doi: [10.1111/bcp.12717](https://doi.org/10.1111/bcp.12717)] [Medline: [26147850](https://pubmed.ncbi.nlm.nih.gov/26147850/)]
19. Cesare N, Grant C, Nsoesie E. Understanding demographic bias and representation in social media health data. In: *Proceedings of the Companion Publication of the 10th ACM Conference on Web Science*. 2019 Presented at: WebSci '19: 11th ACM Conference on Web Science; Jun 30- Jul 3, 2019; Boston Massachusetts USA. [doi: [10.1145/3328413.3328415](https://doi.org/10.1145/3328413.3328415)]
20. Bate A, Reynolds RF, Caubel P. The hope, hype and reality of Big Data for pharmacovigilance. *Ther Adv Drug Saf* 2018 Jan 31;9(1):5-11 [FREE Full text] [doi: [10.1177/2042098617736422](https://doi.org/10.1177/2042098617736422)] [Medline: [29318002](https://pubmed.ncbi.nlm.nih.gov/29318002/)]
21. Berrewaerts J, Delbecq L, Orban P, Deseilles M. Patient participation and the use of eHealth tools for pharmacovigilance. *Front Pharmacol* 2016;7:90 [FREE Full text] [doi: [10.3389/fphar.2016.00090](https://doi.org/10.3389/fphar.2016.00090)] [Medline: [27148052](https://pubmed.ncbi.nlm.nih.gov/27148052/)]
22. Bulcock A, Hassan L, Giles S, Sanders C, Nenadic G, Campbell S, et al. Public perspectives of using social media data to improve adverse drug reaction reporting: a mixed-methods study. *Drug Saf* 2021 May 13;44(5):553-564 [FREE Full text] [doi: [10.1007/s40264-021-01042-6](https://doi.org/10.1007/s40264-021-01042-6)] [Medline: [33582973](https://pubmed.ncbi.nlm.nih.gov/33582973/)]
23. Hargittai E. Is bigger always better? Potential biases of big data derived from social network sites. *ANNALS Am Acad Political Soc Sci* 2015 Apr 09;659(1):63-76. [doi: [10.1177/0002716215570866](https://doi.org/10.1177/0002716215570866)]
24. Hargittai E. Whose space? Differences among users and non-users of social network sites. *J Comput Commun* 2007;13(1):276-297.

25. Blank G, Lutz C. Representativeness of social media in great Britain: investigating Facebook, LinkedIn, twitter, Pinterest, Google+, and Instagram. *Am Behav Sci* 2017 Jun 29;61(7):741-756. [doi: [10.1177/0002764217717559](https://doi.org/10.1177/0002764217717559)]
26. Koiranen I, Keipi T, Koivula A, Räsänen P. Changing patterns of social media use? A population-level study of Finland. *Univ Access Inf Soc* 2019 Jun 5;19(3):603-617. [doi: [10.1007/s10209-019-00654-1](https://doi.org/10.1007/s10209-019-00654-1)]
27. Hargittai E. Potential biases in big data: omitted voices on social media. *Soc Sci Comput Rev* 2018 Jul 30;38(1):10-24. [doi: [10.1177/0894439318788322](https://doi.org/10.1177/0894439318788322)]
28. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/> [accessed 2021-06-12]
29. Han JY, Kim J, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. *J Health Commun* 2012;17(3):356-371 [FREE Full text] [doi: [10.1080/10810730.2011.585696](https://doi.org/10.1080/10810730.2011.585696)] [Medline: [22085215](https://pubmed.ncbi.nlm.nih.gov/22085215/)]
30. Han JY, Hou J, Kim E, Gustafson DH. Lurking as an active participation process: a longitudinal investigation of engagement with an online cancer support group. *Health Commun* 2014;29(9):911-923 [FREE Full text] [doi: [10.1080/10410236.2013.816911](https://doi.org/10.1080/10410236.2013.816911)] [Medline: [24345206](https://pubmed.ncbi.nlm.nih.gov/24345206/)]
31. den Hollander D, Dirkson AR, Verberne S, Kraaij W, van Oortmerssen G, Gelderblom H, et al. Symptoms reported by gastrointestinal stromal tumour (GIST) patients on imatinib treatment: combining questionnaire and forum data. *Support Care Cancer* 2022 Jun 02;30(6):5137-5146 [FREE Full text] [doi: [10.1007/s00520-022-06929-3](https://doi.org/10.1007/s00520-022-06929-3)] [Medline: [35233640](https://pubmed.ncbi.nlm.nih.gov/35233640/)]
32. Dirkson A, Verberne S, Kraaij W, van Oortmerssen G, Gelderblom H. Automated gathering of real-world data from online patient forums can complement pharmacovigilance for rare cancers. *Sci Rep* 2022 Jun 20;12(1):10317 [FREE Full text] [doi: [10.1038/s41598-022-13894-8](https://doi.org/10.1038/s41598-022-13894-8)] [Medline: [35725736](https://pubmed.ncbi.nlm.nih.gov/35725736/)]
33. Baeza-Yates R. Bias on the web. *Commun ACM* 2018 May 23;61(6):54-61. [doi: [10.1145/3209581](https://doi.org/10.1145/3209581)]
34. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *J Med Internet Res* 2008 Jun 30;10(2):e18 [FREE Full text] [doi: [10.2196/jmir.992](https://doi.org/10.2196/jmir.992)] [Medline: [18653442](https://pubmed.ncbi.nlm.nih.gov/18653442/)]
35. McKenna KY, Bargh JA. Coming out in the age of the internet: identity "demarginalization" through virtual group participation. *J Pers Soc Psychol* 1998 Sep;75(3):681-694. [doi: [10.1037/0022-3514.75.3.681](https://doi.org/10.1037/0022-3514.75.3.681)]
36. van de Poll-Franse LV, Horevoorts N, van Eenbergen M, Denollet J, Roukema JA, Aaronson NK, Profiles Registry Group. The patient reported outcomes following initial treatment and long term evaluation of survivorship registry: scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *Eur J Cancer* 2011 Sep;47(14):2188-2194 [FREE Full text] [doi: [10.1016/j.ejca.2011.04.034](https://doi.org/10.1016/j.ejca.2011.04.034)] [Medline: [21621408](https://pubmed.ncbi.nlm.nih.gov/21621408/)]
37. Charlson ME, Pompei P, Ales KL, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987 Jan;40(5):373-383. [doi: [10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)]
38. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, 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](https://doi.org/10.1093/jnci/85.5.365)] [Medline: [8433390](https://pubmed.ncbi.nlm.nih.gov/8433390/)]
39. Fayers P, Bottomley A. Quality of life research within the EORTC—the EORTC QLQ-C30. *Euro J Cancer* 2002 Mar;38:125-133. [doi: [10.1016/s0959-8049\(01\)00448-8](https://doi.org/10.1016/s0959-8049(01)00448-8)]
40. EORTC QLQ-30 Scoring Manual. EORTC Quality of Life Group. URL: <https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf> [accessed 2021-09-11]
41. Buuren SV, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Soft* 2011;45(3):1-67. [doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03)]
42. Little R, Rubin D. *Statistical Analysis with Missing Data*. New York City, USA: John Wiley & Sons, Ltd; 1987.
43. Eekhout I, van de Wiel MA, Heymans M. Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: power and applicability analysis. *BMC Med Res Methodol* 2017 Aug 22;17(1):129 [FREE Full text] [doi: [10.1186/s12874-017-0404-7](https://doi.org/10.1186/s12874-017-0404-7)] [Medline: [28830466](https://pubmed.ncbi.nlm.nih.gov/28830466/)]
44. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B (Methodological)* 2018 Dec 05;57(1):289-300. [doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)]
45. Kraut R, Kiesler S, Boneva B, Cummings J, Helgeson V, Crawford A. Internet paradox revisited. *J Social Issues* 2002 Jan;58(1):49-74. [doi: [10.1111/1540-4560.00248](https://doi.org/10.1111/1540-4560.00248)]
46. Søreide K, Sandvik OM, Søreide JA, Giljaca V, Jureckova A, Bulusu VR. Global epidemiology of gastrointestinal stromal tumours (GIST): a systematic review of population-based cohort studies. *Cancer Epidemiol* 2016 Feb;40:39-46 [FREE Full text] [doi: [10.1016/j.canep.2015.10.031](https://doi.org/10.1016/j.canep.2015.10.031)] [Medline: [26618334](https://pubmed.ncbi.nlm.nih.gov/26618334/)]
47. Aymé S, Kole A, Groft S. Empowerment of patients: lessons from the rare diseases community. *Lancet* 2008 Jun;371(9629):2048-2051. [doi: [10.1016/S0140-6736\(08\)60875-2](https://doi.org/10.1016/S0140-6736(08)60875-2)]
48. Goettsch WG, Bos SD, Breekveldt-Postma N, Casparie M, Herings RM, Hogendoorn PC. Incidence of gastrointestinal stromal tumours is underestimated: results of a nation-wide study. *Eur J Cancer* 2005 Dec;41(18):2868-2872. [doi: [10.1016/j.ejca.2005.09.009](https://doi.org/10.1016/j.ejca.2005.09.009)] [Medline: [16293410](https://pubmed.ncbi.nlm.nih.gov/16293410/)]

49. Golder S, Norman G, Loke YK. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *Br J Clin Pharmacol* 2015 Oct;80(4):878-888 [FREE Full text] [doi: [10.1111/bcp.12746](https://doi.org/10.1111/bcp.12746)] [Medline: [26271492](https://pubmed.ncbi.nlm.nih.gov/26271492/)]

Abbreviations

ADR: adverse drug response
GIST: gastrointestinal stromal tumor
LR: likelihood ratio
NCR: Netherlands Cancer Registry
QoL: quality of life
TKI: tyrosine kinase inhibitor

Edited by A Mavragani; submitted 24.01.22; peer-reviewed by S Matsuda, P Wicks; comments to author 09.06.22; revised version received 21.09.22; accepted 22.09.22; published 15.12.22.

Please cite as:

Dirkson A, den Hollander D, Verberne S, Desar I, Husson O, van der Graaf WTA, Oosten A, Reyners AKL, Steeghs N, van Loon W, van Oortmerssen G, Gelderblom H, Kraaij W

Sample Bias in Web-Based Patient-Generated Health Data of Dutch Patients With Gastrointestinal Stromal Tumor: Survey Study
JMIR Form Res 2022;6(12):e36755

URL: <https://formative.jmir.org/2022/12/e36755>

doi: [10.2196/36755](https://doi.org/10.2196/36755)

PMID: [36520526](https://pubmed.ncbi.nlm.nih.gov/36520526/)

©Anne Dirkson, Dide den Hollander, Suzan Verberne, Ingrid Desar, Olga Husson, Winette T A van der Graaf, Astrid Oosten, Anna K L Reyners, Neeltje Steeghs, Wouter van Loon, Gerard van Oortmerssen, Hans Gelderblom, Wessel Kraaij. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 15.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>